

George Acquaah

**Principles  
of Plant  
Genetics  
and Breeding**

## **Principles of Plant Genetics and Breeding**

**Dedication**  
To my parents  
Shiloh and Ernestina  
With love and admiration

# **Principles of Plant Genetics and Breeding**

**George Acquaah**



Copyright © 2007 by George Acquaah

BLACKWELL PUBLISHING  
350 Main Street, Malden, MA 02148-5020, USA  
9600 Garsington Road, Oxford OX4 2DQ, UK  
550 Swanston Street, Carlton, Victoria 3053, Australia

The right of George Acquaah to be identified as the Author of this Work  
has been asserted in accordance with the UK Copyright, Designs, and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored  
in a retrieval system, or transmitted, in any form or by any means,  
electronic, mechanical, photocopying, recording or otherwise, except as  
permitted by the UK Copyright, Designs, and Patents Act 1988, without the  
prior permission of the publisher.

First published 2007 by Blackwell Publishing Ltd

1 2007

*Library of Congress Cataloging-in-Publication Data*

Acquaah, George.  
Principles of plant genetics and breeding / George Acquaah.  
p. cm.  
Includes bibliographical references and index.  
ISBN-13: 978-1-4051-3646-4 (hardback : alk. paper)  
ISBN-10: 1-4051-3646-4 (hardback : alk. paper)  
1. Plant breeding. 2. Plant genetics. I. Title.  
SB123.A334 2007  
631.5'233—dc22  
2006004754

A catalogue record for this title is available from the British Library.

Set in 10/12pt Galliard  
by Graphicraft Limited, Hong Kong  
Printed and bound in UK  
by TJ International Ltd

The publisher's policy is to use permanent paper from mills that operate a  
sustainable forestry policy, and which has been manufactured from pulp  
processed using acid-free and elementary chlorine-free practices.  
Furthermore, the publisher ensures that the text paper and cover board  
used have met acceptable environmental accreditation standards.

For further information on  
Blackwell Publishing, visit our website:  
[www.blackwellpublishing.com](http://www.blackwellpublishing.com)

# Contents

Industry highlights boxes, vii  
Industry highlights box authors, ix  
Preface, xi  
Acknowledgments, xiii

## Part I Underlying science and methods of plant breeding, 1

### Section 1 Historical perspectives and importance of plant breeding, 2

1 History and role of plant breeding in society, 3

### Section 2 General biological concepts, 16

2 The art and science of plant breeding, 17  
3 Plant cellular organization and genetic structure: an overview, 35  
4 Plant reproductive systems, 55

### Section 3 Germplasm issues, 74

5 Variation: types, origin, and scale, 75  
6 Plant genetic resources for plant breeding, 87

### Section 4 Genetic analysis in plant breeding, 108

7 Introduction to concepts of population genetics, 109  
8 Introduction to quantitative genetics, 121  
9 Common statistical methods in plant breeding, 146

### Section 5 Tools in plant breeding, 163

10 Sexual hybridization and wide crosses in plant breeding, 164  
11 Tissue culture and the breeding of clonally propagated plants, 181  
12 Mutagenesis in plant breeding, 199  
13 Polyploidy in plant breeding, 214  
14 Biotechnology in plant breeding, 231  
15 Issues in the application of biotechnology in plant breeding, 257

### Section 6 Classic methods of plant breeding, 281

16 Breeding self-pollinated species, 282  
17 Breeding cross-pollinated species, 313  
18 Breeding hybrid cultivars, 334

### Section 7 Selected breeding objectives, 351

19 Breeding for physiological and morphological traits, 352  
20 Breeding for resistance to diseases and insect pests, 367  
21 Breeding for resistance to abiotic stresses, 385  
22 Breeding compositional traits and added value, 404

### Section 8 Cultivar release and commercial seed production, 417

23 Performance evaluation for crop cultivar release, 418  
24 Seed certification and commercial seed multiplication, 435

- 25 International plant breeding efforts, 450
- 26 Emerging concepts in plant breeding, 462

Part II Breeding selected crops, 471

- 27 Breeding wheat, 472
- 28 Breeding corn, 485
- 29 Breeding rice, 498
- 30 Breeding sorghum, 509
- 31 Breeding soybean, 519
- 32 Breeding peanut, 529
- 33 Breeding potato, 537
- 34 Breeding cotton, 546

Glossary, 556

Appendix 1: Internet resources, 561

Appendix 2: Conversion rates, 563

Index, 564

# Industry highlights boxes

## Chapter 1

Normal Ernest Borlaug: the man and his passion  
George Acquaah

## Chapter 2

Introduction and adaptation of new crops  
Jaime Prohens, Adrián Rodríguez-Burrueto, and Fernando Nuez

## Chapter 3

No box

## Chapter 4

Maize × *Tripsacum* hybridization and the transfer of apomixis: historical review  
Bryan Kindiger

## Chapter 5

No box

## Chapter 6

Plant genetic resources for breeding  
K. Hammer, F. Heuser, K. Khoshbakht, and Y. Teklu

## Chapter 7

No box

## Chapter 8

Recurrent selection with soybean  
Joe W. Burton

## Chapter 9

Multivariate analyses procedures: applications in plant breeding, genetics, and agronomy  
A. A. Jaradat

## Chapter 10

The use of the wild potato species, *Solanum etuberosum*, in developing virus- and insect-resistant potato varieties  
Richard Novy

## Chapter 11

Haploids and doubled haploids: their generation and application in plant breeding  
Sergey Chalyk

## Chapter 12

Current apple breeding programs to release apple scab-resistant scion cultivars  
F. Laurens

## Chapter 13

Application of tissue culture for tall wheatgrass improvement  
Kanyand Matand and George Acquaah

## Chapter 14

Bioinformatics for sequence and genomic data  
Hugh B. Nicholas, Jr., David W. Deerfield II, and Alexander J. Ropelewski

## Chapter 15

The intersection of science and policy in risk analysis of genetically engineered plants  
David A. Lee and Laura E. Bartley

## Chapter 16

Barley breeding in the United Kingdom  
W. T. B. Thomas

## Chapter 17

Developing a new cool-season perennial grass forage: interspecific hybrids of *Poa arachnifera* × *Poa secunda*  
Bryan Kindiger

## Chapter 18

Pioneer Hi-Bred International, Inc.: bringing seed value to the grower  
Jerry Harrington

## Chapter 19

Bringing Roundup Ready® technology to wheat  
Sally Metz

## Chapter 20

Genetic improvement of cassava through biotechnology  
Nigel J. Taylor

## Chapter 21

Discovering genes for drought adaptation in sorghum  
Andrew Borrell, David Jordan, John Mullet, Patricia

Klein, Robert Klein, Henry Nguyen, Darrell Rosenow, Graeme Hammer, and Bob Henzell

Chapter 22

QPM: enhancing protein nutrition in sub-Saharan Africa  
Twumasi Afriyie

Chapter 23

MSTAT: a software program for plant breeders  
Russell Freed

Chapter 24

Public release and registration of "Prolina" soybean  
Joe W. Burton  
*and*  
Plant variety protection in Canada  
B. Riché and D. J. Donnelly

Chapter 25

Plant breeding research at ICRISAT  
P. M. Gaur, K. B. Saxena, S. N. Nigam, B. V. S. Reddy,  
K. N. Rai, C. L. L. Gowda, and H. D. Upadhyaya

Chapter 26

An example of participatory plant breeding: barley at  
ICARDA  
S. Ceccarelli and S. Grando

Chapter 27

Bringing genomics to the wheat fields  
K. A. Garland-Campbell, J. Dubcovsky, J. A. Anderson,  
P. S. Baenziger, G. Brown-Guedira, X. Chen, E. Elias,

A. Fritz, B. S. Gill, K. S. Gill, S. Haley, K. K. Kidwell,  
S. F. Kianian, N. Lapitan, H. Ohm, D. Santra, M. Sorrells,  
M. Soria, E. Souza, and L. Talbert

Chapter 28

Hybrid breeding in maize  
F. J. Betrán

Chapter 29

Breeding rice  
Anna Myers McClung

Chapter 30

Sorghum breeding  
William Rooney

Chapter 31

Estimating inheritance factors and developing cultivars  
for tolerance to charcoal rot in soybean  
James R. Smith

Chapter 32

Peanut (*Arachis hypogaea* L.) breeding and root-knot  
nematode resistance  
Charles Simpson

Chapter 33

The breeding of potato  
John E. Bradshaw

Chapter 34

Cotton breeding  
Don L. Keim

# Industry highlights box authors

- Acquaah, G., Department of Agriculture and Natural Resources, Langston University, Langston, OK 73050, USA
- Afriyie, T., International Maize and Wheat Improvement Center (CIMMYT), PO Box 5689, Addis Ababa, Ethiopia
- Anderson, J. A., Department of Agronomy and Plant Genetics, University of Minnesota, Twin Cities, St Paul, MN 55108, USA
- Baenziger, P. S., Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583, USA
- Bartley, L. E., USDA-APHIS Biotechnology Regulatory Services, Riverdale, MD 20737, USA
- Betrán, F. J., Texas A&M University, College Station, TX 77843, USA
- Borrell, A., Department of Primary Industries and Fisheries, Hermitage Research Station, Warwick, Queensland 4370, Australia
- Bradshaw, J. E., Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK
- Brown-Guedira, G., USDA-ARS Plant Science Research Unit, North Carolina State University, Raleigh, NC 27606, USA
- Burton, J. W., USDA Plant Science Building, 3127 Ligon Street, Raleigh, NC 27607, USA
- Ceccarelli, S., International Center for Agricultural Research in the Dry Areas (ICARDA), PO Box 5466, Aleppo, Syria
- Chalyk, S., 12 Goldfinch Court, Apt. 1007, Toronto M2R 2C4, Canada
- Chen, X., USDA-ARS Wheat Genetics, Quality, Physiology, and Disease Research Unit, Washington State University, Pullman WA 99164, USA
- Deerfield, D. W. II, Pittsburgh Supercomputing Center, Pittsburgh, PA 15213, USA
- Donnelly, D. J., Plant Science Department, McGill University, Ste. Anne de Bellevue, QC H9X 3V9, Canada
- Dubcovsky, J., Department of Agronomy and Range Science, University of California at Davis, Davis, CA 95616, USA
- Elias, E., Department of Plant Sciences, North Dakota State University, Fargo, ND 58105, USA
- Freed, R., Department of Crop and Soil Science, Michigan State University, East Lansing, MI 48824, USA
- Fritz, A., Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA
- Garland-Campbell, K. A., USDA-ARS Wheat Genetics, Quality, Physiology, and Disease Research Unit, Washington State University, Pullman, WA 99164, USA
- Gaur, P. M., International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, AP, India
- Gill, B. S., Wheat Genetics Resource Center, Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA
- Gill, K. S., Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164, USA
- Gowda, C. L. L., International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, AP, India
- Grando, S., International Center for Agricultural Research in the Dry Areas (ICARDA), PO Box 5466, Aleppo, Syria
- Haley, S., Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO 80526, USA
- Hammer, G., School of Land and Food, University of Queensland, Queensland 4072, Australia
- Hammer, K., Institute of Crop Science, Agrobiodiversity Department, University Kassel, D-37213 Witzenhausen, Germany
- Harrington, J., Pioneer Hi-Bred International, Des Moines, IA 50307, USA
- Henzell, R., Department of Primary Industries and Fisheries, Hermitage Research Station, Warwick, Queensland 4370, Australia
- Heuser, F., Institute of Crop Science, Agrobiodiversity Department, University Kassel, D-37213 Witzenhausen, Germany
- Jaradat, A. A., USDA-ARS, Morris, 56267 MN, USA
- Jordan, D., Department of Primary Industries and Fisheries, Hermitage Research Station, Warwick, Queensland 4370, Australia
- Keim, D. L., Delta and Pine Land Company, One Cotton Row, PO Box 157, Scott, MS 38772, USA
- Khoshbakht, K., Institute of Crop Science, Agrobiodiversity Department, University Kassel, D-37213 Witzenhausen, Germany
- Kianian, S. F., Department of Plant Sciences, North Dakota State University, Fargo, ND 58105, USA
- Kidwell, K. K., Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164, USA
- Kindiger, B., USDA-ARS Grazinglands Research Laboratory, El Reno, OK 73036, USA
- Klein, P., Institute for Plant Genomics and Biotechnology, Texas A&M University, College Station, TX 77843, USA
- Klein, R., USDA-ARS Southern Agricultural Research Station, College Station, TX 77843, USA
- Lapitan, N., Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO 80526, USA
- Laurens, F., UMR Génétique et Horticulture (GenHort) (INRA/INH/UA), INRA Centre d'Angers, 49070 Beaucouzé, France
- Lee, D. A., EPA Office of Research and Development, 8623N, Washington, DC 20460, USA
- Matand, K., Department of Agriculture and Applied Sciences, Langston University, Langston, OK 73050, USA

- McClung, M. A., USDA-ARS Dale Bumpers National Rice Research Center and Beaumont Rice Research Unit, 1509 Aggie Drive, Beaumont, TX 77713, USA
- Metz, S., Monsanto Corporation, 800 North Lindbergh Blvd, St Louis, MO 63167, USA
- Mullet, J., Institute for Plant Genomics and Biotechnology, Texas A&M University, College Station, TX 77843, USA
- Nguyen, H., Plant Sciences Unit and National Center for Soybean Biotechnology, University of Missouri, Columbia, MO 65211, USA (previously Texas Tech University, Lubbock, USA)
- Nicholas, H. B., Jr., Pittsburgh Supercomputing Center, Pittsburgh, PA 15213, USA
- Nigam, S. N., International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, AP, India
- Novy, R., USDA-ARS, Aberdeen, ID 83210, USA
- Nuez, F., Instituto para la Conservación y Mejora de la Agrobiodiversidad Valenciana, Universidad Politécnica de Valencia, 46022 Valencia, Spain
- Ohm, H., Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA
- Prohens, J., Instituto para la Conservación y Mejora de la Agrobiodiversidad Valenciana, Universidad Politécnica de Valencia, 46022 Valencia, Spain
- Rai, K. N., International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, AP, India
- Reddy, B. V. S., International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, AP, India
- Riché, B., Plant Science Department, McGill University, Ste. Anne de Bellevue, QC H9X 3V9, Canada
- Rodríguez-Burrueto, A., Instituto para la Conservación y Mejora de la Agrobiodiversidad Valenciana, Universidad Politécnica de Valencia, 46022 Valencia, Spain
- Rooney, W., Texas A&M University, College Station, TX 77843, USA
- Ropelewski, A. J., Pittsburgh Supercomputing Center, Pittsburgh, PA 15213, USA
- Rosenow, D., Texas A&M Agricultural Research and Extension Center, Lubbock, TX 79403, USA
- Santra, D., Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164, USA
- Saxena, K. B., International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, AP, India
- Simpson, C., Texas A&M University, College Station, TX 77843, USA
- Smith, J. R., USDA-ARS Crop Genetics and Production Research Unit, Stoneville, MS 38776, USA
- Soria, M., Department of Agronomy and Range Science, University of California at Davis, Davis, CA 95616, USA
- Sorrells, M., Department of Plant Breeding, Cornell University, Ithaca, NY 14853, USA
- Souza, E., Aberdeen Research and Extension Center, University of Idaho, Aberdeen, ID 83210, USA
- Talbert, L., Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717, USA
- Taylor, N. J., International Laboratory for Tropical Agricultural Biotechnology (ILTAB), Donald Danforth Plant Science Center, St Louis, MO 63132, USA
- Teklu, Y., Institute of Crop Science, Agrobiodiversity Department, University Kassel, D-37213 Witzenhausen, Germany
- Thomas, W. T. B., Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK
- Upadhyaya, H. D., International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, AP, India

# Preface

Plant breeding is an art and a science. May be it should be added that it is also a business. Modern plant breeding is a discipline that is firmly rooted in the science of genetics. As an applied science, breeders are offered opportunities to apply principles and technologies from several scientific disciplines to manipulate plants for specific purposes.

This textbook, *Principles of Plant Genetics and Breeding*, is designed to present plant breeding in a balanced, comprehensive, and current fashion to students at the upper undergraduate level to early graduate level. It is divided into two parts. Part I is devoted to discussing the underlying science, and principles and concepts of plant breeding, followed by a detailed discussion of the methods of breeding. Part II is devoted to discussing the applications of the principles and concepts learned in Part I to breeding eight major field crops. The principles and concepts discussed are generally applicable to breeding all plants. However, most of the examples used in the book are drawn from the breeding of field crops.

The book has several very unique components, some of them never before presented in traditional plant breeding textbooks at this level:

- The principles and concepts of genetics are presented in more detail in scope and depth than obtains in other textbooks written at this level. But, more importantly, the student is shown how the principles are applied in plant breeding. As much as possible, specific examples of application in plant breeding are always given.
- Genetic variation is indispensable to plant breeding. The issue of germplasm in plant breeding is discussed in detail, including genetic vulnerability in crops, and germplasm collection and maintenance.
- The latest most versatile and most controversial tools in the tool kit of plant breeders are the technologies of biotechnology, especially genetic engineering technologies. The underlying principles of genetic engineering are discussed in detail. This is followed by the application of biotechnology in breeding, including molecular breeding of crops.
- Because of the controversial nature of genetic engineering, the book discusses in detail the issues of risk, regulation, and public perception of biotechnology as applied in plant breeding.
- A significant subject that is rarely discussed in plant breeding books is the issue of intellectual property (IP) and ethics. These issues are important because they protect the breeder from abuse of their inventions and provide incentive for research and development of new cultivars. IP is thoroughly discussed in the book, with particular reference to plant breeding.
- Some of the important yet often ignored subjects in plant breeding books are prebreeding (or germplasm enhancement) and heterotic groups. These concepts are effectively discussed.
- Both the conventional methods and contemporary methods of plant breeding are discussed in detail, pointing out their strengths and weaknesses, but more importantly emphasizing their complementary use in modern plant breeding.
- Breeding objectives in plant breeding are as diverse as plant breeders. Breeding objectives are discussed according to effective themes. The presentation is unique in that it includes discussions of the sources of germplasm, and the genetics and progress in breeding specific traits. Breeding for environmental stresses is especially uniquely presented in this book.
- The discussion on breeding for disease and pest resistance is comprehensive, incorporating the current applications of genetic engineering in the development of genetically modified breeding materials.
- The cultivar release process is discussed to a good depth and scope.
- The book is well illustrated to help students better understand the principles and concepts discussed in the book.
- Plant breeding methods have remained fairly unchanged over the years. This book takes a bold step in introducing, for the first time in a plant breeding textbook at this level, the emerging concepts of decentralized participatory breeding and organic plant breeding.
- Perhaps the most unique aspect of this book is the incorporation of contributions from plant breeding professionals. Industry professionals were invited to present practical applications of plant breeding principles and concepts. In this way students are able to see how the principles and concepts of breeding are applied in real life to address specific plant breeding

problems. The professionals were given the latitude to make their presentation in the format of their choosing, without being too technical. Each participant has provided a significant list of references that will be of special interest to graduate students who wish to further investigate the problems discussed.

The style of presentation throughout the book is easy to follow and comprehend. Students are constantly re-

minded of previous topics of relevance to current topics being discussed. This book is not only an excellent teaching tool, but it is also suitable as a reference source for professionals.

For instructors, if you did not receive an artwork CD-ROM with your comp copy, please contact this email address: [artworkcd@bos.blackwellpublishing.com](mailto:artworkcd@bos.blackwellpublishing.com)

# Acknowledgments

The author expresses sincere gratitude to the professionals from all over the world who contributed highlights to the chapters to significantly enhance the instructional value of the textbook. Also, gratitude and appreciation are hereby extended to the reviewers whose comments and suggested were incorporated in the preparation of the final manuscript. The reviewers were: William Berzonsky, North Dakota State University; Paul Bilsborrow, University of Newcastle; Dennis Decoteau, Pennsylvania State University; Majid R.

Foolad, Pennsylvania State University; Vernon Gracen, Cornell University; Sean Mayes, Nottingham University; Habibur Rahman, University of Alberta; Andrew Riseman, University of British Columbia; Lee Tarpley, Texas A&M University; David Weaver, Auburn University; and Todd C. Wehner, North Carolina State University.

Finally, the author reserves the highest gratitude for Dr. J. C. El Shaddai, for ideas, guidance, encouragement, and support throughout the duration of the project.



# Part I

## Underlying science and methods of plant breeding

Part I of this book deals with the underlying basic science of plant breeding, emphasizing both classic and contemporary principles and concepts. All the pertinent genetic concepts that are needed to understand and conduct plant breeding are discussed. These include a consideration of Mendelian, population, and quantitative genetic principles and concepts. More importantly, the discussions show how plant breeders use these scientific principles and concepts in their work. Examples of plant breeding applications are provided.

The chapters also present a detailed discussion of the tools of plant breeding, including both classic and non-classic or contemporary tools, and emphasizing the complementarity of the two sets of tools. Then, a discussion of the methods of breeding follows, indicating how plant breeding tools are used for breeding self-pollinated, cross-pollinated, and clonally propagated species. Part I concludes with sections on the cultivar release process, international plant breeding efforts, and emerging concepts in plant breeding. Two contemporary breeding approaches that have never before been formally addressed in plant breeding textbooks are discussed. It is important for students to be introduced to the most recent issues in the field of plant improvement so that they may participate in the discussion.

# **Section 1**

# **Historical perspectives and importance of plant breeding**

## **Chapter 1 History and role of plant breeding in society**

---

Plant breeding as a human endeavor has its origins in antiquity, starting off simply as discrimination among plant types to select and retain plants with the most desirable features. Remarkably, the practice of selection remains the primary strategy for crop improvement, even though many technologically advanced techniques have been added to the arsenal of the modern plant breeder. Plant breeding differs from evolution in that the former is planned and purposeful. The student needs to appreciate how this formal activity of plant manipulation started and the advances made over the ages. More importantly, the student needs to appreciate the achievements of plant breeding and its impact on society. This section introduces the student to the field of plant breeding, highlighting its development, importance, and approaches.

---



# 1

# History and role of plant breeding in society

---

## Purpose and expected outcomes

*Agriculture is the deliberate planting and harvesting of plants and herding animals. This human invention has, and continues to, impact on society and the environment. Plant breeding is a branch of agriculture that focuses on manipulating plant heredity to develop new and improved plant types for use by society. People in society are aware and appreciative of the enormous diversity in plants and plant products. They have preferences for certain varieties of flowers and food crops. They are aware that whereas some of this variation is natural, humans with special expertise (plant breeders) create some of it. Generally, also, there is a perception that such creations derive from crossing different plants. The tools and methods used by plant breeders have been developed and advanced through the years. There are milestones in plant breeding technology as well as accomplishments by plant breeders over the years. This introductory chapter is devoted to presenting a brief overview of plant breeding, including a brief history of its development, how it is done, and its benefits to society. After completing this chapter, the student should have a general understanding of:*

- 1 The historical perspectives of plant breeding.
  - 2 The need and importance of plant breeding to society.
  - 3 The goals of plant breeding.
  - 4 Trends in plant breeding as an industry.
  - 5 Milestones in plant breeding.
  - 6 The accomplishments of plant breeders.
  - 7 The future of plant breeding in society.
- 

## What is plant breeding?

Plant breeding is a deliberate effort by humans to nudge nature, with respect to the heredity of plants, to an advantage. The changes made in plants are permanent and heritable. The professionals who conduct this task are called **plant breeders**. This effort at adjusting the *status quo* is instigated by a desire of humans to improve certain aspects of plants to perform new roles or enhance existing ones. Consequently, the term “plant breeding” is often used synonymously with “plant improvement” in modern society. It needs to be emphasized that the

goals of plant breeding are focused and purposeful. Even though the phrase “to breed plants” often connotes the involvement of the sexual process in effecting a desired change, modern plant breeding also includes the manipulation of asexually reproducing plants (plants that do not reproduce through the sexual process). Breeding is hence about manipulating plant attributes, structure, and composition, to make them more useful to humans. It should be mentioned at the onset that it is not every plant character or trait that is amenable to manipulation by breeders. However, as technology advances, plant breeders are increasingly able to

accomplish astonishing plant manipulations, needless to say not without controversy, as is the case involving the development and application of **biotechnology** to plant genetic manipulation. One of the most controversial of these modern technologies is **transgenesis**, the technology by which gene transfer is made across natural biological barriers.

Plant breeders specialize in breeding different groups of plants. Some focus on field crops (e.g., soybean, cotton), horticultural crops (e.g., vegetables), ornamentals, fruit trees (e.g., citrus, apple), forage crops (e.g., alfalfa, grasses), or turf species. More importantly, breeders tend to focus on specific species in these groups. This way, they develop the expertise that enables them to be most effective in improving the species of their choice. The principles and concepts discussed in this book are generally applicable to breeding all species. However, most of the examples supplied are from breeding field crops.

### Goals of plant breeding

The plant breeder uses various technologies and methodologies to achieve targeted and directional changes in the nature of plants. As science and technology advance, new tools are developed while old ones are refined for use by breeders. Before initiating a breeding project, clear breeding objectives are defined based on factors such as producer needs, consumer preferences and needs, and environmental impact. Breeders aim to make the crop producer's job easier and more effective in various ways. They may modify plant structure so it can resist lodging and thereby facilitate mechanical harvesting. They may develop plants that resist pests so the farmer does not have to apply pesticides or can apply smaller amounts of these chemicals. Not applying pesticides in crop production means less environmental pollution from agricultural sources. Breeders may also develop high-yielding varieties (or **cultivars**) so the farmer can produce more for the market to meet consumer demands while improving his or her income. The term cultivar is reserved for variants deliberately created by plant breeders and will be introduced more formally later in the book. It will be the term of choice in this book.

When breeders think of consumers, they may, for example, develop foods with higher nutritional value and that are more flavorful. Higher nutritional value means reduced illnesses in society (e.g., nutritionally related ones such as blindness or ricketsia) caused by the

consumption of nutrient-deficient foods, as obtains in many developing regions where staple foods (e.g., rice, cassava) often lack certain essential amino acids or nutrients. Plant breeders may also target traits of industrial value. For example, fiber characteristics (e.g., strength) of fiber crops such as cotton can be improved, while oil crops can be improved to yield high amounts of specific fatty acids (e.g., the high oleic content of sunflower seed). The latest advances in technology, specifically genetic engineering technologies, are being applied to enable plants to be used as bioreactors to produce certain pharmaceuticals (called **biopharming** or simply **pharming**).

The technological capabilities and needs of societies of old, restricted plant breeders to achieving modest objectives (e.g., product appeal, adaptation to production environment). It should be pointed out that these "older" breeding objectives are still important today. However, with the availability of sophisticated tools, plant breeders are now able to accomplish these genetic alterations in novel ways that are sometimes the only option, or are more precise and more effective. Furthermore, as previously indicated, they are able to undertake more dramatic alterations that were impossible to attain in the past (e.g., transferring a desirable gene from a bacterium to a plant!). Some of the reasons why plant breeding is important to society are summarized next.

### Concept of genetic manipulation of plant attributes

The work of Gregor Mendel and the further advances in science that followed his discoveries established that plant characteristics are controlled by hereditary factors or **genes** that consist of DNA (deoxyribose nucleic acid, the hereditary material). These genes are expressed in an environment to produce a trait. It follows then that in order to change a trait or its expression, one may change the *nature* or its genotype, and/or modify the *nurture* (environment in which it is expressed). Changing the environment essentially entails modifying the growing or production conditions. This may be achieved through an agronomic approach, for example, the application of production inputs (e.g., fertilizers, irrigation). Whereas this approach is effective in enhancing certain traits, the fact remains that once these supplemental environmental factors are removed, the expression of the plant trait reverts to the *status quo*. On the other hand, plant breeders seek to modify plants with respect

to the expression of certain attributes by modifying the genotype (in a desired way by targeting specific genes). Such an approach produces an alteration that is permanent (i.e., transferable from one generation to the next).

### Why breed plants?

The reasons for manipulating plant attributes or performance change according to the needs of society. Plants provide food, feed, fiber, pharmaceuticals, and shelter for humans. Furthermore, plants are used for aesthetic and other functional purposes in the landscape and indoors.

#### Addressing world food, feed, and nutritional needs

Food is the most basic of human needs. Plants are the primary producers in the **ecosystem** (a community of living organisms including all the non-living factors in the environment). Without them, life on earth for higher organisms would be impossible. Most of the crops that feed the world are cereals (Table 1.1). Plant breeding is needed to enhance the value of food crops, by improving their yield and the nutritional quality of their products, for healthy living of humans. Certain plant foods are deficient in certain essential nutrients to the extent that where these foods constitute the bulk of a staple diet, diseases associated with nutritional deficiency are often common. Cereals tend to be low in lysine and threonine, while legumes tend to be low in cysteine and methionine (both sulfur-containing amino acids). Breeding is needed to augment the nutritional quality of food crops. Rice, a major world food, lacks

**Table 1.1** The 25 major food crops of the world, ranked according to total tonnage produced annually.

1	Wheat	11	Sorghum	21	Apple
2	Rice	12	Sugarcane	22	Yam
3	Corn	13	Millet	23	Peanut
4	Potato	14	Banana	24	Watermelon
5	Barley	15	Tomato	25	Cabbage
6	Sweet potato	16	Sugar beet		
7	Cassava	17	Rye		
8	Grape	18	Orange		
9	Soybean	19	Coconut		
10	Oat	20	Cottonseed oil		

Source: J.R. Harlan. 1976. Plants and animals that nourish man. In: Food and agriculture, A Scientific American Book. W.H. Freeman and Company, San Francisco.

pro-vitamin A (the precursor of vitamin A). The “Golden Rice” project, currently underway at the International Rice Research Institute (IRRI) in the Philippines and other parts of the world, is geared towards developing, for the first time ever, a rice cultivar with the capacity to produce pro-vitamin A. An estimated 800 million people in the world, including 200 million children, suffer chronic undernutrition, with its attendant health issues. Malnutrition is especially prevalent in developing countries.

Breeding is also needed to make some plant products more digestible and safer to eat by reducing their toxic components and improving their texture and other qualities. A high lignin content of the plant material reduces its value for animal feed. Toxic substances occur in major food crops, such as alkaloids in yam, cyanogenic glucosides in cassava, trypsin inhibitors in pulses, and steroidal alkaloids in potatoes. Forage breeders are interested, among other things, in improving feed quality (high digestibility, high nutritional profile) for livestock.

#### Addressing food needs for a growing world population

In spite of a doubling of the world population in the last three decades, agricultural production rose at an adequate rate to meet world food needs. However, an additional 3 billion people will be added to the world population in the next three decades, requiring an expansion in world food supplies to meet the projected needs. As the world population increases, there would be a need for an agricultural production system that is apace with population growth. Unfortunately, arable land is in short supply, stemming from new lands that have been brought into cultivation in the past, or surrendered to urban development. Consequently, more food will have to be produced on less land. This calls for improved and high-yielding varieties to be developed by plant breeders. With the aid of plant breeding, the yields of major crops have dramatically changed over the years. Another major concern is the fact that most of the population growth will occur in developing countries where food needs are currently most serious, and where resources for feeding people are already most severely strained, because of natural or human-made disasters, or ineffective political systems.

#### The need to adapt plants to environmental stresses

The phenomenon of global climatic change that is occurring over the years is partly responsible for

modifying the crop production environment (e.g., some regions of the world are getting drier and others saltier). This means that new cultivars of crops need to be bred for new production environments. Whereas developed economies may be able to counter the effects of unseasonable weather by supplementing the production environment (e.g., by irrigating crops), poor countries are easily devastated by even brief episodes of adverse weather conditions. For example, the development and use of drought-resistant cultivars is beneficial to crop production in areas of marginal or erratic rainfall regimes. Breeders also need to develop new plant types that can resist various biotic (diseases, insect pests) and other abiotic (e.g., salt, drought, heat, cold) stresses in the production environment. Crop distribution can be expanded by adapting crops to new production environments (e.g., adapting tropical plants to temperate regions). The development of photoperiod-insensitive crop cultivars would allow the expansion in production of previously photoperiod-sensitive species.

### The need to adapt crops to specific production systems

Breeders need to produce plant cultivars for different production systems to facilitate crop production and optimize crop productivity. For example, crop cultivars must be developed for rain-fed or irrigated production, and for mechanized or non-mechanized production. In the case of rice, separate sets of cultivars are needed for upland production and for paddy production. In organic production systems where pesticide use is highly restricted, producers need insect- and disease-resistant cultivars in crop production.

### Developing new horticultural plant varieties

The ornamental horticultural production industry thrives on the development of new varieties through plant breeding. Aesthetics is of major importance to horticulture. Periodically, ornamental plant breeders release new varieties that exhibit new colors and other morphological features (e.g., height, size, shape). Also, breeders develop new varieties of vegetables and fruits with superior yield, nutritional qualities, adaptation, and general appeal.

### Satisfying industrial and other end-use requirements

Processed foods are a major item in the world food supply system. Quality requirements for fresh produce

meant for the table are different from those used in the food processing industry. For example, there are table grapes and grapes bred for wine production. One of the reasons why the first **genetically modified (GM)** crop (produced by using genetic engineering tools to incorporate foreign DNA) approved for food, the FlavrSavr® tomato, did not succeed was because the product was marketed as a table or fresh tomato, when in fact the gene of interest was placed in a genetic background for developing a processing tomato variety. Other factors contributed to the demise of this historic product. Different markets have different needs that plant breeders can address in their undertakings. For example, the potato is a versatile crop used for food and industrial products. Different varieties are bred for baking, cooking, fries (frozen), chipping, and starch. These cultivars differ in size, specific gravity, and sugar content, among other properties. A high sugar content is undesirable for frying or chipping because the sugar caramelizes under high heat to produce undesirable browning of fries and chips.

### Plant breeding through the ages

Plant breeding as a conscious human effort has ancient origins.

### Origins of agriculture and plant breeding

In its primitive form, plant breeding started after the invention of agriculture, when people of primitive cultures switched from a lifestyle of hunter-gatherers to sedentary producers of selected plants and animals. Views of agricultural origins range from the mythological to ecological. This lifestyle change did not occur overnight but was a gradual process during which plants were transformed from being independent, wild progenitors, to fully dependent (on humans) and domesticated varieties. Agriculture is generally viewed as an invention and discovery. During this period, humans also discovered the time-honored and most basic plant breeding technique – **selection**, the art of discriminating among biological variation in a population to identify and pick desirable variants. Selection implies the existence of variability. In the beginnings of plant breeding, the variabilities exploited were the naturally occurring variants and wild relatives of crop species. Furthermore, selection was based solely on the intuition, skill, and judgment of the operator. Needless to say, this form of selection is practiced to date by farmers

in poor economies, where they save seed from the best-looking plants or the most desirable fruit for planting the next season. These days, scientific techniques are used in addition to the aforementioned qualities to make the selection process more precise and efficient. Even though the activities described in this section are akin to some of those practiced by modern plant breeders, it is not being suggested that primitive crop producers were necessarily conscious of the fact that they were nudging nature to their advantage as modern breeders do.

### Plant breeding past (pre-Mendelian)

Whereas early plant breeders did not deliberately create new variants, modern plant breeders are able to create new variants that previously did not occur in natural populations. It is difficult to identify the true beginnings of modern plant breeding. However, certain early observations by certain individuals helped to lay the foundation for the discovery of the modern principles of plant breeding. It has been reported that archaeological records indicate that the Assyrians and Babylonians artificially pollinated date palm, at least 700 BC. R. J. Camerarius (aka Rudolph Camerer) of Germany is credited with first reporting sexual reproduction in plants in 1694. Through experimentation, he discovered that pollen from male flowers was indispensable to fertilization and seed development on female plants. His work was conducted on monoecious plants (both sexes occur on separate parts of the plant, e.g., spinach and maize). However, it was Joseph Koelreuter who conducted the first known systematic investigations into plant hybridization (crossing of genetically dissimilar parents) of a number of species, between 1760 and 1766. Similarly, in 1717, Thomas Fairchild, an Englishman, conducted an **interspecific cross** (a cross between two species) between sweet william (*Dianthus barbatus*) and *D. caryophyllis*, to obtain what became known as Fairchild's sweet william. Another account describes an observation in 1716 by an American, Cotton Mather, to the effect that ears from yellow corn grown next to blue or red corn had blue and red kernels in them. This suggested the occurrence of natural cross-pollination. Maize is one of the crops that has received extensive breeding and genetic attention in the scientific community. As early as 1846, Robert Reid of Illinois was credited with developing what became known as "Reid's Yellow Dent". The landmark work by Swedish botanist, Carolus Linnaeus (1707–1778), which culminated in the **binomial systems of classification** of plants, is

invaluable to modern plant breeding. In 1727, Louis Leveque de Vilmorin of the Vilmorin family of seed growers founded the Vilmorin Breeding Institute in France as the first institution dedicated to plant breeding and the production of new cultivars. There, another still commonly used breeding technique – **progeny test** (growing the progeny of a cross for the purpose of evaluating the genotype of the parent) – was first used to evaluate the breeding value of a single plant. Selected milestones in plant breeding are presented in Table 1.2.

### Plant breeding present (post-Mendelian)

Modern plant breeding depends on the principles of genetics, the science of heredity to which Gregor Mendel made some of its foundational contributions. Mendel's original work on the garden pea was published in 1865. It described how factors for specific traits are transmitted from parents to offspring and through subsequent generations. His work was rediscovered in 1900, with confirmation by E. von Tschermak, C. Correns, and H. de Vries. These events laid the foundation for modern genetics. Mendel's studies gave birth to the concept of genes (and the discipline of **genetics**), factors that encode traits and are transmitted through the sexual process to the offspring. Further, his work resulted in the formulation of the basic rules of heredity that are called Mendel's laws.

One of the earliest applications of genetics to plant breeding was made by the Danish botanist, Wilhelm Johannsen. In 1903, Johannsen developed the **pure-line theory** while working on the garden bean. His work confirmed an earlier observation by others that the techniques of selection could be used to produce uniform, true-breeding cultivars by selecting from the progeny of a single self-pollinated crop (through repeated selfing) to obtain highly homozygous lines (true breeding), which he later crossed. Previously, H. Nilson had demonstrated that the unit of selection was the plant. The products of the crosses (called **hybrids**) yielded plants that outperformed either parent with respect to the trait of interest (the concept of **hybrid vigor**). Hybrid vigor (or heterosis) is the foundation of modern hybrid crop production programs.

In 1919, D. F. Jones took the idea of a single cross further by proposing the double-cross concept, which involved a cross between two single crosses. This technique made the commercial production of hybrid corn seed economical. The application of genetics in crop improvement has yielded spectacular successes over the

**Table 1.2** Selected milestones in plant breeding

9000 BC	First evidence of plant domestication in the hills above the Tigris river
3000 BC	Domestication of all important food crops in the Old World completed
1000 BC	Domestication of all important food crops in the New World completed
700 BC	Assyrians and Babylonians hand pollinate date palms
1694	Camerarius of Germany first to demonstrate sex in plants and suggested crossing as a method to obtain new plant types
1716	Mather of USA observed natural crossing in maize
1719	Fairchild created first artificial hybrid (carnation × sweet william)
1727	Vilmorin Company of France introduced the pedigree method of breeding
1753	Linnaeus published <i>Species plantarum</i> . Binomial nomenclature born
1761–1766	Koelreuter of Germany demonstrated that hybrid offspring received traits from both parents and were intermediate in most traits; produced first scientific hybrid using tobacco
1847	“Reid’s Yellow Dent” maize developed
1866	Mendel published his discoveries in <i>Experiments in plant hybridization</i> , cumulating in the formulation of laws of inheritance and discovery of unit factors (genes)
1899	Hopkins described the ear-to-row selection method of breeding in maize
1900	Mendel’s laws of heredity rediscovered independently by Correns of Germany, de Vries of Holland, and von Tschermark of Austria
1903	The pure-line theory of selection developed
1904–1905	Nilsson-Ehle proposed the multiple factor explanation for inheritance of color in wheat pericarp
1908–1909	Hardy of England and Weinberg of Germany developed the law of equilibrium of populations
1908–1910	East published his work on inbreeding
1909	Shull conducted extensive research to develop inbreds to produce hybrids
1917	Jones developed first commercial hybrid maize
1926	Pioneer Hi-bred Corn Company established as first seed company
1934	Dustin discovered colchicines
1935	Vavilov published <i>The scientific basis of plant breeding</i>
1940	Harlan used the bulk breeding selection method in breeding
1944	Avery, MacLeod, and McCarty discovered DNA is hereditary material
1945	Hull proposed recurrent selection method of breeding
1950	McClintock discovered the <i>Ac-Ds</i> system of transposable elements
1953	Watson, Crick, and Wilkins proposed a model for DNA structure
1970	Borlaug received Nobel Prize for the Green Revolution
1994	Berg, Cohen, and Boyer introduced the recombinant DNA technology
1995	“FlavrSavr” tomato developed as first genetically modified food produced for the market
1996	<i>Bt</i> corn developed
2004	Roundup Ready® soybean introduced
	Roundup Ready® wheat developed

years, one of the most notable being the development of dwarf, environmentally responsive cultivars of wheat and rice for the subtropical regions of the world. These new plant materials transformed food production in these regions in a dramatic fashion, and in the process became dubbed the **Green Revolution**. This remarkable achievement in food production is discussed below.

Mutagenesis (the induction of mutations using mutagenic agents (mutagens) such as radiation or chemicals) became a technique for plant breeding in the 1920s when researchers discovered that exposing plants to X-rays increased the variation in plants. Mutation breeding

accelerated after World War II, when scientists included nuclear particles (e.g., alpha, protons, and gamma) as mutagens for inducing mutations in organisms. Even though very unpredictable in outcome, mutagenesis has been successfully used to develop numerous mutant varieties.

In 1944, DNA was discovered to be the genetic material. Scientists then began to understand the molecular basis of heredity. New tools (molecular tools) are being developed to facilitate plant breeding. Currently, scientists are able to circumvent the sexual process to transfer genes from one parent to another. In fact, genes

can now be transferred from virtually any organism to another. This newest tool, specifically called **genetic engineering**, has its proponents and distractors. Current successes include the development of insect resistance in crops such as maize by incorporating a gene from the bacterium *Bacillus thuringiensis*. Cultivars containing an alien gene for insect resistance from this particular organism are called *Bt* cultivars, diminutive of the scientific name of the bacterium. The products of the application of this alien gene transfer technology are generally called genetically modified (GM) or **transgenic** products. **Plant biotechnology**, the umbrella name for the host of modern plant manipulation techniques, has produced, among other things, molecular markers to facilitate the selection process in plant breeding.

### Achievements of modern plant breeders

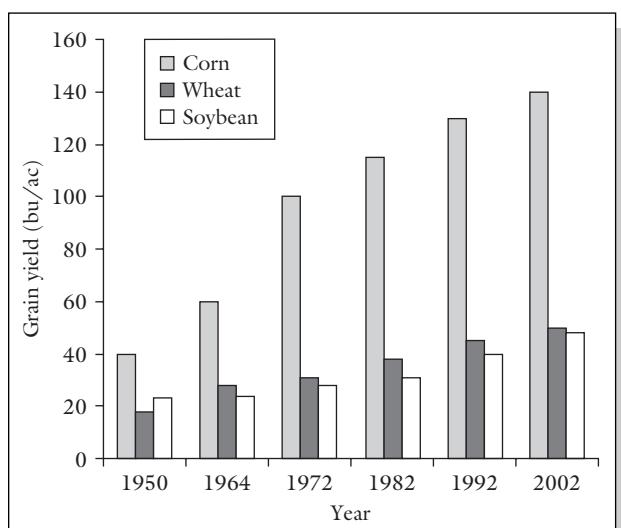
The achievements of plant breeders are numerous, but may be grouped into several major areas of impact – yield increase, enhancement of compositional traits, crop adaptation, and the impact on crop production systems.

#### **Yield increase**

Yield increase in crops has been accomplished in a variety of ways including targeting yield *per se* or its components, or making plants resistant to economic diseases and insect pests, and breeding for plants that are responsive to the production environment. Yields of major crops (e.g., corn, rice, sorghum, wheat, soybean) have significantly increased in the USA over the years (Figure 1.1). For example, the yield of corn rose from about 2,000 kg/ha in the 1940s to about 7,000 kg/ha in the 1990s. In England, it took only 40 years for wheat yields to rise from 2,000 to 6,000 kg/ha. These yield increases are not totally due to the genetic potential of the new crop cultivars but also due to improved agronomic practices (e.g., application of fertilizer, irrigation). Crops have been armed with disease resistance to reduce yield loss. Lodging resistance also reduces yield loss resulting from harvest losses.

#### **Enhancement of compositional traits**

Breeding for plant compositional traits to enhance nutritional quality or to meet an industrial need are major plant breeding goals. High protein crop varieties (e.g., high lysine or quality protein maize) have been



**Figure 1.1** The yield of major world food crops is steadily rising, as indicated by the increasing levels of crops produced in the US agricultural system. A significant portion of this rise is attributable to the use of improved crop cultivars by crop producers. bu/ac, bushels per acre. Source: Drawn with data from the USDA.

produced for use in various parts of the world. For example, different kinds of wheat are needed for different kinds of products (e.g., bread, pasta, cookies, semolina). Breeders have identified the quality traits associated with these uses and have produced cultivars with enhanced expression of these traits. Genetic engineering technology has been used to produce high oleic sunflower for industrial use, while it is also being used to enhance the nutritional value of crops (e.g., pro-vitamin A "Golden Rice"). The shelf-life of fruits (e.g., tomato) has been extended through the use of genetic engineering techniques to reduce the expression of compounds associated with fruit deterioration.

#### **Crop adaptation**

Crop plants are being produced in regions to which they are not native, because breeders have developed cultivars with modified physiology to cope with variations, for example, in the duration of day length (photoperiod). Photoperiod-insensitive cultivars will flower and produce seed under any day length conditions. The duration of the growing period varies from one region of the world to another. Early maturing cultivars of crop plants enable growers to produce a crop during a short window of opportunity, or even to produce two crops in

one season. Furthermore, early maturing cultivars can be used to produce a full season crop in areas where adverse conditions are prevalent towards the end of the normal growing season. Soils formed under arid conditions tend to accumulate large amounts of salts. In order to use these lands for crop production, salt-tolerant (saline and aluminum tolerance) crop cultivars have been developed for certain species. In crops such as barley and tomato, there are commercial cultivars in use, with drought, cold, and frost tolerance.

### The Green Revolution

Producing enough food to feed the world's ever increasing population has been a lingering concern of modern societies. Perhaps the most notable essay on food and population dynamics was written by Thomas Malthus in 1798. In this essay, "Essay on the principles of population", he identified the geometric role of natural population increase in outrunning subsistence food supplies. He observed that unchecked by environmental or social constraints it appears that human populations double every 25 years, regardless of the initial population size. Because population increase, according to this observation, was geometric, whereas food supply at best was arithmetic, there was implicit in this theory pessimism about the possibility of feeding ever growing populations. Fortunately, mitigating factors such as technological advances, advances in agricultural production, changes in socioeconomics, and political thinking of modern society, has enabled this dire prophesy to remain unfulfilled.

Unfortunately, the technological advances in the 20th century primarily benefited the industrial countries, leaving widespread hunger and malnutrition to persist in most developing countries. Many of these nations depend on food aid from industrial countries for survival. In 1967, a report by the US President's Science Advisory Committee came to the grim conclusion that "the scale, severity and duration of the world food problem are so great that a massive, long-range, innovative effort unprecedented in human history will be required to master it". The Rockefeller and Ford Foundations, acting on this challenge, proceeded to establish the first international agricultural system to help transfer the agricultural technologies of the developed countries to the developing countries. These humble beginnings led to a dramatic impact on food production in the third world, especially Asia, which would be dubbed the Green Revolution, a term coined in 1968 by the USAID Administrator, William S. Gaud.

The Green Revolution started in 1943 when the Mexican government and the Rockefeller Foundation co-sponsored a project, the Mexican Agricultural Program, to increase food production in Mexico. The first target crop was wheat, and the goal was to increase wheat production by a large margin. Using an interdisciplinary approach, the scientific team headed by Norman Borlaug, a wheat breeder at the Rockefeller Foundation, started to assemble genetic resources (germplasm) of wheat from all over the world (East Africa, Middle East, South Asia, Western Hemisphere). The key genotypes used by Norman Borlaug in his breeding program were the Japanese "Norin" dwarf genotypes supplied by Burton Bayles of the United States Department of Agriculture (USDA) and a segregating ( $F_2$ ) population of "Norin 10" crossed with "Brevor", a Pacific Northwest wheat, supplied by Orville Vogel of the USDA. These introductions were crossed with indigenous (Mexican) wheat that had adaptability (to temperature, photoperiod) to the region and were disease resistant, but were low yielding and prone to lodging. The team was able to develop lodging-resistant cultivars through introgression of dwarf genes from semidwarf cultivars from North America. This breakthrough occurred in 1953. Further crossing and selection resulted in the release of the first Mexican semidwarf cultivars, "Penjamo 62" and "Pitic 62". Together with other cultivars, these two hybrids dramatically transformed wheat yields in Mexico, eventually making Mexico a major wheat exporting country. The successful wheat cultivars were introduced into Pakistan, India, and Turkey in 1966, with similar results of outstanding performance. During the period, wheat production increased from 300,000 to 2.6 million tons/year; yields per unit area increased from 750 to 3,200 kg/ha.

The Mexican model (interdisciplinary approach, international team effort) for agricultural transformation was duplicated in rice in the Philippines in 1960. This occurred at the IRRI. The goal of the IRRI team was to increase productivity of rice in the field. Rice germplasm was assembled. Scientists determined that, like wheat, a dwarf cultivar that was resistant to lodging, amenable to high density crop stand, responsive to fertilization and highly efficient in partitioning of photosynthates or dry matter to the grain, was the cultivar to breed.

In 1966, the IRRI released a number of dwarf rice cultivars to farmers in the Philippines. The most success was realized with IR8, which was early maturing (120 days), thus allowing double cropping in certain regions. The key to the high yield of the IR series was their



## Industry highlights

### *Normal Ernest Borlaug: the man and his passion*

George Acquaah

Department of Agriculture and Natural Resources, Langston University, Langston, OK 73050, USA

*For more than half a century, I have worked with the production of more and better wheat for feeding the hungry world, but wheat is merely a catalyst, a part of the picture. I am interested in the total development of human beings. Only by attacking the whole problem can we raise the standard of living for all people, in all communities, so that they will be able to live decent lives. This is something we want for all people on this planet.*

**Norman E. Borlaug**

Dr Norman E. Borlaug has been described in the literature in many ways, including as "the father of the Green Revolution", "the forgotten benefactor of humanity", "one of the greatest benefactors of human race in modern times", and "a distinguished scientist-philosopher". He has been presented before world leaders and received numerous prestigious academic honors from all over the world. He belongs to an exclusive league with the likes of Henry Kissinger, Elie Wiesel, and President Jimmy Carter – all Nobel Peace laureates. Yet, Dr Borlaug is hardly a household name in the USA. But, this is not a case of a prophet being without honor in his country. It might be more because this outstanding human being chooses to direct the spotlight on his passion, rather than his person. As previously stated in his own words, Dr Borlaug has a passion for helping to achieve a decent living status for the people of the world, starting with the alleviation of hunger. To this end, his theatre of operation is the third world countries, which are characterized by poverty, political instability, chronic food shortages, malnutrition, and the prevalence of preventable diseases. These places are hardly priority sources for news for the first world media, unless an epidemic or catastrophe occurs.

Dr Borlaug was born on March 25, 1914, to Henry and Clara Borlaug, Norwegian immigrants in the city of Saude, near Cresco, Iowa. He holds a BS degree in Forestry, which he earned in 1937. He pursued an MS in Forest Pathology, and later earned a PhD in Pathology and Genetics in 1942 from the University of Minnesota. After a brief stint with the E. I. du Pont de Nemours in Delaware, Dr Borlaug joined the Rockefeller Foundation team in Mexico in 1944, a move that would set him on course to achieve one of the most notable accomplishments in history. He became the director of the Cooperative Wheat Research and Production Program in 1944, a program initiated to develop high-yielding cultivars of wheat for producers in the area.

In 1965, the Centro Internationale de Mejoramiento de Maiz y Trigo (CIMMYT) was established in Mexico, as the second of the currently 16 International Agricultural Research Centers (IARCs) by the Consultative Group on International Agricultural Research (CGIAR). The purpose of the center was to undertake wheat and maize research to meet the production needs of developing countries. Dr Borlaug served as the director of the Wheat Program at CIMMYT until 1979 when he retired from active research, but not until he had accomplished his landmark achievement, dubbed the Green Revolution. The key technological strategies employed by Dr Borlaug and his team were to develop high-yielding varieties of wheat, and an appropriate agronomic package (fertilizer, irrigation, tillage, pest control) for optimizing the yield potential of the varieties. Adopting an interdisciplinary approach, the team assembled germplasm of wheat from all over the world. Key contributors to the efforts included Dr Burton Bayles and Dr Orville Vogel, both of the USDA, who provided the critical genotypes used in the breeding program. These genotypes were crossed with Mexican genotypes to develop lodging-resistant, semidwarf wheat varieties that were adapted to the Mexican production region (Figure 1). Using the improved varieties and appropriate agronomic packages, wheat production in Mexico increased dramatically from its low 750 kg/ha to about 3,200 kg/ha. The successful cultivars were introduced into other parts of the world, including Pakistan, India, and Turkey in 1966, with equally dramatic results. So successful was the effort in wheat that the model was duplicated in rice in the Philippines in 1960. In 1970, Dr Norman Borlaug was honored with the Nobel Peace Prize for contributing to curbing hunger in Asia and other parts of the world where his improved wheat varieties were introduced (Figure 2).

Whereas the Green Revolution was a life-saver for countries in Asia and some Latin American countries, another part of the world that is plagued by periodic food shortages, the sub-Saharan Africa, did not benefit from this event. After retiring from CIMMYT in 1979, Dr Borlaug focused his energies on alleviating hunger and promoting the general well-being of the people on the continent of Africa. Unfortunately, this time around, he had to go without the support of these traditional allies, the Ford Foundation, the Rockefeller Foundation, and the World Bank. It appeared the activism of powerful environmental groups in the developed world had managed to persuade these donors from supporting what, in their view, was an environmentally intrusive practice advocated by people such as Dr Borlaug. These environmentalists promoted the notion that high-yield agriculture for Africa, where the agronomic package included inorganic fertilizers, would be ecologically disastrous.

Incensed by the distractions of "green politics", which sometimes is conducted in an elitist fashion, Dr Borlaug decided to press on undeterred with his passion to help African farmers. At about the same time, President Jimmy Carter was collaborating with the



**Figure 1** Dr Norman Borlaug working in a wheat crossing block.

late Japanese industrialist, Ryoichi Sasakawa, in addressing some of the same agricultural issues dear to Dr Borlaug. In 1984, Mr Sasakawa persuaded Dr Borlaug to come out of retirement to join them to vigorously pursue food production in Africa. This alliance gave birth to the Sasakawa Africa Association, presided over by Dr Borlaug. In conjunction with Global 2000 of The Carter Center, Sasakawa-Global 2000 was born, with a mission to help small-scale farmers to improve agricultural productivity and crop quality in Africa. Without wasting time, Dr Borlaug selected an initial set of countries in which to run projects. These included Ethiopia, Ghana, Nigeria, Sudan, Tanzania, and Benin (Figure 3). The crops targeted included popular staples such as corn, cassava, sorghum, and cowpeas, as well as wheat. The most spectacular success was realized in Ethiopia, where the country recorded its highest ever yield of major crops in the 1995–1996 growing season.

Sasakawa-Global 2000 operates in some 12 African nations. Dr Borlaug is still associated with CIMMYT and also holds a faculty position at Texas A&M University, where he teaches international agriculture in the fall semester. On March 29, 2004, in commemoration of his 90th birthday, Dr Borlaug was honored by the USDA with the establishment of the Norman E. Borlaug International Science and Technology Fellowship Program. The fellowship is designed to bring junior and mid-ranking scientists and policy-makers from African, Asian, and Latin American countries to the United States to learn from their US counterparts.



DET NORSKE STORTINGS  
NOBELKOMITÉ  
HAR OVERENSSTEMMENDE MED  
REGLENE I DET AV  
**ALFRED NOBEL**  
DEN 27. NOVEMBER 1895 OPPRETTEDE  
TESTAMENTE TILDELT  
NORMAN ERNEST BORLAUG  
NOBELS FREDSPRIS FOR 1970

OSLO 10. DESEMBER 1970  
Hele en Rektorat i full etterskrift  
Gjennomført

**Figure 2** A copy of the actual certificate presented to Dr Norman Borlaug as part of the 1970 Nobel Peace Prize Award he received.



**Figure 3** Dr Twumasi Afriyie, CIMMYT Highland Maize Breeder and a native of Ghana, discusses the quality protein maize he was evaluating in a farmer's field in Ghana with Dr Borlaug.

#### Further reading

- Borlaug, N.E. 1958. The impact of agricultural research on Mexican wheat production. *Trans. New York Acad. Sci.* 20:278–295.
- Borlaug, N.E. 1965. Wheat, rust, and people. *Phytopathology* 55:1088–1098.
- Borlaug, N.E. 1968. Wheat breeding and its impact on world food supply. Public lecture at the 3rd International Wheat Genetics Symposium, August 5–9, 1968. Australian Academy of Science, Canberra, Australia.
- Brown, L.R. 1970. Seeds of change: The Green Revolution and development in the 1970s. Praeger, New York.
- Byerlee, D., and P. Moya. 1993. Impacts of international wheat breeding research in the developing world. Mexico CIMMYT, El Batán, Mexico.
- Dalrymple, D.G. 1986. Development and spread of high-yielding rice varieties in developing countries. Agency for International Development, Washington, DC.
- Haberman, F.W. 1972. Nobel lectures, 1951–1970. Nobel lectures, peace. Elsevier Publishing Company, Amsterdam.
- Wharton, C.R. Jr. 1969. The Green Revolution: cornucopia or Pandora's box? *Foreign Affairs* 47:464–476.

responsiveness to heavy fertilization. The short, stiff stalk of the improved dwarf cultivar resisted lodging under heavy fertilization. Unimproved indigenous genotypes experienced severe lodging under heavy fertilization, resulting in drastic reduction in grain yield.

Similarly, cereal production in Asia doubled between 1970 and 1995, as the population increased by 60%. Unfortunately, the benefits of the Green Revolution barely reached sub-Saharan Africa, a region of the world with perennial severe food shortages, partly because of the lack of appropriate infrastructure and limited resources. Dr Norman Borlaug received the 1970 Nobel Prize for Peace for his efforts at curbing global hunger.

Three specific strategies were employed in the Green Revolution:

- 1 **Plant improvement.** The Green Revolution centered on the breeding of high-yielding, disease-resistant, and environmentally responsive (adapted, responsive to fertilizer, irrigation, etc.) cultivars.
- 2 **Complementary agronomic package.** Improved cultivars are as good as their environment. To realize the full potential of the newly created genotype, a certain production package was developed to complement the improved genotype. This agronomic package included tillage, fertilization, irrigation, and pest control.

**3 Favorable returns on investment in technology.** A favorable ratio between the cost of fertilizer and other inputs and the price the farmer received for using this product was an incentive for farmers to adopt the production package.

Not unexpectedly, the Green Revolution has been the subject of some intensive discussion to assess its socio-logical impacts and identify its shortcomings. Incomes of farm families were raised, leading to an increase in demand for goods and services. The rural economy was energized. Food prices dropped. Poverty declined as agricultural growth increased. However, critics charge that the increase in income was inequitable, arguing that the technology package was not scale neutral (i.e., owners of larger farms were the primary adopters because of their access to production inputs – capital, seed, irrigation, fertilizers, etc.). Furthermore, the Green Revolution did not escape the accusations often leveled at high-yielding agriculture – environmental degradation from improper or excessive use of

agrochemicals. Recent studies have shown that many of these charges are overstated.

### **Future of plant breeding in society**

For as long as the world population is expected to continue to increase, there will continue to be a demand for more food. However, with an increasing population comes an increasing demand for land for residential, commercial, and recreational uses. Sometimes, farm lands are converted to other uses. Increased food production may be achieved by increasing production per unit area or bringing new lands into cultivation. Some of the ways in which society will affect and be affected by plant breeding in the future are as follow:

- 1 New roles of plant breeding.** The traditional roles of plant breeding (food, feed, fiber, and ornamentals) will continue to be important. However, new roles are gradually emerging for plants. The technology for using plants as bioreactors to produce pharmaceuticals will advance; this technology has been around for over a decade. Strategies are being perfected for use of plants to generate pharmaceutical antibodies, engineering antibody-mediated pathogen resistance, and altering plant phenotypes by immunomodulation. Successes that have been achieved include the incorporation of *Streptococcus* surface antigen in tobacco, and the herpes simplex virus in soybean and rice.
- 2 New tools for plant breeding.** New tools will be developed for plant breeders, especially, in the areas of the application of biotechnology to plant breeding. New marker technologies continue to be developed and older ones advanced. Tools that will assist breeders to more effectively manipulate quantitative traits will be enhanced.
- 3 Training of plant breeders.** As discussed elsewhere in the book, plant breeding programs have experi-

enced a slight decline in graduates in recent past. Because of the increasing role of biotechnology in plant genetic manipulation, graduates who combine skills and knowledge in both conventional and molecular technologies are in high demand. It has been observed that some commercial plant breeding companies prefer to hire graduates with training in molecular genetics, and then provide them with the needed plant breeding skills on the job.

- 4 The key players in plant breeding industry.** The last decade saw a fierce race by multinational pharmaceutical corporations to acquire seed companies. There were several key mergers as well. The modern technologies of plant breeding are concentrated in the hands of a few of these giant companies. The trend of acquisition and mergers are likely to continue in the future.
- 5 Yield gains of crops.** With the dwindling of arable land and the increase in policing of the environment by activists, there is an increasing need to produce more food or other crop products on the same piece of land in a more efficient and environmentally safer manner. High-yielding cultivars will continue to be developed, especially in crops that have received less attention from plant breeders. Breeding for adaptation to environmental stresses (e.g., drought, salt) will continue to be important, and will enable more food to be produced on marginal lands.
- 6 The biotechnology debate.** It is often said that these modern technologies for plant genetic manipulation benefit the developing countries the most since they are in dire need of food, both in quantity and nutritional value. On the other hand, the intellectual property that covers these technologies is owned by the giant multinational corporations. Efforts will continue to be made to negotiate fair use of these technologies. Appropriate technology transfer and support to the poor third world nations will continue, to enable them to develop capacity for the exploitation of these modern technologies.

### **References and suggested reading**

- Charles, D., and B. Wilcox. 2002. Lords of the harvest: Biotechnology, big money and the future of food. Perseus Publishing, Cambridge, MA.
- Duvick, D.N. 1986. Plant breeding: Past achievements and expectations for the future. Econ. Bot. 40:289–297.
- Frey, K.J. 1971. Improving crop yields through plant breeding. Am. Soc. Agron. Spec. Publ. 20:15–58.
- International Food Policy Research Institute. 2002. Green Revolution – Curse or blessing? IFPRI, Washington, DC.
- Solheim, W.G., II. 1972. An earlier agricultural revolution. Sci. Am. 226(4):34–41.
- Zirkle, C. 1932. Some forgotten records of hybridization and sex in plants. J. Heredity 23:443–448.

**Outcomes assessment****Part A**

Please answer the following questions true or false:

- 1 Plant breeding causes permanent changes in plant heredity.
- 2 Rice varieties were the first products of the experiments leading to the Green Revolution.
- 3 Rice is high in pro-vitamin A.
- 4 The IR8 was the rice variety released as part of the Green Revolution.
- 5 Wilhelm Johannsen developed the pure-line theory.

**Part B**

Please answer the following questions:

- 1 ..... won the Nobel Peace Prize in ..... for being the chief architect of the .....
- 2 Define plant breeding.
- 3 Give three specific objectives of plant breeding.
- 4 Discuss plant breeding before Mendel's work was discovered.
- 5 Give the first two major wheat cultivars to come out of the Mexican Agricultural Program initiated in 1943.

**Part C**

Please write a brief essay on each of the following topics:

- 1 Plant breeding is an art and a science. Discuss.
- 2 Discuss the importance of plant breeding to society.
- 3 Discuss how plant breeding has changed through the ages.
- 4 Discuss the role of plant breeding in the Green Revolution.
- 5 Discuss the impact of plant breeding on crop yield.
- 6 Plant breeding is critical to the survival of modern society. Discuss.

# **Section 2**

## **General biological concepts**

### **Chapter 2 The art and science of plant breeding**

### **Chapter 3 Plant cellular organization and genetic structure: an overview**

### **Chapter 4 Plant reproductive systems**

---

This section introduces the student to the fundamental concepts and rationale of plant genetic manipulation. It is instructive for the student to have an overview of the historical perspectives of plant genetic manipulation, how it all began and how things have changed (or stayed the same) over time, hence the discussion of domestication, evolution, and their relationship to plant breeding. Before attempting to genetically manipulate plants, it is important to understand their fundamental biological and genetic structure. It is important to know the natural tendencies of plants before attempting to modify their behaviors. Plant breeders seek to make modifications in plants that are permanent and can be inherited from generation to generation. Modern technologies allow plant breeders to manipulate plants at all levels of biological organization, from molecular, to cellular, to whole-plant levels. The student needs to understand cellular structure as well as DNA structure. In addition, the student should understand basic plant reproductive biology, for it is not only an avenue for genetic manipulation of sexually reproducing plants, but also it is the means by which materials are increased for release to producers.

---



# 2

## The art and science of plant breeding

---

### Purpose and expected outcomes

*As indicated in Chapter 1, plant breeders want to cause specific and permanent alterations in the plants of interest. They use various technologies and methodologies to accomplish their objectives. Certain natural processes can also cause permanent genetic changes to occur in plants. In this chapter, three processes that bring about such heritable changes – evolution, domestication, and plant breeding – are discussed, drawing parallels among them and pointing out key differences. After studying this chapter, the student should be able to:*

- 1 Define the terms evolution, domestication, and plant breeding.
  - 2 Discuss the impact of domestication on plants.
  - 3 Compare and contrast domestication and evolution.
  - 4 Compare and contrast evolution and plant breeding.
  - 5 Discuss plant breeding as an art.
  - 6 Discuss plant breeding as a science.
  - 7 Present a brief overview of the plant breeding industry.
- 

### Concept of evolution

**Evolution** is a population phenomenon. Populations, not individuals, evolve. Evolution is concerned with the effect of changes in the frequency of alleles within a gene pool of a population, such changes leading to changes in genetic diversity and the ability of the population to undergo evolutionary divergence. Simply stated, evolution is descent with modification. Proposed by Charles Darwin in 1859, there are certain key features of the concept or **theory of evolution**. Variation exists in the initial population of organisms, both plants and animals. As Darwin stated, variation is a feature of natural populations. More individuals are produced each generation than can be supported by or survive in the environment. Environmental stresses place certain individuals in the population at a disadvantage. The individuals with the best genetic fitness for the specific environment will

survive and reproduce more successfully and become more competitive than other individuals. The more competitive individuals will leave more offspring to participate in the next generation. Such a trend, where the advantageous traits increase, will continue each generation, with the result that the population will be dominated by these favored individuals and is said to have evolved. The discriminating force, called **natural selection** by Darwin, is the final arbiter in deciding which individuals are advanced. When individuals in the original population become reproductively isolated, new species will eventually form.

Patterns of such evolutionary changes have been identified and exploited by plant breeders in the development of new cultivars. Scientists have been able to identify relationships between modern cultivars and their wild and weedy progenitors. Further, adaptive variations in geographic races of crops have been

discovered. As will be discussed in detail later in the book, scientists collect, process, and store this natural variation in germplasm banks for use by breeders in their breeding programs.

The process of evolution has parallels in plant breeding. Darwin's theory of evolution through natural selection can be summed up in three principles that are at the core of plant breeding. These are the principles of:

- 1 Variation.** Variation in morphology, physiology, and behavior exist among individuals in a natural population.
- 2 Heredity.** Offspring resemble their parents more than they resemble unrelated individuals.
- 3 Selection.** Some individuals in a group are more capable of surviving and reproducing than others (i.e., more fit).

A key factor in evolution is time. The changes in evolution occur over *extremely long* periods of time.

Plant breeders depend on biological variation as a source of desired alleles. Induced mutation and hybridization for recombination are major sources of variation. Once variation has been assembled, the breeder imposes a selection pressure (**artificial selection** in this case) to discriminate among the variation to advance only desired plants. Plant breeding may be described as directed or targeted and accelerated evolution, because the plant breeder, with a breeding objective in mind, deliberately and genetically manipulates plants (wild or domesticated) to achieve a stated goal, but in a *very short* time. Conceptually, breeding and evolution are the same, a key difference being the duration of the processes. Plant breeding has been described as evolution directed by humans. Compared to evolution, a plant breeding process is completed in a twinkle of an eye! Also, unlike evolution, plant breeders do not deal with closed populations. They introgress new variability from different genotypes of interest, and, for practical and economic purposes, deal with limited population sizes.

## Domestication

**Domestication** is the process by which genetic changes (or shifts) in wild plants are brought about through a selection process imposed by humans. It is an evolutionary process in which selection (both natural and artificial) operates to change plants genetically, morphologically, and physiologically. The results of domestication

are plants that are adapted to supervised cultural conditions, and possessing characteristics that are preferred by producers and consumers. In some ways, a domesticated plant may be likened to a tamed wild animal that has become a pet.

There are degrees of domestication. Species that become completely domesticated often are unable to survive when reintroduced into the wild. This is so because the selection process that drives domestication strips plants of natural adaptive features and mechanisms that are critical for survival in the wild, but undesirable according to the needs of humans.

Like evolution, domestication is also a process of genetic change in which a population of plants can experience a shift in its genetic structure in the direction of selection imposed by the domesticator. New plant types are continually selected for as domesticates as new demands are imposed, thereby gradually moving the selected individuals farther away (genetically, morphologically, and physiologically) from their wild progenitors. Both wild and domesticated populations are subject to evolution.

### Patterns of plant domestication

Domestication has been conducted for over 10,000 years, and ever since agriculture was invented. Archeological and historical records provide some indications as to the period certain crops may have been domesticated, even though such data are not precise. Archeological records from arid regions are better preserved than those from the humid regions of the world.

### Concepts of domestication

As G. Ladizinsky points out in discussing patterns of domestication, the challenge is to determine whether the domesticate evolved under wild conditions, or was discovered and then cultivated by humans, or whether cultivation preceded the selection of domesticates. This is a subject of debate. For example, seed dormancy is a problem in wild legumes, and hence would have hindered their use in cultivation. It is likely that the domesticates evolved in the wild before being used in cultivation. However, in most cereal species, most experts believe that domestication occurred after cultivation. In wheat and barley, for example, a tough rachis, which is resistant to natural seed dispersal, and characterizes domesticates, would have been selected for during cultivation.

Two categories of crop plants are identified, with respect to domestication, as **primary crops** or **secondary**

**crops.** Primary crops are those whose wild progenitors were deliberately cultivated by humans, genetic changes occurring in their new environments. Secondary crops are those that evolved from weeds that arose in cultivated fields. For example, the common oat (*Avena sativa*) evolved from the hexaploid wild oats (*A. sterilis* and *A. fatua*). The domestication of vegetables, root and tuber crops, and most fruit trees is described as gradual domestication. This is because it is difficult to use a single characteristic to differentiate between wild and cultivated species of these horticultural plants. These crops are commonly vegetatively propagated, hence evolution under cultivation would occur mainly from variation originating from somatic mutations. Seed crops have the advantage of genetic recombination through sexual reproduction to create new variability more rapidly.

#### *Centers of plant domestication*

**Centers of plant domestication** are of interest to researchers from different disciplines including botany, genetics, archeology, anthropology, and plant breeding. Plant breeders are interested in centers of plant domestication as regions of genetic diversity, variability being critical to the success of crop improvement. De Candolle was the first to suggest in 1886 that a crop plant originates from the area where its wild progenitor occurs. He considered archeological evidence to be the direct proof of the ancient existence of a crop species in a geographic area.

Several scientists, notably N. Vavilov of Russia and J. R. Harlan of the USA have provided the two most enduring views of plant domestication. Vavilov, on his plant explorations around the world in the 1920s and 1930s, noticed that extensive genetic variability within a crop species occurred in clusters within small geographic regions separated by geographic features such as mountains, rivers, and deserts. For example, whereas he found different forms of diploid, tetraploid, and hexaploid species of wheat in the Middle East, he observed that only hexaploid cultivars were grown in Europe and Asia. Vavilov proposed the concept of **centers of diversity** to summarize his observations. He defined the **center of origin of a crop plant** as the geographic area(s) where it exhibits maximum diversity (i.e., where the greatest number of races and botanical varieties occur). He identified eight major centers of diversity, some of which were subdivided (subcenters). These centers, with examples of associated plants, were:

- 1 China (e.g., lettuce, rhubarb, soybean, turnip).
- 2 India (e.g., cucumber, mango, rice, oriental cotton).
- 2a Indochina (e.g., banana, coconut, rice).
- 3 Central Asia (north India, Afghanistan, Turkmenistan) (e.g., almond, flax, lentil).
- 4 The Near East (e.g., alfalfa, apple, cabbage, rye).
- 5 Mediterranean Sea, coastal and adjacent regions (e.g., celery, chickpea, durum wheat).
- 6 Ethiopia (e.g., coffee, grain sorghum, pearl millet).
- 7 Southern Mexico and Middle America (e.g., lima bean, maize, papaya, upland cotton).
- 8 Northeastern South America, Bolivia, Ecuador, and Peru (e.g., Egyptian cotton, potato, tomato).
- 8a Isles of Chile (e.g., potato).

Furthermore, he associated over 500 Old World crops and about 100 New World crops with these centers. Most (over 400) of the Old World crops were located in Southern Asia.

Vavilov noticed that even though one species or one genus was associated with a center of diversity, often it occurred also at a few other centers. However, whenever this was the case, the types were often distinguishable from place to place. He called the centers where maximum diversity occurs **primary centers**, and the places where types migrate to, the **secondary centers**. For example, the primary center of corn is Mexico, but China is a secondary center of waxy types of corn. Vavilov associated these centers of diversity with the centers of origin of these crops, proposing that the variability was *predominantly* caused by mutations and their accumulation in the species over a long period of time. These variations were preserved through the domestication process.

Other scientists of that era, notably Jack Harlan, disagreed with the association of centers of diversity with the centers of crop origin. He argued that the origin of a cultivated plant was diffuse both in time and space. This opposing view was arrived at from his observations that plant diversity appeared to exhibit hybrid features, indicating they likely arose from recombination (i.e., centers of recombination). He proposed the new concept of centers and non-centers as summarized below:

Centers	Non-centers
(Temperate and geographically restricted)	(Corresponding tropical areas)
A1 Near East	A2 Africa
B1 North China	B2 South East Asia
C1 Mesoamerica	C2 South America

Each center had a corresponding non-center. The centers contained wild relatives of many crop plants, whose

antiquity is established by archeological evidence. It is from these centers that the crops diffused to their geographically less restricted corresponding non-centers. Other scientists including C. D. Darlington and I. H. Burkhill suggested that some variability could be attributed to shifts in civilizations that brought about migrations of crops, changes in selection pressure, and opportunities for recombination.

Vavilov made other unique observations from his plant explorations. He found that the maximum amount of variability and the maximum concentration of dominant genes for crops occurred at the center and decreased toward the periphery of the cluster of diversity. Also, he discovered there were parallelisms (common features) in variability among related species and genera. For example, various cotton species, *Gossypium hirsute* and *G. barbadense*, have similar pubescence, fiber

color, type of branching, color of stem, and other features. Vavilov called this the **law of homologous series** in heritable variation (or parallel variation). In other words, species and genera that are genetically closely related are usually characterized by a similar series of heritable variations such that it is possible to predict what parallel forms would occur in one species or genera, from observing the series of forms in another related species. The breeding implication is that if a desirable gene is found in one species, it likely would occur in another related species. Through comparative genomic studies, the mapping of molecular markers has revealed significant homology regarding the chromosomal location of DNA markers among species of the Poaceae family (specifically, rice, corn, sorghum, barley, wheat), a condition called **synteny**, the existence of highly conserved genetic regions of the chromosome.



## Industry highlights

### *Introduction and adaptation of new crops*

Jaime Prohens, Adrián Rodríguez-Burrueto, and Fernando Nuez

Instituto para la Conservación y Mejora de la Agrodiversidad Valenciana, Universidad Politécnica de Valencia, 46022 Valencia, Spain

*The greatest service which can be rendered any country is to add a useful plant to its culture.*

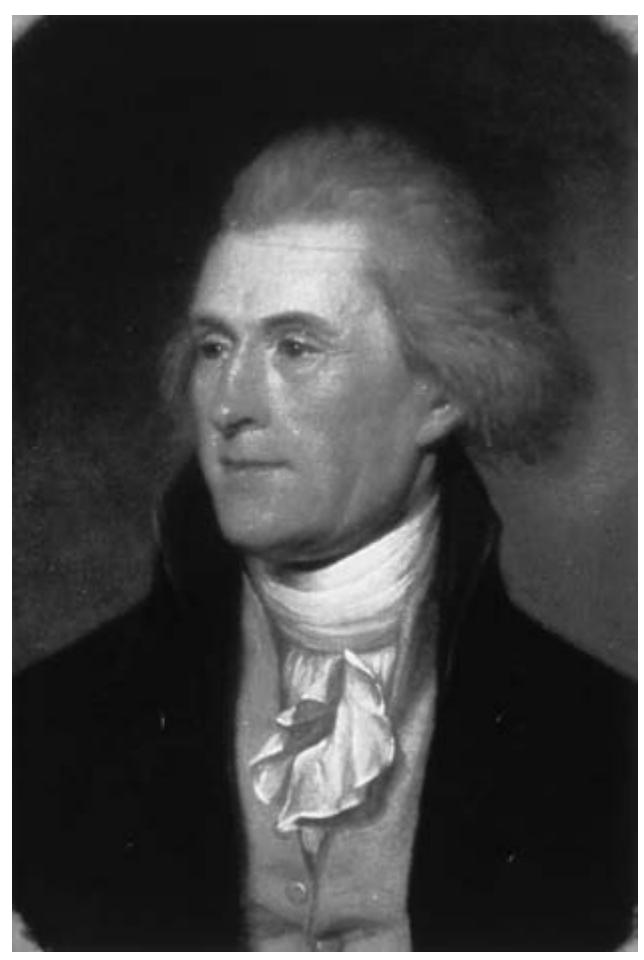
**Thomas Jefferson (c. 1800; Figure 1)**

Since the domestication of the first crops, societies that practice agriculture have been attracted to new crops because they present opportunities for improving crop production and food supply. In fact, most of the relevant crops grown in a particular region are usually native to other regions. Thus, any cultivated species grown in an area different to its center of origin was, at one time, a new crop. Just to cite a few examples, soybean, wheat, rice, beans, tomato, or citrus, which are important crops in Europe and USA are not native to these regions.

Diversification of crop production through the introduction of new crops is desirable for several reasons. New crops represent an alternative to growers and markets with produces that have a high value and for which usually there is no overproduction. They also may contribute to a sustainable horticulture because an increase in diversity reduces the problems caused by pests and diseases caused by monocrop and allows a higher efficiency in the use of production factors. A greater diversity of crops also favors the stability of production and growers' incomes because the cultivation of a higher number of species decreases risks against unpredictable environmental and market changes. Finally, new crops contribute in improving ethnobotanical knowledge, which is a substantial part of folk culture.

Historically, the introduction of new crops has taken place thanks to the movement of plant material through trade routes or by contacts among cultures. The discovery of America was one of the most important events in the adaptation of new crops, which resulted in an enormous exchange of species between the Old World and the New World. Nowadays it is estimated that 40% of economically relevant crops originated in America, and it is difficult to imagine the present Old World's culture and gastronomy without many American-originated crops. For example, corn, sunflower, potato, tobacco, peanut, cocoa, beans, squash, pumpkin and gourds, tomato, capsicum pepper, and many others originated in the New World and all of them were "new crops" in the Old World a few centuries ago. On the other hand, many Old World crops adapted well in America and this continent has become the main producing area for some of them, e.g. soybean (from China), coffee (from Africa and Arabia), or banana (from South East Asia).

A great effort in the attempt to adapt foreign species took place during the 18th and 19th centuries. There were several outstanding stories in this endeavor, such as the establishment of rubber plantations in South East Asia, after seeds and plants were smuggled from Amazon plantations; the expeditions in search of breadfruit, which is native to Polynesia and



**Figure 1** Thomas Jefferson, third president of the USA, and a great promoter of the introduction of new crops.

the new crop, like a satisfactory performance under the new agroclimatic conditions and an easy adaptation to the cultural practices commonly used in the cultivation of the main crops of the new region. Growers will be attracted to a new crop if it adapts well to the existing crop.

There are few cases of immediate success in the introduction of new crops. In this respect, many crops were introduced into the Old World after the discovery of America, although their acceptance differed and some of them did not succeed at first. For example, *Capsicum* pepper had an early acceptance and its cultivation was fully established a few years after being introduced. At that time, hot peppers became an alternative to black pepper and that surely contributed to its fast worldwide spread. On the contrary, tomato needed much more time before being fully accepted. It was brought into Europe a few years after the discovery of America. However, although there was some consumption in Spain and Italy, the rest of European countries rejected it (perhaps because of its red colored skin, usually an indication of toxicity in nature, and also because many Old World Solanaceae are toxic) and it was just used as an ornamental until the 19th century.

Nowadays, scientific and technological advances can make the introduction of a new crop a much shorter process than centuries ago because of our knowledge in genetics, breeding, biotechnology, plant physiology, pathology, and other disciplines. Breeding for adaptation has been a research field that has had a tremendous impact in the success of the introduction of new crops. For example, the selection and development of materials insensitive to the photoperiod has allowed the introduction of wheat into tropical areas. Also, adapted materials resistant to colder or warmer conditions, or shorter growing seasons, have been obtained in several crops by a gradual and long process of progressive adaptation. For example, in corn – a tropical plant – the natural and artificial selection on genetically diverse populations has allowed its

was going to be a food supply for slaves in the West Indies (Figure 2), and was described in the famous Bounty mutiny (brought to the cinema in the famous film *Mutiny on the Bounty*); or the introduction of cinchona in the colonies of Africa and India from South America, due to the medicinal importance of quinine, obtained from cinchona bark, against malaria.

Throughout history, the introduction of new crops has contributed to an increase in the diversity of the plants cultivated; however, the trend during the last century, associated with industrial agriculture, has led to a reduction in the number of crops grown. In this respect, although around 3,000 species are known to have been used as a source of food by humans, at present only 11 species (wheat, rice, corn, barley, sorghum, millet, potato, sweet potato, yam, sugarcane, soybean) contribute more than 75% of world human food supply. More worryingly, 60% of the calories consumed in the world are based in just three crops (rice, corn, wheat), and the trend is towards a concentration of production in fewer and fewer crops.

Among the huge number of domesticated species, there are many little-known species that only have local relevance or have been neglected that could be very interesting as “new crops”. Although the denomination “new crop” seems to be more appropriate for recently domesticated plants, it usually refers to exotic crops. Actually, most of these “new crops” were domesticated thousands years ago, although there are examples of recent domestication (in the 19th and 20th centuries) such as several berries belonging to the genus *Rubus* that are currently being introduced and improved in Europe.

Not all crops have the same opportunities of succeeding when introduced in a certain region. Success will depend on several characteristics of



**Figure 2** The breadfruit (*Artocarpus altilis*), a crop that was the subject of some fascinating expeditions to the Polynesian islands in order to obtain propagation material to introduce it as a new crop in the Western Indies.

The next step is to conduct preliminary field plot research. The goal is to test or develop genotypes or varieties with satisfactory adaptation and to obtain basic information about the production practices and pests and diseases affecting the new crop. A critical aspect deals with the use of sufficient genetic variation in the trials. Many attempts to adapt a new crop to a new region have failed because of the use of limited genetic variation (one or two cultivars). In this way, different genotypes show different behaviors under the same environmental conditions, and this may allow for the selection of individuals or populations with the most satisfactory behavior (i.e., exploiting genotype  $\times$  environment interaction) either for direct cultivation or as a starting point for breeding programs. Another key point is identifying growing techniques that can improve the productive potential of the new crop.

After this, a more extensive evaluation should be conducted. This usually needs the involvement of growers and industry and the technical assistance of research centers. Basically, it deals with trials to evaluate the performance of adapted plant material at different locations of the potential production area, as well conducting postharvest research and marketing studies in order to determine the best marketing channels. Finally, if results are promising, the product can be released.

The development of a new crop is a slow and complex process with uncertain results. Several cases show that investment in the introduction and adaptation of new crops may be highly profitable and returns in new crop research are, as a whole, many times higher than the investment. The introduction of soybean in the USA from China is the story of one such success. Nowadays, the USA is the main producer of soybean in the world. This plant was introduced in the 18th century and its interest as a crop began at the end of the 19th century in several agricultural experiment stations. The development of soybean as a new crop cost American taxpayers US\$5 million from 1912 to 1941. However, US soybean export trade in 2000 alone was estimated at \$6.6 billion. Another example comes from kiwifruit introduction into New Zealand. This exotic and half-domesticated plant was first introduced into New Zealand from Chinese forests at the beginning of the 20th century and was cultivated as an ornamental until the 1950s. Finally, New Zealand growers decided to exploit its potential as an exotic fruit in the 1960s and 1970s. From that moment on, this crop has provided "kiwi" growers with very high profits, particularly in the 1970s and 1980s, when kiwifruit production and marketing were performed exclusively by New Zealand. Currently, kiwifruit is the biggest horticultural export in New Zealand with a total value of about NZ\$600 million (US\$250 million). These are only two examples of how research on new crops has been very profitable, but many others exist.

#### *Further reading*

- Janick, J. (ed.). 1996. Progress in new crops. ASHS Press, Arlington, VA.  
National Research Council. 1989. Lost crops of the Incas: Little-known plants of the Andes with promise for worldwide cultivation. National Academy Press, Washington, DC. 428 pp.  
Vietmeyer, N.D. 1986. Lesser-known plants of potential use in agriculture and forestry. Science 232:1379–1384.

cultivation in areas as far north as Canada or Scandinavian countries, which have a very short warm season. After several years in one experimental or breeding station, adapted populations can be moved northwards for adaptation to a shorter summer. In this way, new varieties of corn, adapted to these new environments, and with a very short cycle length, have been developed. Many other examples exist, and the gardens or experimental stations for adaptation have played a major role in the successful introduction of new crops.

Not all crops have the same possibilities of being introduced as a new crop into a region or country. The introduction needs a previous study of the suitability of the new crop to the new conditions. It is essential to evaluate its adaptation to agroecological conditions and the potential market, and it is also important to collect information on the management of the crop in its region of origin. All this information will be useful in identifying potential growing areas because, frequently, a crop displays its optimum performance under a limited range of environmental conditions.

### Roll call of domesticated plants

It is estimated that 230 crops have been domesticated, belonging to 180 genera and 64 families. Some families, such as Gramineae (Poaceae), Leguminosae (Fabaceae), Cruciferae, and Solanaceae, have yielded more domesticates than others. Further, culture plays a role in the types of crops that are domesticated. For example, the major world tuber and root crops – Irish potato, sweet potato, yam, cassava, and aroids – have similar cultural uses or purposes but represent distinct taxonomic groups. Four general periods of domestication were proposed by N. W. Simmonds as: (i) ancient (7000–5000 BC); (ii) early (5000–0 BC); (iii) late (AD 0–1750); and (iv) recent (after AD 1750). Early domesticates were made by peasant farmers who selected and advanced desirable plants suited to their cultural practices and food needs.

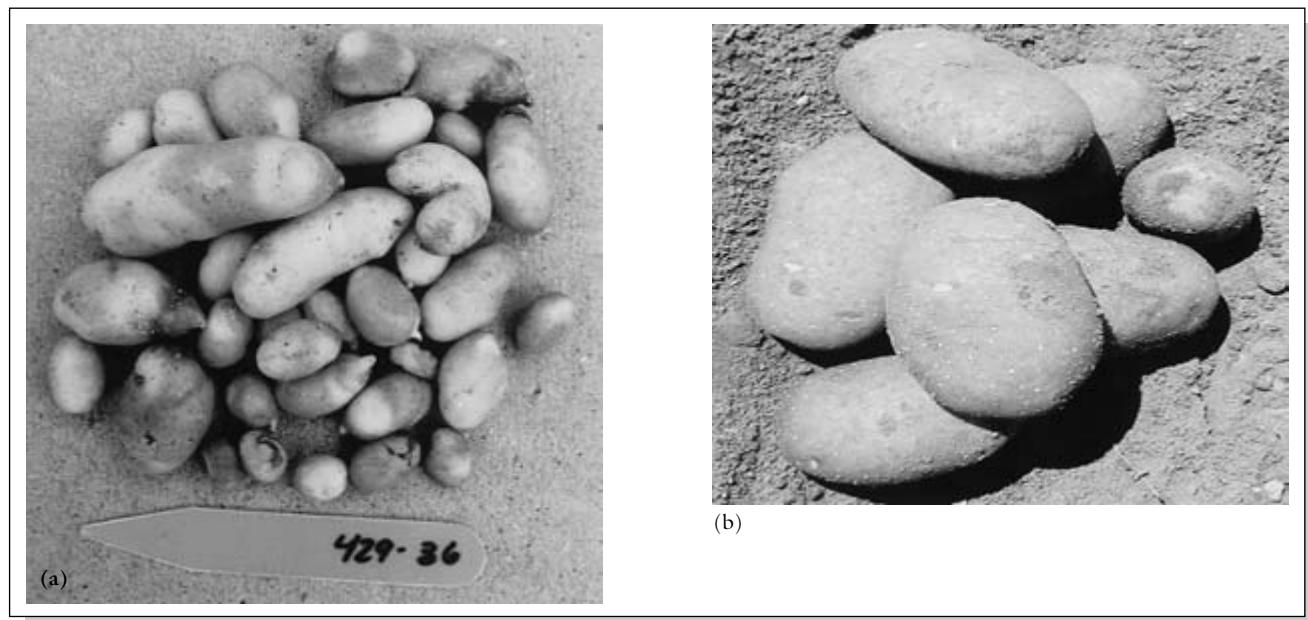
### Changes accompanying domestication

Selection exerted by humans on crop plants during the domestication process causes changes in the plants as they transit from wild species to domesticates (Figure 2.1). The assortments of morphological and physiological

traits that are modified in the process and differentiate between the two types of plants were collectively called the **domestication syndrome** by J. R. Harlan. Although the exact composition of the domestication syndrome traits depends on the particular species, certain basic characteristics are common (Table 2.1.). These traits are selected at three stages in the domestication process – seedling, reproductive, and at or after harvest.

At the seedling stage, the goal of domestication is to get more seeds to germinate. This entails a loss of seed dormancy as well as increased seedling vigor. At the reproductive stage, the goal of domestication includes a capacity for vegetative reproduction and increased selfing rate. Plant traits modified at harvest or after the harvest stage include elimination of seed dispersal (no shattering), uniform seed maturity, more compact plant architecture, and modification in photoperiod sensitivity. Modifications targeted at the consumer include fruit size, color, taste, and reduction in toxic substances.

The genetic control of the traits comprising the domestication syndrome has been studied in many crops. Generally, these traits are controlled by a few qualitative genes or quantitative genes with major phenotypic effects. For example, quantitative trait locus (QTL) research has indicated that a few loci control



**Figure 2.1** Tubers of domesticated tuberous species are larger and have well-defined shape, as compared to their wild ancestors as shown in these photos of a) wild potato and b) domesticated potato. (Courtesy of Jonathan Withworth, USDA-ARS, University of Idaho, and Peggy Bain, University of Idaho, respectively.)

**Table 2.1** Characteristics of domestication syndrome traits.

General effect	Specific traits altered
Increased seedling vigor (more plants germinating)	Loss of seed or tuber dormancy Large seeds
Modified reproductive system	Increased selfing Vegetatively reproducing plants Altered photoperiod sensitivity
Increased number of seeds harvested	Non-shattering Reduced number of branches (more fruits per branch)
Increased appeal to consumers	Attractive fruit/seed colors and patterns Enhanced flavor, texture, and taste of seeds/fruits/tubers (food parts) Reduced toxic principles (safer food) Larger fruits Reduced spikiness
Altered plant architecture and growth habit	Compact growth habit (determinacy, reduced plant size, dwarfism) Reduced branching

traits such as flowering time, seed size, and seed dispersal in maize, rice, and sorghum; and growth habit, photoperiod sensitivity, and dormancy in common bean. Furthermore, linkage blocks of adaptation traits have been found in some species. A study by E. M. K. Koinange and collaborators indicated that the domestication syndrome genes in common bean were primarily clustered in three genomic locations, one for growth habit and flowering time, a second for seed dispersal and dormancy, and a third for pod and seed size.

The domestication process essentially makes plants more dependent on humans for survival. Consequently, a difference between domesticates and their wild progenitors is the lack of traits that ensure survival in the wild. Such traits include dehiscence, dormancy, and thorns. Plants that dehisce their seeds can invade new areas for competitive advantage. However, in modern cultivation, dehiscence or shattering is undesirable because seeds are lost to harvesting when it occurs. Some directions in the changes in plant domesticates have been dictated by the preferences of consumers. Wild tomato (*Pinpenifolium*) produces numerous tiny and hard fruits that are advantageous in the wild for survival. However, consumers prefer succulent and juicy fruits. Consequently, domesticated tomato (whether small or large fruited) is juicy and succulent. Thorns protect against predators in the wild, but are a nuisance to modern uses of plants. Hence, varieties of ornamentals such as roses that are grown for cut flowers are thornless.

## The art and science of plant breeding

The early domesticators relied solely on experience and intuition to select and advance plants they thought had superior qualities. As knowledge abounds and technology advances, modern breeders are increasingly depending on science to take the guesswork out of the selection process, or at least to reduce it. At the minimum, a plant breeder should have a good understanding of genetics and the principles and concepts of plant breeding, hence the emphasis of both disciplines in this book.

### *Art and the concept of the "breeder's eye"*

Plant breeding is an applied science. Just like other non-exact science disciplines or fields, art is important to the success achieved by a plant breeder. It was previously stated in Chapter 1 that early plant breeders depended primarily on intuition, skill, and judgment in their work. These attributes are still desirable in modern day plant breeding. This book discusses the various tools available to plant breeders. Plant breeders may use different tools to tackle the same problem, the results being the arbiter of the wisdom in the choices made. In fact, it is possible for different breeders to use the same set of tools to address the same kind of problem with different results, due in part to the difference in skill and experience. As will be discussed later in the book, some breeding methods depend on phenotypic selection. This calls for the proper design of the field test to minimize the

misleading effect of a variable environment on the expression of plant traits. Selection may be likened to a process of informed “eye-ball” to discriminate among variability.

A good breeder should have a keen sense of observation. Several outstanding discoveries were made just because the scientists who were responsible for these events were observant enough to spot unique and unexpected events. Luther Burbank selected one of the most successful cultivars of potato, the “Burbank” potato, from among a pool of variability. He observed a seed ball on a vine of the “Early Rose” cultivar in his garden. The ball contained 23 seeds, which he planted directly in the field. At harvest time the following fall, he dug up and kept the tubers from the plants separately. Examining them, he found two vines that were unique, bearing large smooth and white potatoes. Still, one was superior to the others. The superior one was sold to a producer who named it Burbank. The “Russet Burbank” potato is produced on about 50% of all lands devoted to potato production in the USA.

Breeders often have to discriminate among hundreds and even tens of thousands of plants in a segregating population to select only a small fraction of promising plants to advance in the program. Visual selection is an art, but it can be facilitated by selection aids such as **genetic markers** (simply inherited and readily identified traits that are linked to desirable traits that are often difficult to identify). Morphological markers (not biochemical markers) are useful when visual selection is conducted. A keen eye is advantageous even when markers are involved in the selection process. As will be emphasized later in this book, the breeder ultimately adopts a holistic approach to selection, evaluating the overall worth or desirability of the cultivar, not just the character targeted in the breeding program.

### Scientific disciplines and technologies of plant breeding

The science and technology component of modern plant breeding is rapidly expanding. Whereas a large number of science disciplines directly impact plant breeding, several are closely associated with it. These are plant breeding, genetics, agronomy, cytogenetics, molecular genetics, botany, plant physiology, biochemistry, plant pathology, entomology, statistics, and tissue culture. Knowledge of the first three disciplines is applied in all breeding programs. Special situations (e.g., wide crosses – crosses involving different species or distantly related genotypes) and the application of

biotechnology in breeding, involve the latter two disciplines.

The technologies used in modern plant breeding are summarized in Table 2.2. These technologies are discussed in varying degrees in this book. The categorization is only approximate and generalized. Some of these tools are used to either generate variability directly or to transfer genes from one genetic background to another to create variability for breeding. Some technologies facilitate the breeding process through, for example, identifying individuals with the gene(s) of interest.

#### *Genetics*

Genetics is the principal scientific basis of modern plant breeding. As previously indicated, plant breeding is about targeted genetic modification of plants. The science of genetics enables plant breeders to predict to varying extents the outcome of genetic manipulation of plants. The techniques and methods employed in breeding are determined based on the genetics of the trait of interest, regarding, for example, the number of genes coding for it and gene action. For example, the size of the segregating population to generate in order to have a chance of observing that unique plant with the desired combination of genes depends on the number of genes involved in the expression of the desired trait.

#### *Botany*

Plant breeders need to understand the reproductive biology of their plants as well as their taxonomic attributes. They need to know if the plants to be hybridized are cross-compatible, as well as the fine detail about flowering habits, in order to design the most effective crossing program.

#### *Plant physiology*

Physiological processes underlie the various phenotypes we observe in plants. Genetic manipulation alters plant physiological performance, which in turn impacts the plant performance in terms of the desired economic product. Plant breeders manipulate plants for optimal physiological efficiency so that dry matter is effectively partitioned in favor of the economic yield. Plants respond to environmental factors, both biotic (e.g., pathogens) and abiotic (e.g., temperature, moisture). These factors are sources of physiological stress when they occur at unfavorable levels. Plant breeders need to understand these stress relationships in order to

**Table 2.2** An operational classification of technologies of plant breeding.

Technology/tool	Common use of the technology/tool
<b>Classic/traditional tools</b>	
Emasculation	Making a complete flower female; preparation for crossing
Hybridization	Crossing unidentical plants to transfer genes or achieve recombination
Wide crossing	Crossing of distantly related plants
Selection	Primary tool for discriminating among variability
Chromosome counting	Determination of ploidy characteristics
Chromosome doubling	For manipulating ploidy for fertility
Male sterility	To eliminate need for emasculation in hybridization
Triploidy	To achieve seedlessness
Linkage analysis	For determining association between genes
Statistical tools	For evaluation of germplasm
<b>Relatively advanced tools</b>	
Mutagenesis	To induce mutations to create new variability
Tissue culture	For manipulating plants at the cellular or tissue level
Haploidy	Used for creating extremely homozygous diploid
Isozyme markers	To facilitate the selection process
<i>In situ</i> hybridization	To detect successful interspecific crossing
<b>More sophisticated tools</b>	
DNA markers	
RFLP	More effective than protein markers (isozymes)
RAPD	PCR-based molecular marker
<b>Advanced technology</b>	
Molecular markers	SSR, SNPs, etc.
Marker-assisted selection	To facilitate the selection process
DNA sequencing	Ultimate physical map of an organism
Plant genomic analysis	Studying the totality of the genes of an organism
Bioinformatics	Computer-based technology for prediction of biological function from DNA sequence data
Microarray analysis	To understand gene expression and for sequence identification
Primer design	For molecular analysis of plant genome
Plant transformation	For recombinant DNA work

PCR, polymerase chain reaction; RAPD, random amplified polymorphic DNA; RFLP, restricted fragment length polymorphism; SNP, single nucleotide polymorphism; SSR, simple sequence repeat.

develop cultivars that can resist them for enhanced productivity.

### Agronomy

Plant breeders conduct their work in both controlled (greenhouse) and field environments. An understanding of agronomy (the art and science of producing crops and managing soils) will help the breeder to provide the appropriate cultural conditions for optimal plant growth and development for successful hybridization and selection in the field. An improved cultivar is only as good as its cultural environment. Without the proper nurturing, the genetic potential of an improved cultivar would not

be realized. Sometimes, breeders need to modify the plant-growing environment to identify individuals to advance in a breeding program to achieve an objective (e.g., withholding water in breeding for drought resistance).

### Pathology and entomology

Disease-resistance breeding is a major plant breeding objective. Plant breeders need to understand the biology of the insect pest or pathogen against which resistance is being sought. The kind of cultivar to breed, the methods to use in breeding, and evaluation all depend on the kind of pest (e.g., its races or variability, pattern

of spread, life cycle, and most suitable environment). Sometimes, the pest must be controlled to avoid interfering with the breeding program.

### *Statistics*

Plant breeders need to understand the principles of research design and analysis. This knowledge is essential for effectively designing field and laboratory studies (e.g., for heritability, inheritance of a trait, combining ability), and evaluating genotypes for cultivar release at the end of the breeding program. Familiarity with computers is important for record keeping and data manipulation. Statistics is indispensable to plant breeding programs. This is because the breeder often encounters situations in which predictions about outcomes, comparison of results, estimation of response to a treatment, and many more, need to be made. Genes are not expressed in a vacuum but in an environment with which they interact. Such interactions may cause certain outcomes to deviate from the expected. Statistics is needed to analyze the variance within a population to separate real genetic effects from environmental effects. The application of statistics in plant breeding can be as simple as finding the mean of a set of data, to complex estimates of variance and multivariate analysis.

### *Biochemistry*

In this era of biotechnology, plant breeders need to be familiar with the molecular basis of heredity. They need to be familiar with the procedures of plant genetic manipulation at the molecular level, including the development and use of molecular markers and gene transfer techniques.

### **The plant breeder as a decision-maker**

Modern plant breeding is a carefully planned and executed activity. It is expensive and time-consuming to breed a new cultivar. Consequently, the breeder should make sound decisions, some of which are scientific (e.g., type of cultivar to breed, germplasm to use, breeding methods), whereas others are socioeconomic or even political.

Some of the key specific decisions in a plant breeding program are discussed next. Because these elements are interdependent, the breeder should integrate the decisions to form a harmonious and continuous sequence, from inception to cultivar release. A breeder should have good management skills. Experts have identified four dimensions of management as follows:

**1 Organization design.** The breeder should plan the physical structure of the project as pertains to personnel, equipment, field, greenhouse, nurseries, and other needs.

**2 Planning and control.** Planning entails defining clear objectives and strategies for accomplishing them, while control entails establishing an effective system of data management (collection, storage, retrieval, processing) to provide reliable and accurate information for decision-making at all the steps in the plant breeding project.

**3 Behavioral process.** The team engaged in the project should work and relate well with each other (teamwork).

**4 Decision-making.** The plant breeder is a decision-maker. Critical decisions are made throughout the breeding program. This is the most important aspect of management in a breeding project. It entails first identifying the problem, then analyzing it to find the root causes and effects. Next, the breeder should develop alternative solutions, evaluate them, and then choose and implement the most desirable solution.

Some of the specific decisions made in a breeding project are as follow.

### *Breeding objectives*

The breeder must first define a clear breeding objective and ascertain its importance, feasibility, and cost-effectiveness. As previously noted plant breeding is expensive to conduct and hence a breeding objective should be economically viable or of significant social benefit. Furthermore, every problem is not amenable to genetic manipulation through breeding. Breeding objectives vary among crops (see Part II). Where multiple objectives are identified, they should be prioritized. Keeping in close touch with crop producers and consumers will allow the breeder to gain insight into what ameliorations are likely to be acceptable to them. Growers will not grow what they cannot sell. Long-term plant breeding programs are usually formulated to address the key problems that producers face.

### *Germplasm*

The plant material used to initiate a breeding program is critical to its success. The parents used in a cross should supply the gene(s) for the trait of interest. Sometimes, germplasm may have to be imported for developing new cultivars or evaluated for adaptation to a specific environment. Advanced breeding programs maintain

elite germplasm or advanced breeding lines from previous activities, which serve as a source of materials for initiating future breeding projects. Breeders have access to an enormous amount of germplasm maintained in repositories all over the world (see Chapter 6). Sometimes, certain sources of germplasm are protected by intellectual property rights and may require a fee to use. Using wild germplasm introduces a unique set of problems into a breeding program, stemming from the unadapted genetic background introduced.

#### Breeding strategy

The plant breeder should select the most effective breeding method and use the most effective techniques to accomplish the breeding objective. Hybrids may be best for certain situations, whereas synthetics (a type of variety developed by open-pollination of selected parents) may be more practical in other areas. To speed up the breeding program, the breeder may include, for example, a winter nursery where applicable, or use selection aids (e.g., genetic markers or **marker-assisted selection**). A number of standard techniques and methods (with variations) have been developed for use by breeders to tackle breeding problems (see Section 6). New technologies (e.g., biotechnology) are available to address some breeding issues that could not be adequately addressed by conventional tools. The breeder may utilize multiple technologies and methods at different stages in the breeding program.

#### Type of cultivar

The breeder decides what type of cultivar to breed (e.g., hybrid, synthetic, blend). A decision also needs to be made about whether a cultivar has to be developed for use over a broad region or a very specific production area. The type of cultivar being bred determines how to conduct yield trials prior to release of a commercial cultivar for use by the consumer.

#### Market

Some products are developed for processing while others are developed for the fresh market. The parents used in a breeding program are selected based on the type of market product needed. Some markets prefer uniformity in the plant product, whereas others (e.g., canning industry) can tolerate some variation in quality with respect to a specific trait. For example, potato for the fresh market is appealing to the consumer if

the tuber shape is attractive and uniform. On the other hand, producers of potato starch do not mind processing potatoes that may have a little blemish, provided it has the appropriate industrial quality for starch production.

#### Evaluation

It is said that plant breeding is a numbers game. A large segregating population is created in the early stages of the program. The numbers are steadily reduced with time (e.g., from 10,000 to 1,000, to 100, to 10, and then to one cultivar released in the end). A decision has to be made at each stage as to what genotypes to keep and what to discard. It has been suggested by N. F. Jensen that two basic questions are critical in a plant breeding decision-making process: Does the plant or line have the potential to become a cultivar? Does the plant have any other possible uses (e.g., as a parent in future projects)? If the answer is no to these questions, the plant should be discarded; otherwise, it should be kept for another season for further evaluation. The breeder has to decide where to evaluate the genotypes, and for how long (i.e., locations, seasons, years).

#### Cultivar release

This is the climax of a breeding program. The decisions at this stage include using information from stability analysis to select the most desirable genotype to release as a cultivar. The process also includes assigning a name, and seeking legal protection, among other actions.

### Conducting plant breeding

As previously stated, modern plant breeding is a planned activity. There are standard approaches to breeding. The breeder may choose from a variety of methods for conducting a plant breeding program, based on factors including the mode of reproduction of the plant, the type of cultivar to be developed, and the resources available.

#### Basic approaches

Plant breeding has come a long way from the cynical view of “crossing the best with the best and hoping for the best” to carefully planned and thought-out strategies to develop high-performance cultivars. Plant breeding methods and tools keep changing as technology advances. Consequently, plant breeding approaches may be categorized into two general types: **conventional**

and **unconventional**. This categorization is only for convenience.

#### *Conventional approach*

Conventional breeding is also referred to as **traditional** or **classic breeding**. This approach entails the use of tried, proven, and older tools. Crossing two plants (hybridization) is the primary technique for creating variability in flowering species. Various breeding (selection) methods are then used to discriminate among the variability to identify the most desirable recombinant. The selected genotype is increased and evaluated for performance before release to producers. Plant traits controlled by many genes (quantitative traits) are more difficult to breed. Age notwithstanding, the conventional approach remains the workhorse of the plant breeding industry. It is readily accessible to the average breeder and is relatively easy to conduct, compared to the unconventional approach.

#### *Unconventional approach*

The unconventional approach to breeding entails the use of cutting-edge technologies, to create new variability that is sometimes impossible to achieve with conventional methods. However, this approach is more involved, requiring special technical skills and knowledge. It is also expensive to conduct. The advent of recombinant DNA (rDNA) technology gave breeders a new set of powerful tools for genetic analysis and manipulation. Gene transfer can now be made across natural biological barriers, circumventing the sexual process (e.g., the *Bt* products that consist of bacterial genes transferred into crops to confer resistance to the European corn borer). Molecular markers are available for aiding the selection process to make the process more efficient and effective.

Even though two basic breeding approaches have been described, it should be pointed out that they are best considered as complementary rather than independent approaches. Usually, the molecular tools are used to generate variability for selection, or to facilitate the selection process. After genetically modifying plants using molecular tools, they may be used as parents in subsequent crosses to transfer the desirable genes into adapted and commercially desirable genetic backgrounds, using conventional tools. Whether developed by conventional or molecular approaches, the genotypes are evaluated in the field by conventional methods, and

then advanced through the standard seed certification process before the farmer can have access to the seed for planting a crop. The unconventional approach to breeding tends to receive more attention from funding agencies than the conventional approach, partly because of its novelty and advertized potential, as well as the glamour of the technologies involved.

#### **Overview of the basic steps in plant breeding**

Regardless of the approach, a breeder follows certain general steps in conducting a breeding project. As previously discussed, a breeder should have a comprehensive plan for a breeding project that addresses the following steps.

#### *Objectives*

The breeder should first define a clear objective for initiating the breeding program. This may be for the benefit of the producer (e.g., high yield, disease resistance, early maturity, lodging resistance) or the consumer (e.g., high nutritional quality, enhanced processing quality).

#### *Germplasm*

Once the objectives have been determined, the breeder then assembles the germplasm to be used to initiate the breeding program. Sometimes, new variability is created through crossing of selected parents, inducing mutations, or using biotechnological techniques. Whether used as such or recombined through crossing, the base population used to initiate a breeding program must of necessity include the gene(s) of interest. That is, you cannot breed for disease resistance, if the gene conferring resistance to the disease of interest does not occur in the base population.

#### *Selection*

After creating or assembling variability, the next task is to discriminate among the variability to identify and select individuals with the desirable genotype to advance and increase to develop potential new cultivars. This calls for using standard selection or breeding methods suitable for the species and the breeding objective(s).

#### *Evaluation*

The potential cultivars are evaluated in the field, sometimes at different locations and over several years, to

identify the most promising one for release as a commercial cultivar.

#### *Certification and cultivar release*

Before a cultivar is released, it is processed through a series of steps, called the seed certification process, to increase the experimental seed, and to obtain approval for release from the designated crop certifying agency in the state or country. These steps in plant breeding are discussed in detail in this book.

### **Qualifications of a plant breeder**

Some plant breeding can be undertaken by farmers with little education, lots of intuition, and keen observation. As previously discussed, early domesticators observed and selected plants, saving seed from the current season for planting the next season's crop. Modern commercial plant breeding is more technical and science-based, requiring the breeder to have some formal training to be successful.

Plant breeders, as previously discussed, are involved in genetically manipulating plants to accomplish a pre-determined objective. Furthermore, it was previously indicated that plant breeding is an art and a science. Consequently, the breeder should have knowledge in certain scientific disciplines in order to be able to conduct modern plant breeding. The key disciplines, as previously discussed, includes genetics, biochemistry, botany, pathology, physiology, agronomy, statistics, biotechnology, and computer science. Whereas it is not critical to master all these disciplines to be successful, a breeder needs, at least, to have a strong background in plant genetics and the principles of plant breeding. Breeding is about causing a heritable change to occur in a desired direction. Consequently, a breeder should understand the principles and concepts of heredity (or transmission genetics). To be able to use some of the modern sophisticated technologies, the breeder should understand molecular genetics and other techniques of biotechnology such as tissue culture. Basic and pertinent genetic principles and concepts are discussed in this book to facilitate the understanding of breeding principles. Biotechnological applications in plant breeding are also discussed.

It should be pointed out that a breeder may take advantage of workshops and short courses offered by national institutes (e.g., National Institute of Health in the USA) and universities, to acquire the new skills

necessary to use new techniques in a breeding project. Collaborating with experts in the use of certain techniques is also a way classically trained plant breeders may pursue to accomplish a breeding objective that requires the use of molecular techniques. It is also possible to contract or outsource a technical part of an unconventional breeding project to competent service providers.

The other issue that needs to be addressed is the level of qualification required to be a successful plant breeder. As stated in the preface, this book is designed for upper undergraduate to early graduate students. A firm grasp of the genetics and plant breeding concepts discussed should adequately equip the student to conduct plant breeding upon graduation. Having said that, graduate studies in plant breeding provide opportunities for acquiring advanced knowledge in genetics and research methodologies. Usually, the undergraduate course in plant breeding offers limited opportunities for research and hands-on exposure (especially in the conventional methods of plant breeding). Further, leaders of plant breeding programs in both the public and private sectors usually have advanced degrees, preferably, a PhD. However, BS or MS degree holders are also employed in the breeding industry.

### **The plant breeding industry**

Commercial plant breeding is undertaken in both the private and public sectors. Breeding in the private sector is primarily for profit. It should be pointed out these companies operate under the umbrella of giant multinational corporations such as Monsanto, Pioneer/Dupont, Novartis/Syngenta, and Advanta Seed Group, through mergers and acquisitions (see Chapter 24). Products from private seed companies are proprietary.

#### **Private sector plant breeding**

Four factors are deemed by experts to be critical in determining the trends in investment in plant breeding by the private sector.

#### *Cost of research innovation*

Modern plant breeding technologies are generally expensive to acquire and use. Consequently, the cost of research and development of new cultivars by these technologies are exorbitant. However, some of these innovations result in increased product quality and yield, and sometimes facilitate the production of the

crop by the producer. Also, some innovations eventually reduce the duration of the cumulative research process.

### *Market structure*

Private companies are more likely to invest in plant breeding where the potential size of the seed market is large and profitable. Further, the attraction to enter into plant breeding will be greater if there are fixed costs in marketing the new cultivars to be developed.

### *Market organization of the seed industry*

Conventional wisdom suggests that the more concentrated a seed market, the greater the potential profitability a seed production enterprise would be. However, contemporary thought on industrial organization suggests that the ease of entry into an existing market would depend on the contestability of the specific market, and would subsequently decide the profitability to the company. Plant breeding is increasingly becoming a technology-driven industry. Through research and development, a breakthrough may grant a market monopoly to an inventor of a technology or product, until another breakthrough occurs that grants a new monopoly in a related market. For example, Monsanto, the developer of Roundup Ready® technology is also the developer of the Roundup® herbicide that is required for the technology to work.

### *Ability to appropriate the returns to research and distribution of benefits*

The degree to which a seed company can appropriate returns to its plant breeding inventions is a key factor in the decision to enter the market. Traditionally, cross-pollinated species (e.g., corn) that are amenable to hybrid breeding and high profitability have been most attractive to private investors. Public sector breeding develops most of the new cultivars in self-pollinated species (e.g., wheat, soybean). However, the private sector interest in self-pollinated species is growing. This shift is occurring for a variety of reasons. Certain crops are associated in certain cropping systems. For example, corn-soybean rotations are widely practiced. Consequently, producers who purchase improved corn are likely to purchase improved soybean seed. In the case of cotton, the shift is for a more practical reason. Processing cotton to obtain seed entails ginning and delinting, which are more readily done by seed companies than farmers.

Another significant point that needs to be made is that the for-profit private breeding sector is obligated not to focus only on profitability of a product to the company, but they must also price their products such that the farmer can use them profitably. Farmers are not likely to adopt a technology that does not significantly increase their income.

### **Public sector plant breeding**

#### *The US experience*

Public sector breeding in the USA is conducted primarily by land grant institutions and researchers in the federal system (e.g., the US Department of Agriculture, USDA). The traditional land grant institutional program is centered on agriculture, and is funded by the federal government and the various states, often with support from local commodity groups. The plant research in these institutions is primarily geared towards improving field crops and horticultural and forest species of major economic importance to a state's agriculture. For example, the Oklahoma State University, an Oklahoma land grant university, conducts research on wheat, the most important crop in the state. A fee is levied on produce presented for sale at the elevator by producers, and is used to support agricultural research pertaining to wheat.

In addition to its in-house research unit, the Agricultural Research Service (ARS), USDA often has scientists attached to land grant institutions to conduct research of benefit to a specific state as well as the general region. For example, the Grazinglands Research Laboratory at El Reno, Oklahoma, is engaged in forage research for the benefit of the Great Plains of the USA. Research output from land grant programs and the USDA is often public domain and often accessible to the public. However, just like the private sector, inventions may be protected by obtaining plant variety protection or a patent.

#### *The UK experience*

Information regarding the UK experience has been obtained through personal communication with W. T. B. Thomas of the Scottish Crop Research Institute, Invergowrie, UK. The equivalent of a land grant system does not operate in the UK but, up to the 1980s, there were a number of public sector breeding programs at research institutes such as the Plant Breeding Institute (PBI) (now part of John Innes Centre), Scottish Crop

Research Institute (SCRI), Welsh Plant Breeding Station (now Institute of Grassland and Environmental Research, IGER), and National Vegetable Research Station (now Horticultural Research International, HRI) with the products being marketed through the National Seed Development Organization (NSDO). In addition, there were several commercial breeding programs producing successful finished cultivars, especially for the major crops. Following a review of "Near Market Research", the plant breeding program at PBI and the whole portfolio of NSDO were sold to Unilever and traded under the brand PBI Cambridge, later to become PBI Seeds. The review effectively curtailed the breeding activities in the public sector, especially of the major crops. Plant breeding in the public sector did continue at IGER, HRI, and SCRI but was reliant on funding from the private sector for at least a substantial part of the program. Two recent reviews of crop science research in the UK have highlighted the poor connection between much public sector research and the needs of the plant breeding and end-user communities. The need for good public plant breeding was recognized in the Biotechnology and Biological Sciences Research Council (BBSRC) Crop Science Review to translate fundamental research into deliverables for the end-user and is likely to stimulate prebreeding activity, at the very least, in the public sector.

### International plant breeding

There are other private sector efforts that are supported by foundations and world institutions such as the Food and Agricultural Organization (FAO), Ford Foundation, and Rockefeller Foundation. These entities tend to address issues of global importance, and also support the improvement of the so-called "orphaned crops" (crops that are of importance to developing countries, but not of enough economic value to attract investment by multinational corporations). Developing countries vary in their capabilities for modern plant breeding research. Some countries such as China, India, Brazil, and South Africa have advanced plant breeding research programs. Other countries have national research stations that devote efforts to the breeding of major national crops or plants, such as the Crops Research Institute in Ghana, where significant efforts have led to the country being a world-leading adopter of quality protein maize (QPM). A chapter has been devoted to international plant breeding efforts (see Chapter 25).

Public sector breeding is disadvantaged in an increasingly privatized world. The issues of intellectual property protection, globalization, and the constraints on public budgets in both developed and developing economies are responsible for the shift in the balance of plant breeding undertakings from the public to the private sector. This shift in balance has occurred over a period of time, and differs from one country to another, as well as from one crop to another. The shift is driven primarily by economic factors. For example, corn breeding in developed economies is dominated by the private sector. However, the trends in wheat breeding are variable in different parts of the world and even within regions in the same country. Public sector plant breeding focuses on problems that are of great social concern, even though they may not be of tremendous economic value (having poor market structure), whereas private sector breeding focuses on problems of high economic return. Public sector breeders can afford to tackle long-term research while the private sector, for economic reasons, prefers to have quicker returns on investment. Public sector breeders also engage in minor crops in addition to the principal crops of importance to various states (in the case of the land grant system of the USA). A great contribution of public sector research is the training of plant breeders who work in both public and private sectors. Also, the public sector is primarily responsible for germplasm conservation and preservation. Hence, private sector breeding benefits tremendously from public sector efforts.

It has been suggested by some that whereas scientific advances and cost of research are relevant factors in the public sector breeding programs, plant breeding investment decisions are not usually significantly impacted directly by the market structure and organization of the seed industry.

A major way in which private and public breeding efforts differ is on the returns to research. Public sector breeders are primarily not profit-oriented and can afford to exchange and share some of their inventions more freely. However, it must be pointed out that access to some public germplasm and technologies is now highly restricted, requiring significant protocol and fees to be paid for their use. The public sector plays a critical role in important activities such as education and training of plant breeders, development of new methods of breeding, and germplasm preservation and enhancement. These activities are generally long term and less profitable, at least in the short run, and hence less attractive to the private sector.

## Resource investment

### *Human capital*

In 1996, K. J. Frey of Iowa State University conducted a survey to determine the number of science person-years devoted to plant breeding research and development in the USA. He observed that of the 2,241 science person-years devoted to plant breeding, 1,499 (67%) were in the private sector. Of the remainder, 529 were in the State Agricultural Experimental Station system, while 177 were in the USDA-ARS and 36 in the USDA Plant Materials Centers. Private breeders dominate the crops that are produced primarily as hybrids. Of 545 breeders in field corn development, 510 were in the private sector. Similarly, 41 of the 56 sorghum breeders were in the private sector. On the other hand, 77 of the total of 131 wheat breeders were in the public sector, while 41 of 50 potato breeders were in the public sector. Other crops of breeding interest are soybean, cotton, and tomato.

Frey also observed that 1,571 (71%) of plant breeders in the USA were engaged in breeding agronomic crops with 634 (29%) breeding horticultural crops. About 75% of public breeders were engaged in agronomic crops versus 25% in the private sector. In the US, 100% of all maize production in 1997 was derived from private sector cultivars, compared to about 24% from wheat. In soybean, only about 8% of the acreage was planted to the crop in 1980, while 70–90% of the crop acreage in 1997 was planted to private sector seed. About 93% of

cotton acreage was also planted to private sector seed. These trends indicate the surging role of genetically modified (GM) cultivars in the production of these crops.

## Duration and cost of plant breeding programs

It is estimated that it takes about 7–10 years (or even longer) to complete (cultivar release) a breeding program for annual cultivars such as corn, wheat, and soybeans, and much longer for tree crops. The use of molecular techniques to facilitate the selection process may reduce the time for plant breeding in some cases. The use of tissue culture can reduce the length of breeding programs of perennial species. Nonetheless, the development of new cultivars may cost from hundreds of thousands of dollars to even several million dollars. The cost of cultivar development can be much higher if proprietary material is involved. Genetically engineered parental stock attracts a steep fee to use because of the costs involved in their creation. The cost of breeding also depends on where and by whom the activity is being conducted. Because of high overheads, similar products produced by breeders in developed and developing economies, are produced at dramatically higher cost in the former. Cheap labor in developing countries can allow breeders to produce hybrids of some self-pollinated species less expensively, because they can afford to pay for hand pollination (e.g., cotton in India).

## References and suggested reading

- Allard, R.W. 1988. Genetic changes associated with the evolution of adaptedness in cultivated plant and wild progenitors. *J. Heredity* 79:225–238.
- Burke, J.M., S. Tang, S.J. Knapp, and L.H. Rieseberg. 2002. Genetics analysis of sunflower domestication. *Genetics* 161:1257–1267.
- Frary, A., and S. Doganla. 2003. Comparative genetics of crop plant domestication and evolution. *Turk. J. Agric. Forest.* 27:59–69.
- Gepts, P. 2003. A comparison between crop domestication, classical plant breeding, and genetic engineering. *Crop Sci.* 42:1780–1790.
- Harlan, J.R. 1992. Crops and man. American Society of Agronomy and Crop Science Society of America, Madison, WI.
- Jansen, N.F. 1983. Crop breeding as a decision science. In: *Crop breeding* (Woods, D.R., ed.), pp. 38–64. American Society of Agronomy and Crop Science Society of America. Madison, WI.
- Koinange, E.M.K., S.P. Singh, and P. Gepts. 1996. Genetic control of the domestication syndrome in common bean. *Crop Sci.* 36:1038–1045.
- Koornneef, M.P., and P. Stam. 2001. Changing paradigms in plant breeding. *Plant Physiol.* 125:150–159.
- Ladizinsky, G. 1998. Plant evolution under domestication. Kluwer Academic Publishers, Boston.
- Zizumbo-Villarreal, D., P. Colunga-GarciaMarin, E.P. de la Cruz, P. Delgado-Valerio, and P. Gepts. 2005. Population structure and evolutionary dynamics of wild-weedy-domesticated complexes of common bean in a Mesoamerican region. *Crop Sci.* 45:1073–1083.
- Zohary, D., and M. Hopf. 1988. Domestication of plants in the Old World. Clarendon, Oxford, UK.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Evolution is a population phenomenon.
- 2 Land grant institutions conduct plant breeding in the private sector in the USA.
- 3 Traditional plant breeding tools are obsolete.
- 4 Plant breeding causes heritable changes in plants.

### Part B

Please answer the following questions:

- 1 ..... is the arbiter of evolution.
- 2 ..... is the process by which wild plants are genetically changed through human selection.
- 3 Compare and contrast evolution and plant breeding.
- 4 Give four specific examples of ways in which domesticated plants may differ from their wild progenitors.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the concept of breeder's eye.
- 2 Discuss the concept of the breeder as a decision-maker.
- 3 Discuss the general steps in a plant breeding program.
- 4 Discuss the qualifications of a plant breeder.
- 5 Distinguish between public sector and private sector plant breeding.
- 6 Discuss the molecular and classic plant breeding approaches as complementary approaches in modern plant breeding.



# 3

## Plant cellular organization and genetic structure: an overview

---

### Purpose and expected outcomes

*Plant breeders manipulate the genotype of plants to create heritable modifications in plant structure and function. Sometimes plants are genetically modified to change their shape, height, or size, to facilitate a production operation or system. For example, dwarf plants are more environmentally responsive and tend to resist lodging. Changes in stature and size impact their metabolic activities. To undertake such modifications, it is important for breeders to understand the fundamental plant structure and organization at the molecular, cellular, and whole-plant levels. Manipulating plants by some biotechnological procedures occurs at the cellular level, hence the need to understand DNA and cellular structure and function. Because the goal of plant breeding is to create modifications that are permanent and heritable, it is important to also understand the genetic architecture of plants and how genes condition plant traits. This chapter is designed to present an overview of concepts pertaining to the cellular organization and genetic structure of plants. After studying this chapter, the student should be able to:*

- 1 Briefly describe plant cell structure and organization.
  - 2 Briefly discuss nuclear division processes.
  - 3 Discuss Mendelian concepts.
  - 4 Discuss DNA structure and function.
  - 5 Distinguish between phenotype and genotype.
  - 6 Discuss the role of genetic linkage in plant breeding.
  - 7 More importantly, the student should be able to discuss the role of these plant structures and processes in plant breeding.
- 

### Units of organization of living things

The **cell** is the fundamental unit of organization of living things. Some organisms consist entirely of one cell (called **unicellular**) whereas others consist of numerous cells working together (called **multicellular**). Except for a bacterium, which lacks cellular compartmentalization into discrete functional units called **organelles**, and is called a **prokaryote**, all other cells have a membrane-

bound nucleus and several other membrane-enclosed organelles and are called **eukaryotes**.

The cell can be a unit for selection in breeding if, for example, molecular tools are used. The technology of genetic engineering targets single cells for manipulation. After successfully transferring foreign genes into the cell, it is isolated and nurtured into a full plant. On the other hand, when conventional tools are used, the whole plant is the unit of selection. It should be pointed

out that when plants are manipulated by molecular techniques, they eventually have to be evaluated via conventional selection process in the field using whole plants as the unit of selection.

### Levels of eukaryotic organization

A eukaryote may also be structurally organized at various levels of complexity: whole organism, organs, tissues, cells, organelles, and molecules, in order of descending complexity. Plant breeding of sexually reproducing species by conventional tools is usually conducted at the whole-plant level by crossing selected parents. Flowers are the units for crossing. The progeny of the cross is evaluated to select those with the desired combination of parental traits. The use of molecular tools allows plant breeders to directly manipulate the DNA, the hereditary material, and thereby circumvent the sexual process. Also, other biotechnological tools (e.g., tissue culture, cell culture, protoplast culture) enable genetic manipulation to be made below the whole-plant level.

### Plant cells and tissue

The plant cell consists of several organelles and structures with distinct as well as interrelated functions (Table 3.1). Some organelles occur only in plants while others occur only in animals. The nucleus is the most prominent organelle in the cell. The extranuclear region is called the **cytoplasm**. For the plant breeder, the organelles of special interest are those directly associated with plant heredity, as discussed next.

There are three basic cell and tissue types – **parenchyma**, **collenchyma**, and **sclerenchyma** – with increasing thickness in the cell wall. Cells aggregate to form tissues of varying complexity and functions. Parenchyma cells have thin walls and occur in actively growing parts of the plant and extensively in herbaceous plants. The fleshy and succulent parts of fruits and other swollen parts of plants (e.g., tubers, roots) contain parenchyma cells. Collenchyma cells have a thick primary wall and play a role in the mechanical support system of plants by forming strengthening tissues. Like parenchyma cells, collenchyma cells occur in regions where active growth occurs so as to provide the plant some protection from damage. Sclerenchyma cells have both primary and secondary cell walls. The short types are called sclereids, and the long cells, fibers.

**Table 3.1** A summary of the structures of plant cells and their functions.

Plasma membrane	This differentially permeable cell boundary delimits the cell from its immediate external environment. The surface may contain specific receptor molecules and may elicit an immune response
Nucleus	It contains DNA and proteins that are condensed in strands called chromosomes (called chromatin when uncoiled)
Cytoplasm	The part of the cell excluding the nucleus and enclosed by the plasma membrane. It is made up of a colloidal material called cytosol and contains various organelles
Endoplasmic reticulum	A membranous structure of two kinds – smooth (no ribosomes) and rough (has ribosomes). It increases the surface area for biochemical synthesis
Ribosomes	Organelles that contain RNA and are the sites of protein synthesis
Mitochondria	Organelles that are the sites of respiration; they contain DNA
Chloroplasts	Contain DNA and chlorophyll; they are the sites of photosynthesis
Cell wall	A rigid boundary outside the plasma membrane
Golgi apparatus	Also called dictyosomes. It has a role in cell wall formation
Vacoules	These are storage regions of the cell for undesirable compounds. They help to regulate water pressure in the cell and maintain cell rigidity

Sclerenchyma occurs abundantly in plants that yield fiber (e.g., cotton, kenaf, flax, hemp).

### Plant genome

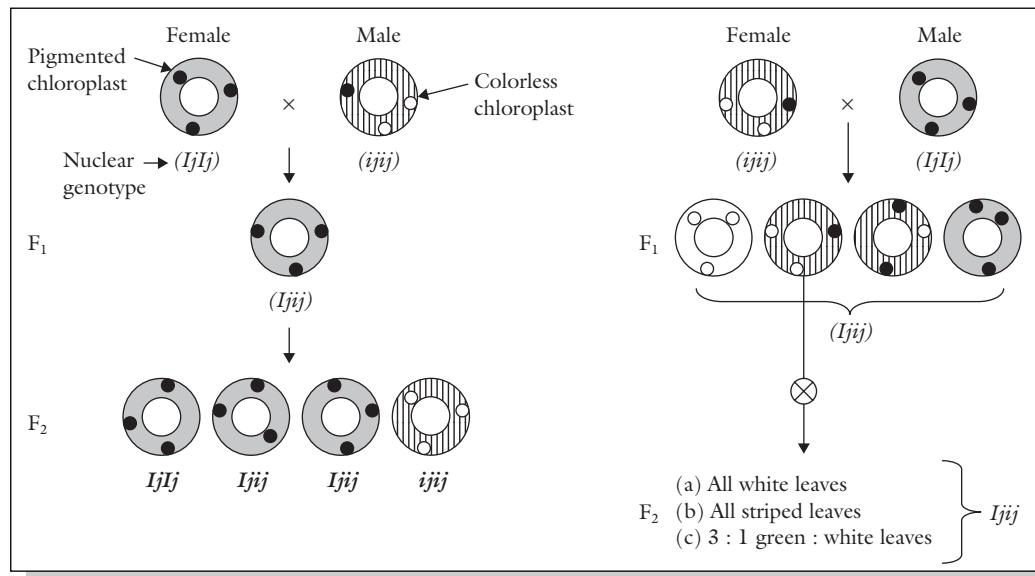
A **genome** may be defined as the set of chromosomes (or genes) within a gamete of a species. As previously stated, DNA is the hereditary material of organisms. Most of the DNA (hence most of the genes) in plants occurs in the nucleus in linear structures called **chromosomes**. The nuclear genes are subject to **Mendelian inheritance** (are transmitted according to the laws of Mendel through the processes of nuclear division)

(discussed next). In addition to the nucleus, DNA occurs in some **plastids** (organelles that are capable of dividing, growing, and differentiating into different forms). These plastids are chloroplasts. DNA also occurs in the mitochondria. The DNA in these organelles is not subject to Mendelian inheritance but follows what is called **cytoplasmic** (or extrachromosomal or extranuclear) **inheritance**. The distribution of DNA into gametes following nuclear division is unpredictable and not equitable. Molecular techniques may be used to separate nuclear DNA from non-nuclear DNA during DNA extraction from a tissue, for independent analysis. Some extranuclear genes are of special importance to plant breeding. Some male sterility genes are located in the mitochondria. As will be described later, cytoplasmic male sterility (CMS) is used in the breeding of corn and many other species. It is used to eliminate the need for emasculation (a time-consuming and tedious operation to prepare plants for crossing by removing the anthers). Also, because genes occur in the cytoplasm but pollen grains (plant male sex units) lack cytoplasm, it is important in a hybrid program which of the two parents is used as female (provides both nuclear genes and cytoplasmic genes) and which as male (provides only nuclear genes). Genes carried in the maternal cytoplasm may influence the hybrid phenotype, an effect called the **maternal effect** (Figure 3.1). When uncertain about the presence

of any special beneficial genes in the cytoplasm, some breeders conduct reciprocal crossing in which the parents take turns in being used as the female parent.

### Chromosomes and nuclear division

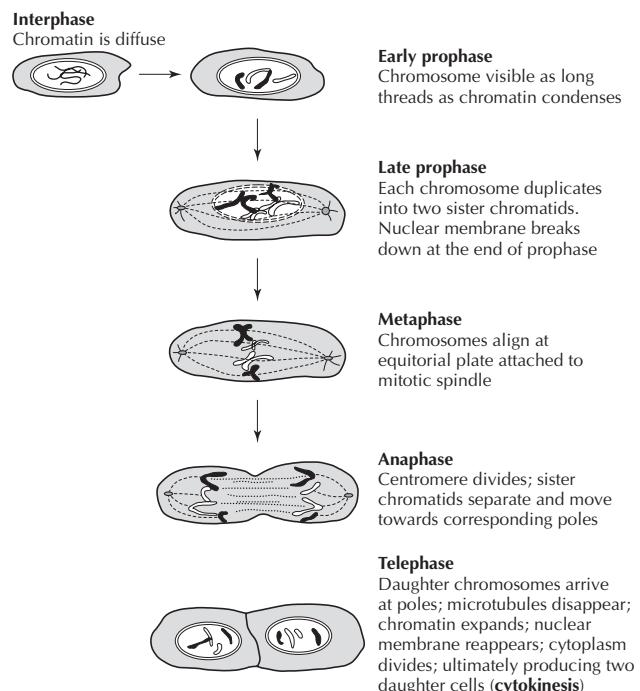
Genes (DNA sequences) are arranged in linear fashion in chromosomes, which may be visible as strands in the condensed stage as the cell prepares for nuclear division. Each species is characterized by a set of chromosomes per cell (Table 3.2). On the basis of the number of chromosomes, there are two kinds of cells in a sexually reproducing plant. Cells in the **gametes (gametic cells)** of the plant (pollen grains, eggs) contain half the set of chromosomes in the cells in other parts of the body (**somatic cells**). The somatic chromosome number is called the **diploid number** ( $2n$ ), while the gametic cells contain the **haploid number** ( $n$ ). Further, the somatic chromosomes can be arranged in pairs called **homologous chromosomes**, based on morphological features (size, length, centromere position). In sexually reproducing plants, one member of each pair is derived from the maternal parent (through the egg) and the other from the paternal parent (through the pollen). This occurrence is called **biparental inheritance** and as a result each diploid cell contains two forms of each gene



**Figure 3.1** Maternal inheritance of the *iojap* (*ij*) gene in maize. The wild type gene is *Ij*. The green color of the leaf is caused by the chloroplasts, which are maternally inherited. The appearance of the leaf color is determined solely by the maternal phenotype. (Adapted from Klug W.S., and M.R. Cummings. 1997. Concepts of genetics, 5th edn. Prentice Hall.)

**Table 3.2** Number of chromosomes per cell possessed by a variety of plant species.

Species	Scientific name	Chromosome number ( $2n$ )
Broad bean	<i>Vicia faba</i>	24
Potato	<i>Solanum tuberosum</i>	48
Maize	<i>Zea mays</i>	20
Bean	<i>Phaseolus vulgaris</i>	22
Cucumber	<i>Cucumis sativus</i>	28
Wheat	<i>Triticum aestivum</i>	42
Rice	<i>Oryza sativa</i>	24
Tobacco	<i>Nicotiana tabacum</i>	48
Soybean	<i>Glycine max</i>	40
Peanut	<i>Arachis hypogaea</i>	40
Cotton	<i>Gossypium hirsutum</i>	52
Alfalfa	<i>Medicago sativa</i>	32
Sugar beet	<i>Beta vulgaris</i>	18
Sunflower	<i>Helianthus annus</i>	34
Bermudagrass	<i>Cynodon dactylon</i>	18, 36



**Figure 3.2** Diagrammatic presentation of mitosis in a cell with a diploid number of 4. The male and female chromosomes are presented in black and white. Mitosis produces genetically identical daughter cells.

(called **alleles**). At various stages in the plant life cycle, a cell nucleus may divide according to one of two processes – **mitosis** and **meiosis**.

### Mitosis

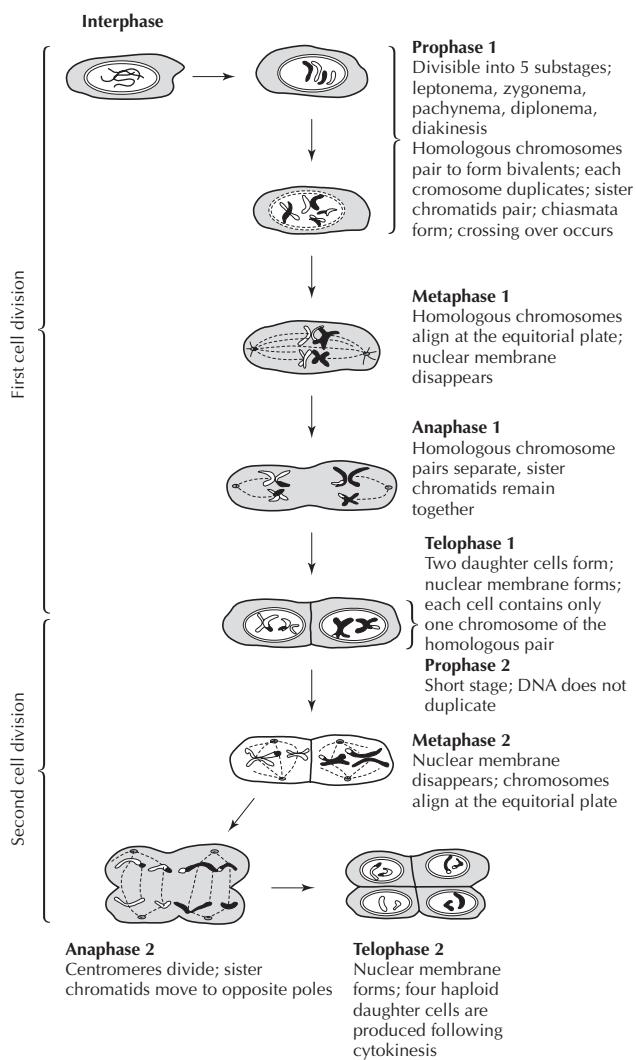
Mitosis occurs only in somatic cells and is characterized by a division of the nucleus (karyokinesis) into two so that each daughter nucleus contains the same number of chromosomes as the mother cell (Figure 3.2). The cytoplasm divides (cytokinesis) so that the mitotic products are genetically identical (equational division). This conservative process produces new cells for growth and maintenance of the plant. Cells in tissue culture divide mitotically. Through the application of appropriate chemicals and other suitable environmental conditions, plant cells can be made to proliferate into an amorphous mass called **callus**. Callus is an undifferentiated mass of cells (cells with no assigned functions). It is a material used in genetic engineering to receive and incorporate foreign DNA into cells.

The nuclear division process may be disrupted (e.g., using a chemical called colchicine) on purpose by scientists, by interfering with the **spindle fibers** (the structures that pull the chromosomes to opposite poles of the cell). The consequence of this action is that the chromosomes fail to separate properly into the daughter cells. Instead, a mitotic product may contain a duplication of

all or some of the original set of chromosomes (ploidy modification; see Chapter 13).

### Meiosis

Meiosis occurs only in specialized tissues in flowers of plants and produces daughter cells that contain the haploid number of chromosomes (Figure 3.3). This nuclear division is responsible for producing gametes or spores. A meiotic event called **crossing over** occurs in the diplonema stage, resulting in genetic exchange between non-sister chromatids. This event is a major source of genetic variability in flowering plants. It is responsible for the formation of new combinations of genetic material (**recombinants**) for use by plant breeders. Closely linked genes may also undergo recombination to separate them. Hence, plant breeders sometimes take advantage of this phenomenon of recombination to attempt to break undesirable genetic linkages through repeated crossing, and more importantly to forge desirable linkage blocks. Meiosis is also critical in the life cycle of flowering species as it pertains to the maintenance of



**Figure 3.3** Diagrammatic presentation of meiosis in a cell with a diploid number of 4. The process has two distinct cell divisions. Prophase I consists of five distinguishable stages; the most genetically significant event of crossing over occurs in the fourth stage, diplonema.

the ploidy level of the species. By reducing the diploid number to a haploid number before fertilization, the diploid number is restored thereafter.

### Mendelian concepts in plant breeding

As previously stated, genetics is the principal science that underlies plant breeding. Gregor Mendel made significant contributions to the development of the

discipline of genetics, albeit *in absentia*. He derived several postulates or principles of inheritance, which are often couched as Mendel's laws of inheritance.

### Mendelian postulates

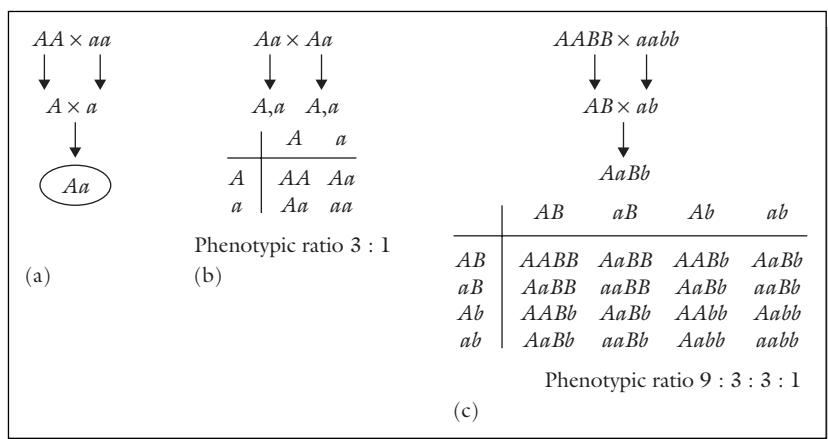
Because plant breeders transfer genes from one source to another, an understanding of **transmission genetics** is crucial to a successful breeding effort. The method of breeding used depends upon the heredity of the trait being manipulated, among other factors. According to Mendel's results from his hybridization studies in pea, traits are controlled by heritable factors that are passed from parents to offspring, through the reproductive cells. Each of these unit factors occurs in pairs in each cell (except reproductive cells or gametes).

In his experiments, Mendel discovered that in a cross between parents displaying two contrasting traits, the hybrid ( $F_1$ ) expressed one of the traits to the exclusion of the other. He called the expressed trait **dominant** and the suppressed trait **recessive**. This is the phenomenon of **dominance** and **recessivity**. When the hybrid seed was planted and self-pollinated, he observed that both traits appeared in the second generation ( $F_2$ ) (i.e., the recessive trait reappeared), in a ratio of 3 : 1 dominant : recessive individuals (Figure 3.4). Mendel concluded that the two factors that control each trait do not blend but remain distant throughout the life of the individual and segregate in the formation of gametes. This is called the **law of segregation**. In further studies in which he considered two characters simultaneously, he observed that the genes for different characters are inherited independently of each other. This is called the **law of independent assortment**. In summary, the two key laws are as follows:

**Law I Law of segregation:** paired factors segregate during the formation of gametes in a random fashion such that each gamete receives one form or the other.

**Law II Law of independent assortment:** when two or more pairs of traits are considered simultaneously, the factors for each pair of traits assort independently to the gametes.

Mendel's pairs of factors are now known as **genes**, while each factor of a pair (e.g.,  $HH$  or  $hh$ ) is called an **allele** (i.e., the alternative form of a gene:  $H$  or  $h$ ). The specific location on the chromosome where a gene resides is called a **gene locus** or simply a **locus** (loci for plural).



**Figure 3.4** Mendel's postulates: (a) dominance, (b) segregation, and (c) independent assortment.

### Concept of genotype and phenotype

The term **genotype** is used to describe the totality of the genes of an individual. Because the totality of an individual's genes is not known, the term, in practice, is usually used to describe a very small subset of genes of interest in a breeding program or research. Conventionally, a genotype is written with an uppercase letter ( $H, G$ ) indicating the **dominant allele** (expressed over the alternative allele), while a lower case letter ( $h, g$ ) indicates the **recessive allele**. A plant that has two identical alleles for genes is **homozygous** at that locus (e.g.,  $AA$ ,  $aa$ ,  $GG$ ,  $gg$ ) and is called a **homozygote**. If it has different alleles for a gene, it is **heterozygous** at these loci (e.g.,  $Aa$ ,  $Gg$ ) and is called a **heterozygote**. Certain plant breeding methods are designed to produce products that are homozygous (breed true – most or all of the loci are homozygous) whereas others (e.g., hybrids) depend on heterozygosity for success.

The term **phenotype** refers to the observable effect of a genotype (the genetic makeup of an individual). Because genes are expressed in an environment, a phenotype is the result of the interaction between a genotype and its environment (i.e., phenotype = genotype + environment, or symbolically,  $P = G + E$ ). At a later time in this book, a more complete form of this equation will be introduced as  $P = G + E + GE$ , where  $GE$  represents the interaction between the environment and the genotype. This interaction effect helps plant breeders in the cultivar release decision-making process (see Chapter 23).

### Predicting genotype and phenotype

Based upon Mendel's laws of inheritance, statistical probability analysis can be applied to determine the outcome of a cross, given the genotype of the parents and gene action (dominance/recessivity). A genetic grid called a **Punnett square** facilitates the analysis (Figure 3.5). For example, a monohybrid cross in which the genotypes of interest are  $AA \times aa$ , where  $A$  is dominant over  $a$ , will produce a hybrid genotype  $Aa$  in the  $F_1$  (first filial generation) with an  $AA$  phenotype. However, in the  $F_2$  ( $F_1 \times F_1$ ), the Punnett square shows a genotypic ratio of  $1AA : 2Aa : 1aa$ , and a phenotypic ratio of  $3 : 1$ , because of dominance. A **dihybrid cross** (involving simultaneous analysis of two different genes) is more complex but conceptually like a **monohybrid cross** (only one gene of interest) analysis. An analysis of a dihybrid cross  $AABB \times aabb$ , using the Punnett square is illustrated in Figure 3.5. An alternative method of genetic analysis of a cross is by the **branch diagram** or forked line method (Figure 3.6).

Predicting the outcome of a cross is important to plant breeders. One of the critical steps in a hybrid program is to authenticate the  $F_1$  product. The breeder must be certain that the  $F_1$  truly is a successful cross and not a product of selfing. If a selfed product is advanced, the breeding program will be a total waste of resources. To facilitate the process, breeders may include a genetic marker in their program. If two plants are crossed, for example, one with purple flowers and the other with white flowers, we expect the  $F_1$  plant to have purple

		Pollen	
Egg		$\frac{1}{2}A$	$\frac{1}{2}a$
$\frac{1}{2}A$		$\frac{1}{4}AA$	$\frac{1}{4}Aa$
$\frac{1}{2}a$		$\frac{1}{4}Aa$	$\frac{1}{4}aa$

Phenotypic ratio of 3 : 1  $A-$  :  $aa$

(a)

		Pollen			
Egg		$\frac{1}{4}AB$	$\frac{1}{4}Ab$	$\frac{1}{4}aB$	$\frac{1}{4}ab$
$\frac{1}{4}AB$		$\frac{1}{16}AABB$	$\frac{1}{16}AABb$	$\frac{1}{16}AaBB$	$\frac{1}{16}AaBb$
$\frac{1}{4}Ab$		$\frac{1}{16}AABb$	$\frac{1}{16}AAbb$	$\frac{1}{16}AaBB$	$\frac{1}{16}Aabb$
$\frac{1}{4}aB$		$\frac{1}{16}aABB$	$\frac{1}{16}AaBb$	$\frac{1}{16}aaBB$	$\frac{1}{16}aaBb$
$\frac{1}{4}ab$		$\frac{1}{16}AaBb$	$\frac{1}{16}Aabb$	$\frac{1}{16}aaBb$	$\frac{1}{16}aabb$

Phenotypic ratio of 9 : 3 : 3 : 1       $A-B-$  :  $A-ab$  :  $aaB-$  :  $aabb$

(b)

**Figure 3.5** The Punnett square procedure may be used to demonstrate the events that occur during hybridization and selfing in (a) a monohybrid cross, and (b) a dihybrid cross, showing the proportions of genotypes in the  $F_2$  population and the corresponding Mendelian phenotypic and genotypic ratios.

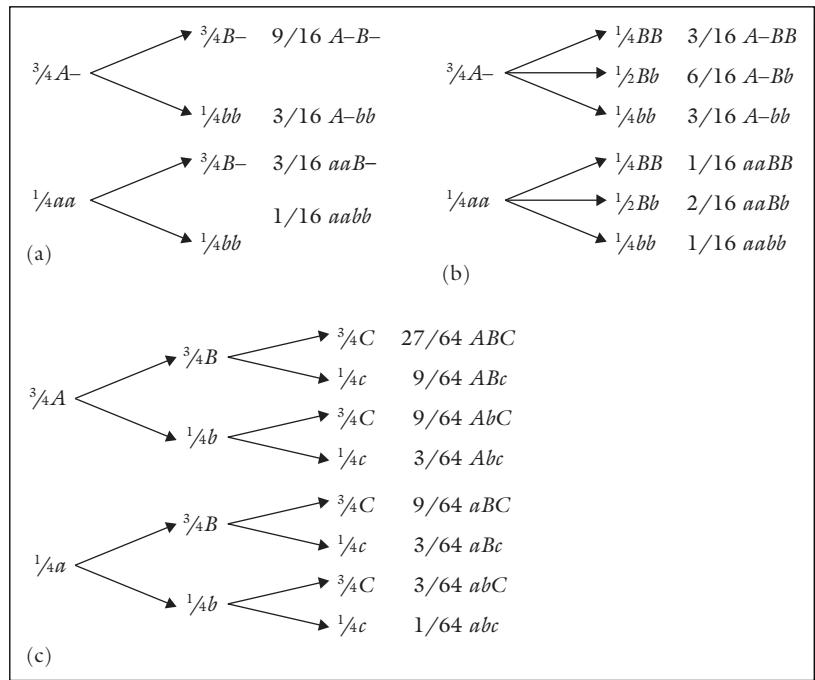
flowers because of dominance of purple over white flowers. If the  $F_1$  plant has white flowers, it is proof that the cross was unsuccessful (i.e., the product of the “cross” is actually from selfing).

### Distinguishing between heterozygous and homozygous individuals

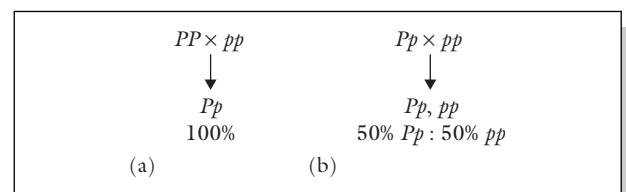
In a segregating population where genotypes  $PP$  and  $Pp$  produce the same phenotype (because of dominance), it is necessary, sometimes, to know the exact genotype of a plant. There are two procedures that are commonly used to accomplish this task.

#### Testcross

Developed by Mendel, a **testcross** entails crossing the plant with the dominant allele but unknown genotype with a homozygous recessive individual (Figure 3.7). If the unknown genotype is  $PP$ , crossing it with the genotype  $pp$  will produce all  $Pp$  offspring. However, if the unknown is  $Pp$  then a testcross will produce offspring segregating 50 : 50 for  $Pp : pp$ . The testcross also



**Figure 3.6** The branch diagram method may also be used to predict the phenotypic and genotypic ratios in the  $F_2$  population. (a) Two genes with dominance at both loci. (b) Two genes with dominance at one locus. (c) F<sub>2</sub> trihybrid phenotypic ratio.



**Figure 3.7** The testcross. (a) Crossing a homozygous dominant genotype with a homozygous recessive genotype always produces all heterozygotes. (b) However, crossing a heterozygote with a homozygous recessive produces both homozygotes and heterozygotes.

supports Mendel's postulate that separate genes control purple and white flowers.

#### Progeny test

Unlike a testcross, a **progeny test** does not include a cross with a special parent but selfing of the  $F_2$ . Each  $F_2$  plant is harvested and separately bagged, and then subsequently planted. In the  $F_3$  stage, plants that are homozygous dominant will produce progenies that are uniform for the trait, whereas plants that are heterozygous will produce a segregating progeny row.

Plant breeders use the progeny test for a number of purposes. In breeding methodologies in which selection is based on phenotype, a progeny test will allow a breeder to select superior plants from among a genetically mixed population. Following an environmental stress, biotic or abiotic, a breeder may use a progeny test to identify superior individuals and further ascertain if the phenotypic variation is due to genetic effects or just caused by environment factors.

### Complex inheritance

Just how lucky was Mendel in his experiments that yielded his landmark results? This question has been widely discussed among scientists over the years. Mendel selected traits whose inheritance patterns enabled him to avoid certain complex inheritance patterns that would have made his results and interpretations more challenging. Traits such as those studied by Mendel are described as **simple (simply inherited) traits**, or having **Mendelian inheritance**. There are other numerous traits that have complex inheritance patterns that cannot be predicted by Mendelian ratios. Several factors are

responsible for the observation of non-Mendelian ratios as discussed next.

#### Incomplete dominance and codominance

Mendel worked with traits that exhibited complete dominance. Post-Mendelian studies revealed that, frequently, the masking of one trait by another is only partial (called **incomplete dominance** or **partial dominance**). A cross between a red-flowered ( $RR$ ) and white-flowered ( $rr$ ) snapdragon produces pink-flowered plants ( $Rr$ ). The genotypic ratio remains  $1 : 2 : 1$ , but a lack of complete dominance also makes the phenotypic ratio  $1 : 2 : 1$  (instead of the  $3 : 1$  expected for complete dominance).

Another situation in which there is no dominance occurs when both alleles of a heterozygote are expressed to equal degrees. The two alleles code for two equally functional and detectable gene products. Commonly observed and useful examples for plant breeding technology are **allozymes**, the production of different forms of the same enzyme by different alleles at the same locus. Allozymes catalyze the same reaction. This pattern of inheritance is called **codominant inheritance** and the gene action **codominance**. Some molecular markers are codominant. Whereas incomplete dominance produces a blended phenotype, codominance produces distinct and separate phenotypes.

#### Multiple alleles of the same gene

The concept of **multiple alleles** can be studied only in a population. Any individual diploid organism can, as previously stated, have at most two homologous gene loci that can be occupied by different alleles of the same gene. However, in a population, members of a species can have many alternative forms of the same gene. A diploid by definition can have only two alleles at each locus (e.g.,  $C_1C_1$ ,  $C_7C_{10}$ ,  $C_4C_6$ ). However, mutations may cause additional alleles to be created in a population. Multiple alleles of allozymes are known to occur. The mode of inheritance by which individuals have access to three or more alleles in the population is called **multiple allelism** (the set of alleles is called an **allelic series**). A more common example of multiple allelism that may help the reader better understand the concept is the **ABO** blood group system in humans. An allelic series of importance in plant breeding are the **S alleles** that condition self-incompatibility (inability of a flower to be fertilized by its own pollen). Self-incompatibility is a constraint to sexual biology and can be used as a tool in plant breeding as discussed in detail in Chapter 4.

## Multiple genes

Just as a single gene may have multiple alleles that produce different forms of one enzyme, there can be more than one gene for the same enzyme. The same enzymes produced by different genes are called **isozymes**. Isozymes are common in plants. For example, the enzyme phosphoglucomutase in *Helianthus debilis* is controlled by two nuclear genes and two chloroplast genes. As discussed in detail in Chapter 14, isozymes and allozymes were the first molecular markers developed for use in plant and animal genetic research.

## Polygenic inheritance

Mendelian genes are also called **major genes** (or **oligo-genes**). Their effects are easily categorized into several or many non-overlapping groups. The variation is said to be discrete. Some traits are controlled by several or many genes that have effects too small to be individually distinguished. These traits are called **polygenes** or **minor genes** and are characterized by non-discrete (or continuous) variation, because the effects of the environment on these genes make their otherwise discrete segregation difficult to be readily observed. Scientists use statistical genetics to distinguish between genetic variation due to the segregation of polygenes and environmental variation (see Chapter 9). Many genes of interest to plant breeders exhibit polygenic inheritance.

## Concept of gene interaction and modified Mendelian ratios

Mendel's results primarily described discrete (discontinuous) variation even though he observed continuous variation in flower color. Later studies established that the genetic influence on the phenotype is complex, involving the interactions of many genes and their products. It should be pointed out that genes do not necessarily interact directly to influence a phenotype, but rather, the cellular function of numerous gene products work together in concert to produce the phenotype.

Mendel's observation of dominance/recessivity is an example of an interaction between alleles of the same gene. However, interactions involving non-allelic genes do occur, a phenomenon called **epistasis**. There are several kinds of epistatic interactions, each modifying the expected Mendelian ratio in a characteristic way. Instead of the  $9 : 3 : 3 : 1$  dihybrid ratio for dominance at two loci, modifications of the ratio include  $9 : 7$  (complementary genes),  $9 : 6 : 1$  (additive genes),  $15 : 1$

(duplicate genes),  $13 : 3$  (suppressor genes),  $12 : 3 : 1$  (dominant epistasis), and  $9 : 3 : 4$  (recessive epistasis) (Figure 3.8). Other possible ratios are  $6 : 3 : 3 : 4$  and  $10 : 3 : 3$ . To arrive at these conclusions, researchers test data from a cross against various models, using the chi-square statistical method. Genetic linkage (discussed next), cytoplasmic inheritance, mutations, and transposable elements (see Chapter 5) are considered the most common causes of non-Mendelian inheritance.

## Pleiotropy

Sometimes, one gene can affect multiple traits, a condition called **pleiotropy**. It is not hard to accept this fact when one understands the complex process of development of an organism in which the event of one stage is linked to those before (i.e., correlated traits). That is, genes that are expressed early in the development of a trait are likely to affect the outcome of the developmental process. In sorghum, the gene *hl* causes the high lysine content of seed storage proteins to increase as well as causing the endosperm to be shrunken. Declaring genes to be pleiotropic is often not clear-cut, since closely associated or closely linked (see next section) genes can behave this way. Conducting a large number of crosses may produce a recombinant, thereby establishing that linkage, rather than pleiotropy, exists.

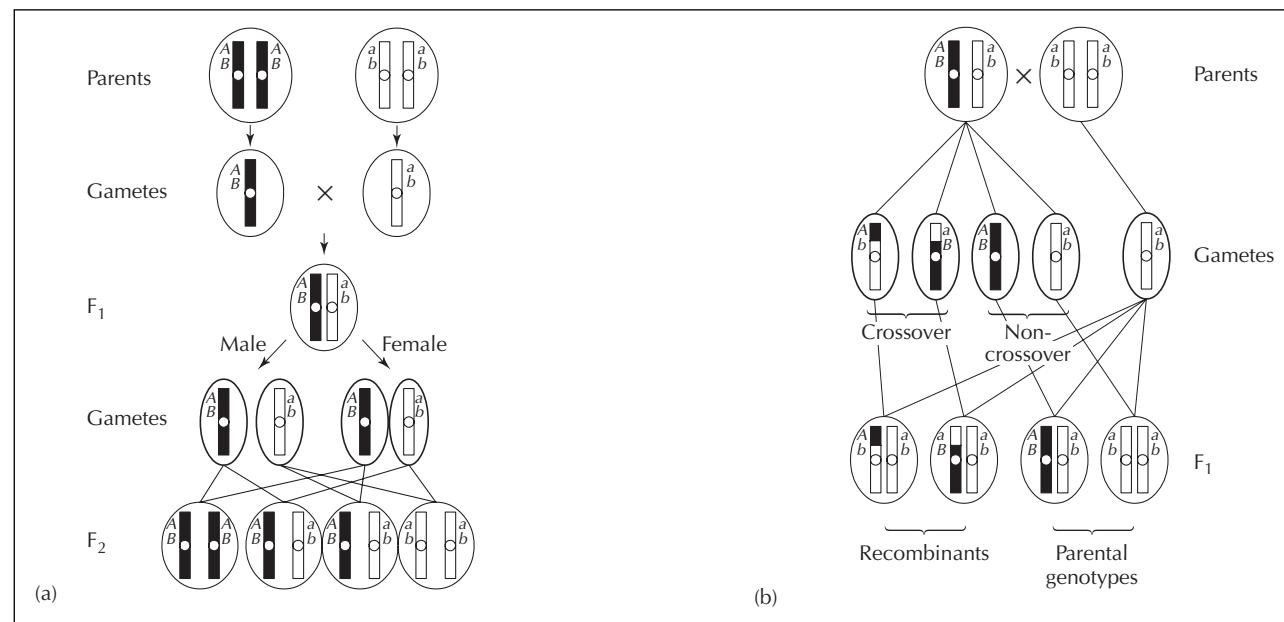
## Genetic linkage and its implications

First reported in sweet pea (*Lathyrus odoratus*) by Cambridge University geneticists, **genetic linkage** is the phenomenon whereby certain genes tend to be inherited together. Because chromosomes are allocated to gametes during nuclear division, the genes they contain tend to be inherited together, an event that violates Mendel's postulate of independent assortment of genes. Genes within a single chromosome constitute a **linkage group**. Consequently, the number of genetic linkage groups in a species corresponds to the haploid number of chromosomes.

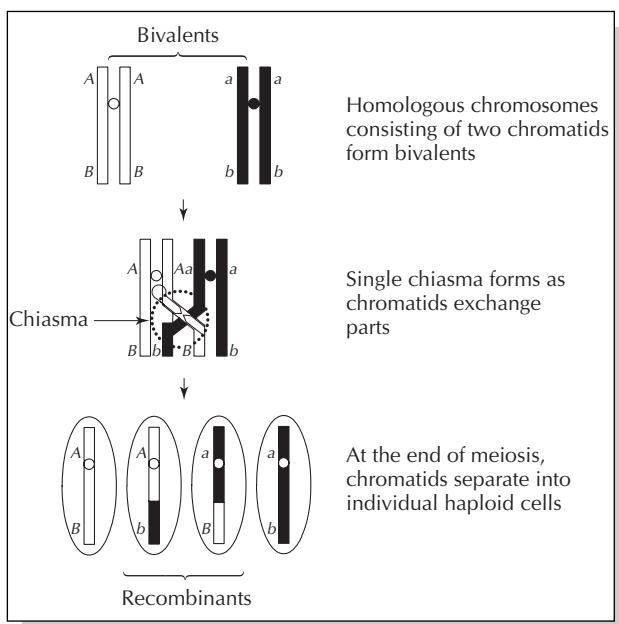
Genes on separate chromosomes as well as genes on the same chromosomes can assort independently. When genes on the same chromosome do not assort independently, they are said to be linked (Figure 3.9). In this example, genes *A* and *B* are transmitted as one gene (a gene block). The consequence of this linkage (called **complete linkage**) is that, instead of nine different genotypes (as would be expected with Mendelian inheritance), only three different genotypes are produced in

$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$
$AABB$ $AAbb$ $AaBb$ $AaBb$	$AABB$ $AAbb$ $AaBb$ $Aabb$	$AaBB$ $AaBb$ $aaBB$ $aaBb$	$AaBb$ $Aabb$ $aaBb$ $aabb$
$AABB$ $AAbb$ $AaBb$ $AaBb$	$AABB$ $AAbb$ $AaBb$ $Aabb$	$AaBB$ $AaBb$ $aaBB$ $aaBb$	$AaBb$ $Aabb$ $aaBb$ $aabb$
(a) Complementary genes 9 : 7			
$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$
$AABB$ $AAbb$ $AaBb$ $AaBb$	$AABB$ $AAbb$ $AaBb$ $Aabb$	$AaBB$ $AaBb$ $aaBB$ $aaBb$	$AaBb$ $Aabb$ $aaBb$ $aabb$
(b) Additive genes 9 : 6 : 1			
$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$
$AABB$ $AAbb$ $AaBb$ $AaBb$	$AABB$ $AAbb$ $AaBb$ $Aabb$	$AaBB$ $AaBb$ $aaBB$ $aaBb$	$AaBb$ $Aabb$ $aaBb$ $aabb$
(c) Duplicate genes 15 : 1			
$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$
$AABB$ $AAbb$ $AaBb$ $AaBb$	$AABB$ $AAbb$ $AaBb$ $Aabb$	$AaBB$ $AaBb$ $aaBB$ $aaBb$	$AaBb$ $Aabb$ $aaBb$ $aabb$
(d) Dominant epistasis 12 : 3 : 1			
$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$	$AB$ $Ab$ $aB$ $ab$
$AABB$ $AAbb$ $AaBb$ $AaBb$	$AABB$ $AAbb$ $AaBb$ $Aabb$	$AaBB$ $AaBb$ $aaBB$ $aaBb$	$AaBb$ $Aabb$ $aaBb$ $aabb$
(e) Recessive epistasis 9 : 3 : 4			

**Figure 3.8** Epistasis or non-Mendelian inheritance is manifested in a variety of ways, according to the kinds of interaction. Some genes work together while other genes prevent the expression of others.



**Figure 3.9** Genetic linkage. (a) Linked genes  $AB/ab$  are transmitted intact from one generation to the next. (b) Genetic linkage may be broken by the process of recombination. A testcross may be used to reveal the occurrence of recombination. Recombinants are the individuals that are derived from gametes with crossovers.



**Figure 3.10** Crossover is preceded by the formation of bivalents, the pairing of homologous chromosomes. Adjacent chromatids physically exchange parts during the formation of a characteristic x-configuration, called the chiasma.

the  $F_2$ , in the ratio of 1( $AABB$ ) : 2( $AaBb$ ) : 1( $aabb$ ). The meiotic products are either parental or non-crossover gametes. In the example in Figure 3.10, the phenomenon of crossing over that occurs in meiosis has caused some alteration in linkage (called **incomplete linkage**). In the absence of linkage, the testcross products would segregate in the genotypic ratio of 1 : 1 : 1 : 1 for the four products,  $AaBb$ ,  $Aabb$ ,  $aaBb$ , and  $aabb$ . However, in this example, the presence of linkage allowed most gametes to inherit parental genotypes ( $AaBb$ ,  $aabb$ ), as a result of normal gamete formation. Crossing over created new genotypes ( $Aabb$ ,  $aaBb$ ; non-parental), called **recombinants** (because they are products of recombination). When the genes of interest are arranged in a homolog such that one chromosome has both dominant alleles (in this example) while the other has both recessive alleles ( $AB/ab$ ), the condition is described as linkage in the **coupling phase**. However, when the arrangement is  $Ab/aB$ , the linkage is in the **repulsion phase**.

Again, in this example, the numbers (frequency) of parental gametes were roughly equal, and so were the numbers for the recombinants. The proportion of recombinant gametes produced in meiosis in the multiple

hybrid is called the **recombination frequency** (RF). If two genes are completely linked,  $RF = 0$ . Detection of linkage is accomplished by using the chi-square test (see Chapter 9):

Genotype	Observed frequency ( $O$ )	Expected frequency ( $E$ )	$(O-E)^2/E$
$A-B-$	284	214.3	22.67
$A-bb$	21	71.4	35.58
$aaB-$	21	71.4	35.58
$aabb$	5	23.8	40.9
	381	380.9	134.72

Degrees of freedom ( $df$ ) = 3; chi-square at  $\alpha = 0.05$  is 7.82. Since the calculated  $\chi^2$  is greater than tabulated, we reject the null hypothesis and declare the presence of linkage.

When a cross involves three gene pairs (a **trihybrid cross**),  $ABC$ , there may be recombination between  $A$  and  $B$ ,  $A$  and  $C$ , and  $B$  and  $C$ . This cross is called a **three-point cross**. The most common genetic types are the parental types, and the least common, the double crossovers. A testcross should reveal eight genotypes in the progeny. The order of the genes can be deduced from a three-point cross because one gene in the middle will be the one that apparently changes places in going from the parental to the double crossover type. For example:

Recombinational events	Gametes	Testcross data
No crossover	$ABC$	401
	$abc$	409
Crossover in $AB$ region	$Abc$	32
	$aBC$	28
Crossover in $BC$ region	$ABc$	61
	$abC$	64
Crossover in both regions (double crossovers)	$AbC$	2
	$aBc$	3
		1,000

Recombination between  $A$  and  $B$  is calculated for:

$$\text{Parental types } (ABC, ABc, abC, abc) = (401 + 61) + (64 + 409) = 935 = 93.5\%$$

$$\text{Recombinant types } (AbC, Abc, aBC, aBc) = (2 + 32) + (28 + 3) = 65 = 6.5\%$$

Recombination between  $B$  and  $C$  can be similarly calculated.

Fortunately, for the plant breeder, genes in a chromosome are not completely linked. If this were so, the

lifeblood of plant breeding, genetic variation, would be very limited. However, during meiosis, as was previously indicated, the phenomenon of crossing over causes recombination or shuffling of linked genes to occur, thereby producing gametes that are unlike the mother cell. Genetic recombination is the most common source of variation in flowering species. Along with independent assortment of genes, these two phenomena ensure that all offspring will contain a diverse mixture of both maternal and paternal alleles.

Whereas breaking linkages is desirable for the creation of the much-needed variation, plant breeders would sometimes rather have certain linkages left intact. This is the case when several desirable genes are tightly linked. On the other hand, there are some occasions when a desirable gene is linked to an undesirable gene, in which case breeders would like to break the association. The probability of breaking a linkage depends on how close the genes are in the group or block. A tight linkage (close association) is more difficult to break than a loose linkage. An opportunity for crossover occurs whenever meiosis occurs.

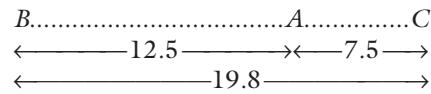
### Chromosome mapping

Plant breeders develop and use “biological maps” to guide them in their work. The two basic types of maps are the physical and genetic maps. Genetic maps are constructed based on the linkage relationship between genes. The degree of crossing over between any two genes or loci on a single chromosome is proportional to the distance between them. This correlation information is used to construct chromosome maps.

**Chromosome maps** provide information about gene locations, gene order, and the relative position of various genes, according to genetic distances. Linkage maps may be used by plant breeders to aid the selection process. If a desired gene is closely linked with a genetic marker, the breeder may use the marker to indirectly select for the desired gene. In the example of a dihybrid cross, it is possible to calculate the genetic distance between the two genes (or markers), but one cannot tell the order of the genes (i.e., whether *A* comes before *B* or *B* before *A*). A trihybrid cross is needed for this determination, as previously stated.

The distance between two genes is defined as the recombination frequency between them. The unit of measure is the **map unit** or **centimorgan** (cM), which is defined as 1% of crossover. In a dihybrid cross, the percent crossover (e.g., between genes *A* and *B*) is

calculated as the percentage of recombinant offspring produced in a cross. For example, for 50 recombinants out of 400 offspring, it is calculated as  $(50/400) \times 100 = 12.5\% = 12.5$  map units. If the crossover between *A* and *C* is calculated as 7.5% and between *B* and *C* as 19.8%, then the gene order is *BAC*.

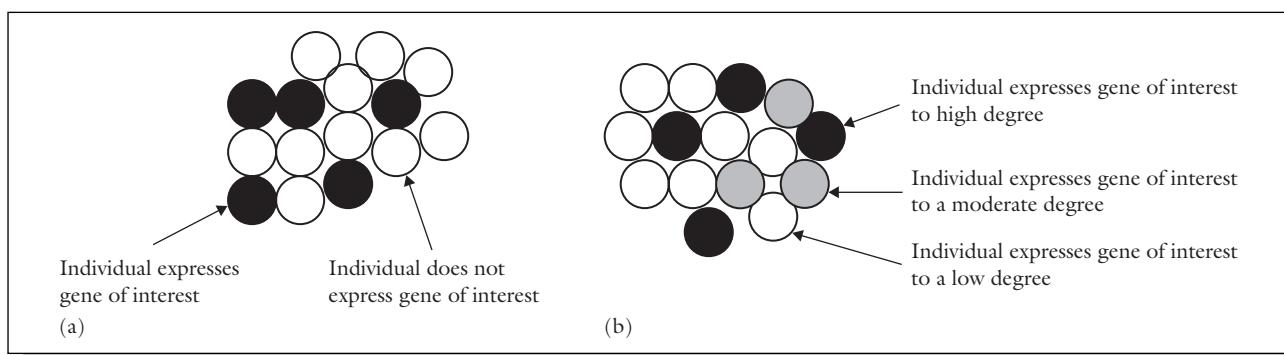


A low frequency of double crossover between *B* and *C* will give the parental genotype, so that the crossover units will be less than the sum of those between *B* and *A*, and *A* and *C* combined. Further, genes that are separated by 50 or more crossover units are essentially non-linked and will assort independently.

**Physical maps** are constructed based on nucleotides, the building blocks of DNA. Genetic distance on a linkage map expressed in centimorgans is not directly correlated with the physical distance expressed in nucleotides.

### Penetrance and expressivity

It has previously been said that the environment in which a gene occurs influences how it is expressed. The source of this environmental effect could be as close and intimate as the immediate cellular environments, or as remote as the general plant external environment. In plants, breeders may transfer genes from one genetic background into another through hybridization. Sometimes, they encounter a situation in which the gene may be successfully transferred, but the desired effect is not observed. A researcher can quantitatively study the degree of expression of a trait. For example, in one case, a disease-resistance gene may offer resistance in one plant but fail to do the same in another plant from the same population. This phenomenon is described as **variable gene penetrance** and measures the percentage of individuals that show some degree of expression of the genotype of interest. If 20% of plants show the desired resistance trait, the resistance gene is said to have 80% penetrance (Figure 3.11a). Sometimes, changes in the plant environment may cause the same plant to produce different phenotypes or degrees of expression of a trait under these different conditions. For example, the hibiscus plant normally produces single flowers (a flower with one set of petals). A double-flowered mutant (additional petals added to the primary set) has



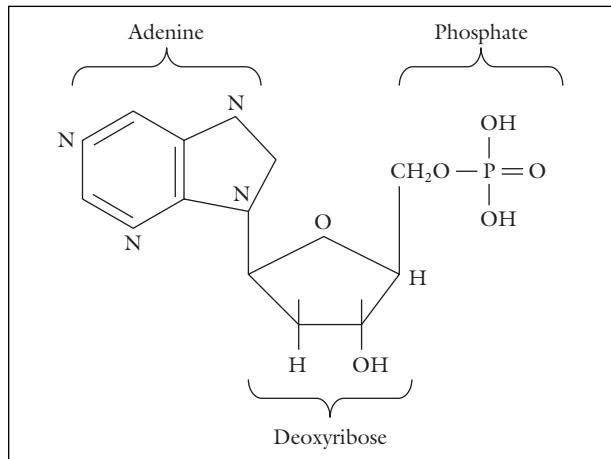
**Figure 3.11** Diagrammatic presentation of (a) penetrance and (b) expressivity.

been developed. However, the number of petals of the double flower is influenced by temperature. When grown at between 1.5 and 10°C, the double flower characteristic is lost or diminished, and the flowers produce fewer petals. This gene interaction is called **variable gene expressivity**, and describes the *range* of expression of the genotype of interest (Figure 3.11b).

The effect of genetic background on the expression of a phenotype is often difficult to assess. The expression of other genes in a genome may affect the phenotype observed, a phenomenon called **genetic suppression**. Suppression of genes is known to modify the effect of primary genes. Sometimes, relocation of a gene in the genome can influence the expression of the gene, a phenomenon called **position effect**. This may occur when chromosomal mutations such as translocations and inversions occur (a region of the chromosome is relocated to another part of the chromosome).

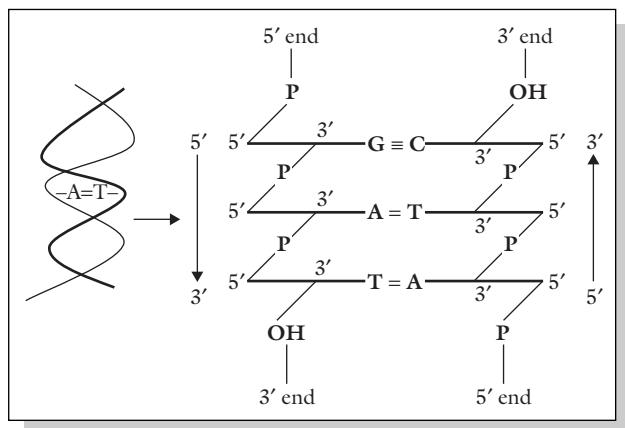
### Nucleic acids: structure and function

**Nucleic acids** are polymers of **nucleotides**. There are two kinds of nucleic acids: **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. A nucleotide consists of three basic components: pentose sugar, nitrogenous base, and a phosphate group. The sugar is a cyclic five-carbon sugar and is ribose in RNA and deoxyribose in DNA. Similarly, there are two kinds of bases: **purines** and **pyrimidines**. There are two purines, adenine (A) and guanine (G), and three pyrimidines, cytosine (C), thymine (T), and uracil (U). Thymine occurs only in DNA, while uracil occurs only in RNA. The letters A, C, T, G, are casually referred to as the alphabets of life.



**Figure 3.12** The basic chemical structure of a nucleotide molecule, showing its three constituents: a sugar, a nitrogenous base, and a phosphate group.

When a base is linked to a sugar, the product is called a **nucleoside**. A nucleoside linked to a phosphate forms a **nucleotide** (Figure 3.12). Two nucleotides may be linked by a phosphodiester group to form a dinucleotide. Shorter chains (consisting of less than 20 nucleotides) are called **oligonucleotides** while longer chains are called **polynucleotides**. A single nucleoside is also called nucleoside monophosphate (NMP), while two nucleosides form a nucleoside diphosphate (NDP). Triphosphates are important in cellular bioenergetics, especially adenosine triphosphate (ATP) and guanosine triphosphate (GTP). When these compounds are hydrolyzed, inorganic phosphate is produced, accompanied by the release of energy (e.g.,  $\text{ATP} \rightarrow \text{ADP} + \text{energy}$ ).



**Figure 3.13** The DNA molecule has a double helix structure comprising a sugar-phosphate backbone and horizontal rungs of nitrogenous bases. The two chains are antiparallel. The helix has minor grooves alternating with major grooves.

### Structure of DNA

DNA is the universal, hereditary material (except in certain viruses – RNA viruses). The most powerful direct evidence for DNA being the hereditary material is currently provided by the cutting-edge technology of **recombinant DNA (rDNA)**. The structure of the DNA molecule is a **double helix** (Figure 3.13). The key features about the DNA molecule are as follows:

- 1 It consists of two polynucleotide chains coiled around a central axis in a spiral fashion. The most common natural form of DNA is a right-handed double helix of diameter 2.0 nm, called the **B-DNA**. A left-handed form (**Z-DNA**) and an **A-form** of DNA also occur under certain conditions.
- 2 The polynucleotide chains are **antiparallel**; one chain runs in the 5' to 3' orientation and the other 3' to 5' (carbon atoms of a sugar are conventionally numbered from the end closest to the aldehyde or ketone).
- 3 The two bases in each base pair lie in the same plane. Each plane is perpendicular to the axis of the helix. There are 10 base pairs per helical turn.
- 4 The helix has two kinds of alternating external grooves: a deep groove (called the **major groove**) and a shallow groove (called the **minor groove**).
- 5 The nitrogenous bases on one strand pair with those on the other strand in **complementary** fashion (A always pairs with T, while G pairs with C).

In addition to these features described above, certain implications deserve emphasis:

- 1 Complementary base pairing means that the replicate of each strand is given the base sequence of its complementary strand when DNA replicates.
- 2 Because the strands are antiparallel, when two nucleotides are paired, the sugar portions of these molecules lie in opposite directions (one upward and the other downward along the chain).
- 3 Because the strands are antiparallel, the convention for writing the sequence of bases in a strand is to start from the 5'-P terminus at the left (e.g., GAC refers to a trinucleotide 5'-P-GAC-3'-OH).
- 4 The conventional way of expressing the base composition of an organism is by the percentage of [G] + [C]. This value is approximately 50% for most eukaryotes with only minor variations among species. In simpler organisms, there are significant variations (e.g., 27% for *Clostridium*, 50% for *Escherichia coli*, and 76% for *Sarcina*, all of these organisms being bacteria).
- 5 The chains of the double helix are held together by hydrogen bonds between base pairs in opposite strands. The bond between A and T is a double bond, while the bond between G and C is a triple hydrogen bond.

### Structure of RNA

RNA is similar in structure to DNA. However, there are significant differences, the key ones being:

- 1 RNA consists of ribose sugar (in place of deoxyribose) and uracil in place of thymine.
- 2 Most RNA is predominantly single stranded (except in some viruses). Sometimes, the molecule folds back on itself to form double-stranded regions.
- 3 Certain animal and plant viruses use RNA as their genetic material.
- 4 A typical cell contains about 10 times more RNA than DNA.
- 5 Whereas DNA stores genetic information, RNA most often functions in the expression of the genetic information.
- 6 There are three major classes of RNA known to be involved in gene expression: ribosomal RNA (rRNA), messenger RNA (mRNA), and transfer RNA (tRNA). The site of protein synthesis, the ribosome, contains rRNA.

#### Messenger RNA structure

**Messenger RNA (mRNA)** is the molecular carrier of genetic information from the DNA to ribosomes, where this DNA transcript or template is translated (the genetic information of DNA transcript is expressed)

into proteins. Because genes vary in size (number of nucleotides) the mRNA species are variable in length.

#### Transfer RNA structure

The structure of tRNA is very unique among the three key RNA molecules in the cell. These molecules are small in size and very stable. tRNA molecules range in size from 75 to 90 nucleotides. Single stranded, a tRNA molecule is able to fold back onto itself and undergo complementary base pairing in short stretches to form double strands. This folding also creates four characteristic loops and a cloverleaf 2D structure (Figure 3.14). Of the four loops, three are involved in translating the message of the mRNA. The anticodon loop (or simply anticodon) consists of a sequence of three bases that are complementary to the sequence of a codon on the mRNA. The stop codons do not have tRNA with anticodons for them. Another feature of the tRNA molecule is the occurrence of the sequence pCpCpA-3' at the 3' end. The terminal adenine residue is the point of attachment for an amino acid and hence is called the amino acid attachment (or binding) site. During protein synthesis, the amino acid corresponding to a particular mRNA codon that base pairs with the tRNA anticodon is attached to this terminal and transported to the appropriate segment of the mRNA.

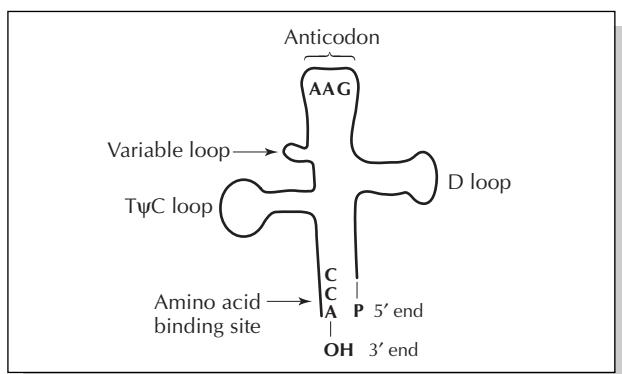
#### Ribosomal structure

**Ribosomes** are the sites (“factories”) of **polypeptide synthesis** (or **protein synthesis**). A bacterial cell may contain about 1,000 ribosomes. A ribosome consists

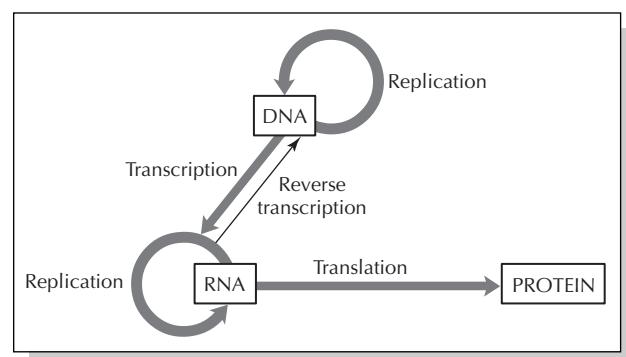
of two subunits, which together form the monosome. The ribosomal particles are classified according to their sedimentation coefficient or rate (S). Monosomes of bacteria are 70S (70S ribosomes) whereas eukaryotic monosomes are about 80S. Because sedimentation coefficients are not additive, a 70S monosome in actuality comprises two subunits that are 50S and 30S, while an 80S monosome consists of 60S and 40S subunits. A ribosome subunit consists of molecules of rRNA and proteins. For example, the 50S subunit contains one 55 rRNA molecule, one 235 rRNA molecule, and 32 different ribosomal proteins.

#### Central dogma of molecular biology

The genetic information of the DNA is changed into biological material principally through proteins, according to the **central dogma of molecular biology**. The dogma states that genetic information flow is generally unidirectional from DNA to proteins, except in special cases (Figure 3.15). This flow, mediated by **transcription** (copying of the DNA template by synthesizing the RNA molecule) and **translation** (synthesis of a polypeptide using the genetic information encoded in an mRNA molecule), and preceded by **replication** (the process of DNA synthesis), can now be reversed *in vitro* (in the test tube) by scientists. Thus, once a protein is known, the nucleotide sequence in the prescribing DNA strand can be determined and synthesized (the product is called a **complementary DNA** or **cDNA**). Production of cDNA is a technique used in genetic engineering (see Chapter 14).



**Figure 3.14** A tRNA molecule has a cloverleaf shape. Two parts of special interest are the anticodon and the amino acid binding sites that are critical in polypeptide or protein synthesis.



**Figure 3.15** The central dogma of molecular genetics, showing the information flow involving DNA, RNA, and proteins within a cell. Simply stated, DNA makes RNA, which in turn makes proteins.

## Expression of genetic information

A key question in genetics is how the information of the DNA is interpreted to produce protein. The expression of genetic information involves complex molecular events as summarized next.

### The genetic code

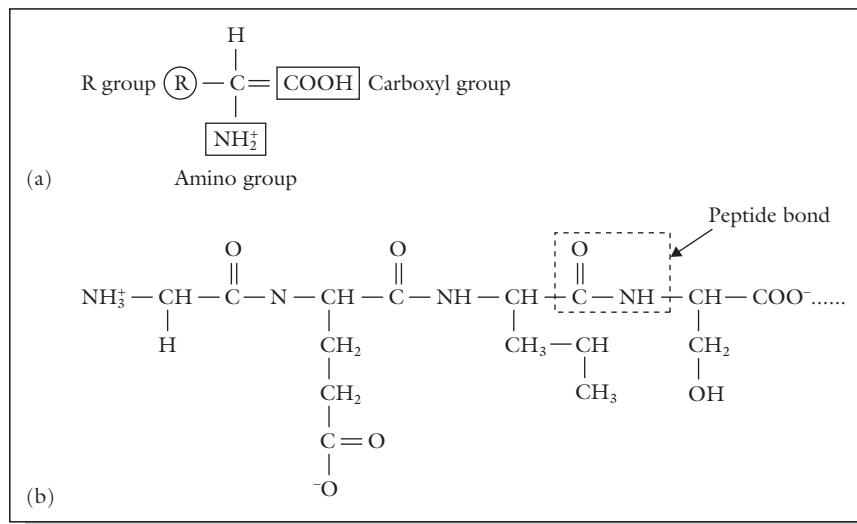
The sequence of bases in the polynucleotide chain holds the key to DNA function. The sequence is critical because it represents the **genetic code** (the set of rules giving the correspondence between mRNA and amino acids in a protein) for the synthesis of corresponding amino acids that constitute proteins. DNA does not code for adult traits directly, there being no genes for adult traits as such. Instead, genes code for various developmental processes. The variety of protein products in a cell undertake catalytic and structural activities that eventually result in an adult phenotype.

There are about 20 commonly occurring amino acids. According to the prescribed sequence (based on the genetic code), amino acids are joined together by **peptide bonds** to form polypeptide chains (Figure 3.16). The genetic code is a triplet code. Three adjacent bases form a code for an amino acid. Each trinucleotide sequence is called a **codon** (Figure 3.17). The genetic code is read from a fixed starting point of the DNA strand.

The genetic code is said to be degenerate because nearly all amino acids are specified by at least two codons. Some (serine, arginine, leucine) are encoded by six different codons. Only tryptophan and methionine are encoded by single codons. Further, for a set of codons encoding the same amino acid, the first two letters in the figure are the same, with only the third being different (called the **wobble hypothesis**). Consequently, at least 30 different tRNA species are required to account for the 61 different triplets in the coding dictionary in Figure 3.17 (the three remaining triplets include termination codons or signals – UAG, UAA, UGA).

### Transcription: RNA synthesis

The genetic information of the DNA template is copied by the process of **transcription** (or RNA synthesis) to produce an RNA sequence (mRNA). The DNA strand that is transcribed is called the **template strand**. The process starts with a recognition of a special DNA sequence (called a **promoter**) and binding to it by an enzyme, a process called **template binding**. The RNA chain then grows (chain elongation) in the 3' direction. The first product of transcription in eukaryotes is called **pre-mRNA**, part of a group of molecules called **heterogeneous nuclear DNA (hnRNA)**. This molecule undergoes severe alterations to remove non-coding parts (**introns**) of the sequence, leaving the coding parts



**Figure 3.16** (a) The basic structure of an amino acid consists of three units – an amino group, a carboxyl group, and a side chain (R) – that distinguish among the different amino acids. (b) A polypeptide chain is formed by linking many amino acids together; adjacent amino acids are linked by a peptide bond.

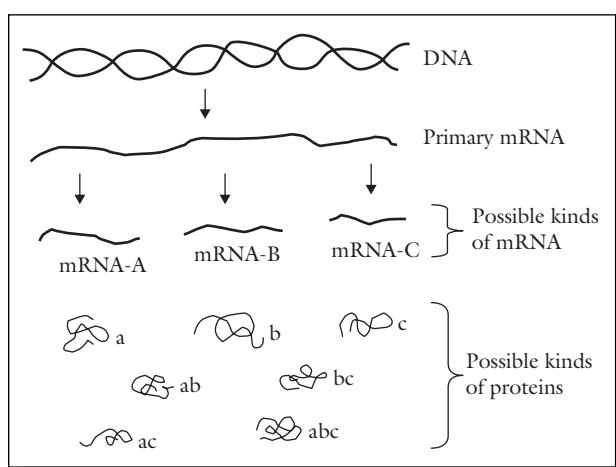
First base	Second base								Third base
	U	C	A	G					
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys					U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys					C
	UUA Leu	UCA Ser	[UAA] stop	[UGA] stop					A
	UUG Leu	UCG Ser	[UAG] stop	UGG Trp					G
C	CUU Leu	CCU Pro	CAU His	CGU Arg					U
	CUC Leu	CCC Pro	CAC His	CGC Arg					C
	CUA Leu	CCA Pro	CAA Gin	CGA Arg					A
	CUG Leu	CCG Pro	CAG Gin	CGG Arg					G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser					U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser					C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg					A
	AUG Met	ACG Thr	AAG Lys	AGG Arg					G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly					U
	GUC Val	GCC Ala	GAC Asp	GGC Gly					C
	GUU Val	GCA Ala	GAA Glu	GGA Gly					A
	GUG Val	GCG Ala	GAG Glu	GGG Gly					G

**Figure 3.17** The genetic code may be likened to a coding dictionary for constructing polypeptide chains. The triplets UAG, UAA, and UGA are termination signals and do not code for amino acids. Of the remaining codes, all amino acids are encoded by at least two codes (up to six in some), except for tryptophan.

(exons) to produce the mRNA, which is typically about 25% of the original length of the pre-mRNA. After removing the introns, the splicing or linking of the pieces results in different types of mRNA (called **alternative splicing**). Consequently, different kinds of proteins may be encoded by the same gene (Figure 3.18). The mRNA is transported to the ribosomes.

### Translation: protein synthesis

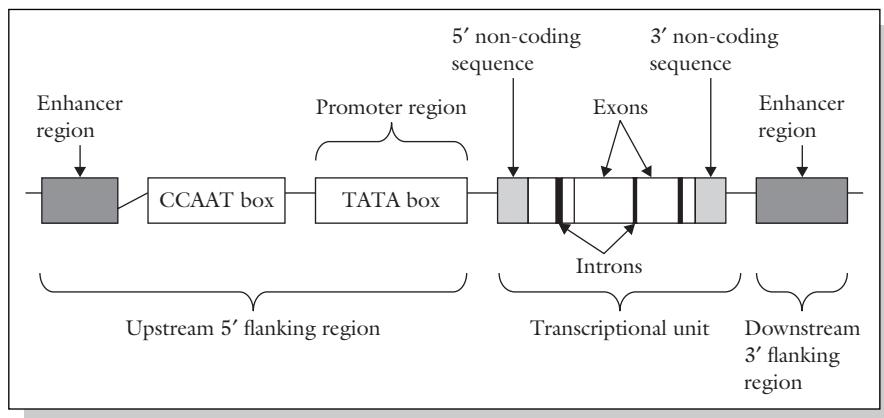
Protein synthesis consists of three steps – **initiation**, **elongation**, and **termination**. Translation starts with the formation of an initiation complex that includes initiation factors that bind to the small rRNA subunit and then to the mRNA. The next step is to set the reading frame for accurate translation. The AUG triplet is usually the initiation point. The large subunit binds to the complex. The sequence of the next triplet determines which charged tRNA (with an amino acid attached) will be attached. The process is repeated until the whole mRNA is translated, adjacent amino acids being linked by peptide bonds. The termination of translation occurs when the elongation process encounters a **stop codon** or termination codon. The interval between the start and stop codons that encodes an amino acid for inser-



**Figure 3.18** Alternative splicing of the mRNA in eukaryotes to remove introns and joining exons results in the production of different mature mRNAs and consequently different protein products.

tion into a polypeptide chain is called the **open reading frame (ORF)**.

Each gene codes for one polypeptide. Some proteins comprise more than one polypeptide (have multiple



**Figure 3.19** Diagrammatic presentation of a typical eukaryotic gene showing the three basic regions – the upstream 5' flanking regions, the transcriptional unit, and the downstream 3' flanking region – and their constitution.

subunits). All genes do not code for proteins, and further, all genes in a cell are not actively transcribing mRNA all of the time. Also, most enzymes are proteins, but all proteins are not enzymes.

### Protein structure

Polypeptides are precursors of proteins. Once produced, they fold to assume 3D forms, the functional stage that becomes proteins. There are four basic levels of protein structure – **primary, secondary, tertiary, and quaternary**. The primary structure of proteins is the sequence of the amino acids in the linear backbone of the polypeptide. The next fold (exemplified by the DNA molecule), is an  $\alpha$ -helix, a spiral chain of amino acids stabilized by hydrogen bonds. The secondary structure describes the arrangement of amino acids within certain areas of the polypeptide chain. The tertiary structure is a 3D conformation of the entire chain in space. Proteins with more than one polypeptide chain may exhibit the quaternary protein structure through aggregations of the polypeptides.

### Regulation of gene expression

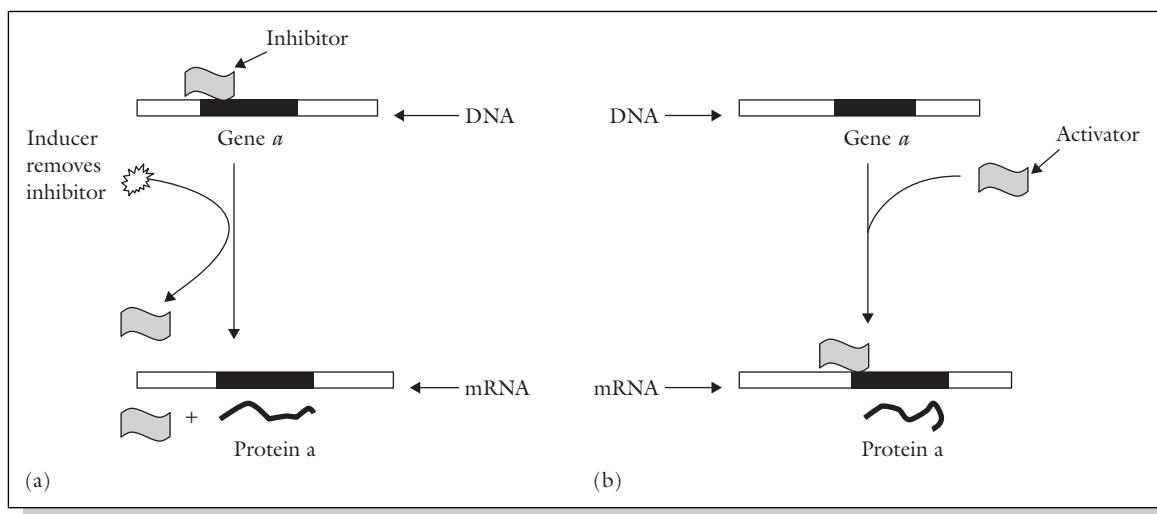
**Gene regulation** is a critical activity performed by plants for proper growth and development. It is not important for a gene just to be expressed, but its expression must be regulated such that it is expressed at the right time only and to the desired extent. Regulation entails the “turning on” and “turning off”

of genes. It is through regulation of gene expression that cellular adaptation, variation, differentiation, and development occur. Some genes are turned on all the time (called **constitutive expression**), while others are turned on only some of the time (called **differential expression**).

The underlying principle of gene regulation is that there are regulatory molecules that interact with nucleic acid sequences to control the rate of transcription or translation. Six potential levels for regulation of gene expression exist in eukaryotes – the regulation of: (i) transcription; (ii) RNA processing; (iii) mRNA transport; (iv) mRNA stability; (v) translation; and (vi) protein activity. Transcription is temporarily and spatially separated from translation in eukaryotes.

A typical eukaryotic gene is shown in Figure 3.19. Unlike that of a **monocistronic** gene (lacks introns; has one transcriptional unit and one translational unit) as occurs in bacteria, eukaryotic genes are **polycistronic** (split genes with introns). Genes that encode the primary structures of proteins required by all cells for enzymatic or structural functions are called **structural genes**. In prokaryotes, these genes are organized into clusters that are transcribed as a single unit (coordinately controlled). The mRNA is called polycistronic mRNA, coding for multiple proteins involved in the same regulatory pathway (e.g., the *lac* operon).

There are two basic categories of gene regulation – negative and positive (Figure 3.20). In negative regulation, an inhibitor that is bound to a DNA (gene) must be removed in order for transcription to occur. In positive regulation, gene transcription occurs when an activator binds to the DNA. One of the main ways in



**Figure 3.20** Schematic representation of the regulation of gene expression showing (a) negative gene regulation and (b) positive gene regulation.

which scientists genetically engineer organisms is by manipulating the gene expression.

### Synteny and plant breeding

In Chapter 2, Vavilov's law of homologous series was introduced. Gene order in chromosomes is conserved over wide evolutionary distances. In some comparative studies, scientists discovered that large segments of chromosomes, or even entire chromosomes in some cases, had the same order of genes. However, the spacing between the mapped genes was not always proportional. The term **colinearity** is used to refer to the conservation of the gene order within a chromosomal

segment between different species. The term **synteny** is technically used to refer to the presence of two or more loci on the same chromosome that may or may not be linked. Modern definition of the term has been broadened to include the homoeology (homoeologous chromosomes are located in different species or in different genomes in polyploid species and originate from a common ancestral chromosomes) of originally completely homologous chromosomes. Whole-genome comparative maps have been developed for many species, but are most advanced in the Gramineae family (Poaceae). Some researchers have attempted to clone a gene in one plant species based on the detailed and sequence information (microsynteny) in a homoeologous region of another genus.

### References and suggested reading

- Clark, A.G. 2002. Limits to the prediction of phenotypes from knowledge of genotypes. *Evol. Biol.* 32:205–224.  
 Crick, F.H.C. 1962. The genetic code. *Sci. Am.* 207:66–77.  
 Felsenfeld, G. 1985. DNA. *Sci. Am.* 253:58–78.  
 Holland, J.B. 2001. Epistasis and plant breeding. *Plant Breed. Rev.* 21:27–92.  
 Klug, W.S., and M.R. Cummings. 1996. *Essentials of genetics*, 2nd edn. Prentice Hall, Upper Saddle, NJ.  
 Klug, W.S., and M.R. Cummings. 1997. *Concepts of genetics*, 5th edn. Prentice Hall, Upper Saddle, NJ.  
 Stahl, F.W. 1979. Genetic recombination. *Sci. Am.* 256:91–101.  
 Weinberg, R.A. 1985. The molecules of life. *Sci. Am.* 253:48–57.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 A whole organism can be raised by nurturing a single cell.
- 2 The diploid chromosome number has half the number of chromosomes in the gametic cell.
- 3 Mitosis produces identical daughter cells.
- 4 A heterozygote has identical alleles of a gene at the locus.
- 5 Thymine occurs in DNA but not in RNA.

### Part B

Please answer the following questions:

- 1 The four nucleic acids of DNA are ..... , ..... , ..... , and .....
- 2 Distinguish between DNA and RNA molecules.
- 3 Discuss the levels of organization of eukaryotes.
- 4 In plants DNA occur in the nucleus, ..... , and .....
- 5 A nucleic acid consists of a base, ..... , and .....
- 6 Define epistasis.
- 7 What is an mRNA?
- 8 Give the three basic types of tissue and distinguish among them.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the laws of Mendel.
- 2 Discuss the use of a testcross in plant breeding.
- 3 Discuss genetic linkage and its importance in plant breeding.
- 4 Distinguish between the phenomenon of variable gene penetrance and variable gene expressivity.
- 5 Discuss the regulation of gene expression.



# 4

## Plant reproductive systems

### Purpose and expected outcomes

*Reproduction is the process by which plants multiply themselves. It is not only important to plant producers but also to plant breeders. The mode of reproduction determines the method of breeding the species and how the product of breeding is maintained for product identity preservation. It is important to add that, whereas in the past plant breeding methods were fairly distinct for self-pollinated species and for cross-pollinated species, such a clear distinction does not currently exist. Rather, the methods of plant breeding for the two groups tend to overlap. After studying this chapter, the student should be able to:*

- 1 Discuss the types of plant life cycles and their implication in breeding.
- 2 Describe the basic types of floral morphology.
- 3 Discuss the mechanisms of pollination and fertilization.
- 4 Discuss the breeding implications of self- and cross-pollination.
- 5 Describe the constraints to pollination and their implication in breeding.
- 6 Discuss the genetics and applications of male sterility in breeding.

### Importance of mode of reproduction to plant breeding

Plant breeders need to understand the reproductive systems of plants for the following key reasons:

- 1 The genetic structure of plants depends on their mode of reproduction. Methods of breeding are generally selected such that the natural genetic structure of the species is retained in the cultivar. Otherwise, special efforts will be needed to maintain the newly developed cultivar in cultivation.
- 2 In flowering species, artificial hybridization is needed to conduct genetic studies to understand the inheritance of traits of interest, and for transfer of genes of interest from one parent to another. To accomplish this, the breeder needs to thoroughly understand the floral biology and other factors associated with flowering in the species.

- 3 Artificial hybridization requires an effective control of pollination so that only the desired pollen is allowed to be involved in the cross. To this end, the breeder needs to understand the reproductive behavior of the species. Pollination control is critical to the hybrid seed industry.
- 4 The mode of reproduction also determines the procedures for multiplication and maintenance of cultivars developed by plant breeders.

### Overview of reproductive options in plants

Four broad contrasting pairs of reproductive mechanisms or options occur in plants.

- 1 **Hermaphrodity versus unisexuality.** Hermaphrodites have both male and female sexual organs and hence

may be capable of self-fertilization. On the other hand, unisexuals, having one kind of sexual organ, are compelled to cross-fertilize. Each mode of reproduction has genetic consequences, hermaphrodity promoting a reduction in genetic variability, whereas unisexuality, through cross-fertilization, promotes genetic variability.

- 2 **Self-pollination versus cross-pollination.** Hermaphrodites that are self-fertile may be self-pollinated or cross-pollinated. In terms of pollen donation, a species may be **autogamous** (pollen comes from the same flower – selfing), or **allogamous** (pollen comes from a different flower). There are finer differences in these types. For example, there may be differences between the time of pollen shed and stigma receptivity.
- 3 **Self-fertilization versus cross-fertilization.** Just because a flower is successfully pollinated does not necessarily mean fertilization will occur. The mechanism of self-incompatibility causes some species to reject pollen from their own flowers, thereby promoting outcrossing.
- 4 **Sexuality versus asexuality.** Sexually reproducing species are capable of providing seed through sexual means. Asexuality manifests in one of two ways – vegetative reproduction (in which no seed is produced) or agamospermy (in which seed is produced).

### Types of reproduction

Plants are generally classified into two groups based on mode of reproduction as either **sexually reproducing** or **asexually reproducing**. Sexually reproducing plants produce seed as the primary propagules. Seed is produced after sexual union (fertilization) involving the fusion of sex cells or **gametes**. Gametes are products of meiosis and, consequently, seeds are genetically variable. Asexual or vegetative reproduction mode entails the use of any vegetative part of the plant for propagation. Some plants produce modified parts such as creeping stems (stolons or rhizomes), bulbs, or corms, which are used for their propagation. Asexual reproduction is also applied to the condition whereby seed is produced without fusion of gametes (called apomixis). It should be pointed out that some plants can reproduce by either the sexual or asexual mode. However, for ease of either propagation or product quality, one mode of reproduction, often the vegetative mode, is preferred. Such is the case in flowering species such as potato (propagated by tubers or stem cuttings) and sugarcane (propagated by stem cuttings).

### Sexual reproduction

#### Sexual life cycle of a plant (alternation of generations)

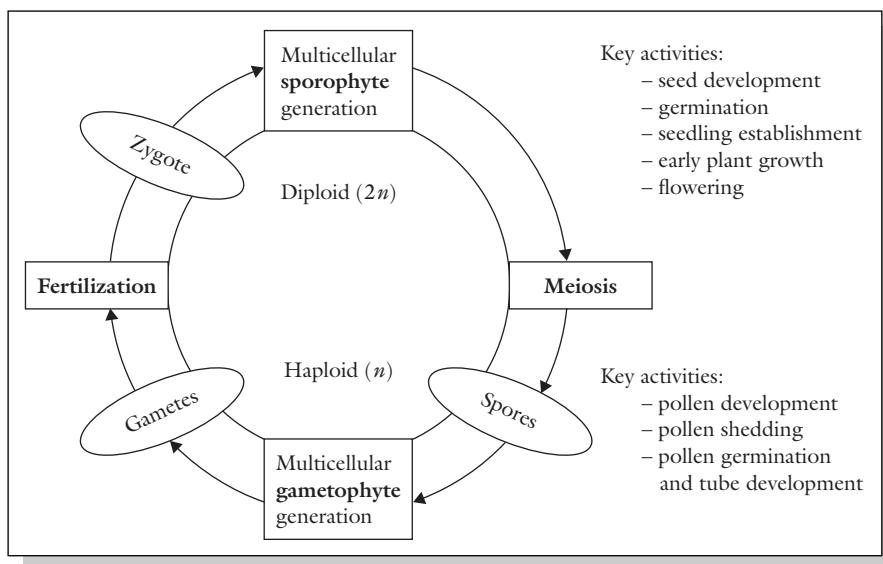
The normal sexual life cycle of a flowering plant may be simply described as consisting of events from establishment to death (from seed to seed in seed-bearing species). A flowering plant goes through two basic growth phases: **vegetative** and **reproductive**, the former preceding the latter. In the vegetative phase, the plant produces vegetative growth only (stem, branches, leaves, etc., as applicable). In the reproductive phase, flowers are produced. In some species, exposure to a certain environmental factor (e.g., temperature, photoperiod) is required to switch from the vegetative to reproductive phase. The duration between phases varies among species and can be manipulated by modifying the growing environment.

In order for sexual reproduction to occur, two processes must occur in sexually reproducing species. The first process, **meiosis**, reduces the chromosome number of the diploid ( $2n$ ) cell to the haploid ( $n$ ) number. The second process, **fertilization**, unites the nuclei of two gametes, each with the haploid number of chromosomes, to form a diploid. In most plants, these processes divide the life cycle of the plant into two distinct phases or generations, between which the plant alternates (called **alternation of generations**) (Figure 4.1). The first phase or generation, called the **gametophyte generation**, begins with a haploid spore produced by meiosis. Cells derived from the gametophyte by mitosis are haploid. The multicellular gametophyte produces gametes by mitosis. The sexual reproductive process unites the gametes to produce a **zygote** that begins the diploid **sporophyte generation** phase.

In lower plants (mosses, liverworts), the sporophyte is small and dependent upon the gametophyte. However, in higher plants (ferns, gymnosperms, angiosperms), the male gametophyte generation is reduced to a tiny pollen tube and three haploid nuclei (called the microgametophyte). The female gametophyte (called the megagametophyte) is a single multinucleated cell, also called the **embryo sac**. The genotype of the gametophyte or sporophyte influences sexual reproduction in species with self-incompatibility problems. This has implications in the breeding of certain plants as discussed further in this chapter.

#### Duration of plant growth cycles

The plant breeder should know the life cycle of the plant to be manipulated. The strategies for breeding are



**Figure 4.1** Schematic representation of the alternation of generations in flowering plants. The sporophyte generation is diploid, and often the more conspicuous phase of the plant life cycle. The gametophyte is haploid.

influenced by the duration of the plant growth cycle. Angiosperms (flowering plants) may be classified into four categories based on the duration of their growth cycle as follows (Figure 4.2).

#### Annuals

Annual plants (or annuals) complete their life cycle in one growing season. Examples of such plants include corn, wheat, and sorghum. Annuals may be further categorized into **winter annuals** or **summer annuals**. Winter annuals (e.g., wheat) utilize parts of two seasons. They are planted in fall and undergo a critical physiological inductive change called **vernalization** that is required for flowering and fruiting in spring. In cultivation, certain non-annuals (e.g., cotton) are produced as though they were annuals.

#### Biennials

A biennial completes its life cycle in two growing seasons. In the first season, it produces basal roots and leaves; then it grows a stem, produces flowers and fruits, and dies in the second season. The plant usually requires a special environmental condition or treatment (e.g., vernalization) to be induced to enter the reproductive phase. For example, sugar beet grows vegetatively in the first season. In winter, it becomes vernalized and starts reproductive growth in spring.

#### Perennials

Perennials are plants that have the ability to repeat their life cycles indefinitely by circumventing the death stage. They may be herbaceous, as in species with underground vegetative structures called **rhizomes** (e.g., indiangrass) or aboveground structures called **stolons** (e.g., buffalograss). They may also be woody as in shrubs, vines (grape), and trees (orange).

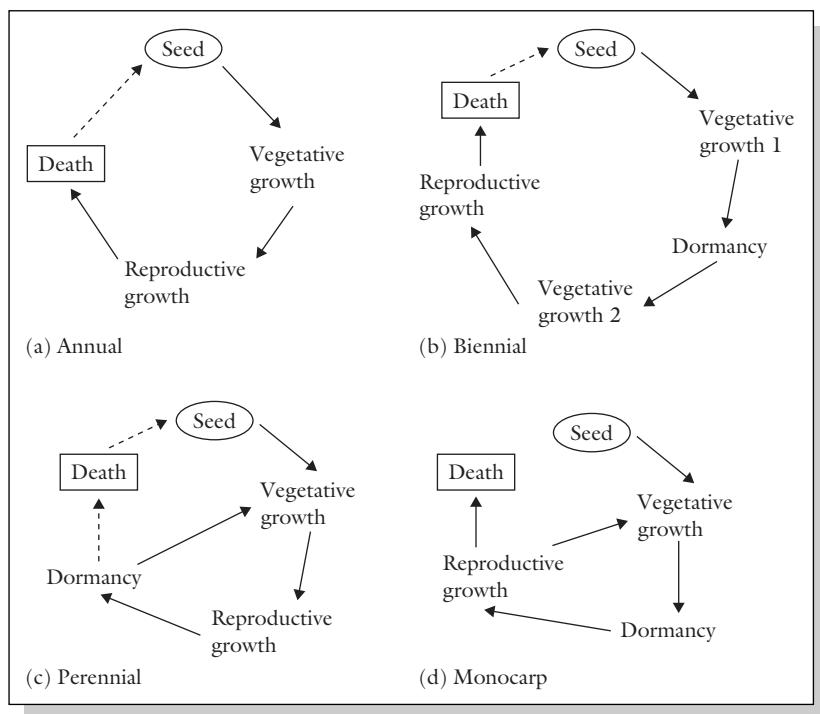
#### Monocarps

Monocarps are characterized by repeated, long vegetative cycles that may go on for many years without entering the reproductive phase. Once flowering occurs, the plant dies. Common examples are bromeliads. The top part dies, so that new plants arise from the root system of the old plant.

It should be pointed out that certain plants that may be natural biennials or perennials are cultivated by producers as annuals. For example, sugar beet, a biennial, is commercially produced as an annual for its roots. For breeding purposes it is allowed to bolt to produce flowers for crossing, and subsequently to produce seed.

#### Structure of the flower

Genetic manipulation of flowering plants by conventional tools is accomplished by the technique of crossing,

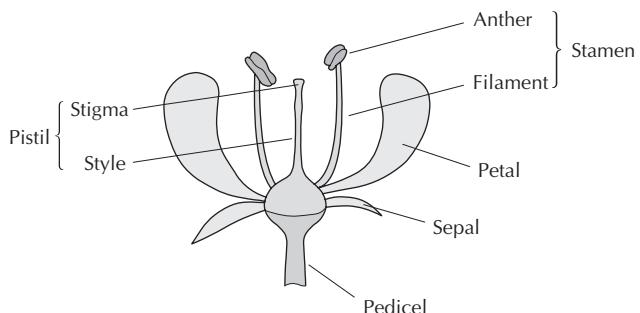


**Figure 4.2** Flowering plants have one of four life cycles – annual, biennial, perennial, and monocarp. Variations occur within each of these categories, partly because of the work of plant breeders.

which involves flowers. To be successful, the plant breeder should be familiar with the flower structure, regarding the parts and their arrangement. Flower structure affects the way flowers are emasculated (prepared for crossing by removing the male parts to make the flower female). The size of the flower affects the kinds of tools and techniques that can be used for crossing.

### General reproductive morphology

Four major parts of a flower are generally recognized: **petal**, **sepal**, **stamen**, and **pistil**. These form the basis of flower variation. Flowers vary in the color, size, numbers, and arrangement of these parts. Typically, a flower has a receptacle to which these parts are attached (Figure 4.3). The male parts of the flower, the **stamen**, comprise a stalk called a **filament** to which is attached a structure consisting of four pollen-containing chambers that are fused together (**anther**). The stamens are collectively called the **androeum**. The center of the flower is occupied by a pistil, which consists of the style, stigma, and ovary (which contains the carpels). The pistil is also called the **gynoecium**. Sepals are often leaf-like structures that enclose the flower in its bud stage.



**Figure 4.3** The typical flower has four basic parts – the petals, sepals, pistil, and stamen. The shape, size, color, and other aspects of these floral parts differ widely among species.

Collectively, sepals are called the **calyx**. The showiest parts of the flower are the petals, collectively called the **corolla**.

### Types of flowers

When a flower has all the four major parts, it is said to be a **complete flower** (e.g., soybean, tomato, cotton,

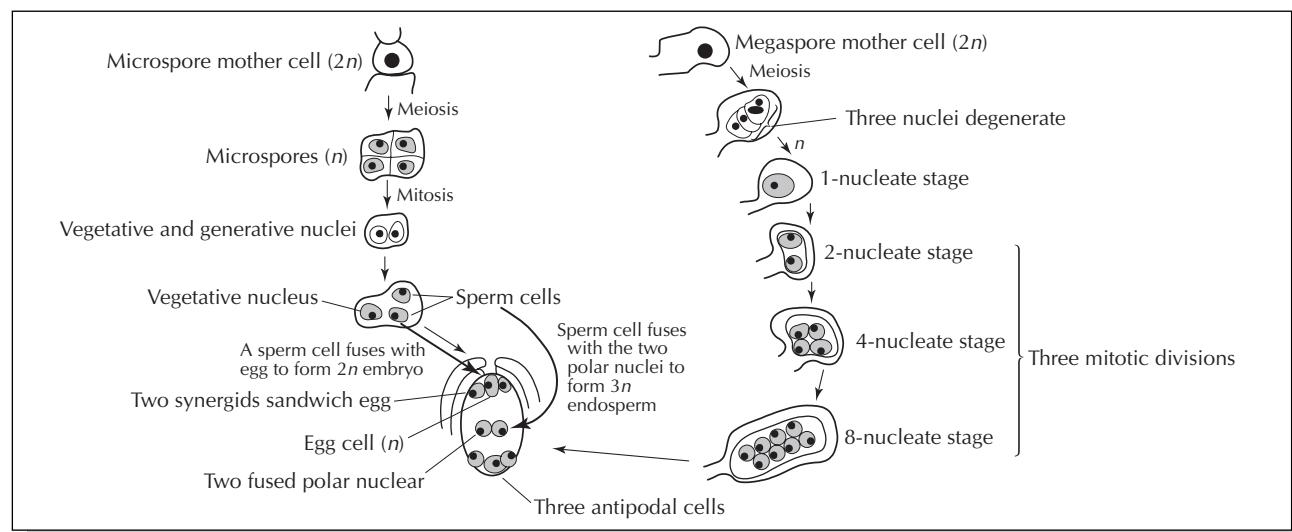
tobacco). However if a flower lacks certain parts (often petals or sepals), as is the case in many grasses (e.g., rice, corn, wheat), it is said to be an **incomplete flower**. Some flowers either have only stamens or a pistil, but not both. When both stamens and a pistil occur in the same flower, the flower is said to be a **perfect flower** (bisexual), as in wheat, tomato, and soybean. Some flowers are unisexual (either stamens or pistil may be absent) and are called **imperfect flowers**. If imperfect flowers have stamens they are called **stamine flowers**. When only a pistil occurs, the flower is a **pistilate flower**. A plant such as corn bears both stamine (tassel) and pistillate (silk) flowers on the same plant and is said to be a **monoecious plant**. However in species such as asparagus and papaya, plants may either be pistillate (female plant) or stamine (male plant) and are said to be **dioecious plants**. Flowers may either be **solitary** (occur singly or alone) or may be grouped together to form an **inflorescence**. An inflorescence has a primary stalk (peduncle) and numerous secondary smaller stalks (pedicels). The most common inflorescence types in crop plants are the cyme and raceme. A branched raceme is called a **panicle** (e.g., oats) while a raceme with sessile (short pedicels) is called a **spike** (e.g., wheat). From the foregoing, it is clear that a plant breeder should know the specific characteristics of the flower in order to select the appropriate techniques for crossing.

## Gametogenesis

Sexual reproduction entails the transfer of gametes to specific female structures where they unite and are then transformed into an embryo, a miniature plant. Gametes are formed by the process of gametogenesis. They are produced from specialized diploid cells called **microspore mother cells** in anthers and **megasporangium mother cells** in the ovary (Figure 4.4). Microspores derived from the mother cells are haploid cells each dividing by mitosis to produce an immature **male gametophyte** (pollen grain). Most pollen is shed in the two-cell stage, even though sometimes, as in grasses, one of the cells later divides again to produce two sperm cells. In the ovule, four megasporangia are similarly produced by meiosis. The nucleus of the functional megasporangium divides three times by mitosis to produce eight nuclei, one of which eventually becomes the egg. The **female gametophyte** is the seven-celled, eight-nucleate structure. This structure is also called the **embryo sac**. Two free nuclei remain in the sac. These are called polar nuclei because they originate from opposite ends of the embryo sac.

## Pollination and fertilization

**Pollination** is the transfer of pollen grains from the anther to the stigma of a flower. This transfer is achieved



**Figure 4.4** Gametogenesis in plants results in the production of pollen and egg cells. Pollen is transported by agents to the stigma of the female flower, from which it travels to the egg cell to unite with it.

**Table 4.1** Pollination mechanisms in plants.

Pollination vector	Flower characteristics
Wind	Tiny flowers (e.g., grasses); dioecious species
Insects	
Bees	Bright and showy (blue, yellow); sweet scent; unique patterns; corolla provides landing pad for bees
Moths	White or pale color for visibility at night time; strong penetrating odor emitted after sunset
Beetles	White or dull color; large flowers; solitary or inflorescence
Flies	Dull or brownish color
Butterflies	Bright colors (often orange, red); nectar located at base of long slender corolla tube
Bats	Large flower with strong fruity pedicels; dull or pale colors; strong fruity or musty scents; thick copious nectar
Birds	Bright colors (red, yellow); odorless; thick copious nectar

**Table 4.2** Examples of predominantly self-pollinated species.

Common name	Scientific name
Barley	<i>Hordeum vulgare</i>
Chickpea	<i>Cicer arietinum</i>
Clover	<i>Trifolium</i> spp.
Common bean	<i>Phaseolus vulgaris</i>
Cotton	<i>Gossypium</i> spp.
Cowpea	<i>Vigna unguiculata</i>
Eggplant	<i>Solanum melongena</i>
Flax	<i>Linum usitatissimum</i>
Jute	<i>Corchorus esculularis</i>
Lettuce	<i>Lupinus</i> spp.
Oat	<i>Avena sativa</i>
Pea	<i>Pisum sativum</i>
Peach	<i>Prunus persica</i>
Peanut	<i>Arachis hypogaea</i>
Rice	<i>Oryza sativa</i>
Sorghum	<i>Sorghum bicolor</i>
Soybean	<i>Glycine max</i>
Tobacco	<i>Nicotiana tabacum</i>
Tomato	<i>Lycopersicon esculentum</i>
Wheat	<i>Triticum aestivum</i>

## Self-pollination

### Mechanisms that promote self-pollination

Certain natural mechanisms promote or ensure self-pollination. **Cleistogamy** is the condition in which the flower fails to open. The term is sometimes extended to mean a condition in which the flower opens only after it has been pollinated (as occurs in wheat, barley, and lettuce), a condition called **chasmogamy**. Some floral structures such as those found in legumes, favor self-pollination. Sometimes, the stigma of the flower is closely surrounded by anthers, making it prone to selfing.

Very few species are completely self-pollinated. The level of self-pollination is affected by factors including the nature and amount of insect pollination, air current, and temperature. In certain species, pollen may become sterilized when the temperature dips below freezing. Any flower that opens prior to self-pollination is susceptible to some cross-pollination. A list of predominantly self-pollinated species is presented in Table 4.2.

### Genetic and breeding implications of self-pollination

Self-pollination is considered the highest degree of inbreeding a plant can achieve. It promotes homozygosity

through a vector or pollination agent. The common pollination vectors are wind, insects, mammals, and birds. Flowers have certain features that suit the various pollination mechanisms (Table 4.1): insect-pollinated flowers tend to be showy and exude strong fragrances, whereas birds are attracted to red and yellow flowers. When compatible pollen falls on a receptive stigma, a pollen tube grows down the style to the micropylar end of the embryo sac, carrying two sperms or male gametes. The tube penetrates the sac through the micropyle. One of the sperms unites with the egg cell, a process called **fertilization**. The other sperm cell unites with the two polar nuclei (called triple fusion). The simultaneous occurrence of two fusion events in the embryo sac is called **double fertilization**.

On the basis of pollination mechanisms, plants may be grouped into two **mating systems**: **self-pollinated** or **cross-pollinated**. Self-pollinated species accept pollen primarily from the anthers of the same flower (autogamy). The flowers, of necessity, must be bisexual. Cross-pollinated species accept pollen from different sources. In actuality, species express varying degrees of cross-pollination, ranging from lack of cross-pollination to complete cross-pollination.

of all gene loci and traits of the sporophyte. Consequently, should there be cross-pollination the resulting heterozygosity is rapidly eliminated. To be classified as self-pollinated, cross-pollination should not exceed 4%. The genotypes of gametes of a single plant are all the same. Further, the progeny of a single plant is homogeneous. A population of self-pollinated species, in effect, comprises a mixture of homozygous lines. Self-pollination restricts the creation of new gene combinations (no introgression of new genes through hybridization). New genes may arise through mutation, but such a change is restricted to individual lines or the progenies of the mutated plant. The proportions of different genotypes, not the presence of newly introduced types, define the variability in a self-pollinated species. Another genetic consequence of self-pollination is that mutations (which are usually recessive) are readily exposed through homozygosity, for the breeder or nature to apply the appropriate selection pressure on.

Repeated selfing has no genetic consequence in self-pollinated species (no **inbreeding depression** or loss of vigor following selfing). Similarly, self-incompatibility does not occur. Because a self-pollinated cultivar is generally one single genotype reproducing itself, breeding self-pollinated species usually entails identifying one superior genotype (or a few) and multiplying it. Specific breeding methods commonly used for self-pollinated species are pure-line selection, pedigree breeding, bulk populations, and backcross breeding (see Chapter 16).

### Cross-pollinating species

#### Mechanisms that favor cross-pollination

Several mechanisms occur in nature by which cross-pollination is ensured, the most effective being **dioecy**. As previously noted, dioecious species are those in which a plant is either female or male but not a hermaphrodite (e.g., hemp, date, palm). When such species are cultivated from grain or fruit, it is critical that the producer provides pollinator rows. A less stringent mechanism is **monoecy** (i.e., monoecious plants). Monoecious species can receive pollen from their own male flowers. **Dichogamy** occurs in hermaphroditic flowers, whereby cross-pollination may be enforced when the stamens mature before the pistil is mature and receptive (a condition called **protandry**) or the reverse (called **protogyny**). Sometimes, the pollen from a flower is not tolerated by its own stigma, a condition known as **self-incompatibility**. Male sterility, the condition whereby

**Table 4.3** Examples of predominantly cross-pollinated species.

Common name	Scientific name
Alfalfa	<i>Medicago sativa</i>
Annual ryegrass	<i>Lolium multiflorum</i>
Banana	<i>Musa</i> spp.
Birdsfoot trefoil	<i>Lotus corniculatus</i>
Cabbage	<i>Brassica oleracea</i>
Carrot	<i>Ducus carota</i>
Cassava	<i>Manihot esculentum</i>
Cucumber	<i>Cucumis sativus</i>
Fescue	<i>Festuca</i> spp.
Kentucky bluegrass	<i>Poa pratensis</i>
Maize	<i>Zea mays</i>
Muskmelon	<i>Cucumis melo</i>
Onion	<i>Allium</i> spp.
Pepper	<i>Capsicum</i> spp.
Potato	<i>Solanum tuberosum</i>
Radish	<i>Raphanus sativus</i>
Rye	<i>Secale cereale</i>
Sugar beet	<i>Beta vulgaris</i>
Sunflower	<i>Helianthus annuus</i>
Sweet potato	<i>Ipomea batatas</i>
Watermelon	<i>Citrullus lanatus</i>

the pollen of the male is sterile, compels the plant to receive pollen from different flowers. Similarly, a condition called **heterostyly** is one in which significant difference in the lengths of the stamen and pistil makes it less likely for self-pollination to occur, and thereby promotes cross-pollination. Cross-pollinated species depend on agents of pollination, especially wind and insects. A partial list of cross-pollinated species is presented in Table 4.3.

#### Genetic and breeding implications of cross-pollination

The genotype of the sporophytic generation is heterozygous while the genotypes of gametes of a single plant are all different. The genetic structure of a cross-pollinated species is characterized by heterozygosity. Self-incompatibility occurs in such species. Unlike self-pollinated species in which new gene combinations are prohibited, cross-pollinated species share a wide gene pool from which new combinations are created to form the next generation. Furthermore, when cross-pollinated species are selfed, they suffer inbreeding depression. Deleterious recessive alleles that were suppressed because of heterozygous advantage have opportunities to be

homozygous and therefore become expressed. However, such depression is reversed upon cross-pollination. **Hybrid vigor** (the increase in vigor of the hybrid over its parents resulting from crossing unlike parents) is exploited in hybrid seed production (see Chapter 18). In addition to hybrid breeding, population-based improvement methods (e.g., mass selection, recurrent selection, synthetic cultivars) are common methods of breeding cross-pollinated species.

### Asexual reproduction

Asexual reproduction may be categorized into two – **vegetative propagation** and **apomixis**. Asexual reproduction is also called **clonal propagation** because the products are genetically identical to the propagules.

#### Vegetative propagation

As previously indicated, certain species may be reproduced by using various vegetative parts including bulbs, corms, rhizomes, stems, and buds. Vegetative propagation is widely used in the horticultural industry. Pieces of vegetative materials called **cuttings** are obtained from parts of the plant (e.g., root, stem, leaf) for planting. Potato, cassava, sugarcane, rose, grape, and some perennial grasses are frequently propagated by stem cuttings. Methods such as grafting and budding are used for propagating tree crops, where two different plant parts are united by attaching one to the other and securing with a tape. Healing of the graft junction permanently unites the two parts into one plant.

A variety of sophisticated techniques are used to vegetatively propagate high value plants. Numerous plantlets may be generated from a small piece of vegetative material (e.g., a segment of a leaf) by the technique of **micro-propagation**. The **tissue culture** technique is used to rapidly multiply planting material under aseptic conditions (see Chapter 11 for more details). Perennial horticultural species tend to be clonally propagated, whereas annuals and biennials tend to be propagated by seed.

Clonally propagated crops may be divided into two broad categories on the basis of economic use:

- Those cultivated for a vegetative product.** Important species vegetatively cultivated for a vegetative product include sweet potato, yam, cassava, sugarcane, and Irish potato. These species tend to exhibit certain reproductive abnormalities. For example, flowering is reduced, and so is fertility. Some species such as

potatoes have cytoplasmic male sterility. Sometimes, flowering is retarded (e.g., by chemicals) in production (e.g., in sugarcane).

- Those cultivated for a fruit or reproductive product.** Plants in this category include fruit trees, shrubs, and cane fruits. Examples include apple, pear, grape, strawberry, and banana.

#### Breeding implications of vegetative propagation

There are certain characteristics of clonal propagation that have breeding implications.

- Clonal species with viable seed and high pollen fertility can be improved by hybridization.
- Unlike hybridization of sexual species, which often requires additional steps to fix the genetic variability in a genotype for release as a cultivar (except for hybrid cultivars), clonal cultivars can be released immediately following a cross, provided a desirable genotype combination has been achieved. Clonal breeding is hence quick.
- When improving species whose economic parts are vegetative products, it is not important for the hybrid to be fertile.
- Because of the capacity to multiply from vegetative material (through methods such as cuttings or micropropagation), the breeder only needs to obtain a single desirable plant to be used as stock.
- Heterosis (hybrid vigor), if it occurs, is fixed in the hybrid product. That is, unlike hybrid cultivars in seed-producing species, there is no need to reconstitute the hybrid. Once bred, heterozygosity is maintained indefinitely.
- It is more difficult to obtain large quantities of planting material from clones in the short term.
- Plant species that are vegetatively parthenocarpic (e.g., banana) cannot be improved by hybrid methodology.
- Species such as mango and citrus produce polyembryonic seedlings. This reproductive irregularity complicates breeding because clones of the parent are mixed with hybrid progeny.
- Clonal crops are perennial outcrossers and intolerant of inbreeding. They are highly heterozygous.
- Unlike sexual crop breeding in which the genotype of the cultivar is determined at the end of the breeding process (because it changes with inbreeding), the genotype of a family is fixed and determined at the outset.
- Both general combining ability (GCA) and specific combining ability (SCA) (that is, performance in crosses) can be fully exploited with appropriate breeding method and population sizes.

## Apomixis

Seed production in higher plants that are sexually propagated species normally occurs after a sexual union in which male and female gametes fuse to form a zygote, which then develops into an embryo. However, some species have the natural ability to develop seed without fertilization, a phenomenon called **apomixis**. The consequence of this event is that apomictically produced seeds are clones of the mother plant. That is, apomixis is the asexual production of seed. Unlike sexual reproduction, there is no opportunity in apomixis for new recombination to occur to produce diversity in the offspring.

### Occurrence in nature

Apomixis is widespread in nature, having been found in unrelated plant families. However, it is infrequent in occurrence. About 10% of the estimated 400 plant families and a mere 1% of the estimated 40,000 species they comprise exhibit apomixis. The plant families with the highest frequency of apomixis are Gramineae (Poaceae), Compositae, Rosaceae, and Asteraceae. Many species of citrus, berries, mango, perennial forage grasses, and guayule reproduce apomictically.

Some species can produce both sexual and apomictic seeds and are called **facultative apomicts** (e.g., bluegrass, *Poa pratensis*). Species such as bahiagrass (*Paspalum notatum*) reproduce exclusively or nearly so by apomixis and are called **obligate apomicts**. There are several indicators of apomixis. When the progeny from a cross in a cross-pollinated (heterozygous) species fails to segregate, appearing uniform and identical to the mother plant, this could indicate obligate apomixis. Similarly, when plants expected to exhibit high sterility (e.g., aneuploids, triploids) instead show significantly high fertility, apomixis could be the cause. Obligate apomicts may display multiple floral features (e.g., multiple stigmas and ovules per floret, double or fused ovaries), or multiple seedlings per seed. Facultative apomixis may be suspected if the progeny of a cross shows an unusually high number of identical homozygous individuals that resemble the mother plant in addition to the presence of individuals that are clearly different (hybrid products). Using such morphological indicators to discover apomicts requires keen observation and familiarity with the normal breeding behavior of the species.

The indicators suggested are by no means conclusive evidence of apomixis. To confirm the occurrence of apomixis and discovery of its mechanisms requires

additional progeny tests as well as cytological tests of megasporogenesis and embryo sac development.

### Benefits of apomixis

The benefits of apomixis may be examined from the perspectives of the plant breeder and the crop producer.

**Benefits to the plant breeder** Apomixis is a natural process of cloning plants through seed. As a breeding tool, it allows plant breeders to develop hybrids that can retain their original genetic properties indefinitely with repeated use, without a need to reconstitute them. In other words, hybrid seed can be produced from hybrid seed. The plant breeder does not need to make crosses each year to produce the hybrid. This advantage accelerates breeding programs and reduces development costs of hybrid cultivars. Apomixis is greatly beneficial when uniformity of product is desired. Breeders can use this tool to quickly fix superior gene combinations. That is, vigor can be duplicated, generation after generation without decline. Furthermore, commercial hybrid production can be implemented for species without fertility control mechanisms (e.g., male-sterility system), neither is there a need for isolation in  $F_1$  hybrid seed production. There is no need to maintain and increase parental genotypes. Cultivar evaluation can proceed immediately following a cross.

Apart from these obvious benefits, it is anticipated that plant breeders will divert the resources saved (time, money) into other creative breeding ventures. For example, cultivars could be developed for smaller and more specific production environments. Also, more parental stock could be developed to reduce the risk of genetic vulnerability through the use of a few elite genetic stocks as parents in hybrid development.

There are some plant breeding concerns associated with apomixis. Species that exhibit facultative apomixis are more challenging to breed because they produce both sexual and apomictic plants in the progeny. Obligate apomicts are easier to breed by conventional methods provided compatible (sexually reproducing) counterparts can be found.

**Benefits to the producer** The most obvious benefit of apomictic cultivars to crop producers is the ability to save seed from their field harvest of hybrid cultivars for planting the next season. Because apomixis fixes hybrid vigor, the farmer does not need to purchase fresh hybrid seed each season. This especially benefits the producer in poor economies, who often cannot afford the high

price of hybrid seed. Apomixis, as previously indicated, accelerates plant breeding. This could translate into less expensive commercial seed for all producers. Realistically, such benefits will materialize only if commercial breeders can make an acceptable profit from using the technology.

**Impact on the environment** Some speculate that apomixis has the potential to reduce biodiversity because it produces clonal cultivars and hence uniform populations that are susceptible to disease epidemics. However, others caution that the suspected reduction in biodiversity would be minimal since apomixis occurs naturally in polyploids, which occur less frequently than diploids.

#### Mechanisms of apomixis

Apomixis arises by a number of mechanisms of which four major ones that differ according to origin (the cell that undergoes mitosis to produce the embryo) are discussed next. Seed formation without sexual union is called **agamospermy**, a mechanism that can be summarized into two categories: **gametophytic apomixis** and **adventitious apomixis**. There are two types of gametophytic apomixes: apospory and diplosropy.

**1 Apospory.** This is the most common mechanism of apomixis in higher plants. It is a type of agamospermy

that involves the nucellar. The somatic cells of the ovule divide mitotically to form unreduced ( $2n$ ) embryonic sacs. The megasporangium or young embryo sac aborts, as occurs in species such as Kentucky bluegrass.

**2 Diplosropy.** An unreduced megasporangium mother cell produces embryo sacs following mitosis instead of meiosis. This cytological event occurs in species such as *Tripsacum*.

**3 Adventitious embryo.** Unlike apospory and diplosropy in which an embryo sac is formed, no embryo sac is formed in adventitious embryony. Instead, the source of the embryo could be somatic cells of the ovule, integuments, or ovary wall. This mechanism occurs commonly in citrus but rarely in other higher plants.

**4 Parthenogenesis.** This mechanism is essentially equivalent to haploidy. The reduced ( $n$ ) egg nucleus in a sexual embryo sac develops into a haploid embryo without fertilization by the sperm nucleus.

Other less common mechanisms of apomixis are **androgenesis** (development of a seed embryo from the sperm nucleus upon entering the embryo sac) and **semigamy** (sperm nucleus and egg nucleus develop independently without uniting, leading to a haploid embryo). The resulting haploid plants contain sectors of material from both maternal and paternal origin.



## Industry highlights

### *Maize × Tripsacum hybridization and the transfer of apomixis: historical review*

Bryan Kindiger

USDA-ARS Grazinglands Research Laboratory, El Reno, OK 73036, USA

Research in maize–*Tripsacum* hybridization is extensive and encompasses a period of more than 60 years of collective research. The publication *The origin of Indian corn and its relatives* describes some of the initial research in this area (Mangelsdorf & Reeves 1939) and is recommended reading for anyone interested in this area of research. Since this historic publication, an abundance of literature has been developed with regard to the various facets of this type of hybridization ranging from agronomy, plant disease, cytogenetics, and breeding, to genetic analysis. As a consequence, no single article can cover all the research relevant to this topic. This report will only briefly highlight a specific series of experiments that address an attempt to investigate the transfer of apomixis from *Tripsacum dactyloides* to *Zea mays*.

One of the most interesting instances of intergeneric hybridization involves hybridizing maize (*Z. mays* L.) ( $2n = 2x = 20$ ) with its distant relative eastern gamagrass (*T. dactyloides*) ( $2n = 4x = 72$ ). Regardless of their complete difference in chromosome number, plant phenotype, and environmental niche, hybrids are relatively easy to generate. The  $F_1$  hybrids are completely pollen-sterile and microsporogenesis is associated with a varying array of meiotic anomalies (Kindiger 1993). The hybrids vary in seed fertility ranging from completely sterile to highly seed-fertile (Harlan & de Wet 1977). To date, all seed-fertile hybrids generated from tetraploid *T. dactyloides* exhibit some level of apomictic expression; however, backcrossing with maize commonly results in the loss of apomixis.

### Potential pathways for apomixis gene introgression

Research strongly suggests that there is little homeology between the genomes of *Tripsacum* and maize. Maguire (1962) and Galinat (1973), each utilizing a set of recessive phenotypic maize markers, suggested that only maize chromosomes 2, 5, 8, and 9 have potentials for pairing and recombination and for gene introgression with *Tripsacum*. Additional research has confirmed the conservation of loci specific to pistil development between maize and *Tripsacum* genomes (Kindiger et al. 1995; Li et al. 1997). Genomic *in situ* hybridization (GISB) studies have also strongly suggested that only three regions of maize chromosomes have homeology with the *Tripsacum* genome: the subterminal regions of Mz2S, Mz6L, and Mz8L (Poggio et al. 1999). Though there is little chromosome homeology, there is some hope for apomixis transfer from *Tripsacum* to maize. Two approaches that have been successful in transferring components of apomixis from *Tripsacum* to maize can be detailed in two particular backcross pathways (Harlan & de Wet 1977).

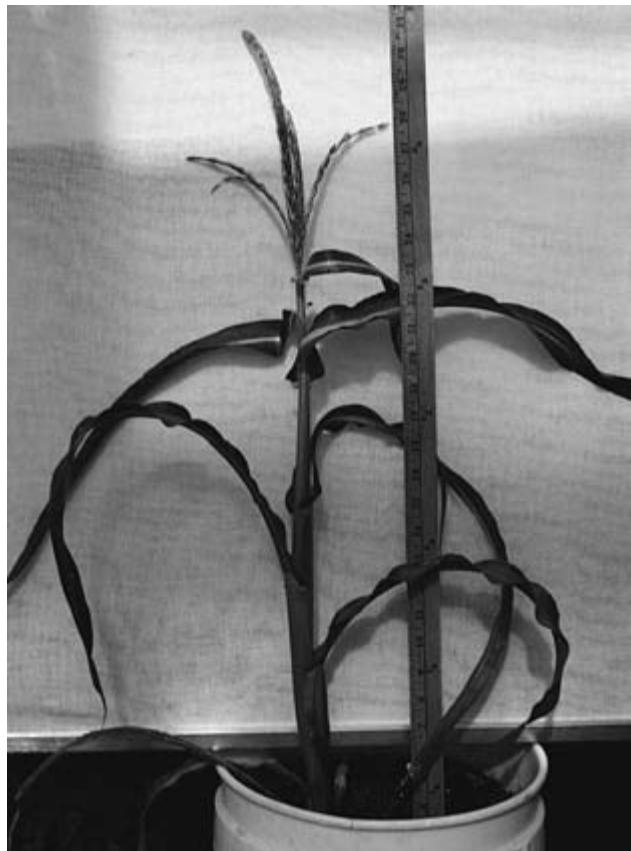
The first approach is called the **28 → 38 apomictic transfer pathway**. This successful approach for apomixis transfer has been described only once and has had little re-examination. In 1958, Dr M. Borovsky (from the Institute of Agriculture, Kishinev, Moldova) performed a series of hybridizations between a diploid popcorn and a sexual diploid ( $2n = 2x = 36$ ) *T. dactyloides* clone with the first maize–*Tripsacum* hybrids being generated in 1960 (Borovsky 1966; Borovsky & Kovarsky 1967). The  $F_1$  hybrids generated from the experiments possessed 28 chromosomes (10Mz + 18Tr). The  $F_1$  plants were completely male-sterile and were highly seed-sterile. Backcrossing with diploid maize identified that some of the  $F_1$  hybrids were approximately 1–1.5% seed-fertile and resulted in the production of progeny possessing 28 chromosomes (10Mz + 18Tr) and 38 chromosomes (20Mz + 18Tr). When the  $F_1$  was backcrossed to the *Tripsacum* parent, the fertile  $F_1$ s generated progeny with 28 chromosomes (10Mz + 18Tr) and 46 chromosomes (10Mz + 18Tr + 18Tr). The complete set of backcrosses with maize and *Tripsacum* resulted in a ratio of approximately 10 (28-chromosome plants) to 1 (38- or 46-chromosome plants). Phenotypic observations suggested that the 28-chromosome progeny were not different from their 28-chromosome parent while the 38- and 46-chromosome progeny were clearly different. Additional evaluations on the 28-chromosome  $F_1$  and its 28-chromosome progeny suggested that these  $F_1$  plants and their progeny were apomictic. This early experiment remains the single incidence where a 28-chromosome  $F_1$  hybrid was maintained by apomixis.

A second pathway whereby apomixis has been introgressed from *Tripsacum* to maize is the **46 → 56 → 38 apomictic transfer pathway**. Though not specifically addressed in the definitive work on maize–*Tripsacum* introgression (Harlan & de Wet 1977), this successful attempt at apomixis transfer requires a brief reiteration. Initially published by Petrov and colleagues as early as 1979, and replicated in similar style by others, a diploid or tetraploid maize line is pollinated by a tetraploid, apomictic *T. dactyloides* clone (Petrov et al. 1979, 1984). If a diploid maize line is utilized, the resultant  $F_1$  46-chromosome hybrid possesses 10Mz and 36Tr chromosomes. Upon backcrossing with diploid maize, both apomictic 46-chromosome and 56-chromosome (20Mz + 36Tr) individuals can be obtained. The 46-chromosome offspring are products of apomixis. The 56-chromosome offspring are products of an unreduced egg being fertilized by the diploid maize pollen source, another  $2n + n$  mating event. Backcrossing the 46-chromosome individuals by maize, repeats the above cycle. Upon backcrossing the 56-chromosome individuals with maize, three types of progeny can be observed. Typically, progeny having 56 chromosomes are generated. However, in some instances,  $2n + n$  matings occur giving rise to individuals possessing 66 chromosomes (30Mz + 36Tr). Occasionally, a reduced egg will be generated and may or may not be fertilized by the available maize pollen. In rare instances of non-fertilization, a 28-chromosome individual is generated (10Mz + 18Tr). In instances whereby the maize pollen fertilizes the reduced egg, 38-chromosome individuals are obtained (20Mz + 18Tr). Generally, individuals possessing 38 chromosomes, rather than 28 chromosomes, are the most common product. What is unique about this pathway is that, occasionally, the 38-chromosome individuals retain all the elements of apomixis that were present in the *Tripsacum* paternal parent and the  $F_1$  and BC<sub>1</sub> individuals (Figure 1). The retention of apomixis to this 38-chromosome level has been well documented and repeated in several laboratories (Petrov et al. 1979, 1984; Leblanc et al. 1996; Kindiger & Sokolov 1997).

Generally, through  $2n + n$  mating events, the 38-chromosome individuals produce only apomictic 38-chromosome progeny and 48-chromosome progeny. Backcrossing the 48-chromosome individuals results in 48-chromosome apomictics and 58-chromosome apomictics. Each of these steps gives rise to a different plant and ear phenotype (Figure 2). This  $2n + n$  accumulation of maize genomes continues until a point is achieved where the additional maize genomes eventually shift the individual from an apomictic to a sexual mode of reproduction. It is extremely difficult to generate and maintain apomixis in hybrids possessing fewer than 18Tr chromosomes. Likely this is due to the expression of apomixis in the 38-chromosome hybrids. However, in one instance an apomictic individual possessing 9Tr chromosomes has been identified (Kindiger et al. 1996b) suggesting that with time and patience, additional *Tripsacum* chromosomes can be removed from these hybrids and still retain the apomixis.

### Recent attempts to transfer apomixis from *Tripsacum* to maize

As of this report, prevailing wisdom suggests that apomixis (at least for *Tripsacum*) is controlled by no more than one or two genes, likely linked on a particular *Tripsacum* chromosome (Leblanc et al. 1995; Grimanelli et al. 1998). Cytogenetic and GISB studies suggest this region may be Tr16L in the vicinity of the nucleolus-organizing region that has homeology with the distal region of Mz6L (Kindiger et al. 1996a; Poggio et al. 1999). Evaluations of materials from the Petrov program and generated through the 46 → 56 → 38 pathway have identified an apomictic line that does not possess an intact Tr16 chromosome or the Mz6L-Tr16L



(a)



(b)

**Figure 1** (a) A highly maize-like 38-chromosome apomictic maize–*Tripsacum* hybrid. This selection has no or few tillers and exhibits a distinct maize phenotype. (b) A top and second ear taken from one of these highly maize-like apomictic individuals. Note the eight rows on the ear are rarely found in other apomictic maize–*Tripsacum* hybrids.



**Figure 2** A series of maize–*Tripsacum* ear types. Left to right: dent corn, apomictic 39-chromosome hybrid, apomictic 38-chromosome hybrid ("Yudin"), apomictic 56-chromosome hybrid, three apomictic 46-chromosome hybrids, and tetraploid *Tripsacum dactyloides*.

translocation (Figure 3). RAPD (random amplified polymorphic DNA) markers previously known to be associated to apomixis continue to be present in this germplasm. Cytological analysis of this particular chromosomal element suggests the chromosome carries the nucleolus-organizing region and the Tr16L satellite. This small isochromosome may indeed possess the loci conferring apomixis in this material.

Regardless of the favorable light academics and researchers alike shed upon the prospects in this area of study, this research endeavor continues to be difficult, time-consuming, and expensive. Though an apomictic maize prototype has been developed (US patent no. 5,710,367) gene transfer through traditional breeding approaches is questionable. The development of apomictic maize through its hybridization with *Tripsacum* offers many opportunities; however, many years of additional research will be required for this to be realized.

### References

- Borovsky, M. 1966. Apomixis in intergeneric maize–*Tripsacum* hybrids. In: Meeting on problems on apomixis in plants. Saratov State University, Saratov, Russia, pp. 8–9.
- Borovsky, M., and A.E. Kovarsky. 1967. Intergenus maize–*Tripsacum* hybridizations. Izvestia Akad. Nauk Moldovaski SSR 11:25–35.
- Galatin, W.C. 1973. Intergenomic mapping of maize, teosinte and *Tripsacum*. Evolution 27:644–655.
- Grimanelli, D., O. Leblanc, E. Espinosa, E. Perotti, D. Gonzalez de Leon, and Y. Savidan. 1998. Mapping diplosporous apomixis in tetraploid *Tripsacum*: One gene or several genes? Heredity 80:33–39.



**Figure 3** (a) The satellite region of Tr16L (arrow), which confers apomixis in the V31 apomictic line. No normal or intact Tr16 is present in this line. (b) An enlargement of the isochromosome-appearing entity with the nucleolus-organizing region (NOR) and satellite regions identified.

- Harlan, J.R., and J.M.J. de Wet. 1977. Pathways of genetic transfer from *Tripsacum* to *Zea mays*. Proc. Natl. Acad. Sci. USA 74:3494–3497.
- Kindiger, B. 1993. Aberrant microspore development in hybrids of maize × *Tripsacum dactyloides*. Genome 36:987–997.
- Kindiger, B., D. Bai, and V. Sokolov. 1996a. Assignment of gene(s) conferring apomixis in *Tripsacum* to a chromosome arm: Cytological and molecular evidence. Genome 39:1133–1141.
- Kindiger, B., C.A. Blakey, and Dewald, C.L. 1995. Sex reversal in maize × *Tripsacum* hybrids: Allelic non-complementation of *ts2* and *gsf1*. Maydica 40:187–190.
- Kindiger, B., and V. Sokolov. 1997. Progress in the development of apomictic maize. Trends Agron. 1:75–94.
- Kindiger, B., V. Sokolov, and C.L. Dewald. 1996b. A comparison of apomictic reproduction in eastern gamagrass (*Tripsacum dactyloides* (L.) and maize–*Tripsacum* hybrids. Genetica 97:103–110.
- Leblanc, O., D. Griminelli, D. Gonzalez-de-Leon, and Y. Savidan. 1995. Detection of the apomictic mode of reproduction in maize–*Tripsacum* hybrids using maize RFLP markers. Theor. Appl. Genetics 90:1198–1203.
- Leblanc, O., D. Griminelli, N. Islan-Faridi, J. Berthaud, and Y. Savidan. 1996. Reproductive behavior in maize–*Tripsacum* polyhaploid plants: Implications for the transfer of apomixis into maize. J. Hered. 87:108–111.
- Li, D., C.A. Blakey, C.L. Dewald, and S.L. Dellaporta. 1997. Evidence for a common sex determination mechanism for pistil abortion in maize and its wild relative *Tripsacum*. Proc. Natl. Acad. Sci. USA 94:4217–4222.
- Maguire, M. 1962. Common loci in corn and *Tripsacum*. J. Hered. 53:87–88.
- Mangelsdorf, P.C., and R.G. Reeves. 1939. The origin of Indian corn and its relatives. Texas Agric. Exp. Stn. Bull. No. 574.
- Petrov, D.F., N.I. Belousova, and E.S. Fokina. 1979. Inheritance of apomixis and its elements in maize × *Tripsacum dactyloides* hybrids. Genetika 15:1827–1836.
- Petrov, D.F., N.I. Belousova, E.S. Fokina, L.I. Laikova, R.M. Yatsenko, and T.P. Sorokina. 1984. Transfer of some elements of apomixis from *Tripsacum* to maize. In: Apomixis and its role in evolution and breeding (Petrov, D.F., ed.), pp. 9–73. Oxonian Press Ltd, New Delhi, India.
- Poggio, L., V. Confalonieri, C. Comas, A. Cuadrado, N. Jouve, and C.A. Naranjo. 1999. Genomic *in situ* hybridization (GISH) of *Tripsacum dactyloides* and *Zea mays* ssp. *mays* with B chromosomes. Genome 42:687–691.

### Constraints of sexual biology in plant breeding

Some constraints of sexual biology are exploited as tools for breeding plants. They were previously mentioned and will be discussed in detail, including how they are

exploited in plant breeding. These include dioecy, monoecy, self-incompatibility, and male sterility.

As previously indicated, crossing is a major procedure employed in the transfer of genes from one parent to another in the breeding of sexual species. A critical aspect of crossing is pollination control to ensure that

only the desired pollen is involved in the cross. In hybrid seed production, success depends on the presence of an efficient, reliable, practical, and economic pollination control system for large-scale pollination. Pollination control may be accomplished in three general ways:

- 1 Mechanical control.** This approach entails manually removing anthers from bisexual flowers to prevent pollination, a technique called emasculation, or removing one sexual part (e.g., detasseling in corn), or excluding unwanted pollen by covering the female part. These methods are time-consuming, expensive, and tedious, limiting the number of plants that can be crossed. It should be mentioned that in crops such as corn, mechanical detasseling is widely used in the industry to produce hybrid seed.
- 2 Chemical control.** A variety of chemicals called chemical hybridizing agents or other names (e.g., male gametocides, male sterilants, pollenocides, androcides) are used to temporally induce male sterility in some species. Examples of such chemicals include Dalapon®, Estrone®, Ethephon®, Hybrex®, and Generis®. The application of these agents induces male sterility in plants, thereby enforcing cross-pollination. The effectiveness is variable among products.
- 3 Genetic control.** Certain genes are known to impose constraints on sexual biology by incapacitating the sexual organ (as in male sterility) or inhibiting the union of normal gametes (as in self-incompatibility). These genetic mechanisms are discussed further next.

### Dioecy and monoecy

As previously discussed, some flowers are complete while others are incomplete. Furthermore, in some species, the sexes are separate. When separate male and female flowers occur on the same plant, the condition is called **monoecy**; when the sexes occur on different plants (i.e., there are female plants and male plants), the condition is called **dioecy**. Examples of dioecious species include date, hops, asparagus, spinach, and hemp. The separation of the sexes means that all seed from dioecious species are hybrid in composition. Where the economic product is the seed or fruit, it is imperative to have female and male plants in the field in an appropriate ratio. In orchards, 3–4 males per 100 females may be adequate. In hops, the commercial product is the female inflorescence. Unfertilized flowers have the highest quality. Consequently, it is not desirable to grow pollinators in the same field when growing hops.

Dioecious crops propagated by seed may be improved by mass selection or controlled hybridization. As previously indicated, the male and female flowers occur on the same plant in monoecious species and even sometimes in different kinds of inflorescence (different locations, as in corn). It is easier and more convenient to self plants when the sexes occur in the same inflorescence. In terms of seed production, dioecy and monoecy are inefficient because not all flowers produce seed. Some flowers produce only pollen.

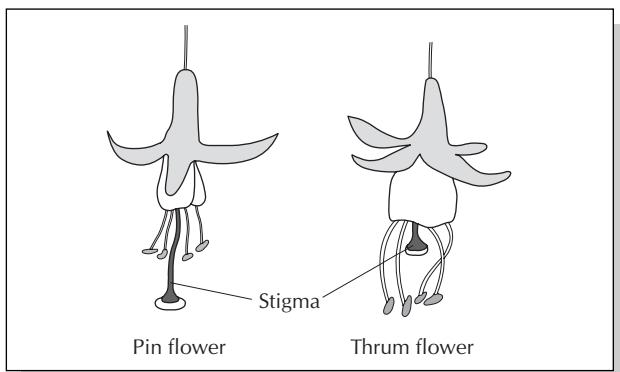
### Self-incompatibility

**Self-incompatibility** (or lack of self-fruitfulness) is a condition in which the pollen from a flower is not receptive on the stigma of the same flower and hence is incapable of setting seed. This happens in spite of the fact that both pollen and ovule development are normal and viable. It is caused by a genetically controlled physiological hindrance to self-fertilization. Self-incompatibility is widespread in nature, occurring in families such as Poaceae, Cruciferae, Compositae, and Rosaceae. The incompatibility reaction is genetically conditioned by a locus designated *S*, with multiple alleles that can number over 100 in some species such as *Trifolium pretense*. However, unlike monoecy and dioecy, all plants produce seed in self-incompatible species.

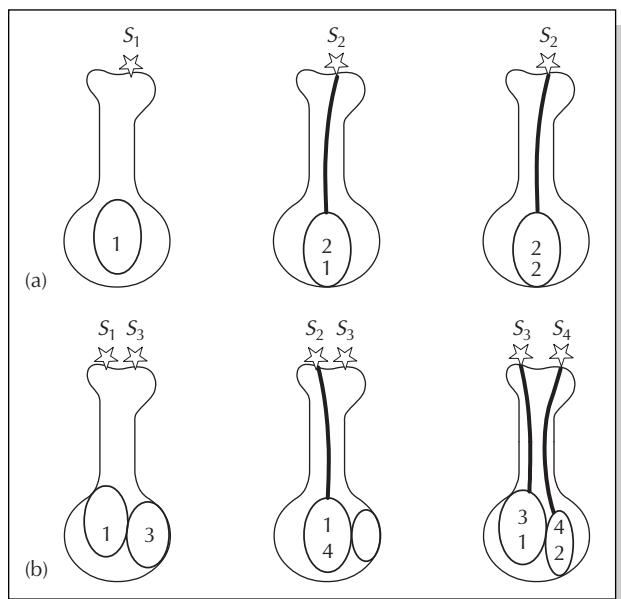
#### *Self-incompatibility systems*

Self-incompatibility systems may be classified into two basic types: **heteromorphic** and **homomorphic**.

- 1 Heteromorphic incompatibility.** This is caused by differences in the lengths of stamens and style (called **heterostyly**) (Figure 4.5). In one flower type called the **pin**, the styles are long while the anthers are short. In the other flower type, **thrum**, the reverse is true (e.g., in *Primula*). The pin trait is conditioned by the genotype *ss* while thrum is conditioned by the genotype *Ss*. A cross of pin (*ss*) × pin (*ss*) as well as thrum (*Ss*) × thrum (*Ss*) are incompatible. However, pin (*ss*) × thrum (*Ss*) or vice versa, is compatible. The condition described is **distyly** because of the two different types of style length of the flowers. In *Lythrum* three different relative positions occur (called **tristyly**).
- 2 Homomorphic incompatibility.** There are two kinds of homomorphic incompatibility – **gametophytic** and **sporophytic** (Figure 4.6).



**Figure 4.5** Heteromorphous incompatibility showing floral modifications in which anthers and pistils are of different lengths in different plants (heterostyly). This type of incompatibility is believed to be always of the sporophytic type. Pin and thrum flowers occurs in flowers such as *Primula*, *Forsythia*, *Oxalis*, and *Silia*.



**Figure 4.6** Types of self-incompatibility: (a) sporophytic and (b) gametophytic. Sporophytic incompatibility occurs in families such as Compositae and Cruciferae. It is associated with pollen grains with two generative nuclei, whereas gametophytic incompatibility is associated with pollen with one generative nucleus in the pollen tube as occurs in various kinds of clover.

- (a) **Gametophytic incompatibility.** In gametophytic incompatibility (originally called the oppositional factor system), the ability of the pollen to function is determined by its own genotype and not the plant that produces it. Gametophytic incompatibility is more widespread than sporophytic incompatibility. Gametophytic incompatibility occurs in species such as red clover, white clover, and yellow sweet clover. Homomorphic incompatibility is controlled by a series of alleles at a single locus ( $S_1, S_2, \dots, S_n$ ) or alleles at two loci in some species. The system is called homomorphic because the flowering structures in both the seed-bearing (female) and pollen-bearing (male) plants are similar. The alleles of the incompatibility gene(s) act individually in the style. They exhibit no dominance. The incompatible pollen is inhibited in the style. The pistil is diploid and hence contains two incompatibility alleles (e.g.,  $S_1S_3, S_3S_4$ ). Reactions occur if identical alleles in both pollen and style are encountered. Only heterozygotes for  $S$  alleles are produced in this system.
- (b) **Sporophytic incompatibility.** In sporophytic incompatibility, the incompatibility characteristics of the pollen are determined by the plant (sporophyte) that produces it. It occurs in species such as broccoli, radish, and kale. The sporophytic system differs from the gametophytic system in that the  $S$  allele exhibits dominance. Also, it may have individual action in both pollen and the style, making this incompatibility system

complex. The dominance is determined by the pollen parent. Incompatible pollen may be inhibited on the stigma surface. For example, a plant with genotype  $S_1S_2$  where  $S_1$  is dominant to  $S_2$ , will produce pollen that will function like  $S_1$ . Furthermore,  $S_1$  pollen will be rejected by an  $S_1$  style but received by an  $S_2$  style. Hence, homozygotes of  $S$  alleles are possible.

Incompatibility is expressed in one of three general ways, depending on the species. The germination of the pollen may be decreased (e.g., in broccoli). Sometimes, removing the stigma allows normal pollen germination. In the second way, pollen germination is normal, but pollen tube growth is inhibited in the style (e.g., tobacco). In the third scenario, the incompatibility reaction occurs after fertilization (e.g., in *Gesneria*). This third mechanism is rare.

#### Changing the incompatibility reaction

Mutagens (agents of mutation) such as X-rays, radioactive sources such as  $P^{32}$ , and certain chemicals have been

used to make a self-inferile genotype self-fertile. Such a change is easier to achieve in gametophytic systems than sporophytic systems. Furthermore, doubling the chromosome number of species with the sporophytic system of incompatibility does not significantly alter the incompatibility reaction. This is because two different alleles already exist in a diploid that may interact to produce the incompatibility effect. Polyploidy only makes more of such alleles available. On the other hand, doubling the chromosome in a gametophytic system would allow the pollen grain to carry two different alleles (instead of one). The allelic interaction could cancel any incompatibility effect to allow selfing to be possible. For example, diploid pear is self-incompatible whereas autotetraploid pear is self-fruitful.

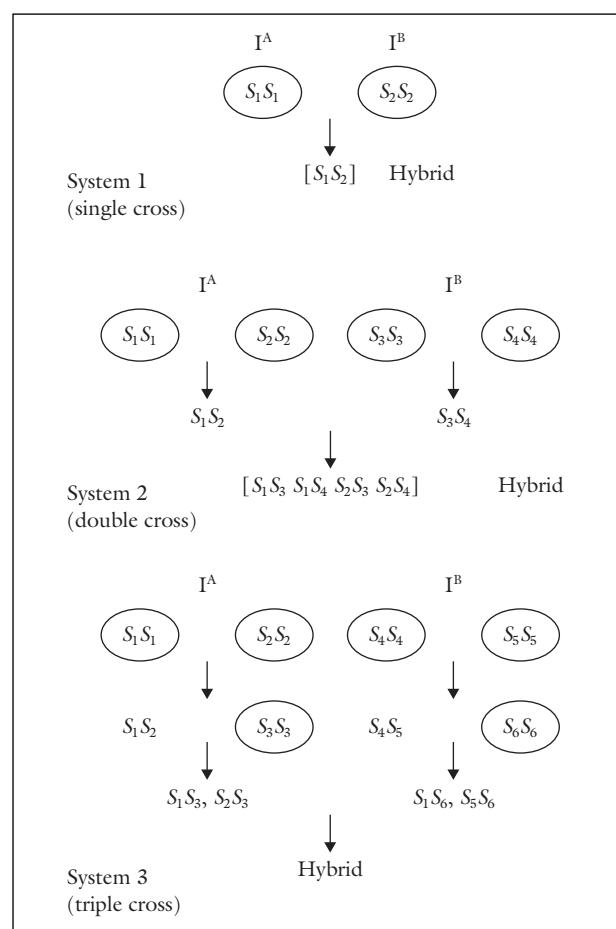
#### *Plant breeding implications of self-incompatibility*

Infertility of any kind hinders plant breeding. However, this handicap may be used as a tool to facilitate breeding by certain methods. Self-incompatibility may be temporarily overcome by techniques or strategies such as the removal of the stigma surface (or application of electric shock), early pollination (before inhibitory proteins form), or lowering the temperature (to slow down the development of the inhibitory substance). Self-incompatibility promotes heterozygosity. Consequently, selfing self-incompatible plants can create significant variability from which a breeder can select superior recombinants. Self-incompatibility may be used in plant breeding (for  $F_1$  hybrids, synthetics, triploids), but first homozygous lines must be developed.

Self-incompatibility systems for hybrid seed production have been established for certain crops (e.g., cabbage, kale) that exhibit sporophytic incompatibility (Figure 4.7). Inbred lines (compatible inbreds) are used as parents. These systems are generally used to manage pollinations for commercial production of hybrid seed. Gametophytic incompatibility occurs in vegetatively propagated species. The clones to be hybridized are planted in adjacent rows.

#### **Male sterility**

**Male sterility** is a condition in plants whereby the anthers or pollen are non-functional. The condition may manifest most commonly as absence, or extreme scarcity, of pollen, severe malformation or absence of flowers or stamens, or failure of pollen to dehisce. Just like self-incompatibility, male sterility enforces cross-pollination. Similarly, it can be exploited as a tool to



**Figure 4.7** Application of self-incompatibility in practical plant breeding. Sporophytic incompatibility is widely used in breeding of cabbage and other *Brassica* species. The single-cross hybrids are more uniform and easier to produce. The topcross is commonly used. A single self-incompatible parent is used as female, and is open-pollinated by a desirable cultivar as the pollen source.

eliminate the need for emasculation for producing hybrid seed. There are three basic kinds of male sterility based on the origin of the abnormality:

- 1 True male sterility.** This is due to unisexual flowers that lack male sex organs (dioecy and monoecy) or to bisexual flowers with abnormal or non-functional microspores (leading to pollen abortion).
- 2 Functional male sterility.** The anthers fail to release their contents even though the pollen is fertile.
- 3 Induced male sterility.** Plant breeders may use chemicals to induce sterility.

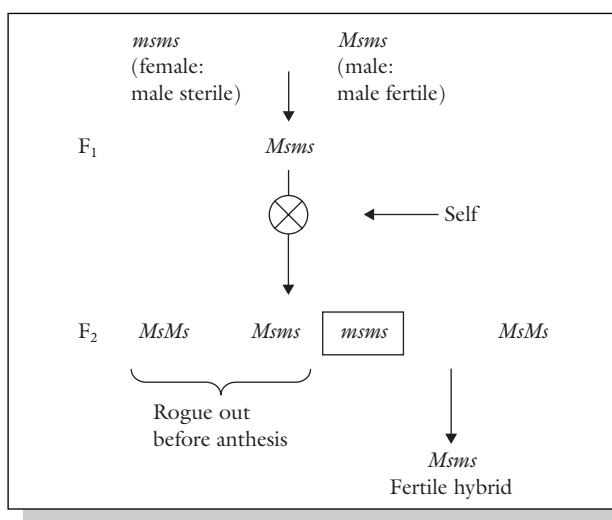
### True male sterility

There are three kinds of pollen sterility – nuclear, cytoplasmic, and cytoplasmic-genetic.

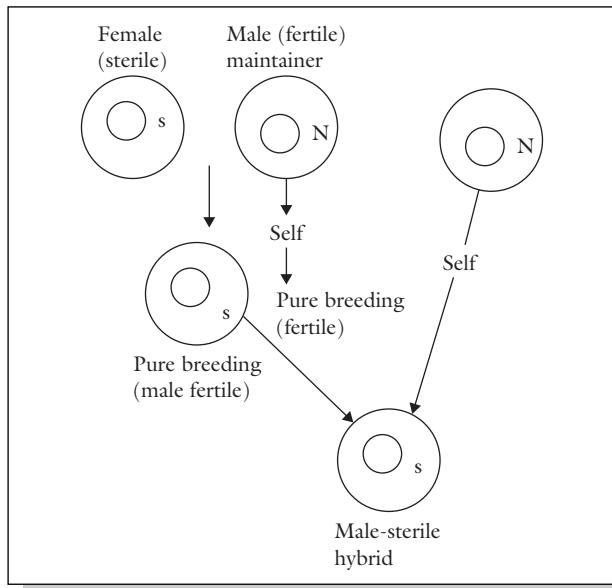
**Genetic male sterility** Genetic (nuclear, genic) male sterility is widespread in plants. The gene for sterility has been found in species including barley, cotton, soybean, tomato, potato, and lima bean. It is believed that nearly all diploid and polyploid plant species have at least one male-sterility locus. Genetic male sterility may be manifested as pollen abortion (pistillody) or abnormal anther development. Genetic male sterility is often conditioned by a single recessive nuclear gene, *ms*, the dominant allele, *Ms*, conditioning normal anther and pollen development. However, male sterility in alfalfa has been reported to be under the control of two independently inherited genes. The expression of the gene may vary with the environment. To be useful for application in plant breeding, the male-sterility system should be stable in a wide range of environments and inhibit virtually all seed production. The breeder cannot produce and maintain a pure population of male-sterile plants. The genetically male-sterile types (*msms*) can be propagated by crossing them with a heterozygous pollen source (*Msms*). This cross will produce a progeny in which 50% of the plants will be male-sterile (*msms*) and 50% male-fertile (*Msms*). If the crossing block is isolated, breeders will always harvest 50% male-sterile plants by harvesting only the male-sterile plants. The use of this system in commercial hybrid production is outlined in Figure 4.8.

Male sterility may be chemically induced by applying a variety of agents. This is useful where cytoplasmic male sterility (CMS) genes have not been found. However, this chemical technique has not been routinely applied in commercial plant breeding, needing further refinement.

**Cytoplasmic male sterility** Sometimes, male sterility is controlled by the cytoplasm (mitochondrial gene), but may be influenced by nuclear genes. A cytoplasm without sterility genes is described as normal (N) cytoplasm, while a cytoplasm that causes male sterility is called a sterile (s) cytoplasm or said to have **cytoplasmic male sterility (CMS)**. CMS is transmitted through the egg only (maternal factor). The condition has been induced in species such as sorghum by transferring nuclear chromosomes into a foreign cytoplasm (in this example, a milo plant was pollinated with kafir pollen and backcrossed to kafir). CMS has been found in

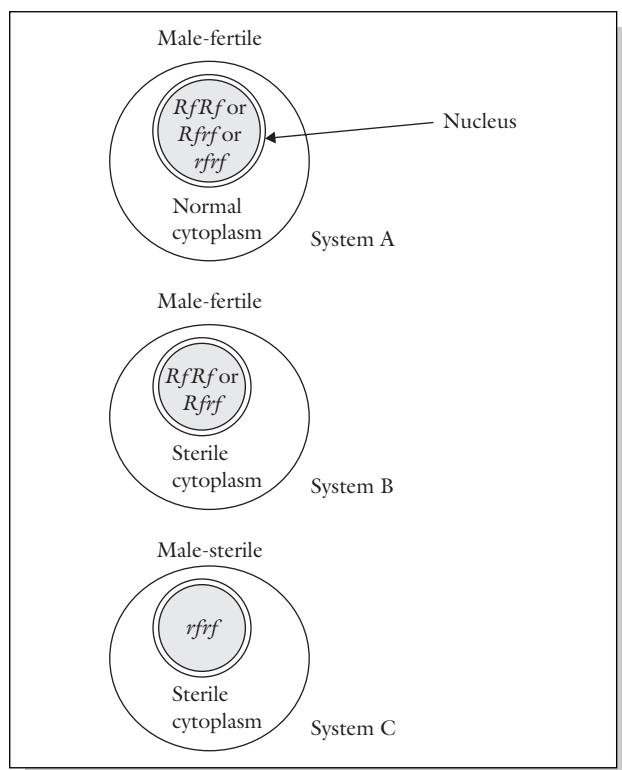


**Figure 4.8** Genetic male sterility as used in practical breeding.



**Figure 4.9** Cytoplasmic male sterility as applied in plant breeding. N, normal cytoplasm; s, sterile cytoplasm.

species including corn, sorghum, sugar beet, carrot, and flax. This system has real advantages in breeding ornamental species because all the offspring are male-sterile, hence allowing them to remain fruitless (Figure 4.9). By not fruiting, the plant remains fresh and in bloom for a longer time.



**Figure 4.10** The three systems of cytoplasmic-genetic male sterility. The three factors involved in CMS are the normal cytoplasm, the male-sterile cytoplasm, and the fertility restorer (*Rf*, *rf*).

**Cytoplasmic-genetic male sterility** CMS may be modified by the presence of fertility-restoring genes in the nucleus. CMS is rendered ineffective when the dominant allele for the fertility-restoring gene (*Rf*) occurs, making the anthers able to produce normal pollen (Figure 4.10). As previously stated, CMS is transmitted only through the egg, but fertility can be restored by *Rf* genes in the nucleus. Three kinds of progeny are possible following a cross, depending on

the genotype of the pollen source. The resulting progenies assume that the fertility gene will be responsible for fertility restoration.

### Exploiting male sterility in breeding

Male sterility is used primarily as a tool in plant breeding to eliminate emasculation in hybridization. Hybrid breeding of self-pollinated species is tedious and time-consuming. Plant breeders use male-sterile cultivars as female parents in a cross without emasculation. Male-sterile lines can be developed by backcrossing.

Using genetic male sterility in plant breeding is problematic because it is not possible to produce a pure population of male-sterile plants using conventional methods. It is difficult to eliminate the female population before either harvesting or sorting harvested seed. Consequently, this system of pollination control is not widely used for commercial hybrid seed production. However, CMS is used routinely in hybrid seed production in corn, sorghum, sunflower, and sugar beet. The application of male sterility in commercial plant hybridization is discussed in Chapter 18.

### Genotype conversion programs

To facilitate breeding of certain major crops, projects have been undertaken by certain breeders to create breeding stocks of male-sterile lines that plant breeders can readily obtain. In barley, over 100 spring and winter wheat cultivars have been converted to male-sterile lines by US Department of Agriculture (USDA) researchers. In the case of CMS, transferring chromosomes into foreign cytoplasm is a method of creating CMS lines. This approach has been used to create male sterility in wheat and sorghum. In sorghum, kafir chromosomes were transferred into milo cytoplasm by pollinating milo with kafir, and backcrossing the product to kafir to recover all the kafir chromosomes as previously indicated.

### References and suggested reading

- Acquaah, G. 2004. Horticulture: Principles and practices, 3rd edn. Prentice Hall, Upper Saddle River, NJ.
- Chaudhury, A.M., L. Ming, C. Miller, S. Craig, E.S. Dennis, and W.J. Peacock. 1997. Fertilization-independent seed development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 94:4223–4228.
- de Nettancourt, D. 1977. Incompatibility in angiosperm. Springer-Verlag, Berlin.
- Edwardson, J.R. 1970. Cytoplasmic male sterility. Bot. Rev. 36:341–420.
- Foskett, D.E. 1994. Plant growth and development. Academic Press, San Diego, CA.

- Kiesselbach, T.A. 1999. The structure and reproduction of corn, 50th anniversary edition. Cold Spring Harbor Press, New York.
- Kindinger, B., D. Bai, and V. Sokolov. 1996. Assignment of a gene(s) conferring apomixis in *Tripsacum* to a chromosome arm: Cytological and molecular evidence. *Genome* 39:1133–1141.
- Koltunow, A.M. 1993. Apomixis: Embryo sacs and embryos formed without meiosis or fertilization in ovules. *Plant Cell* 5:1425–1437.
- Simpson, G.G., A.R. Gendall, and C. Dean. 1999. When to switch to flowering. *Ann. Rev. Cell Dev. Biol.* 99:519–550.
- Stace, C.A. 1989. Plant taxonomy and biosystematics, 2nd edn. Routledge, Chapman & Hall, New York.
- Stern, K.R. 1997. Introductory plant biology, 7th edn. Wm. C. Brown Publishers, Dubuque, IA.
- Uribelarrea, M., J. Carcova, M.E. Otegui, and M.E. Westgate. 2002. Pollen production, pollination dynamics, and kernel set in maize. *Crop Sci.* 42:1910–1918.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Biennial plants complete their life cycle in two growing seasons.
- 2 A staminate flower is a complete flower.
- 3 Self-pollination promotes heterozygosity of the sporophyte.
- 4 The union of egg and sperm is called fertilization.
- 5 A branched raceme is called a panicle.
- 6 The carpel is also called the androecium.

### Part B

Please answer the following questions:

- 1 Plants reproduce by one of two modes, ..... or .....
- 2 Distinguish between monoecy and dioecy.
- 3 ..... is the transfer of pollen grain from the anther to the stigma of a flower.
- 4 What is self-incompatibility?
- 5 Distinguish between heteromorphic self-incompatibility and homomorphic self-incompatibility.
- 6 What is apomixis?
- 7 Distinguish between apospory and diplospory as mechanisms of apomixis.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the genetic and breeding implications of self-pollination.
- 2 Discuss the genetic and breeding implications of cross-pollination.
- 3 Fertilization does not always follow pollination. Explain.
- 4 Discuss the constraints of sexual biology in plant breeding.
- 5 Discuss how cytoplasmic male sterility (CMS) is used in a breeding program.
- 6 Discuss how genetic male sterility is used in a breeding program.

# **Section 3**

## **Germplasm issues**

**Chapter 5 Variation: types, origin, and scale**  
**Chapter 6 Plant genetic resources for plant breeding**

---

Without variability, it is not possible to conduct a plant breeding program! Germplasm is hence the critical first step in initiating a breeding program. Diversity in plants is enormous. It needs to be organized and characterized in order to facilitate its use by plant breeders. Students need to understand how this tremendous variation originates and the manner in which it is organized or classified. They need to know the sources to which breeders may go to find germplasm to initiate their breeding programs, and the rationale for selecting materials for breeding. The way breeders use plant germplasm has consequences in the agricultural production system. Major issues include the vulnerability of existing cultivars to disease epidemics because of the nature in which plant breeders over the years have manipulated and deployed them.

---



# 5

## Variation: types, origin, and scale

### Purpose and expected outcomes

*Biological variation is a fact of nature. No two plants are exactly alike. Plant breeders routinely deal with variability in one shape or form. It is indispensable to plant breeding, and hence breeders assemble or create it as a critical first step in a breeding program. Then, they have to discriminate among the variability, evaluate and compare superior genotypes, and increase and distribute the most desirable genotypes to producers. After completing this chapter, the student should be able to:*

- 1 Discuss the types of variation.
- 2 Discuss the origins of genetic variation.
- 3 Discuss the scale of genetic variation.
- 4 Distinguish between qualitative and quantitative variation.
- 5 Discuss the rules of plant classification.

### Classifying plants

**Plant taxonomy** is the science of classifying and naming plants. Organisms are classified into five major groups (kingdoms) – Animalia, Plantae, Fungi, Protista, and Monera (Table 5.1). Plant breeders are most directly concerned about Plantae, the plant kingdom. However, one of the major objectives of plant breeding is to breed for resistance of the host to diseases and economic destruction caused by organisms in the other four kingdoms that adversely impact plants. Plant breeding depends on plant variation or diversity for success. It is critical that the appropriate plant material is acquired for a breeding program. An international scientific body sets the rules for naming plants. Standardizing the naming of plants eliminates the confusion from the numerous culture-based names of plants. For example corn in the USA is called maize in Europe, not to mention the thousands of other names worldwide.

**Table 5.1** The five kingdoms of organisms as described by Whitaker.

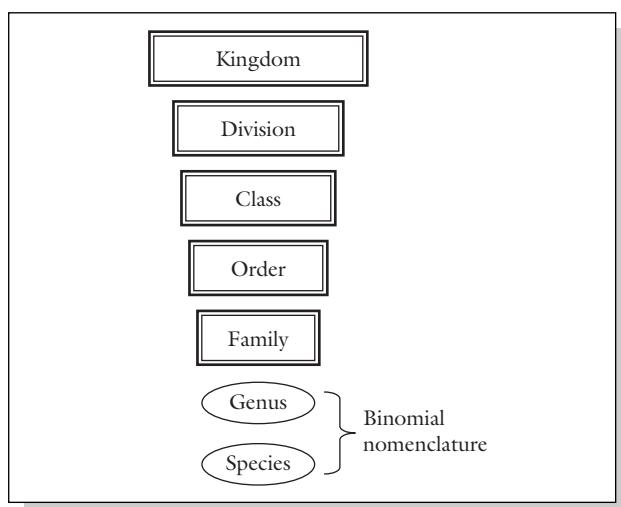
**Monera** (have prokaryotic cells)  
Bacteria

**Protista** (have eukaryotic cells)  
Algae  
Slime molds  
Flagellate fungi  
Protozoa  
Sponges

**Fungi** (absorb food in solution)  
True fungi

**Plantae** (produce own food by the process of photosynthesis)  
Bryophytes  
Vascular plants

**Animalia** (ingest their food)  
Multicellular animals



**Figure 5.1** Taxonomic hierarchy of plants. Plant breeders routinely cross plants without problem within a species. Crosses between species are problematic, and often impossible between genera and beyond.

The **binomial nomenclature** was developed by Carolus Linnaeus and entails assigning two names based on the genus and species, the two bottom taxa in taxonomic hierarchy (Figure 5.1). It is important for the reader to understand that plant breeding by conventional tools alone is possible primarily at the species level. Crosses are possible within species and occasionally between species (but are often problematic). However, plant breeding incorporating molecular tools allows gene transfer from any taxonomic level to another. It is important to emphasize that such a transfer is not routine and has its challenges.

The kingdom Plantae comprises **vascular plants** (plants that contain conducting vessels – xylem and phloem) and **non-vascular plants** (Table 5.2). Vascular plants may be seeded or seedless. Furthermore, seeded plants may be **gymnosperms** (have naked seed) or **angiosperms** (have seed borne in a fruit). Flowering plants may have seed with one cotyledon, called **monocots** (includes grasses such as wheat, barley, and rice), or seed with two cotyledons, called **dicots** (includes legumes such as soybean, pea, and peanut) (Table 5.3). The strategies for breeding flowering species are different from those for non-flowering species. Flowering species (sexually reproducing) can be genetically manipulated through the sexual process by crossing, whereas non-flowering species (asexually reproducing) cannot. Furthermore, even within flowering plants, the method

**Table 5.2** Divisions in the kingdom Plantae.

	Division	Common name
<b>Bryophytes</b> (non-vascular plants)	Hepaticophyta Anthocerotophyta Bryophyta	Liverworts Hornworts Mosses
<b>Vascular plants</b>		
Seedless	Psilotophyta Lycophyta Sphenophyta Pterophyta	Whisk ferns Club mosses Horsetails Ferns
Seeded	Pinophyta Subdivision: Cycadaceae Subdivision: Piniccae Class: Ginkgoatae Class: Pinatae Subdivision: Gneticcae Magnoliophyta Class: Liliopsida Class: Magnoliopsida	Gymnosperms Cycads Conifers <i>Gnetum</i> Flowering plants Monocots Dicots

for breeding differs according to the mode of pollination – self-pollination or cross-pollination.

### Rules of classification of plants

The science of plant taxonomy is coordinated by the **International Board of Plant Nomenclature**, which makes the rules. The Latin language is used in naming plants. Sometimes, the names given reflect specific plant attributes or uses of the plant. For example, some specific epithets indicate color, e.g. *alba* (white), *variegata* (variegated), *rubrum* (red), and *aureum* (golden); others are *vulgaris* (common), *esculentus* (edible), *sativus* (cultivated), *tuberosum* (tuber bearing), or *officinalis* (medicinal). The ending of a name is often characteristic of the taxon. Class names often end in **-opsida** (e.g. Magnoliopsida), orders in **-ales** (e.g. Rosales), and families in **-aceae** (e.g. Rosaceae). There are certain specific ways of writing the binomial name that are strictly adhered to in scientific communication. These rules are as follows:

- 1 It must be underlined or written in italics (being non-English).
- 2 The genus name must start with an upper case letter; the species name always starts with a lower case letter. The term “species” is both singular and plural, and may be shortened to sp. or spp.

**Table 5.3** Important field crop families in the division Magnoliophyta (flowering plants).**Monocots**

Poaceae ( <i>grass family</i> )	In terms of numbers, the grass family is the largest of flowering plants. It is also the most widely distributed Examples of species: wheat, barley, oats, rice, corn, fescues, bluegrass
Aracaceae ( <i>palm family</i> )	The palm family is tropical and subtropical in adaptation Examples of species: oil palm ( <i>Elaeis guineensis</i> ), coconut palm ( <i>Cocos nucifera</i> )
Amaryllidaceae ( <i>amaryllis family</i> )	Plants with tunicate bulbs characterize this family Examples of species: onion, garlic, chives

**Dicots**

Brassicaceae ( <i>mustard family</i> )	The mustard family is noted for its pungent herbs Examples of species: cabbage, radish, cauliflower, turnip, broccoli
Fabaceae ( <i>legume family</i> )	The legume family is characterized by flowers that may be regular or irregular. The species in this family are an important source of protein for humans and livestock Examples of species: dry beans, mung bean, cowpea, pea, peanut, soybean, clover
Solanaceae ( <i>nightshade family</i> )	This family is noted for the poisonous alkaloids many of them produce (e.g., belladonna, nicotine, atropine, solanine) Examples of species: tobacco, potato, tomato, pepper, eggplant
Eupobiaceae ( <i>spurge family</i> )	Members of the spurge family produce milky latex, and include a number of poisonous species Examples of species: cassava ( <i>Manihot esculenta</i> ), castor bean
Asteraceae ( <i>sunflower family</i> )	The sunflower family has the second largest number of flowering plant species Example of species: sunflower, lettuce
Apiaceae ( <i>carrot family</i> )	Plants in this family usually produce flowers that are arranged in umbels Examples of species: carrot, parsley, celery
Cucurbitaceae ( <i>pumpkin family</i> )	The pumpkin or gourd family is characterized by prostrate or climbing herbaceous vines with tendrils and large, fleshy fruits containing numerous seeds Examples of species: pumpkin, melon, watermelon, cucumber

- 3 Frequently, the scientist who first named the plant adds his or her initial to the binary name. The letter L indicates that Linnaeus first named the plant. If revised later, the person responsible is identified after the L, for example, *Glycine max* (L.) Merr. (for Merrill).
- 4 The generic name may be abbreviated and can also stand alone. However, the specific epithet cannot stand alone. Valid examples are *Zea mays*, *Zea*, *Z. mays*, but not *mays*.
- 5 The cultivar or variety name may be included in the binomial name. For example, *Lycopersicon esculentum* Mill cv. "Big Red", or *L. esculentum* "Big Red". The cultivar (cv) name, however, is not written in italics.

**Operational classification systems**

Crop plants may be classified for specific purposes, for example, according to seasonal growth, kinds of stem, growth form, and economic part or agronomic use.

*Seasonal growth cycle*

Plants may be classified according to the duration of their life cycle (i.e., from seed, to seedling, to flowering, to fruiting, to death, and back to seed). On this basis, crop plants may be classified as **annual**, **biennial**, **perennial**, or **monocarp**, as previously discussed in Chapter 4.

*Stem type*

Certain plants have non-woody stems, existing primarily in vegetative form (e.g., onion, corn, sugar beet) and are called **herbs** (or herbaceous plants). **Shrubs** are plants with multiple stems that arise from the ground level (e.g., dogwood, azalea, kalmia), while **trees** (e.g. apple, citrus, palms) have one main trunk or central axis.

*Common growth form*

Certain plants can stand upright without artificial support; others cannot. Based on this characteristic, plants

may be classified into groups. The common groups are as follows.

- 1 Erect.** Erect plants can stand upright without physical support, growing at about a 90° angle to the ground. This feature is needed for mechanization of certain crops during production. Plant breeders develop erect (bush) forms of non-erect (pole) cultivars for this purpose. There are both pole and bush cultivars of crops such as bean (*Phaseolus vulgaris* L.) in cultivation.
- 2 Decumbent.** Plants with decumbent stem growth form, such as peanuts (*Arachis hypogaea*), are extremely inclined with raised tips.
- 3 Creeping (or repent).** Plants in this category, such as strawberry (*Fragaria* spp.), have stems that grow horizontally on the ground.
- 4 Climbing.** Climbers are plants with modified vegetative parts (stems or leaves) that enable them to wrap around a nearby physical support, so they do not have to creep on the ground. Examples are yam (*Dioscorea* spp.) and ivy.
- 5 Despitose (bunch or tufted).** Grass species, such as buffalograss, have a creeping form whereas others, such as tall fescue, have a bunch form and hence do not spread by horizontal growing stems.

#### Agronomic use

Crop plants may be classified according to agronomic use as follows:

- 1 Cereals:** grasses such as wheat, barley, and oats grown for their edible seed.
- 2 Pulses:** legumes grown for their edible seed (e.g., peas, beans).
- 3 Grains:** crop plants grown for their edible dry seed (e.g., corn, soybean, cereals).
- 4 Small grains:** grain crops with small seeds (e.g., wheat, oats, barley).
- 5 Forage:** plants grown for their vegetative matter, which is harvested and used fresh or preserved as animal feed (e.g., alfalfa, red clover).
- 6 Roots:** crops grown for their edible, modified (swollen) roots (e.g., sweet potato, cassava).
- 7 Tubers:** crops grown for their edible, modified (swollen) stem (e.g., Irish potato, yam).
- 8 Oil crops:** plants grown for their oil content (e.g., soybean, peanut, sunflower, oil palm).
- 9 Fiber crops:** crop plants grown for use in fiber production (e.g., jute, flax, cotton).
- 10 Sugar crops:** crops grown for use in making sugar (e.g., sugarcane, sugar beet).

- 11 Green manure crops:** crop plants grown and plowed under the soil while still young and green, for the purpose of improving soil fertility (e.g., many leguminous species).
- 12 Cover crops:** crops grown between regular cropping cycles, for the purpose of protecting the soil from erosion and other adverse weather factors (e.g., many annuals).
- 13 Hay:** grasses or legume plants grown, harvested, and cured for feeding animals (e.g., alfalfa, buffalograss).

#### Adaptation

There are also other operational classifications used by plant scientists. For example, plants may be classified on the basis of temperature adaptation as either cool season or warm season plants.

- 1 Cool season or temperate plants.** These plants, such as wheat, sugar beet, and tall fescue prefer a monthly temperature of between 15 and 18°C (59–64°F) for growth and development.
- 2 Warm season or tropical plants.** These plants, such as corn, sorghum, and buffalograss, require warm temperatures of between 18 and 27°C (64–80°F) during the growing season.

#### Additional classification of horticultural plants

Whereas some of the above operational classifications are applicable, horticultural plants have additional classification systems. These include the following:

- 1 Fruit type:**
  - (a) **Temperate fruits** (e.g., apple, peach) versus **tropical fruits** (e.g., orange, coconut).
  - (b) **Fruit trees**, which have fruits borne on trees (e.g., apple, pear).
  - (c) **Small fruits**, generally woody perennial dicots (e.g., strawberry, blackberry).
  - (d) **Bramble fruits**, non-tree fruits that need physical support (e.g., raspberry).
- 2 Flowering** (e.g., sunflower, pansy) versus **foliage** (non-flowering, e.g., coleus, sansevieria) plants.
- 3 Bedding plants**, annual plants grown in beds (e.g., zinnia, pansy, petunia).
- 4 Deciduous trees** (shed leaves seasonally) versus **evergreen plants** (no leaf shedding).

## Types of variation among plants

As previously indicated, the phenotype (the observed trait) is the product of the genotype and the environment ( $P = G + E$ ). The phenotype may be altered by altering  $G$ ,  $E$ , or both. There are two fundamental sources of change in phenotype – genotype and the environment – and hence two kinds of variation, **genetic** and **environmental**. Later in the book, an additional source of change,  $G \times E$  (interaction of the genotype and the environment) will be introduced.

### Environmental variation

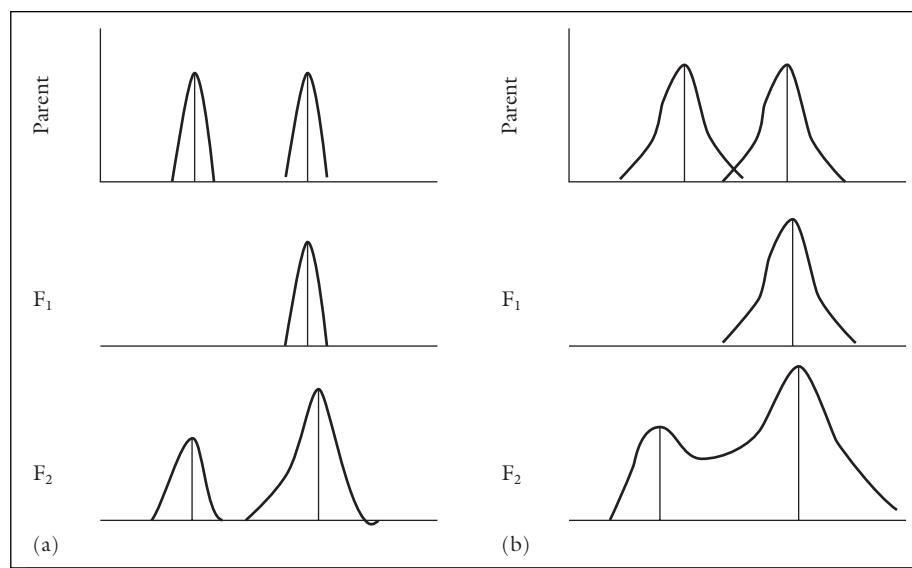
When individuals from a clonal population (i.e., identical genotype) are grown in the field, the plants will exhibit differences in the expression of some traits because of non-uniform environments. The field is often heterogeneous with respect to plant growth factors – nutrients, moisture, light, and temperature. Some fields are more heterogeneous than others. Sometimes, non-growth factors may occur in the environment and impose different intensities of environmental stress on plants. For example, disease and pest agents may not uniformly infect plants in the field. Similarly, plants that occur in more favorable parts of the field or are impacted to a lesser degree by an adverse environmental factor would perform better than disadvantaged plants. That

is, even clones may perform differently under different environments, and inferior genotypes can outperform superior genotypes under uneven environmental conditions. If a breeder selects an inferior genotype by mistake, the progress of the breeding program will be slowed. Consequently, plant breeders use statistical tools and other selection aids to help in reducing the chance of advancing inferior genotypes, and thereby increase progress in the breeding program.

As previously noted, environmental variation is not heritable. However, it can impact on heritable variation (see below). Plant breeders want to be able to select a plant on the basis of its nature (genetics) not nurture (growth environment). To this end, evaluations of breeding material are conducted in as uniform an environment as possible. Furthermore, the selection environment is often similar to the one in which the crop is commercially produced.

### Genetic variability

Variability that can be attributed to genes that encode specific traits, and can be transmitted from one generation to the next, is described as **genetic** or **heritable variation**. Because genes are expressed in an environment, the degree of expression of a heritable trait is impacted by its environment, some more so than others (Figure 5.2). Heritable variability is indispensable to



**Figure 5.2** Environmental effect on gene expression: phenotype = genotype + environment. Some traits are influenced a lot more than others by the environment. Cross (a) has small environmental influence such that the phenotypes are distinguishable in the  $F_2$ ; in cross (b) the environmental influence is strong, resulting in more blurring of the differences among phenotypes in the segregating population.

plant breeding. As previously noted, breeders seek to change the phenotype (trait) permanently and heritably by changing the genotype (genes) that encode it. Heritable variability is consistently expressed generation after generation. For example, a purple-flowered genotype will always produce purple flowers. However, a mutation can permanently alter an original expression. For example, a purple-flowered plant may be altered by mutation to become a white-flowered plant.

Genetic variation can be detected at the molecular as well as the gross morphological level. The availability of biotechnological tools (e.g., DNA markers) allows plant breeders to assess genetic diversity of their materials at the molecular level. Some genetic variation is manifested as visible variation in morphological traits (e.g., height, color, size), while compositional or chemical traits (e.g., protein content, sugar content of a plant part) require various tests or devices for evaluating them. Furthermore, plant breeders are interested in how genes interact with their environment (called genotype  $\times$  environment interaction). This information is used in the decision-making process during cultivar release (see Chapter 23).

### Origins of genetic variability

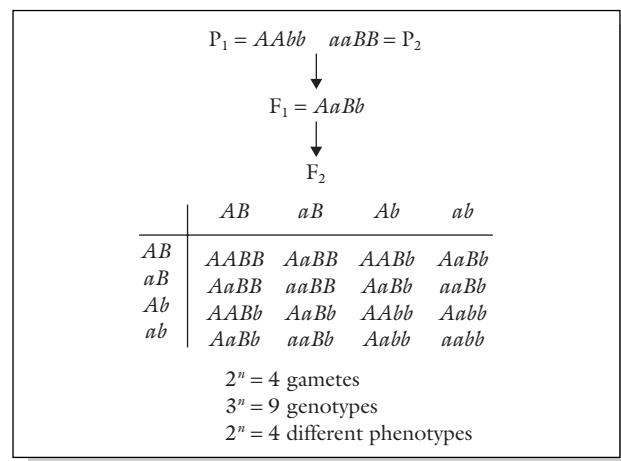
There are three ways in which genetic or heritable variability originates in nature – gene recombination, modifications in chromosome number, and mutations. The significant fact to note is that, rather than wait for them to occur naturally, plant breeders use a variety of techniques and methods to manipulate these three phenomena more and more intensively, as they generate genetic variation for their breeding programs. With advances in science and technology (e.g., gene transfer, somaclonal variation), new sources of genetic variability have become available to the plant breeder. Variability generated from these sources is, however, so far limited.

### Genetic recombination

**Genetic recombination** applies only to sexually reproducing species and represents the primary source of variability for plant breeders in those species. As previously described, genetic recombination occurs via the cellular process of meiosis. This phenomenon is responsible for the creation of non-parental types in the progeny of a cross, through the physical exchange of parts of homologous chromosomes (by breakage fusion). The cytological evidence of this event is the characteristic

crossing (X-configuration or **chiasma**) of the adjacent homologous chromosome strands, as described in Chapter 3, allowing genes that were transmitted together (non-independent assortment) in the previous generation to become independent. Consequently, sexual reproduction brings about gene reshuffling and the generation of new genetic combinations (recombinants). Unlike mutations that cause changes in genes themselves in order to generate variability, recombination generates variability by assembling new combinations of genes from different parents. In doing this, some gene associations are broken.

Consider a cross between two parents of contrasting genotypes  $AAbb$  and  $aaBB$ . A cross between them will produce an  $F_1$  of genotype  $AaBb$ . In the  $F_2$  segregating population, and according to Mendel's law, the gametes ( $AB$ ,  $Ab$ ,  $aB$ , and  $ab$ ) will combine to generate variability, some of which will be old (like the parents: parental), while others will be new (unlike the parents: recombinants) (Figure 5.3). The larger the number of pairs of allelic genes by which the parents differ, the greater the new variability that will be generated. Representing the number of different allelic pairs by  $n$ , the number of gametes produced is  $2^n$ , and the number of genotypes produced in the  $F_2$  following random mating is  $3^n$  with  $2^n$  phenotypes (assuming complete dominance). In this example, two new homozygous genotypes ( $aabb$ ,  $AABB$ ) are obtained.



**Figure 5.3** Genetic recombination results in the production of recombinants in the segregating population. This phenomenon is a primary source of variability in breeding flowering species. The larger the number of genes ( $n$ ) the greater the amount of variability that can be generated from crossing.

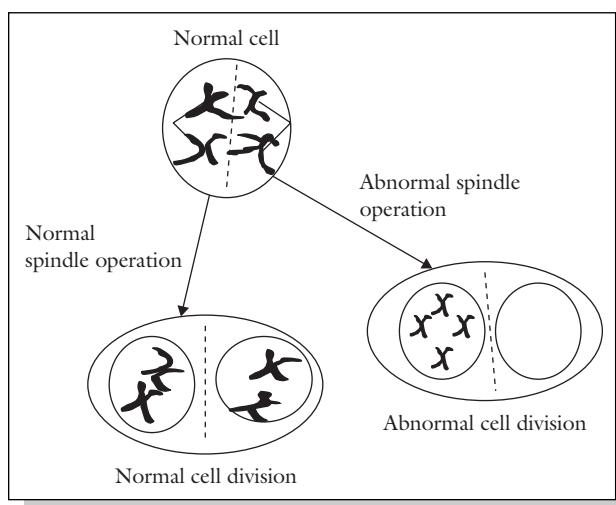
It should be pointed out that recombination only includes genes that are already present in the parents. Consequently, if there is no genetic linkage, the new gene recombination can be predicted. Where linkage is present, knowledge of the distance between gene loci on the chromosomes is needed for estimating their frequencies. As previously discussed in Chapter 3, additional variability for recombination may be observed where intra-allelic and interallelic interactions (epistasis) occur. This phenomenon results in new traits that were not found in the parents. Another source of genetic variability is the phenomenon of **gene transgression**, which causes some individuals in a segregating population from a cross to express the trait of interest outside the boundaries of the parents (e.g., taller than the taller parent, or shorter than the shorter parent). These new genotypes are called **transgressive segregates**. The discussion so far has assumed diploidy in the parents. However, in species of higher ploidy levels (e.g., tetraploid, hexaploid), it is not difficult to see how additional genetic variability could result where allelic interactions occur.

One of the tools of plant breeding is hybridization (crossing of divergent parents), whereby breeders selectively mate plants to allow their genomes to be reshuffled into new combinations to generate variability in which selection can be practiced. By carefully selecting the parents to be mated, the breeder has some control over the nature of the genetic variability to be generated. Breeding methods that include repeated hybridization (e.g., reciprocal selection, recurrent selection) offer more opportunities for recombination to occur.

The speed and efficiency with which a breeder can identify (by selecting among hybrids and their progeny) desirable combinations, is contingent upon the number of genes and linkage relationships that are involved. Because linkage is likely to exist, the plant breeder is more likely to make rapid progress with recombination by selecting plant genotypes with high chiasma frequency (albeit unconsciously). It follows then that the cultivar developed with the desired recombination would also have higher chiasma frequency than the parents used in the breeding program.

### Ploidy modifications

New variability may arise naturally through modifications in chromosome number as a result of hybridization (between unidentical genotypes) or abnormalities in the nuclear division processes (spindle malfunction). Failure of the spindle mechanism, during karyokinesis or



**Figure 5.4** Failure of the genetic spindle mechanism may occur naturally or be artificially induced by plant breeders (using colchicine), resulting in cell division products that inherit abnormal chromosome numbers. Plant breeders deliberately manipulate the ploidy of cells to create polyploids.

even prior to that, can lead to errors in chromosome numbers transmitted to cells, such as **polyploidy** (individuals with multiples of the basic set of chromosomes for the species in their cells) (Figure 5.4). Sometimes, instead of variations involving complete sets of chromosomes, plants may be produced with multiples of only certain chromosomes or deficiencies of others (called **aneuploidy**). Sometimes, plants are produced with half the number of chromosomes in the somatic cells (called **haploids**). Like genetic recombination, plant breeders are able to induce various kinds of chromosome modification to generate variability for breeding. The subject is discussed in detail in Chapter 13.

### Mutation

**Mutation** is the ultimate source of biological variation. Mutations are important in biological evolution as sources of heritable variation. They arise spontaneously in nature as a result of errors in cellular processes such as DNA replication (or duplication), and by **chromosomal aberrations** (deletion, duplication, inversion, translocation). The molecular basis of mutation may be described by mechanisms such as: (i) modification of the structure of DNA or a component base of DNA; (ii) substitution of one base for a different base; (iii) deletion or addition of one base in one DNA strand; (iv) deletion or addition in one or more base pairs in

both DNA strands; and (v) inversion of a sequence of nucleotide base pairs within the DNA molecule. These mechanisms are discussed further in Chapter 12 on mutation breeding.

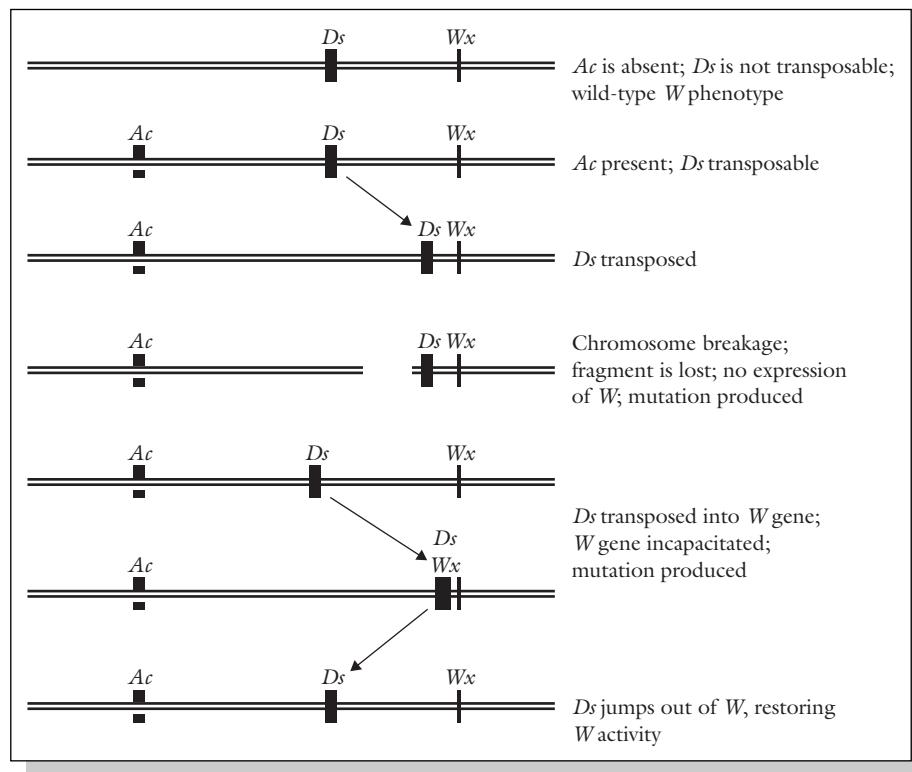
Mutations may also be induced by plant breeders using agents such as irradiation and chemicals. Many useful mutations have been found in nature or induced by plant breeders (e.g., dwarfs, nutritional quality genes). However, many mutations are deleterious to their carriers and are hence selected against in nature or by plant breeding. From the point of view of the breeder, mutations may be useful, deleterious, or neutral. Neutral mutations are neither advantageous nor disadvantageous to the individuals in which they occur. They persist in the population in the heterozygous state as recessive alleles and become expressed only when in the homozygous state, following an event such as selfing.

### Transposable elements

The phenomenon of **transposable elements** (genes with the capacity to relocate within the genome) also creates new variability. **Transposable genetic elements**

(transposable elements, transposons, or “jumping genes”) are known to be nearly universal in occurrence. These mobile genetic units relocate within the genome by the process called **transposition**. The presence of transposable elements indicates that genetic information is not fixed within the genome of an organism. Barbara McClintock, working with corn in the 1940s, was the first to detect transposable elements, which she initially identified as **controlling elements**. This discovery was about 20 years ahead of the discovery of transposable elements in prokaryotes. Controlling elements may be grouped into families. The members of each family may be divided into two classes: autonomous elements or non-autonomous elements. Autonomous elements have the ability to transpose whereas the non-autonomous elements are stable (but can transpose with the aid of an autonomous element through *trans-activation*).

McClintock studied two mutations: **dissociation** (*Ds*) and **activator** (*Ac*). The *Ds* element is located on chromosome 9. *Ac* is capable of autonomous movement, but *Ds* moves only in the presence of *Ac*. *Ds* has the effect of causing chromosome breakage at a point on the chromosome adjacent to its location (Figure 5.5).



**Figure 5.5** In the *Ac–Ds* (activator–dissociation) system of transposable elements in maize, the transposition of the *Ds* to *Wx* causes chromosome breakage, leading to the production of a mutant. In another scenario, the *Ds* is transposed into *Wx*, causing a mutant to be produced.

The *Ac* element has an open reading frame. The activities of corn transposable elements are developmentally regulated. That is, the transposable elements transpose and promote genetic rearrangements at only certain specific times and frequencies during plant development. Transposition involving the *Ac-Ds* system is observed in corn as spots of colored aleurone. A gene required for the synthesis of anthocyanin pigment is inactivated in some cells whereas other cells have normal genes, resulting in spots of pigment in the kernel (genetic mosaicism).

### Biotechnology for creating genetic variability

#### Gene transfer

The **rDNA technology** is state-of-the-art in gene transfer to generate genetic variability for plant breeding. With minor exceptions, DNA is universal. Consequently, DNA from an animal may be transferred to a plant! The tools of biotechnology may be used to incorporate genes from distant sources into adapted cultivars. An increasing acreage of cotton, soybean, and maize are being sown to genetically modified (GM) cultivars, indicating the importance of this technology for creating variability for plant breeding. Economic gene transfers have been made from bacteria to plants to confer disease and herbicide resistance to plants. The most common GM products on the market are Roundup Ready® cultivars (e.g., cotton, soybean) with herbicide tolerance, and *Bt* products (e.g., corn) with resistance to lepidopteran pests. The technique of **site-directed mutagenesis** allows scientists to introduce mutations into specified genes, primarily for the purpose of studying gene function, and not for generating variability for breeding *per se*. Other tissue-culture-based techniques include protoplast fusion, cybrid formation, and the use of transposons. Chapter 14 is devoted to the application of biotechnology in plant breeding.

#### Somaclonal variation

*In vitro* culture of plants is supposed to produce clones (genetically identical derivatives from the parent material). However, the tissue culture environment has been known to cause heritable variation called **somaclonal variation**. The causes cited for these changes include karyotypic changes, cryptic chromosomal rearrangements, somatic crossing over and sister chromatid exchange, transposable elements, and gene amplification.

Some of these variations have been stable and fertile enough to be included in breeding programs.

### Scale of variability

As previously indicated, biological variation can be enormous and overwhelming to the user. Consequently, there is a need to classify it for effective and efficient use. Some variability can be readily categorized by counting and arranging into distinct non-overlapping groups; this is said to be discrete or **qualitative variation**. Traits that exhibit this kind of variation are called **qualitative traits**. Other kinds of variability occur on a continuum and cannot be placed into discrete groups by counting. There are intermediates between the extreme expressions of such traits. They are best categorized by measuring or weighing and are described as exhibiting continuous or **quantitative variation**. Traits that exhibit this kind of variation are called **quantitative traits**.

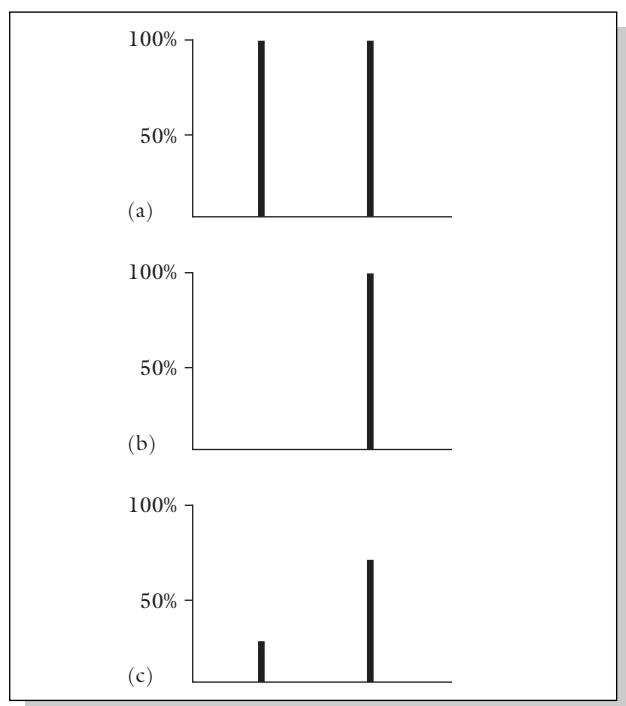
However, there are some plant characters that may be classified either way. Sometimes, for convenience, a quantitative trait may be classified as though it were qualitative. For example, an agronomic trait such as earliness or plant maturity is quantitative in nature. However, it is possible to categorize cultivars into maturity classes (e.g., in soybean, maturity classes range from 000 (very early) to VIII (very late)). Plant height can be treated in a similar fashion, and so can seed coat color (expressed as shades of a particular color).

#### Qualitative variation

Qualitative variation is easy to classify, study, and utilize in breeding. It is simply inherited (controlled by one or a few genes) and amenable to Mendelian analysis (Figure 5.6). Examples of qualitative traits include diseases, seed characteristics, and compositional traits. Because they are amenable to Mendelian analysis, the chi-square statistical procedure may be used to determine the inheritance of qualitative genes. The success of gene transfer using molecular technology so far has involved the transfer of single genes (or a few at best), such as the *Bt* and *Ht* (herbicide-tolerant) products.

#### Breeding qualitative traits

Breeding qualitative traits is relatively straightforward. They are readily identified and selected, although breeding recessive traits is a little different from breeding dominant traits (Figure 5.7). It is important to have



**Figure 5.6** Qualitative variation produces discrete measurements that can be placed into distinct categories: (a) parental phenotype, (b) dominant phenotype in  $F_1$ , and (c) 3 : 1 phenotypic ratio in  $F_2$ .

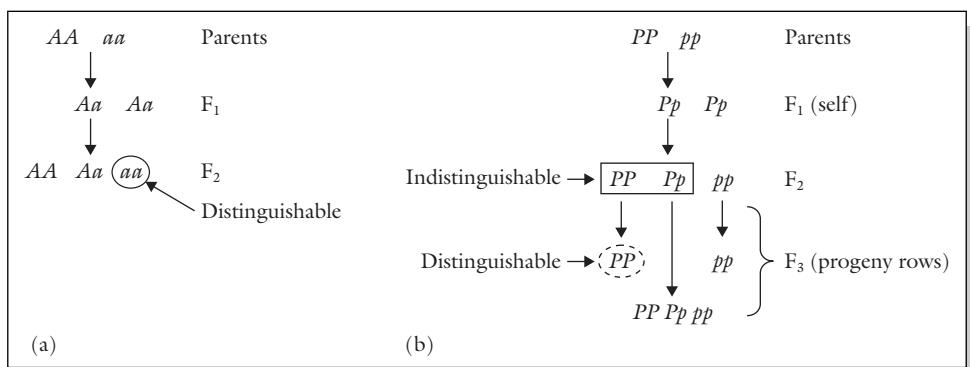
a large segregating population, especially if several loci are segregating, to increase the chance of finding the desired recessive recombinants. For example, if two loci are segregating, a cross between  $AA$  and  $aa$  would produce 25% homozygous recessive individuals in the  $F_2$

( $1AA : 2Aa : 1aa$ ). A minimum of 16 plants would be needed in the  $F_2$  stage to include the desired recombinant. However, if four loci are segregating, at least 256 plants are required in the  $F_2$  in order to observe the desired recombinant. It is important to note that the desired recombinant can be isolated from the  $F_2$  without any further evaluation. In the case of a dominant locus, (e.g., the cross  $PP \times pp$ ), 25% of the  $F_2$  will be homozygous recessive, whereas 75% would be of the heterozygous-dominant phenotype (of which only 25% would be homozygous dominant). The breeder needs to advance the material one more generation to identify individuals that are homozygous dominant.

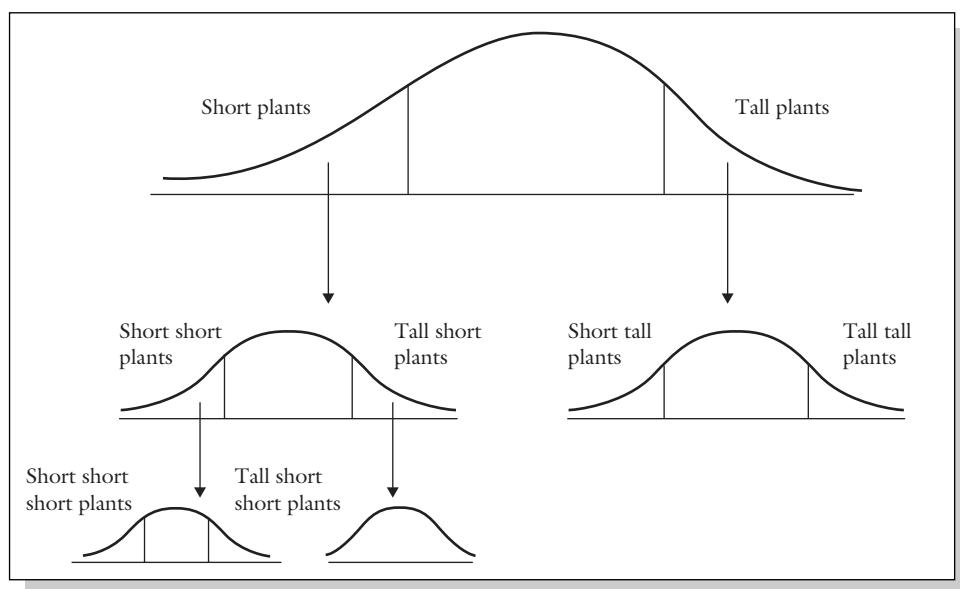
### Quantitative variation

Most traits encountered in plant breeding are quantitatively inherited. Many genes control such traits, each contributing a small effect to the overall phenotypic expression of a trait. Variation in quantitative trait expression is without natural discontinuities, as previously indicated. Traits that exhibit continuous variations are also called **metric traits**. Any attempt to classify such traits into distinct groups is only arbitrary. For example, height is a quantitative trait. If plants are grouped into tall versus short plants, one could find relatively tall plants in the short group and similarly short plants in the tall group (Figure 5.8).

Quantitative traits are conditioned by many to numerous genes (**polygenic inheritance**) with effects that are too small to be individually distinguished. They are sometimes called **minor genes**. Quantitative trait expression is very significantly modified by the variation in environmental factors to which plants in



**Figure 5.7** (a) Breeding a qualitative trait conditioned by a recessive gene. The desired recombinant can be observed and selected in the  $F_2$ . (b) Breeding a qualitative trait conditioned by a dominant gene. The desired trait cannot be distinguished in the  $F_2$ , requiring another generation (progeny row) to distinguish between the dominant phenotypes.



**Figure 5.8** Quantitative traits are influenced to a larger degree by the environment than are qualitative traits. The mean of the  $F_1$  is generally intermediate between the parental means. The  $F_2$  is usually spread within the entire range of the parental values.

the population are subjected. Continuous variation is caused by environmental variation and genetic variation due to the simultaneous segregation of many genes affecting the trait. These effects convert the intrinsically discrete variation to a continuous one. Quantitative genetics is used to distinguish between the two factors that cause continuous variability to occur (see Chapter 8).

#### Breeding quantitative traits

Breeding quantitative traits is more challenging than breeding qualitative traits. A discussion of quantitative genetics will give the reader an appreciation for the nature of quantitative traits and a better understanding of their breeding. Quantitative genetics is discussed in Chapter 8.

---

### References and suggested reading

- Acquaah, G. 2004. Horticulture: Principles and practices, 3rd edn. Prentice Hall, Upper Saddle River, NJ.  
 Falconer, D.S. 1981. Introduction to quantitative genetics, 2nd edn. Longman, London.  
 Klug, W.S., and M.R. Cummings. 1997. Concepts of genetics, 5th edn. Prentice Hall, Upper Saddle, NJ.

---

### Outcomes assessment

#### Part A

Please answer the following questions true or false:

- 1 The binomial nomenclature was discovered by Gregor Mendel.
- 2 Angiosperms have naked seed.
- 3 Environmental variation is heritable.
- 4 Qualitative traits exhibit continuous variation.

**Part B**

Please answer the following questions:

- 1 Define plant taxonomy. Why is it important to plant breeding?
- 2 ..... is the international body responsible for coordinating plant taxonomy.
- 3 Phenotype = ..... + environment.
- 4 What are transposable elements?

**Part C**

Please write a brief essay on each of the following topics:

- 1 Discuss genetic recombination as a source of variability for plant breeding.
- 2 Discuss the nature of qualitative variation.
- 3 Discuss the role of environmental variation in plant breeding.
- 4 Describe the role of biotechnology for creating variability for plant breeding.



# 6

## Plant genetic resources for plant breeding

### Purpose and expected outcomes

*As previously indicated, plant breeders depend on variability for crop improvement. Plant genetic resources (plant germplasm) used in plant breeding are natural resources that are susceptible to erosion from use and abuse. It is important that they be collected, properly used, managed, and conserved to avoid irreparable loss of precious genetic material. After completing this chapter, the student should be able to:*

- 1 Discuss the importance of germplasm to plant breeding.
- 2 Define the types of germplasm.
- 3 Discuss the sources of germplasm for plant breeding.
- 4 Discuss the mechanisms for conservation of germplasm.
- 5 Discuss the international role in germplasm conservation.

### Importance of germplasm to plant breeding

**Germplasm** is the lifeblood of plant breeding without which breeding is impossible to conduct. It is the genetic material that can be used to perpetuate a species or population. It not only has reproductive value, but through genetic manipulation (plant breeding), germplasm can be improved for better performance of the crop. Germplasm provides the materials (parents) used to initiate a breeding program. Sometimes, all plant breeders do is to evaluate plant germplasm and make a selection from existing biological variation. Promising genotypes that are adapted to the production region are then released to producers. Other times, as discussed in Chapter 5, breeders generate new variability by using a variety of methods such as crossing parents, mutagenesis (inducing mutations), and, more recently, gene transfer. This base population is then subjected to appropriate selection methods, leading to the identification and further evaluation of promising genotypes for release as cultivars. When breeders need to improve plants, they

have to find a source of germplasm that would supply the genes needed to undertake the breeding project. To facilitate the use of germplasm, certain entities (germplasm banks) are charged with the responsibility of assembling, cataloguing, storing, and managing large numbers of germplasm. This strategy allows scientists ready and quick access to germplasm when they need it.

### Centers of diversity in plant breeding

The subject of centers of diversity was first discussed in Chapter 2. Whereas the existence of centers of crop origin or domestication is not incontrovertible, the existence of natural reservoirs of plant genetic variability has been observed to occur in certain regions of the world. These centers are important to plant breeders because they represent pools of diversity, especially wild relatives of modern cultivars.

Plant breeding may be a victim of its own success. The consequence of selection by plant breeders in their

programs is the steady erosion or reduction in genetic variability, especially in the highly improved crops. Modern plant breeding tends to focus on a small amount of variability for crop improvement. Researchers periodically conduct plant explorations (or collections) to those centers of diversity where wild plants grow in their natural habitats, to collect materials that frequently yield genes for addressing a wide variety of plant breeding problems, including disease resistance, drought resistance, and chemical composition augmentation.

### Sources of germplasm for plant breeding

Germplasm may be classified into five major types – **advanced (elite) germplasm, improved germplasm, landraces, wild or weedy relatives, and genetic stocks**. The major sources of variability for plant breeders may also be categorized into three broad groups – **domesticated plants, undomesticated plants, and other species or genera**.

#### *Domesticated plants*

Domesticated plants are those plant materials that have been subjected to some form of human selection and are grown for food or other uses. There are various types of such material:

- 1 **Commercial cultivars.** There are two forms of this material – **current cultivars** and **retired or obsolete cultivars**. These are products of formal plant breeding for specific objectives. It is expected that such genotypes would have superior gene combinations, be adapted to a growing area, and have a generally good performance. The obsolete cultivars were taken out of commercial production because they may have suffered a set back (e.g., susceptible to disease) or higher performing cultivars were developed to replace them. If desirable parents are found in commercial cultivars, the breeder has a headstart on breeding since most of the gene combinations would already be desirable and adapted to the production environment.
- 2 **Breeding materials.** Ongoing or more established breeding programs maintain variability from previous projects. These intermediate breeding products are usually genetically narrow-based because they originate from a small number of genotypes or populations. For example, a breeder may release one genotype as a commercial cultivar after yield tests. Many of the genotypes that made it to the final stage or have

unique traits will be retained as breeding materials to be considered in future projects. Similarly, genotypes with unique combinations may be retained.

- 3 **Landraces.** Landraces are farmer-developed and maintained cultivars. They are developed over very long periods of time and have coadapted gene complexes. They are adapted to the growing region and are often highly heterogeneous. Landraces are robust, having developed resistance to the environmental stresses in their areas of adaptation. They are adapted to unfavorable conditions and produce low but relatively stable performance. Landraces, hence, characterize subsistence agriculture. They may be used as starting material in mass selection or pure-line breeding projects.

- 4 **Plant introductions.** The plant breeder may import new, unadapted genotypes from outside the production region, usually from another country (called plant introductions). These new materials may be evaluated and adapted to new production regions as new cultivars, or used as parents for crossing in breeding projects.

- 5 **Genetic stock.** This consists of products of specialized genetic manipulations by researchers (e.g., by using mutagenesis to generate various chromosomal and genomic mutants).

#### *Undomesticated plants*

When desired genes are not found in domesticated cultivars, plant breeders may seek them from wild populations. When wild plants are used in crosses, they may introduce wild traits that have an advantage for survival in the wild (e.g., hard seed coat, shattering, indeterminacy) but are undesirable in modern cultivation. These undesirable traits have been selected against through the process of domestication. Wild germplasms have been used as donors of several important disease- and insect-resistance genes and genes for adaptation to stressful environments. The cultivated tomato has benefited from such introgression by crossing with a variety of wild *Licopersicon* species. Other species such as potato, sunflower, and rice have benefited from wide crosses. In horticulture, various wild relatives of cultivated plants may be used as rootstock in grafting (e.g., citrus, grape) to allow cultivation of the plant in various adverse soil and climatic conditions.

#### *Other species and genera*

Gene transfer via crossing requires that the parents be cross-compatible or cross-fertile. As previously stated,

crossing involving parents from within a species is usually successful and unproblematic. However, as the parents become more genetically divergent, crossing (wide crosses) is less successful, often requiring special techniques (e.g., embryo rescue) for intervening in the process in order to obtain a viable plant. Sometimes, related species may be crossed with little difficulty.

### Concept of gene pools of cultivated crops

J. R. Harlan and J. M. J de Wet proposed a categorization of gene pools of cultivated crops according to the feasibility of gene transfer or gene flow from those species to the crop species. Three categories were defined, primary, secondary, and tertiary gene pools:

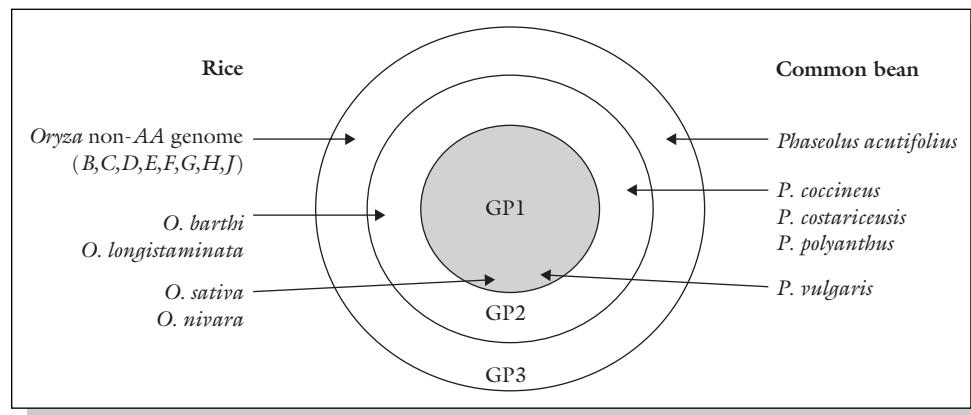
- 1 Primary gene pool (GP1).** GP1 consists of biological species that can be intercrossed easily (interfertile) without any problems with fertility of the progeny. That is, there is no restriction to gene exchange between members of the group. This group may contain both cultivated and wild progenitors of the species.
- 2 Secondary gene pool (GP2).** Members of this gene pool include both cultivated and wild relatives of the crop species. They are more distantly related and have crossability problems. Nonetheless, crossing produces hybrids and derivatives that are sufficiently fertile to allow gene flow. GP2 species can cross with those in GP1, with some fertility of the  $F_1$ , but more difficulty with success.
- 3 Tertiary gene pool (GP3).** GP3 involves the outer limits of potential genetic resources. Gene transfer by hybridization between GP1 and GP3 is very prob-

lematic, resulting in lethality, sterility, and other abnormalities. To exploit germplasm from distant relatives, tools such as embryo rescue and bridge crossing may be used to nurture an embryo from a wide cross to a full plant and to obtain fertile plants.

A classification of dry bean and rice is presented in Figure 6.1 for an illustration of this concept. In assembling germplasm for a plant breeding project, the general rule is to start by searching the domesticated germplasm collection first, before considering other sources, for reasons previously stated. However, there are times when the gene of interest occurs in undomesticated germplasm, or even outside the species. Gene-transfer techniques enable breeders to transfer genes beyond the tertiary gene pool. Whereas all crop plants have a primary gene pool that includes the cultivated forms, all crops do not have wild forms in their GP1 (e.g., broad bean, cassava, and onions whose wild types are yet to be identified). Also, occasionally, the GP1 may contain taxa of other crop plants (e.g., almond belongs to the primary gene pool of peach). Most crop plants have a GP2, which consists primarily of species of the same genus. Some crop plants have no secondary gene pools (e.g., barley, soybean, onion, broad bean).

### Concept of genetic vulnerability

Genetic vulnerability is an important issue in modern plant breeding, brought about largely by the manner in which breeders go about developing new and improved cultivars for modern society.



**Figure 6.1** Crop gene pools. Harlan proposed the crop gene pools to guide the germplasm use by plant breeders. The number of species in each of the pools that plant breeders use varies among crops. Harlan suggested that breeders first utilize the germplasm in GP1 and proceed outwards.

## What is genetic vulnerability?

Genetic vulnerability is a complex problem that involves issues such as crop evolution, trends in breeding, trends in biological technology, decisions by crop producers, demands and preferences of consumers, and other factors. As a result of a combination of the above factors, a certain kind of crop cultivar (genotype) is developed for the agricultural production system. **Genetic vulnerability** is a term used to indicate the genetic homogeneity and uniformity of a group of plants that predisposes it to susceptibility to a pest, pathogen, or environmental hazard of large-scale proportions. A case in point is the 1970 epidemic of southern leaf blight (*Helminthosporium maydis*) in the USA that devastated the corn industry. This genetic vulnerability in corn was attributed to uniformity in the genetic background in corn stemming from the widespread use of T-cytoplasm in corn hybrid seed production.

Genetic uniformity *per se* is not necessarily the culprit in vulnerability of crops. In fact, both producers and consumers sometimes desire and seek uniformity in some agronomic traits. The key issue is *commonality* of genetic factors. Genetically dissimilar crops can share a trait that is simply inherited and that predisposes them to susceptibility to an adverse biotic or abiotic factor. A case in point is the chestnut blight (*Cryphonectria parasitica*) epidemic that occurred in the USA in which different species of the plant were affected.

## Key factors in the susceptibility of crops

The key factors that are responsible for the disastrous epidemics attributable to genetic vulnerability of crops are:

- 1 The desire by breeders or consumers for uniformity in the trait that controls susceptibility to the biotic or abiotic environmental stress.
- 2 The acreage devoted to the crop cultivar and the method of production.

Where uniformity of the susceptible trait is high and cultivars are widely distributed in production (i.e., most farmers use the same cultivars), the risk of disaster will equally be high. Further, where the threat is biotic, the mode of dispersal of the causal agent and the presence of a favorable environment will increase the risk of disasters (e.g., wind mode of dispersal of spores or propagules will cause a rapid spread of the disease). In biotic disasters, the use of a single source of resistance to the pathogen is perhaps the single most important factor in vulnerability. However, the effect can be exacerbated by practices such as intensive and continuous monoculture using one cultivar. Under such production practices, the pathogen only has to overcome one genotype, resulting in rapid disease advance and greater damage to crop production.



## Industry highlights

### *Plant genetic resources for breeding*

K. Hammer, F. Heuser, K. Khoshbakht, and Y. Teklu

Institute of Crop Science, Agrobiodiversity Department, University Kassel, D-37213 Witzenhausen, Germany

#### **Introduction**

In recent years the maintenance of plant genetic resources (PGRs) has attracted growing public and scientific interest as well as political support since it is accepted that there is a close relationship between biological diversity and the health of the biosphere (Callow et al. 1997).

PGRs can be divided into two groups (Evenson et al. 1998). The first one consists of the cultivated plants themselves, their wild forms, and related species. The value of this group for improving crop plants is well recognized and expressed in the investment made by collections, evaluations, and conservation of these PGRs in past and present times. The second group is composed of PGRs from other, non-cultivated plants like weeds and even in species outside the plant kingdom. This group was neglected until the modern methods of biotechnology appeared, which allowed the integration of "alien" genetic material into valuable plant species.

While a great amount of the PGRs of the first group are conserved in gene banks (*ex situ*), there is an increasing requirement for conserving or better preserving the non-cultivated PGRs of the second group *in situ*, meaning to save biodiversity as a whole in the natural habitat. All kinds of plants might have useful properties (in the future) that would be worth maintaining. Frankel (1974) has already pointed out that there is "no doubt [that] primitive and wild gene pools will continue to serve as important sources of

genes for resistance to parasites or for characteristics indicated by advances in science or technology or by changing demands of the consumer". Alongside the rising importance of the PGRs for providing food there also exist traditional uses like medicine, feed, fiber, clothing, shelter, and energy.

Over one 5-year period, 6.5% of all genetic research within the plant breeding and seed industry, resulting in marketed innovation, was concerned with germplasm from wild species and landraces compared with only 2.2% of new germplasm originating from induced mutation (Swanson 1996; Callow et al. 1997).

Since the beginning of agriculture, natural diversity declined due to agricultural domestication, breeding, and distribution of crops (Becker 2000). But in recent years crop species and varieties have also become threatened or even extinct. In agriculture, the widespread adoption of a few varieties leads to a drastic decrease of landraces with their potential valuable genetic resources. Among the cultivated plants, which represent less than 3% of the vascular plants, only about 30 species feed the world (Hammer 2004).

### ***Conservation and monitoring of PGRs***

The monitoring and evaluation of plant material is necessary for the conservation of PGRs. There may be a big difference between the phenotype and the genotype in a population. With improved biotechnology methods, like the assignment of molecular markers, the gene level is of increasing interest.

#### ***Ex situ* conservation**

*Ex situ* conservation stands for all conservation methods in which the species or varieties are taken out of their natural habitat and are kept in surroundings made by humans. Large collections started with the activities of the Russian scientist N. I. Vavilov at the beginning of the last century. Even at that time the employment of *ex situ* measures was necessary because of the rapidly increasing gene erosion of landraces and other plants (Coats 1969). Alongside these collections, plant breeders contributed to maintenance by collecting breeding material. This material was often kept in specific institutions, the first-called "gene banks" in the 1970s. They were established for the collection (Guarino et al. 1995), maintenance, study, and supply of genetic resources of cultivated plants and related wild species. Gene banks maintain the plant material as seeds, *in vivo* (when the storage of seeds is difficult) or *in vitro* (mostly through cryoconservation). In contrast to the cultivation of plants in botanical gardens, the work in gene banks is more engaged in intraspecific variability. Unfortunately, a lot of the material stored in gene banks is not in good condition and urgently needs to be rejuvenated (Hammer 2004).

#### ***In situ* conservation**

As well as *ex situ* conservation, there is also the attempt to save biodiversity and therefore PGRs in ecosystems (*in situ*). This can occur in the natural habitat (especially wild relatives and forestry species) or in locations where the plants (landraces and weeds) have evolved (on farms, in agroecosystems). As opposed to *ex situ* conservation in gene banks where only a section of the whole diversity is covered, the *in situ* approach is able to save larger parts of biological diversity.

Table 1 summarizes the methods of conservation for the different categories of diversity and evaluates their relative importance. It is divided into cultivated plants, wild growing resources, and weeds.

### ***Characterization and evaluation of plant genetic resources***

The yield levels of many crops have reached a plateau due to the narrow genetic base of these crops. Although the results of some surveys (Brown 1983; Chang 1994) indicate that the genetic base of several important crops is beginning to increase in recent years, breeding programs of many important crops continue to include only a small part of the genetic diversity available, and the introduction of new and improved cultivars continues to replace indigenous varieties containing potentially useful germplasm. To widen the genetic base for further improvement, it is necessary to collect, characterize, evaluate, and conserve plant biodiversity, particularly in local, underutilized, and neglected crop plants.

Morphological and agronomic characteristics are often used for basic characterization, because this information is of high interest to users of the genetic diversity. Such characterization requires considerable amounts of human labor, organizational skills, and elaborate systems for data documentation although it can be done by using simple techniques and can reach a high sample throughput. Quantitative agronomic traits can be used to measure the differences between individuals and populations with regard to genetically complex issues such as yield potential and stress tolerance. The diversity of a population, considering such complex issues, can be described by using its mean value and genetic variance in statistical terms. The traits detected are of great interest, but are frequently subject to strong environmental influences, which makes their use as defining units for the measurement of genetic diversity problematic.

Molecular methods can be employed to characterize genetic resources and for the measurement of genetic variation. The major advantage of molecular methods for characterization is their direct investigation of the genotypic situation, which allows them to detect variation at the DNA level, thereby excluding all environmental influences. They can also be employed at very

**Table 1** Conservation methods for different categories of diversity rated by their importance for specific group of diversity (based on Hammer 2004).

Importance	Crops				Wild relatives				Weeds			
	1	2	3	4	1	2	3	4	1	2	3	4
<b><i>Ex situ</i></b>												
Infraspecific diversity			x			x				x		
Diversity of species		x				x				x		
Diversity of ecosystems	x				x				x			
<b><i>On farm</i></b>												
Developing countries					x				x			x
Infraspecific diversity					x				x			x
Diversity of species					x				x			x
Diversity of ecosystems					x				x			x
Developed countries									x		x	
Infraspecific diversity			x			x				x		
Diversity of species			x			x				x		
Diversity of ecosystems		x					x				x	
<b><i>In situ</i></b>												
Infraspecific diversity	x							x		x		
Diversity of species	x						x			x		
Diversity of ecosystems	x						x			x		

1, no importance; 2, low importance; 3, important; 4, very important.

**Table 2** Advantages and disadvantages of several methods of measuring genetic variation (FAO 1996).

Method	1	2	3	4	5	6	7
Morphology	Slight	High	Small number	Medium	Phenotypical characteristic	Qualitative/quantitative	Low
Pedigree analysis	Medium	–	–	Good	Degree of parent relationship	–	Low
Isoenzyme	Medium	Medium	Small number	Medium	Proteins	Codominant	Medium
RFLP (low copy)	Medium	Low	Small number (specific)	Good	DNA	Codominant	High
RFLP (high copy)	High	Low	High number (specific)	Good	DNA	Dominant	High
RAPD	High to medium	High	High number (random)	Slight	DNA	Dominant	Medium
DNA sequencing	High	Slight	Small number (specific)	Good	DNA	Codominant/dominant	High
Seq tag SSRs	High	High	Middle number (specific)	Good	DNA	Codominant	High
AFLPs	Medium to high	High	High number (random)	Medium	DNA	Dominant	High

1, variation found; 2, throughput of samples; 3, examined loci/assay; 4, reproducible on repetition; 5, analyzed character type; 6, inheritance of examined character trait; 7, necessary level of technology.

early growth stages. The advantages and disadvantages of some commonly used techniques for characterization of PGRs are summarized in Table 2 (FAO 1996; see also Hammer 2004, p. 127).

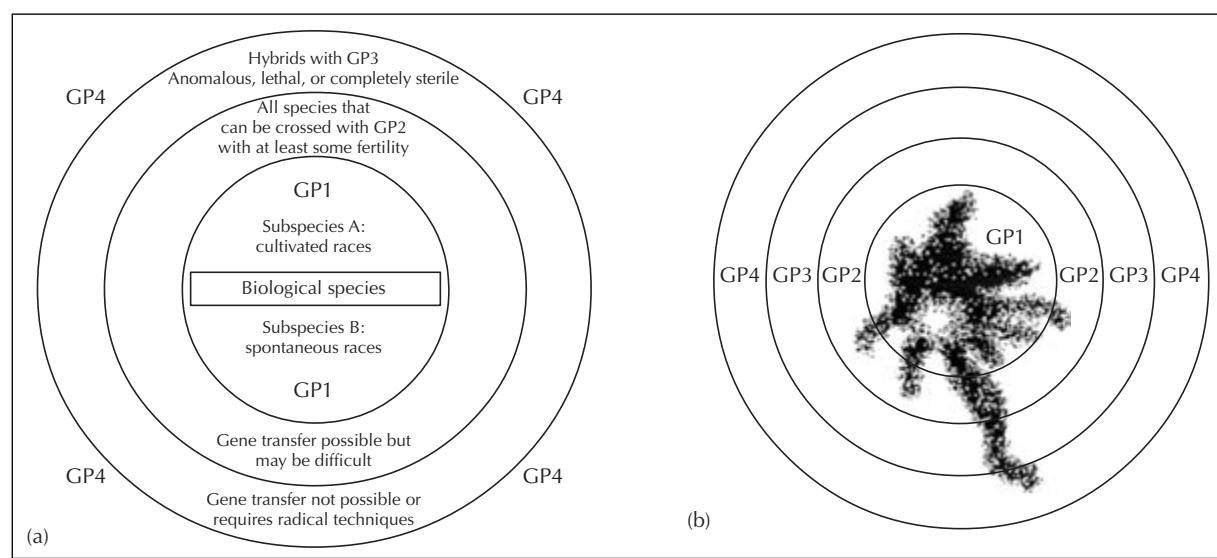
### **Germplasm enhancement**

PGRs are fundamental in improving agricultural productivity. These resources, fortunately stored in gene banks around the world, include an assortment of alleles needed for resistance and tolerance to the diseases, pests, and harsh environments found in their natural habitats. However, only a small amount of this variability has been introgressed to crop species (Ortiz 2002). Most cereal breeders do not make much use of the germplasm of landraces and wild and weedy relatives existing in active collections. The valuable genetic resources are essentially “sitting on the shelf” in what have been dismissively termed “gene morgues” (Hoisington et al. 1999). Germplasm enhancement may be one of the keys to maximizing the utilization of this germplasm. It has become an important tool for the genetic improvement of breeding populations by gene introgression or the incorporation of wild and landrace genetic resources into respective crop breeding pools. The term “germplasm enhancement” or “prebreeding” refers to the early component of sustainable plant breeding that deals with identifying a useful character, “capturing” its genetic diversity, and the transfer or introgression of these genes and gene combinations from non-adapted sources into breeding materials (Peloquin et al. 1989).

The gene pools as defined by Harlan and de Wet (1971) have formed a valid scientific basis for the definition and utilization of plant genetic resources (Figure 1). More recently, however, plant transformation and genomics have led to a new quality that has been defined by Gepts and Papa (2003) as a fourth gene pool, whereas Gladis and Hammer (2002) earlier concluded that information and genes from other species belong to the third gene pool. The fourth gene pool should contain any synthetic strains with nucleic acid frequencies (DNA or RNA) that do *not* occur in nature.

The most widespread application of germplasm enhancement has been in resistance breeding with genetic resources of wild species. Backcross followed by selection has been the most common method for gene introgression from wild germplasm to breeding materials.

However, some problems still remain for genetic enhancement with wild species: linkage drag, sterility, the small sample size of interspecific hybrid population, and restricted genetic recombination in the hybrid germplasm (Ortiz 2002). Transgenesis allows us to bypass sexual incompatibility barriers altogether and introduce new genes into existing cultivars. In recent years,



**Figure 1** (a) The modified gene pool concept, established by Harlan and de Wet (1971). GP1: the biological species, including wild, weedy, and cultivated races. GP2: all species that can be crossed with GP1, with some fertility in individuals of the F<sub>1</sub> generation; gene transfer is possibly but may be difficult. GP3: hybrids with GP1 do not occur in nature; they are anomalous, lethal, or completely sterile; gene transfer is not possible without applying radical techniques. Information from other genes refers to comparative genomic information on gene order and DNA sequence of homologous genes. GP4: any synthetic strains with nucleic acid frequencies (DNA or RNA) that do not occur in nature. (b) Example of an organismoid or hypothetically designed crop with a genome composed of different gene pools and synthetic genes (for further explanation, see Gladis and Hammer (2002)).

**Table 3** Varieties registered from 1973 to 1990 that proved to have been developed with material from the genebank Gatersleben (Hammer 1991).

Crop	Number of varieties
Spring barley	30
Winter barley	3
Spring wheat	1
Winter wheat	12
Dry soup pea	2
Fodder pea	3
Lettuce	1
Vegetable pea	4
Total	56

Developing improved varieties using gene bank materials takes a long time. For instance, when developing disease-resistant material, the resistance must be located with great expenditure of time and effort, from extensive collections. The experience in Gatersleben indicates that it takes roughly 20 years between the first discovery of the material and the launching of a new variety, even if modern breeding methods are employed (Hammer 2004). A positive correlation has been observed between the number of evaluated accessions in gene banks and the number of released varieties on the basis of evaluated material (Hammer et al. 1994).

The use of Turkish wheat to develop genetic resistance to diseases in Western wheat crops was valued in 1995 at US\$50 million per year. Ethiopian barley has been used to protect Californian barley from dwarf yellow virus, saving damage estimated at \$160 million per year. Mexican beans have been used to improve resistance to the Mexican bean weevil, which destroys as much as 25% of stored beans in Africa and 15% in South America (Perrings 1998).

### Conclusion

PGRs are useful for present and future agriculture and horticulture production. They are particularly needed for the genetic improvement of crop plants. Because of their usefulness and their ongoing erosion in the agroecosystems, it was necessary to establish large collections of PGRs. The material in these collections has to be characterized and evaluated in order to introduce it into breeding programs. Prebreeding and germplasm enhancement are necessary as first steps for the introduction of primitive material into modern varieties.

### References

- Becker, H.C. 2000. Einfluß der Pflanzenzüchtung auf die genetische Vielfalt. *Schriftenr. Vegetationskunde* 32:87–94.
- Brown, W.L. 1983. Genetic diversity and genetic vulnerability – An appraisal. *Econ. Bot.* 37:4–12.
- Callow, J.A., B.V. Ford-Lloyd, and H.J. Newbury (eds). 1997. Biotechnology and plant genetic resources – Conservation and use. *Biotechnology in Agriculture No. 19*. CAB International Publishing, New York.
- Chang, T.T. 1994. The biodiversity crisis in Asia crop production and remedial measures. In: *Biodiversity and terrestrial ecosystems* (C.I. Peng, and C.H. Chou, eds), pp. 25–41. Monograph Series No. 14. Institute of Botany, Academia Sinica, Taipei, Taiwan.
- Coats, A. 1969. The quest for plants: A history of the horticultural explorers. Studio Vista, London.
- Evenson, R.E., D. Gollin, and V. Santaniello (eds). 1998. Agricultural values of plant genetic resources. CAB International Publishing, Wallingford, UK.
- FAO. 1996. Report on the state of the world's plant genetic resources for food and agriculture. Food and Agricultural Organization, Rome, 75 pp.
- Frankel, O.H. 1974. Genetic conservation: Our evolutionary responsibility. *Genetics* 78:53–65.
- Gepts, P., and R. Papa. 2003. Possible effects of (trans)gene flow from crops on the genetic diversity from landraces and wild relatives. *Environ. Biosafety Res.* 2:89–103.
- Gladis, T., and K. Hammer. 2002. The relevance of plant genetic resources in plant breeding. *FAL Agriculture Research*, Special Issue 228:3–13.
- Guarino, L., V.R. Ramanathra-Rao, and R. Reid (eds). 1995. Collecting plant genetic diversity: Technical guidelines. CAB International, Wallingford, UK.

transgenic plants have been incorporated as parents of hybrids in US breeding programs for crops such as maize and oilseed rape. Molecular markers are being used to tag specific chromosome segments bearing the desired gene(s) to be transferred (or incorporated) into the breeding lines (or populations).

### Examples of successful uses of plant genetic resources

Over the last few decades, awareness of the rich diversity of exotic or wild germplasm has increased. This has led to a more intensive use of this germplasm in breeding (Kearsey 1997) and thereby yields of many crops have increased dramatically. The introgression of genes that reduce plant height and increase disease and viral resistance in wheat provided the foundation for the Green Revolution and demonstrated the tremendous impact that genetic resources can have on production (Hoisington et al. 1999).

In Germany, PGR material stored in the Gatersleben gene bank has been successfully used for the development of improved varieties (Table 3).

- Hammer, K. 1991. Die Nutzung des Materials der Gaterslebener Genbank für die Resistenzzüchtung – eine Übersicht. Vortr. Pflanzenzüchtung 19:197–206.
- Hammer, K. 2004. Resolving the challenge posed by agrobiodiversity and plant genetic resources – An attempt. Programme des Deutschen Instituts für Tropische und Subtropische Landwirtschaft (DITSL). J. Agric. Rural Dev. Tropics Subtropics 76, 184 pp.
- Hammer, K., H. Gäde, and H. Knüpffer. 1994. 50 Jahre Genbank Gatersleben – eine Übersicht. Vortr. Pflanzenzüchtung 27:333–383.
- Harlan, J.R., and J.M.J. de Wet. 1971. Towards a rational classification of cultivated plants. Taxon 20:509–517.
- Hoisington, D., M. Khairallah, T. Reeves, T.M. Ribaut, B. Skovmand, S. Taba, and M.L. Warburton. 1999. Plant genetic resources: What can they contribute toward increased crop productivity? Proc. Natl. Acad. Sci. USA 96:5937–5943.
- Kearsey, M.J. 1997. Genetic resources and plant breeding (identification, mapping and manipulation of simple and complex traits). In: Biotechnology and plant genetic resources – Conservation and use (J.A. Callow, B.V. Ford-Lloyd, and H.N. Newberry, eds), pp. 175–202. CAB International, Wallingford, UK.
- Ortiz, R. 2002. Germplasm enhancement to sustain genetic gains in crop improvement. In: Managing plant genetic diversity (J.M.M. Engels, V.R. Ramanatha, A.H.D. Brow, and M. Jackson, eds), pp. 275–290. CABI Publishing, Wallingford, UK.
- Peloquin, S.J., G.L. Yerk, J.E. Werner, and E. Darmo. 1989. Potato breeding with haploids and  $2n$  gametes. Genome 31:1000–1004.
- Perrings, C. 1998. The economics of biodiversity loss and agricultural development in low income countries. Paper presented at the American Association of Agricultural Economics International Conference on Agricultural Intensification, Economic Development, and the Environment, Salt Lake City, USA, July 31–August 1.
- Swanson, T. 1996. Global values of biological diversity: The public interest in the conservation of plant genetic resources for agriculture. Plant Genet. Res. News. 105:1.

## What plant breeders can do to address crop vulnerability

As previously indicated, the issue of genetic vulnerability is very complex. However, ultimately, plant breeders are the experts who can effectively address this issue.

### Reality check

First and foremost, plant breeders need to be convinced that genetic vulnerability is a real and present danger. Without this first step, efforts to address the issue are not likely to be taken seriously. A study by D. N. Duvick in 1984, albeit dated, posed the question “How serious is the problem of genetic vulnerability in your crop?” to plant breeders. The responses by breeders of selected crops (cotton, soybean, wheat, sorghum, maize, and others) indicated a wide range of perception of crop vulnerability, ranging from 0–25% thinking it was serious to 25–60% thinking it was not a serious problem (at least at that time). Soybean and wheat breeders expressed the most concern about genetic vulnerability. Their fears are most certainly founded since, in soybean, it is estimated that only six cultivars constitute more than 50% of the genetic base of North American germplasm. Similarly, more than 50% of the acreage of many crops in the USA is planted to less than 10 cultivars per crop.

### Use of wild germplasm

Many of the world’s major crops are grown extensively outside their centers of origin where they coevolved with pests and pathogens. Breeders should make deliberate efforts to expand the genetic base of their crops by exploiting genes from the wild progenitors of their species that are available in various germplasm repositories all over the world.

### Paradigm shift

As D. Tanksley and S. R. McCouch of Cornell University point out, there is a need for a paradigm shift regarding the use of germplasm resources. Traditionally, breeders screen accessions from exotic germplasm banks on a phenotypic basis for clearly defined and recognizable features of interest. Desirable genotypes are crossed with elite cultivars to introgress genes of interest. However, this approach is effective only for the utilization of simply inherited traits (conditioned by single dominant genes). The researchers proposed a shift from the old paradigm of looking for phenotypes to a new paradigm of looking for genes. To accomplish this, the modern techniques of genomics may be used to screen exotic germplasm using a gene-based approach. They propose the use of molecular

linkage maps and a new breeding technique called **advanced backcross QTL** (quantitative trait loci) that allows the breeder to examine a subset of alleles from the wild exotic plant in the genetic background of an elite cultivar.

#### *Use of biotechnology to create new variability*

The tools of modern biotechnology, such as rDNA, cell fusion, somaclonal variation, and others, may be used to create new variability for use in plant breeding. Genetic engineering technologies may be used to transfer desirable genes across natural biological barriers.

#### *Gene pyramiding*

Plant breeders may broaden the diversity of resistance genes as well as introduce multiple genes from different sources into cultivars using the technique of **gene pyramiding**, which allows the breeder to insert more than one resistance gene into one genotype. This approach will reduce the uniformity factor in crop vulnerability.

### **Conservation of plant genetic resources**

Plant breeders manipulate variability in various ways – for example, they assemble, recombine, select, and discard. The preferential use of certain elite genetic stock in breeding programs has narrowed the overall genetic base of modern cultivars. As already noted, pedigree analysis indicates that many cultivars of certain major crops of world importance have common ancestry, making the industry vulnerable to disasters (e.g., disease epidemics, climate changes). National and international efforts have been mobilized to conserve plant genetic resources.

#### **Why conserve plant genetic resources?**

There are several reasons why plant genetic resources should be conserved:

- 1 Plant germplasm is exploited for food, fiber, feed, fuel, and medicines by agriculture, industry, and forestry.
- 2 As a natural resource, germplasm is a depletable resource.
- 3 Without genetic diversity, plant breeding cannot be conducted.
- 4 Genetic diversity determines the boundaries of crop productivity and survival.

- 5 As previously indicated, variability is the life blood of plant breeding. As society evolves, its needs will keep changing. Similarly, new environmental challenges might arise (e.g., new diseases, abiotic stresses) for which new variability might be needed for plant improvement.

When a genotype is unable to respond fully to the cultural environment, as well as to resist unfavorable conditions thereof, crop productivity diminishes. The natural pools of plant genetic resources are under attack from the activities of modern society – urbanization, indiscriminate burning, and the clearing of virgin land for farming, to name a few. These and other activities erode genetic diversity in wild populations. Consequently, there is an urgent need to collect and maintain samples of natural variability. The actions of plant breeders also contribute to genetic erosion as previously indicated. High-yielding, narrow genetic-based cultivars are penetrating crop production systems all over the world, displacing the indigenous high-variability landrace cultivars. Some 20,000 species are listed as endangered species.

#### **Genetic erosion**

**Genetic erosion** may be defined as the decline in genetic variation in cultivated or natural populations largely through the action of humans. Loss of genetic variation may be caused by natural factors, and by the actions of crop producers, plant breeders, curators of germplasm repositories, and others in society at large.

#### *Natural factors*

Genetic diversity can be lost through natural disasters such as large-scale floods, wild fires, and severe and prolonged drought. These events are beyond the control of humans.

#### *Action of farmers*

Right from the beginnings of agriculture, farmers have engaged in activities that promote genetic erosion. These include clearing of virgin land in, especially, germplasm-rich tropical forests, and the choice of planting material (narrow genetic-based cultivars). Farmers, especially in developed economies, primarily grow improved seed, having replaced most or all landraces with these superior cultivars. Also, monoculture tends to narrow genetic diversity as large tracts of land are planted to uniform cultivars. Extending grazing lands

into wild habitats by livestock farmers, destroys wild species and wild germplasm resources.

#### Action of breeders

Farmers plant what breeders develop. Some methods used for breeding (e.g., pure lines, single cross, multi-lines) promote uniformity and a narrower genetic base. When breeders find superior germplasm, the tendency is to use it as much as possible in cultivar development. In soybean, as previously indicated, most of the modern cultivars in the USA can be traced back to about half a dozen parents. This practice causes severe reduction in genetic diversity.

#### Problems with germplasm conservation

In spite of good efforts by curators of germplasm repositories to collect and conserve diversity, there are several ways in which diversity in their custody may be lost. The most obvious loss of diversity is attributed to human errors in the maintenance process (e.g., improper storage of materials leading to loss of variability). Also, when germplasm is planted in the field, natural selection pressure may eliminate some unadapted genotypes. Also, there could be spontaneous mutations that can alter the variability in natural populations. Hybridization as well as genetic drift incidences in small populations are also consequences of periodic multiplication of the germplasm holdings by curators.

#### General public action

As previously indicated, there is an increasing demand on land with increasing populations. Such demands include settlement of new lands, and the demand for alternative use of the land (e.g., for recreation, industry, roads) to meet the general needs of modern society. These actions tend to place wild germplasm in jeopardy. Such undertakings often entail clearing of virgin land where wild species occur.

#### Selected impact of germplasm acquisition

##### *Impact on North American agriculture*

Very few crops have their origin in North America. It goes without saying that North American agriculture owes its tremendous success to plant introductions, which brought major crops such as wheat, barley, soybean, rice, sugar cane, alfalfa, corn, potato, tobacco,

and cotton to this part of the world. North America currently is the world's leading producer of many of these crops. Spectacular contributions by crop introductions to US agriculture include the following (see also Plant introductions, p. 103):

- 1 Avocado: introduced in 1898 from Mexico, this crop has created a viable industry in California.
- 2 Rice: varieties introduced from Japan in 1900 laid the foundation for the present rice industry in Louisiana and Texas.
- 3 Spinach: a variety introduced from Manchuria in 1900 is credited with saving the Virginia spinach industry from blight disaster in 1920.
- 4 Peach: many US peach orchards are established by plants growing on root stalks obtained from collections in 1920.
- 5 Oats: one of the world's most disease-resistant oat varieties was developed from germplasm imported from Israel in the 1960s.

#### *Other parts of the world*

A few examples include dwarf wheat introduced into India, Pakistan, and the Philippines as part of the Green Revolution, and soybean and sunflower into India; these have benefited the agriculture of these countries.

#### **Nature of cultivated plant genetic resources**

Currently five kinds of cultivated plant materials are conserved by concerted worldwide efforts – **landraces** (**folk or primitive varieties**), **obsolete varieties**, **commercial varieties (cultivars)**, **plant breeders' lines**, and **genetic stocks**. Landraces are developed by indigenous farmers in various traditional agricultural systems or are products of nature. They are usually very variable in composition. Obsolete cultivars may be described as "ex-service" cultivars because they are no longer used for cultivation. Commercial cultivars are elite germplasm currently in use for crop production. These cultivars remain in production usually from 5 to 10 years before becoming obsolete and replaced. Breeders' lines may include parents that are inbred for hybrid breeding, genotypes from advanced yield tests that were not released as commercial cultivars, and unique mutants. Genetic stocks are genetically characterized lines of various species. These are advanced genetic materials developed by breeders, and are very useful and readily accessible to other breeders.

## Approaches to germplasm conservation

There are two basic approaches to germplasm conservation – *in situ* and *ex situ*. These are best considered as complementary rather than independent systems.

### *In situ* conservation

This is the preservation of variability in its natural habitat in its natural state (i.e., on site). It is most applicable to conserving wild plants and entails the use of legal measures to protect the ecosystem from encroachment by humans. These protected areas are called by various names (e.g., nature reserves, wildlife refuges, natural parks). Needless to say, there are various socioeconomic and political ramifications in such legal actions by governments. Environmentalists and commercial developers often clash on such restricted use or prohibited use of natural resources. This approach to germplasm conservation is indiscriminatory with respect to species conserved (i.e., all species in the affected area are conserved).

### *Ex situ* conservation

In contrast to *in situ* conservation, *ex situ* conservation entails planned conservation of targeted species (not all species). Germplasm is conserved not in the natural places of origin but under supervision of professionals off site in locations called germplasm or gene banks. Plant materials may be in the form of seed or vegetative materials. The advantage of this approach is that small samples of the selected species are stored in a small space indoors or in a field outdoors, and under intensive management that facilitates their access to breeders. However, the approach is prone to some genetic erosion (as previously indicated) while the evolutionary process is halted. The special care needed is expensive to provide. Other aspects of this approach are discussed later in this chapter.

## Germplasm collection

Planned collections (germplasm explorations or expeditions) are conducted by experts to regions of plant origin. These trips are often multidisciplinary, comprising members with expertise in botany, ecology, pathology, population genetics, and plant breeding. Familiarity with the species of interest and the culture of the regions to be explored are advantageous. Most of the materials

collected are seeds, even though whole plants and vegetative parts (e.g., bulbs, tubers, cuttings, etc.) and even pollen may be collected. Because only a small amount of material is collected, sampling for representativeness of the population's natural variability is critical in the collection process, in order to obtain the maximum possible amount of genetic diversity. For some species whose seed is prone to rapid deterioration, or are bulky to transport, *in vitro* techniques may be available to extract small samples from the parent source. Collectors should bear in mind that the value of the germplasm may not be immediately discernible. Materials should not be avoided for lack of obvious agronomically desirable properties. It takes time to discover the full potential of germplasm.

Seed materials vary in viability characteristics. These have to be taken into account during germplasm collection, transportation, and maintenance in repositories. Based on viability, seed may be classified into two main groups – **orthodox** and **recalcitrant** seed:

- 1 **Orthodox seeds.** These are seeds that can prolong their viability under reduced moisture content and low temperature in storage. Examples include cereals, pulses, and oil seed. Of these, some have superior (e.g., okra) while others have poor (e.g., soybean) viability under reduced moisture cold storage.
- 2 **Recalcitrant seeds.** Low temperature and decreased moisture content are intolerable to these seeds (e.g., coconut, coffee, cocoa). *In vitro* techniques might be beneficial to these species for long-term maintenance.

The conditions of storage differ depending on the mode of reproduction of the species:

- 1 **Seed propagated species.** These seeds are first dried to about 5% moisture content and then usually placed in hermetically sealed moisture-proof containers before storage.
- 2 **Vegetatively propagated species.** These materials may be maintained as full plants for long periods of time in field gene banks, nature reserves, or botanical gardens. Alternatively, cuttings and other vegetative parts may be conserved for a short period of time under moderately low temperature and humidity. For long-term storage, *in vitro* technology is used.

## Types of plant germplasm collections

There are four types of plant genetic resources maintained by germplasm repositories – **base collections**,

**backup collections, active collections, and breeders' or working collections.** These categorizations are only approximate since one group can fulfill multiple functions.

#### *Base collections*

These collections are not intended for distribution to researchers, but are maintained in long-term storage systems. They are the most comprehensive collections of the genetic variability of species. Entries are maintained in the original form. Storage conditions are low humidity at subfreezing temperatures ( $-10$  to  $-18^{\circ}\text{C}$ ) or cryogenic ( $-150$  to  $-196^{\circ}\text{C}$ ), depending on the species. Materials may be stored for many decades under proper conditions.

#### *Backup collections*

The purpose of backup collections is to supplement the base selection. In case of a disaster at a center responsible for a base collection, a duplicate collection is available as insurance. In the USA, the National Seed Storage Laboratory at Fort Collins, Colorado, is a backup collection center for portions of the accessions of the Centro Internationale de Mejoramiento de Maiz y Trigo (CIMMYT) and the International Rice Research Institute (IRRI).

#### *Active collections*

Base and backup collections of germplasm are designed for long-term unperturbed storage. Active collections usually comprise the same materials as in base collections, however, the materials in active collections are available for distribution to plant breeders or other patrons upon request. They are stored at  $0^{\circ}\text{C}$  and about 8% moisture content, and remain viable for about 10–15 years. To meet this obligation, curators of active collections at germplasm banks must increase the amount of germplasm available to fill requests expeditiously. Because the accessions are more frequently increased through field multiplication, the genetic integrity of the accession may be jeopardized.

#### *Working or breeders' collections*

Breeders' collections are primarily composed of elite germplasm that is adapted. They also include enhanced breeding stocks with unique alleles for introgression into these adapted materials. In these times of genetic

engineering, breeders' collections include products of rDNA research that can be used as parents in breeding programs.

### **Managing plant genetic resources**

The key activities of curators of germplasm banks include regeneration of accessions, characterization, evaluation, monitoring seed viability and genetic integrity during storage, and maintaining redundancy among collections. Germplasm banks receive new materials on a regular basis. These materials must be properly managed so as to encourage and facilitate their use by plant breeders and other researchers.

#### *Regeneration*

Germplasm needs to be periodically rejuvenated and multiplied. The regeneration of seed depends on the life cycle and breeding system of the species as well as cost of the activity. To keep costs to a minimum and to reduce loss of genetic integrity, it is best to keep regeneration and multiplication to a bare minimum. It is a good strategy to make the first multiplication extensive so that ample original seed is available for depositing in the base and duplicate or active collections. The methods of regeneration vary for self-pollinated, cross-pollinated, and apomictic species. A major threat to genetic integrity of accessions during regeneration is contamination (from outcrossing or accidental migration), which can change the genetic structure. Other factors include differential survival of alleles or genotypes within the accession, and random drift. The isolation of accessions during regeneration is critical, especially in cross-pollinated species, to maintaining genetic integrity. This is achieved through proper spacing, caging, covering with bags, hand pollination, and other techniques. Regeneration of wild species is problematic because of high seed dormancy, seed shattering, high variability in flowering time, and low seed production. Some species have special environmental requirements (e.g., photoperiod, vernalization) and hence it is best to rejuvenate plants under conditions similar to those in the places of their origin, to prevent selection effect, which can eliminate certain alleles.

#### *Characterization*

Users of germplasm need some basic information about the plant materials to aid them in effectively using these

resources. Curators of germplasm banks characterize their accessions, an activity that entails a systematic recording of selected traits of an accession. Traditionally, these data are limited to highly heritable morphological and agronomic traits. However, with the availability of molecular techniques, some germplasm banks have embarked upon molecular characterization of their holdings. For example, CIMMYT has used the simple sequence repeat (SSR) marker system for characterizing the maize germplasm in their holding. **Passport data** are included in germplasm characterization. These data include an accession number, scientific name, collection site (country, village), source (wild, market), geography of the location, and any disease and insect pests. To facilitate data entry and retrieval, characterization includes the use of **descriptors**. These are specific pieces of information on plant or geographic factors that pertain to the plant collection. The International Plant Genetic Resources Institute (IPGRI) has prescribed guidelines for the categories of these descriptors. Descriptors have been standardized for some species such as rice.

#### *Evaluation*

Genetic diversity is not usable without proper evaluation. Preliminary evaluation consists of readily observable traits. Full evaluations are more involved and may include obtaining data on cytogenetics, evolution, physiology, and agronomy. More detailed evaluation is often done outside of the domain of the germplasm bank by various breeders and researchers using the specific plants. Traits such as disease resistance, productivity, and quality of product are important pieces of information for plant breeders. Without some basic information of the value of the accession, users will not be able to make proper requests and receive the most useful materials for their work.

#### *Monitoring seed viability and genetic integrity*

During storage, vigor tests should be conducted at appropriate intervals to ensure that seed viability remains high. During these tests, abnormal seedlings may indicate the presence of mutations.

#### *Exchange*

The ultimate goal of germplasm collection, rejuvenation, characterization, and evaluation is to make available and facilitate the use of germplasm. There are various

computer-based genetic-resource documentation systems worldwide, some of which are crop-specific. These systems allow breeders to rapidly search and request germplasm information. There are various laws regarding, especially, international exchange of germplasm. Apart from quarantine laws, various inspections and testing facilities are needed at the point of germplasm.

### **Issue of redundancy and the concept of core subsets**

Collections for major crops such as wheat and corn can be very large. Some of these accessions are bound to be duplicates. Because of the cost of germplasm maintenance, it is important for the process to be efficient and effective. Redundancy should be minimized in the collections. However, eliminating duplicates may be as expensive as maintaining them. To facilitate the management of huge accessions, the concept of **core subsets** was proposed. A core subset comprises a sample of the base collection of a germplasm bank that represents the genetic diversity in the crop and its relatives, with minimum redundancy. The core would be well characterized and evaluated for ready access by users. However, some argue that maintaining a core subset might distract from maintaining the balance of the collection, leading possibly to loss of some accessions.

### **Germplasm storage technologies**

Once collected, germplasm is maintained in the most appropriate form by the gene bank with storage responsibilities for the materials. Plant germplasm may be stored in the form of pollen, seed, or plant tissue. Woody ornamental species may be maintained as living plants, as occurs in arboreta. Indoor maintenance is done under cold storage conditions, with temperatures ranging from  $-18$  to  $-196^{\circ}\text{C}$ .

#### *Seed storage*

Seeds are dried to the appropriate moisture content before being placing in seed envelopes. These envelopes are then arranged in trays that are placed on shelves in the storage room. The storage room is maintained at  $-18^{\circ}\text{C}$ , a temperature that will keep most seeds viable for up to 20 years or more. The curator of the laboratory and the staff periodically sample seeds of each accession to conduct a germination test. When germination falls

below a certain predetermined level, the accession is regrown to obtain fresh seed.

#### *Field growing*

Accessions are regrown to obtain fresh seed or to increase existing supplies (after filling orders by scientists and other clients). To keep the genetic purity, the accessions are grown in isolation, each plant covered with a cotton bag to keep foreign sources of pollen out and also to ensure self-pollination.

#### *Cryopreservation*

**Cryopreservation** or freeze-preservation is the storage of materials at extremely low temperatures of between  $-150$  to  $-196^{\circ}\text{C}$  in liquid nitrogen. Plant cells, tissue, or other vegetative material may be stored this way for a long time without loosing regenerative capacity. Whereas seed may also be stored by this method, cryopreservation is reserved especially for vegetatively propagated species that need to be maintained as living plants. Shoot tip cultures are obtained from the material to be stored and protected by dipping in a cryoprotectant (e.g., a mixture of sugar and polyethylene glycol plus dimethylsulfoxide).

#### *In vitro storage*

Germplasm of vegetatively propagated crops is normally stored and distributed to users in vegetative forms such as tubers, corms, rhizomes, and cuttings. However, it is laborious and expensive to maintain plants in these forms. *In vitro* germplasm storage usually involves tissue culture. There are several types of tissue culture systems (suspension cells, callus, meristematic tissues). To use suspension cells and callus materials, there must be an established system of regeneration of full plants from these systems, something that is not available for all plant species yet. Consequently, meristem cultures are favored for *in vitro* storage because they are more stable. The tissue culture material may be stored using the method of slow growth (chemicals are applied to retard the culture temperature) or cryopreservation.

#### *Molecular conservation*

The advent of biotechnology has made it possible for researchers to sequence DNA of organisms. These sequences can be searched (see Bioinformatics in Chapter 14, p. 238) for genes at the molecular level.

Specific genes can be isolated by cloning and used in developing transgenic products.

## **Using genetic resources**

#### **Perceptions and challenges**

Breeders, to varying degrees, acknowledge the need to address the genetic vulnerability of their crops. Further, they acknowledge the presence of large amounts of genetic variation in wild crop relatives. However, much of this variability is not useful to modern plant breeding. In using wild germplasm, there is a challenge to sort out and detect those germplasms that are useful to breeders. Modern cultivars have resulted from years of accumulation of favorable alleles that have been gradually assembled into adapted interacting multilocus combinations. Introgression of unadapted genes may jeopardize these combinations through segregation and recombination. Hence, some breeders are less inclined to use unadapted germplasm. However, there are occasions when the breeder has little choice but to take the risk of using unadapted germplasm (e.g., specific improvement of traits such as new races of disease, quality issues), because alleles for addressing these problems may be non-existent in the adapted materials. Plant breeders engaged in the breeding of plant species that have little or no history of improvement are among the major users of active collections in germplasm banks. For such breeders, they may have no alternative but to evaluate primitive materials to identify those with promise for use as parents in breeding.

Plant breeders may use germplasm collections in one of two basic ways: (i) as sources of cultivars; or (ii) as sources of specific genes. A breeding collection contains alleles for specific traits that breeders can transfer into adapted genotypes using appropriate breeding methods. Accessions must be properly documented to facilitate the search by users. This means, there should be accurate passport and descriptor information for all accessions. Unfortunately, this is not the case for many accessions.

The redundancy in germplasm banks is viewed by some breeders as unacceptable. A study showed that of the 250,000 accessions of barley at that time in repositories, only about 50,000 were unique. Such discrepancy leads to false estimation of the true extent of diversity in the world collection. A large number of the accessions are also obsolete and have little use to modern plant breeding programs. Germplasm evaluation at the level of germplasm banks is very limited, making it more

difficult for users to identify accessions with promise for breeding.

### **Concept of prebreeding**

Plant breeders usually make elite × elite crosses in a breeding program. This practice coupled with the fact that modern crop production is restricted to the use of highly favored cultivars, has reduced crop genetic diversity and predisposed crop plants to disease and pest epidemics. To reverse this trend, plant breeders need to make deliberate efforts to diversify the gene pools of their crops to reduce genetic vulnerability. Furthermore, there are occasions when breeders are compelled to look beyond the advanced germplasm pool to find desirable genes. The desired genes may reside in unadapted gene pools. As previously discussed, breeders are frequently reluctant to use such materials because the desired genes are often associated with undesirable effects (unadapted, unreproductive, yield-reducing factors). Hence, these exotic materials often cannot be used directly in cultivar development. Instead, the materials are gradually introduced into the cultivar development program through crossing and selecting for intermediates with new traits, while maintaining a great amount of the adapted traits.

To use wild germplasm, the unadapted material is put through a preliminary breeding program to transfer the desirable genes into adapted genetic backgrounds. The process of the initial introgression of a trait from an undomesticated source (wild) or agronomically inferior source, to a domesticated or adapted genotype is called **prebreeding** or **germplasm enhancement**. The process varies in complexity and duration, depending on the source, the type of trait, and presence of reproductive barriers. It may be argued that prebreeding is not an entirely new undertaking, considering the fact that all modern crops were domesticated through this process. The difference between then and now, as D. N. Duvick pointed out, is one of demarcation between gene pools. In the beginning of agriculture, there were no discernible differences between highly domesticated and highly selected elite cultivars being deliberately infused with genes from highly undomesticated germplasm. In other words, the early farmer-selectors did what came naturally, discriminating among natural variation without deliberately hybridizing genotypes, and gradually moving them from the wild to adapted domesticated domain.

The traditional techniques used are hybridization followed by backcrossing to the elite parent, or the use of

cyclical population improvement techniques. The issues associated with wide crossing are applicable (e.g., infertility, negative linkage drag, incompatibility), requiring techniques such as embryo rescue to be successful. The modern tools of molecular genetics and other biotechnological procedures are enabling radical gene transfer to be made into elite lines without linkage drag (e.g., transfer of genes from bacteria into plants; see Chapter 14). This new approach to the development of new breeding materials is more attractive and profitable to private investors (for-profit breeding programs). Such creations can be readily protected by patents for commercial exploitation. Further, these technologies are enabling plant breeders not only to develop new and improved highly productive cultivars but also to assign new roles to cultivars (e.g., plants can now be used as bioreactors for producing novel traits such as specialized oils, proteins, and pharmaceuticals).

The major uses of germplasm enhancement may be summarized as follows:

- 1 Preventions of genetic uniformity and the consequences of genetic vulnerability.
- 2 Potential crop yield augmentation. History teaches us that some of the dramatic yield increases in major world food crops, such as rice, wheat, and sorghum, were accomplished through introgression of unadapted genes (e.g., dwarf genes).
- 3 Introduction of new quality traits (e.g., starch, protein).
- 4 Introduction of disease- and insect-resistance genes.
- 5 Introduction of environment-resistance genes (e.g., drought resistance).

Prebreeding can be expensive to conduct and time-consuming as well. With the exception of high value crops, most prebreeding is conducted in the public sector. The Plant Variety Protection Act (see Chapter 15) does not provide adequate financial incentive for for-profit (commercial) breeders to invest resources in germplasm enhancement.

### **Plant explorations and introductions and their impact on agriculture**

#### **Plant explorations**

Plant exploration is an international activity. Recent political developments are making germplasm collections less of an open access activity. Explorers must obtain permission to enter a country to collect plant material. Most of these germplasm-rich places are

located in developing countries, which frequently complain about not reaping adequate benefits from contributing germplasm to plant breeding. Consequently, these nations are increasingly prohibiting free access to their natural resources.

#### *US historical perspectives*

The US Department of Agriculture (USDA) plant germplasm collection efforts began in 1898 under the leadership of David Fairchild. Fairchild collected pima cotton, pistachio, olive, walnut, and other crop materials. Other notable personnel in the plant exploration efforts by the USA include the following: S. A. Knapp, whose rice collection from Japan is credited with making the USA a rice-exporting country; tropical fruits collected by W. Popenoe from South and Central America, also created new industries in the US; and F. N. Meyer made outstanding collections between 1905 and 1918, mainly from Asia and Russia (e.g., alfalfa, apple, barley, melon, elm, dwarf cherry). One of Meyer's most notable collections was the soybean. Prior to his Chinese explorations, there were only eight varieties of soybean grown in the USA, all for forage. This picture changed between 1905 and 1908 when Meyer introduced 42 new soybean varieties into the US, including seed and oil varieties that helped to make the USA a world leader in soybean production. The current US system of plant inventory was established by Fairchild. The first accession, PI 1, was a cabbage accession from Moscow, collected in 1898. PI 600,000 is a pollinator sunflower with dwarf features, developed by ARS (Agricultural Research Service) breeders.

#### *Other efforts*

Potato introduction to Europe and the introduction of maize and millet to Africa and Asia are examples of the impact of plant introductions on world food and agriculture. In fact, the Green Revolution depended on introductions of dwarf wheat and rice into India, Pakistan, and the Philippines.

#### **Plant introductions**

**Plant introduction** is the process of importing new plants or cultivars of well-established plants from the area of their adaptation to another area where their potential is evaluated for suitability for agricultural or horticultural use. First, the germplasm to be introduced is processed through a plant quarantine station at the

entry port, to ensure that no pest and diseases are introduced along with the desired material. Once this is accomplished, the material is released to the researcher for evaluation in the field for adaptation. The fundamental process of plant introductions as a plant breeding approach is acclimatization. The inherent genetic variation in the introduced germplasm serves as the raw material for adaptation to the new environment, enabling the breeder to select superior performers to form the new cultivar.

When the plant introduction is commercially usable as introduced without any modification, it is called a **primary introduction**. However, more often than not, the breeder makes selections from the variable population, or uses the plant introduction as a parent in crosses. The products of such efforts are called **secondary introductions**. Some plant introductions may not be useful as cultivars in the new environment. However, they may be useful in breeding programs for specific genes they carry. Many diseases, plant stature, compositional traits, and genes for environmental stresses have been introduced by plant breeders.

As a plant breeding method, plant introductions have had a significant impact on world food and agriculture, one of the most spectacular stories being the transformation of US agriculture as previously indicated. One of the most successful agricultural nations in the world, US agriculture is built on plant introductions, since very few plants originated on that continent. The USA either leads the world or is among the top nations in the production of major world crops such as wheat, maize, rice, and soybean.

#### **International conservation efforts**

The reality of germplasm transactions is that truly international cooperation is needed for success. No one country is self-sufficient in its germplasm needs. Most of the diversity resides in the tropical and subtropical regions of the world where most developing nations occur. These germplasm-rich nations, unfortunately, lack the resources and the technology to make the most use of this diversity. International cooperation and agreements are needed for the exploitation of these resources for the mutual benefit of donor and recipient countries.

Vavilov collected more than 250,000 plant accessions during the period of his plant collection expeditions. This collection currently resides at the All-Union Institute of Plant Industry in St Petersburg. The Food

and Agricultural Organization (FAO) of the United Nations (UN) is credited with the initial efforts to promote genetic conservation, and assistance in establishing the International Board of Plant Genetic Resources (IBPGR) based in the FAO in Rome, Italy. Founded in 1974, the IBPGR is funded by donor countries, development banks, and foundations. It is a center in the Consultative Group of International Agriculture Research (CGIAR). The primary role of this board is to collect, preserve, evaluate, and assist with the exchange of plant genetic material for specific crops all over the world.

A major sponsor of these genetic conservation activities is the International Agricultural Research Centers (IARCs) strategically located throughout the tropics (see Chapter 25). Gene banks at these centers focus on starchy crops that feed the world (wheat, corn, rice, potato, sorghum). These crops are often grown with high-tech cultivars that have narrow genetic bases as a result of crop improvement.

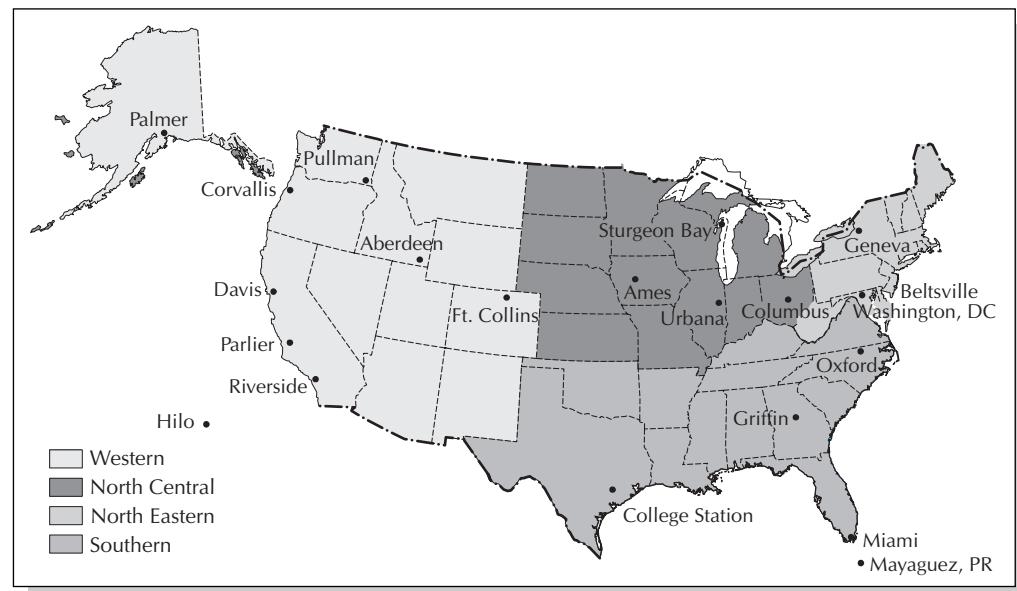
There are other regional- and country-based plant germplasm conservation programs. The EUCARPIA (European Association for Plant Breeding Research), started in 1960, serves Europe and the Mediterranean region. Similarly, the Vegetable Gene Bank at the National Vegetable Research Station in the UK was established in 1981 to conserve vegetable genetic resources.

## An example of a national germplasm conservation system

The US plant genetic conservation efforts are coordinated by the National Plant Germplasm System (NPGS). Over 400,000 accessions exist in the inventory of the NPGS in the form of seed and vegetative material. In August, 2004, the composition of the holdings was 205 families, 1,644 genera, and 10,205 species, for a total of 460,799 accessions. Most of these materials are predominantly landraces and unimproved germplasm from overseas sources. The accessions in the NPGS are estimated to increase at a rate of 7,000–15,000 new entries per year. The system has certain component units with specific functions as follows.

### *Plant introduction*

Located in Beltsville, Maryland, the Plant Introduction Office is part of the Plant Genetics and Germplasm Institute of the USDA-ARS. Each entry is given a plant introduction (PI) number, but this unit does not maintain any plant material collection. The responsibilities to maintain, evaluate, and release plant materials are assigned to four regional Plant Introduction Stations (Western, North Central, North Eastern, and Southern) (Figure 6.2). The Plant Quarantine Facility of USDA



**Figure 6.2** The four regional germplasm jurisdictions defined by the USDA.

and the Animal and Plant Health Inspection Service (APHIS) both operate from the Plant Introduction Station at Glenn Dale, Maryland. These regional centers were established in 1946 for several purposes: (i) to determine the germplasm needs within the region; (ii) to assist with foreign explorations to fill regional needs; (iii) to multiply, evaluate, and maintain new plant and

seeds collections of crops adapted to the regions with minimal loss of genetic variability within the strains; and (iv) to distribute the seed and plant accessions to plant scientists worldwide. These collections come from many countries. For example, in the National Small Grains Collection, the accessions for wheat, barley, and rice come from over 100 countries or regions.

**Table 6.1** Germplasm holdings in germplasm banks in the USA.

Repository	Location	Germplasm	Holding
Barley Genetic Stock Center	Aberdeen, IA	Barley	3,262
Clover Collection	Lexington, KY	Clover	246
Cotton Collection	College Station, TX	Cotton	9,536
Database Management Unit	Beltsville, MD		
Desert Legume Program	Tucson, AZ	Various	2,585
Maize Genetics Stock Center	Urbana, IL	Maize	4,710
National Arctic Plant Genetic Resources Unit	Palmer, AK	Various	515
National Arid Land Plant Genetic Resources Unit	Parlier, CA	Various	1,177
National Center for Genetic Resources Preservation	Fort Collins, CO	Various	23,007
National Clonal Germplasm Repository	Corvallis, OR	Various	12,943
National Clonal Germplasm Repository	Riverside, CA	Citrus, dates	1,167
National Clonal Germplasm Repository	Davis, CA	Tree fruit, nuts, grape	5,397
National Germplasm Resources Laboratory	Beltsville, MD	Various	252
National Small Grains Collection	Aberdeen, IA	Barley, others	126,883
National Temperate Forage Legume Genetic Resources Unit	Prosser, WA	Various	
North Central Regional Plant Introductions Stations	Ames, IA	Various	47,684
Northeast Regional Plant Introduction Station	Geneva, NY	Various	11,690
Ornamental Plant Germplasm Center	Columbus, OH	Various	2,271
Pea Genetic Stock Center	Pullman, WA	Pea	501
Pecan Breeding and Genetics	Brownwood and Somerville, TX	Pecan	881
Plant Genetic Resources Conservation Unit	Griffin, GA	Various	
Plant Genetic Resources Unit	Geneva, NY	Various	5,243
Plant Germplasm Quarantine Office	Beltsville, MD	Various	4,641
Rice Genetic Stock Center	Stuttgart, AR	Rice	19
Southern Regional Plant Introduction Station	Griffin, GA	Various	83,902
Soybean/Maize Germplasm, Pathology and Genetics Research Unit	Urbana, IL	Soybean, maize	20,601
Subtropical Horticulture Research Station	Miami, FL	Various	4,779
Tobacco Collection	Oxford, NC	Tobacco	2,106
Tomato Genetics Resource Center	Davis, CA	Tomato	3,381
Tropical Agriculture Research Station	Mayaguez, Puerto Rico	Various	652
Tropical Plant Genetic Resource Management Unit	Hilo, HI	Various	692
United States Potato Genebank	Sturgeon Bay, WI	Potato	5,648
Western Regional Plant Introduction Station	Pullman, WA	Various	72,190
Wheat Genetic Stocks Center	Aberdeen, ID	Wheat	334
Woody Landscape Plant Germplasm Repository (National Arboretum)	Washington, DC	Various	1,904
Total germplasm holdings as of August 29, 2004			460,799

**Table 6.2** Germplasm holdings at the International Agricultural Research Centers.

International center	Germplasm type	Holdings
International Rice Research Institute (IRRI)	Rice	80,617
Centro Internationale de Mejoramiento de Maiz y Trigo (International Center for the Improvement of Maize and Wheat) (CIMMYT)	Wheat	95,113
	Maize	20,411
International Center for Tropical Agriculture (CIAT)	Forages	18,138
	Bean	31,718
International Institute of Tropical Agriculture (IITA)	Bambara groundnut	2,029
	Cassava	2,158
	Cowpea	15,001
	Soybean	1,901
	Wild <i>Vigna</i>	1,634
	Yam	2,878
International Potato Center (CIP)	Potato	5,057
	Sweet potato	6,413
International Center for Agricultural Research in the Dry Areas (ICARDA)	Andean roots/tubers	1,112
	Barley	24,218
	Chickpea	9,116
	Faba bean	9,074
	Forages	24,581
	Lentil	7,827
	Wheat	30,270
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT)	Chickpea	16,961
	Groundnut	14,357
	Pearl millet	21,250
	Minor millets	9,050
	Pigeon pea	12,698
	Sorghum	35,780
West African Rice Development Association (WARDA)	Rice	14,917
International Center for Research in Agroforestry (ICRAF)	<i>Sesbania</i>	25
International Livestock Research Institute (ILRI)	Forages	11,537
International Plant Genetic Resources Institute (IPGRI)/ International Network for the Improvement of Banana and Plantain (INIBAP)	Musa	931
Total		532,508

### Collections

The base collections of the USA are maintained at the National Seed Storage Laboratory at Fort Collins, Colorado. These collections are seldom regrown to avoid possible genetic changes. The laboratory provides long-term backup storage for the NPGS. In addition to seed, there are National Clonal Repositories for maintaining clonal germplasm. These include Davis, California (for grapes, nuts, and stone fruits) and Maimi, Florida (for subtropical and tropical fruits and sugarcane). The locations and mandates of 35 plant

germplasm conservation programs in the NPGS are presented in Table 6.1. A summary of the germplasm holdings at each location as of August 2004 is provided in Table 6.2. Plant breeders have access to the accessions in these active collections.

### Information

The information on the accessions in the NPGS has been computerized to facilitate its dissemination. The system, Germplasm Resources Information Network, is located at the Beltsville USDA research center.

## References and suggested reading

- Astley, D. 1987. Genetic resource conservation. *Exper. Agric.* 23:245–257.
- Brush, S.B. 1995. *In situ* conservation of landraces in center of crop diversity. *Crop Sci.* 35:346–354.
- Day, P.R. (ed.). 1991. Managing global genetic resources: The US National Plant Germplasm System. National Academic Press, Washington, DC.
- Duvick, D.N. 1984. Genetic diversity in major farm crops on the farm and in reserve. *Econ. Bot.* 38(2):161–178.
- Franco, J., J. Crossa, J.M. Ribaut, J. Betran, M.L. Warburton, and M. Khairallah. 2001. A method of combining molecular markers and phenotypic attributes for classifying plant genotypes. *Theor. Appl. Genet.* 103:944–952.
- Kloppenburg, J.J., and D.L. Kleinman. 1987. The plant germplasm controversy. *BioScience* 37:190–198.
- Mohammadi, S.A., and B.M. Prasanna. 2003. Analysis of genetic diversity in crop plants – Salient statistical tools and considerations. *Crop Sci.* 43:1235–1248.
- National Plant Germplasm System. 2003a. Germplasm Resources Information Network (GRIN) Database Management Unit (DBMU). NPGS, USDA, Beltsville, MD.
- National Plant Germplasm System. 2003b. USDA-ARS Germplasm Resources Information Network database.
- Peterson, A.H., R.K. Boman, S.M. Brown, et al. 2005. Reducing the genetic vulnerability of cotton. *Crop Sci.* 45:1900–1901.
- Stoskopf, N.C. 1993. Plant breeding: Theory and practice. Westview Press, Boulder, CO.
- Warburton, M.L., X. Xianchun, J. Crossa, et al. 2002. Genetic characterization of CIMMYT inbred maize lines and open pollinated populations, using large scale fingerprinting methods. *Crop Sci.* 42:1832–1840.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 The US National Seed Storage Laboratory at Fort Collins maintains a base collection of germplasm.
- 2 Generally, the first source of germplasm considered by a plant breeder is undomesticated plants.
- 3 Without variability, it is impossible to embark upon a plant breeding project.
- 4 CIMMYT is responsible for maintaining wheat and corn germplasm.
- 5 Only seeds are stored at a germplasm bank.

### Part B

Please answer the following questions:

- 1 ..... is the Russian scientist who proposed the concept of centers of diversity.
- 2 What is a landrace?
- 3 Distinguish between a base collection of germplasm and an active collection of germplasm at a gene bank.
- 4 What is Vavilov's law of homologous series in a heritable variation?
- 5 Give specific sources of germplasm erosion.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the importance of plant introductions to US agriculture.
- 2 Discuss the importance of domesticated germplasm to plant breeding.
- 3 Discuss the US germplasm conservation system.
- 4 Discuss the role of the CGIAR in germplasm conservation.
- 5 Discuss the need for germplasm conservation.
- 6 Discuss Vavilov's concept of the centers of diversity of crops.

# **Section 4**

## **Genetic analysis in plant breeding**

**Chapter 7 Introduction to concepts of population genetics**

**Chapter 8 Introduction to quantitative genetics**

**Chapter 9 Common statistical methods in plant breeding**

---

Breeders develop new cultivars by modifying the genetic structure of the base population used to start the breeding program. Students need to have an appreciation of population and quantitative genetics in order to understand the principles and concepts of practical plant breeding. In fact, there is what some call the breeders' equation, a mathematical presentation of a fundamental concept that all breeders must thoroughly understand. This section will help the student understand this and other basic breeding concepts.

Statistics is indispensable in plant breeding. Breeders conduct genetic analysis to understand the inheritance of traits they desire to manipulate. They use statistics to help identify parents to use in crossing, to evaluate the products of crossing, and to guide the selection process, as well to evaluate finished products (cultivars) for release to producers. Some of these analyses are based on the genetic structure of the plant populations.

---



# 7

# Introduction to concepts of population genetics

## Purpose and expected outcomes

*Plant breeders manipulate plants based on the modes of their reproduction (i.e., self- or cross-pollinated). Self-pollinated plants, as previously discussed, are pollinated predominantly by pollen grains from their own flowers, whereas cross-pollinated plants are predominantly pollinated by pollen from other plants. These different reproductive behaviors have implications in the genetic structure of plant populations. In addition to understanding Mendelian genetics, plant breeders need to understand changes in gene frequencies in populations. After all, selection alters the gene frequencies of breeding populations. After studying this chapter, the student should be able to:*

- 1 Define a population.
- 2 Discuss the concept of a gene pool.
- 3 Discuss the concept of gene frequency
- 4 Discuss the Hardy–Weinberg law.
- 5 Discuss the implications of the population concept in breeding.
- 6 Discuss the concept of inbreeding and its implications in breeding.
- 7 Discuss the concept of combining ability.

## Concepts of a population and gene pool

Some breeding methods focus on individual plant improvement, whereas others focus on improving plant populations. Plant populations have certain dynamics, which impact their genetic structure. The genetic structure of a population determines its capacity to be changed by selection (i.e., improved by plant breeding). Understanding population structure is key to deciding the plant breeding options and selection strategies to use in a breeding program.

### Definitions

A **population** is a group of sexually interbreeding individuals. The capacity to interbreed implies that every gene within the group is accessible to all members

through the sexual process. A **gene pool** is the total number and variety of genes and alleles in a sexually reproducing population that are available for transmission to the next generation. Rather than the inheritance of traits, **population genetics** is concerned with how the frequencies of alleles in a gene pool change over time. Understanding population structure is important to breeding by either conventional or unconventional methods. It should be pointed out that the use of recombinant DNA technology, as previously indicated, has the potential to allow gene transfer across all biological boundaries to be made. Breeding of cross-pollinated species tends to focus on improving populations rather than individual plants, as is the case in breeding self-pollinated species. To understand population structure and its importance to plant breeding, it is important to understand the type of variability present, and its

underlying genetic control, in addition to the mode of selection for changing the genetic structure.

### Mathematical model of a gene pool

As previously stated, gene frequency is the basic concept in population genetics. Population genetics is concerned with both the genetic composition of the population as well as the transmission of genetic material to the next generation. The genetic constitution of a population is described by an array of gene frequencies. The genetic properties of a population are influenced in the process of transmission of genes from one generation to the next by four major factors – **population size**, **differences in fertility and viability**, **migration and mutation**, and the **mating system**. Genetic frequencies are subject to sample variation between successive generations. A plant breeder directs the evolution of the breeding population through the kinds of parents used to start the base population in a breeding program, how the parents are mated, and artificial selection.

The genetic constitution of individuals in a population is reconstituted for each subsequent generation. Whereas the genes carried by the population have continuity from one generation to the next, there is no such continuity in the genotypes in which these genes occur. Plant breeders often work with genetic phenomena in populations that exhibit no apparent Mendelian segregation, even though in actuality they obey Mendelian laws. Mendel worked with genes whose effects were categorical (kinds) and were readily classifiable (ratios) into kinds in the progeny of crosses. Breeders, on the other hand, are usually concerned about differences in populations measured in degrees rather than kinds. Population genetics uses mathematical models to attempt to describe population phenomena. To accomplish this, it is necessary to make assumptions about the population and its environment.

#### Calculating gene frequency

To understand the genetic structure of a population, consider a large population in which random mating occurs, with no mutation or gene flow between this population and others, no selective advantage for any genotype, and normal meiosis. Consider also one locus,  $A$ , with two alleles,  $A$  and  $a$ . The frequency of allele  $A_1$  in the gene pool is  $p$ , while the frequency of allele  $A_2$  is  $q$ . Also,  $p + q = 1$  (or 100% of the gene pool). Assume a population of  $N$  diploids (have two alleles at each locus) in which two alleles ( $A, a$ ) occur at one locus. Assuming

dominance at the locus, three genotypes –  $AA$ ,  $Aa$ , and  $aa$  – are possible in an  $F_2$  segregating population. Assume the genotypic frequencies are  $D$  (for  $AA$ ),  $H$  (for  $Aa$ ), and  $Q$  (for  $aa$ ). Since the population is diploid, there will be  $2N$  alleles in it. The genotype  $AA$  has two  $A$  alleles. Hence, the total number of  $A$  alleles in the population is calculated as  $2D + H$ . The proportion or frequency of  $A$  alleles (designated as  $p$ ) in the population is obtained as follows:

$$(2D + H)/2N = (D + \frac{1}{2}H)/N = p$$

The same can be done for allele  $a$ , and designated  $q$ . Further,  $p + q = 1$  and hence  $p = 1 - q$ . If  $N = 80$ ,  $D = 4$ , and  $H = 24$ ,

$$p = (D + \frac{1}{2}H)/N = (4 + 12)/80 = 16/80 = 0.2$$

Since  $p + q = 1$ ,  $q = 1 - p$ , and hence  $q = 1 - 0.2 = 0.8$ .

#### Hardy–Weinberg equilibrium

Consider a random mating population (each male gamete has an equal chance of mating with any female gamete). Random mating involving the previous locus ( $A/a$ ) will yield the following genotypes:  $AA$ ,  $Aa$ , and  $aa$ , with the corresponding frequencies of  $p^2$ ,  $2pq$ , and  $q^2$ , respectively. The gene frequencies must add up to unity. Consequently,  $p^2 + 2pq + q^2 = 1$ . This mathematical relationship is called the **Hardy–Weinberg equilibrium**. Hardy of England and Weinberg of Germany discovered that equilibrium between genes and genotypes is achieved in large populations. They showed that the frequency of genotypes in a population depends on the frequency of genes in the preceding generation, not on the frequency of the genotypes.

Considering the previous example, the genotypic frequencies for the next generation following random mating can be calculated as follows:

$AA = p^2$	$= 0.2^2$	$= 0.04$
$Aa = 2pq$	$= 2(0.2 \times 0.8)$	$= 0.32$
$aa = q^2$	$= 0.8^2$	$= 0.64$
Total = 1.00		

The Hardy–Weinberg equilibrium is hence summarized as:

$$p^2AA + 2pqAa + q^2aa = 1 \text{ (or 100%)}$$

This means that in a population of 80 plants as before, about three plants will have a genotype of  $AA$ , 26 will be

$A\alpha$ , and 51 will be  $\alpha\alpha$ . Using the previous formula, the frequencies of the genes in the next generation may be calculated as:

$$p = [D + \frac{1}{2}H]/N = (3 + 13)/80 = 0.2$$

and

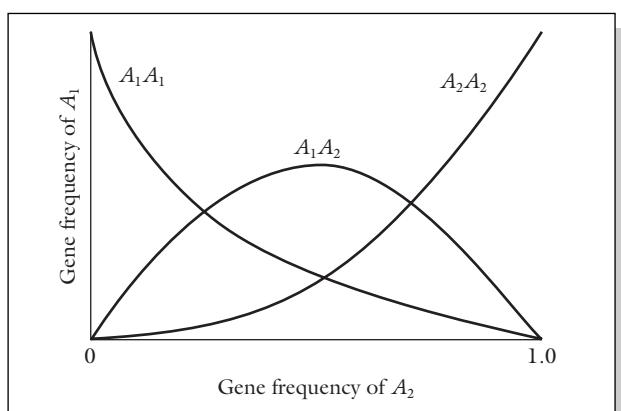
$$q = 1 - p = 0.8$$

The allele frequencies have remained unchanged, while the genotypic frequencies have changed from 4, 24, and 52, to 3, 26, and 51, for  $AA$ ,  $A\alpha$ , and  $\alpha\alpha$ , respectively. However, in subsequent generations, both the genotype and gene frequencies will remain unchanged, provided:

- 1 Random mating occurs in a very large diploid population.
- 2 Allele  $A$  and allele  $\alpha$  are equally fit (one does not confer a more superior trait than the other).
- 3 There is no differential migration of one allele into or out of the population.
- 4 The mutation rate of allele  $A$  is equal to that of allele  $\alpha$ .

In other words, the variability does not change from one generation to another in a random mating population. The maximum frequency of the heterozygote ( $H$ ) cannot exceed 0.5 (Figure 7.1). The Hardy–Weinberg law states that equilibrium is established at any locus after one generation of random mating. From the standpoint of plant breeding, two states of variability are present – two homozygotes ( $AA$ ,  $\alpha\alpha$ ), called “free variability” that can be fixed by selection, and the intermediate heterozygote ( $A\alpha$ ), called “hidden or potential variability” that can generate new variability through segregation. In outcrossing species, the homozygotes can hybridize to generate more heterozygotic variability. Under random mating and no selection, the rate of crossing and segregation will be balanced to maintain the proportion of free and potential variability at 50 : 50. In other words, the population structure is maintained as a dynamic flow of crossing and segregation. However, with two loci under consideration, equilibrium will be attained slowly over many generations. If genetic linkage is strong, the rate of attainment of equilibrium will be even slower.

Most of the important variation displayed by nearly all plant characters affecting growth, development, and reproduction, is quantitative (continuous or polygenic variation, controlled by many genes). Polygenes demonstrate the same properties in terms of dominance,



**Figure 7.1** The relationship between gene frequencies and genotype frequencies in a population in Hardy–Weinberg equilibrium for two alleles. The frequency of the heterozygotes cannot be more than 50%, and this maximum occurs when the gene frequencies are  $p = q = 0.5$ . Further, when the gene frequency of an allele is low, the rare allele occurs predominantly in heterozygotes and there are very few homozygotes. (Adapted from Falconer, D.S. 1981. Introduction to quantitative genetics, 2nd edn. Longman.)

epistasis, and linkage as classic Mendelian genes. The Hardy–Weinberg equilibrium is applicable to these characters. However, it is more complex to demonstrate.

Another state of variability is observed when more than one gene affects the same polygenic trait. Consider two independent loci with two alleles each:  $A$ ,  $a$  and  $B$ ,  $b$ . Assume also the absence of dominance or epistasis. It can be shown that nine genotypes ( $AABB$ ,  $aaAB$ ,  $aabb$ ) and five phenotypes ( $[AABB, 2AaBB] + [2AABb, AAbb, aaBB] + [4AaBb, 2aaBb] + [2Aabb] + [aabb]$ ) in a frequency of 1 : 4 : 6 : 4 : 1, will be produced, following random mating. Again, the extreme genotypes ( $AABB$ ,  $aabb$ ) are the source of completely free variability. However,  $AAbb$  and  $aaBB$ , phenotypically similar but contrasting genotypes, also contain latent variability. Termed homozygotic potential variability, it will be expressed in the free state only when, through crossing, a heterozygote ( $AaBb$ ) is produced, followed by segregation in the  $F_2$ . In other words, two generations will be required to release this potential variability in the free state. Further, unlike the 50 : 50 ratio in the single-locus example, only 1/8 of the variability is available for selection in the free state, the remainder existing as hidden in the heterozygotic or homozygotic potential states. A general mathematical relationship may be derived for any number ( $n$ )

of genes as  $1 : n : n-1$  of free : heterozygotic potential : homozygotic potential.

Another level of complexity may be factored in by considering dominance and non-allelic interactions ( $AA = Aa = BB = Bb$ ). If this is so, the nine genotypes previously observed will produce only three phenotypic classes ( $[AABB, 4AaBb, 2AaBB, 2AABb] + [2Aabb, 2aaBb, AAAb, aaBB] + [aabb]$ , in a frequency of 9 : 6 : 1. A key difference is that 50% of the visible variability is now in the heterozygous potential state that cannot be fixed by selection. The heterozygotes now contribute to the visible variability instead of the cryptic variability. From the plant breeding standpoint, its effect is to reduce the rate of response to phenotypic selection at least in the same direction as the dominance effect. This is because the fixable homozygotes are indistinguishable from the heterozygotes without a further breeding test (e.g., progeny row). Also, the classifications are skewed (9 : 6 : 1) in the positive (or negative) direction.

Key plant breeding information to be gained from the above discussion is that in outbreeding populations, polygenic systems are capable of storing large amounts of cryptic variability. This can be gradually released for selection to act on through crossing, segregation, and recombination. The flow of this cryptic variability to the free state depends on the rate of recombination (which also depends on the linkage of genes on the chromosomes and the breeding system).

Given a recombination value of  $r$  between two linked genes, the segregation in the second generation depends on the initial cross, as M. D. Haywood and E. L. Breese demonstrated:

Initial cross	Free	Homozygous potential
$AABB \times aabb$	$1(1-r)$	$2r$
$AAbb \times aaBB$	$2r$	$2(r-1)$

The second cross shows genes linked in the repulsion phase. The flow of variability from the homozygous potential to the free state depends on how tight a linkage exists between the genes. It will be at its maximum when  $r = 0.5$  and recombination is free, and will diminish with diminishing  $r$ . This illustration shows that with more than two closely linked loci on the same chromosome, the flow of variability would be greatly restricted. In species where selfing is the norm (or when a breeder enforces complete inbreeding), the proportion of heterozygotes will be reduced by 50% in each generation, dwindling to near zero by the eighth or ninth generation.

The open system of pollination in cross-pollinated species allows each plant in the gene pool to have both

homozygous and heterozygous loci. Plant breeders exploit this heterozygous genetic structure of individuals in population improvement programs. In a natural environment, the four factors of genetic change mentioned previously are operational. Fitness or adaptive genes will be favored over non-adaptive ones. Plant breeders impose additional selection pressure to hasten the shift in the population genetic structure toward adaptiveness as well as to increase the frequencies of other desirable genes.

#### *An example of a breeding application of Hardy–Weinberg equilibrium*

In disease-resistance breeding, plant breeders cross an elite susceptible cultivar with one that has disease resistance. Consider a cross between two populations, susceptible  $\times$  resistant. If the gene frequencies of an allele  $A$  in the two populations are represented by  $P_1$  and  $P_2$ , the gene frequency in the  $F_1 = (P_1 + P_2)/2 = p$ . Assuming the frequency of the resistance gene in the resistant cultivar is  $P_1 = 0.7$  and that in the susceptible elite cultivar it is  $P_2 = 0.05$ , the gene frequency in the progeny of the cross  $p$  would be obtained as follows:

$$p = (P_1 + P_2)/2 = (0.7 + 0.05)/2 = 0.375$$

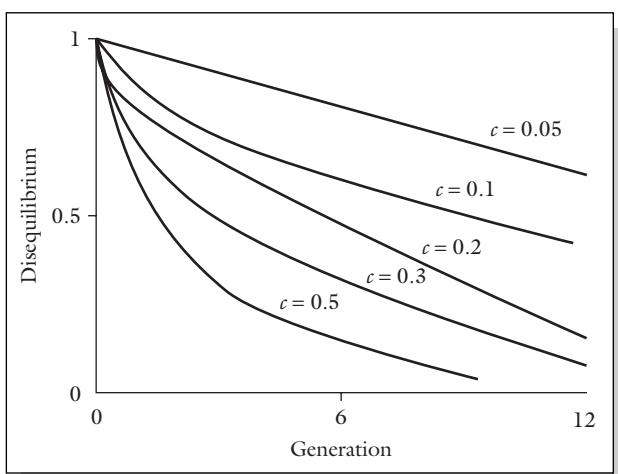
Consequently, the gene frequency for the resistant trait is reduced by about 50% (from 0.7 to 0.375).

### **Issues arising from Hardy–Weinberg equilibrium**

In order for Hardy–Weinberg equilibrium to be true, several conditions must be met. However, some situations provide approximate conditions to satisfy the requirements.

#### **The issue of population size**

The Hardy–Weinberg equilibrium requires a large random mating population (among other factors as previously indicated) to be true. However, in practice, the law has been found to be approximately true for most of the genes in most cross-pollinated species, except when non-random mating (e.g., inbreeding and assortative mating, discussed below on p. 116) occur. Whereas inbreeding is a natural feature of self-pollinated species, assortative mating can occur when cross-pollinated species are closely spaced in the field.



**Figure 7.2** The approach to linkage equilibrium under random mating of two loci considered together. The value of  $c$  gives the linkage frequency between two loci. The effect of linkage is to slow down the rate of approach; the closer the linkage, the slower the rate. For  $c = 0.5$ , there is no linkage. The equilibrium value is approached slowly and is theoretically unattainable.

### The issue of multiple loci

Research has shown that it is possible for alleles at two loci to be in random mating frequencies and yet not in equilibrium with respect to each other. Further, equilibrium between two loci is not attained after one generation of random mating as the Hardy–Weinberg law concluded, but is attained slowly over many generations. Also, the presence of genetic linkage will further slow down the rate of attainment of equilibrium (Figure 7.2). If there is no linkage ( $c = 0.5$ ), the differential between actual frequency and the equilibrium frequency is reduced by 50% in each generation. At this rate, it would take about seven generations to reach approximate equilibrium. However, at  $c = 0.01$  and  $c = 0.001$ , it would take about 69 and 693 generations, respectively, to reach equilibrium. A composite gene frequency can be calculated for genes at the two loci. For example, if the frequency at locus  $Aa = 0.2$  and that for locus  $bb = 0.7$ , the composite frequency of a genotype  $Aabb = 0.2 \times 0.7 = 0.14$ .

### Factors affecting changes in gene frequency

Gene frequency in a population may be changed by one of two primary types of processes – **systematic** or

**dispersive**. A systematic process causes a change in gene frequency that is predictable in both direction and amount. A dispersive process, associated with small populations, is predictable only in amount, not direction. D. S. Falconer listed the systematic processes as **selection**, **migration**, and **mutation**.

### Migration

Migration is important in small populations. It entails the entry of individuals into an existing population from outside. Because plants are sedentary, migration, when it occurs naturally, is via pollen transfer (gamete migration). The impact this immigration will have on the recipient population will depend on the immigration rate and the difference in gene frequency between the immigrants and natives. Mathematically,  $\Delta q = m(q_m - q_o)$ , where  $\Delta q$  = changes in the frequency of genes in the new mixed population,  $m$  = number of immigrants,  $q_m$  = gene frequency of the immigrants, and  $q_o$  = gene frequency of the host. Plant breeders employ this process to change frequencies when they undertake introgression of genes into their breeding populations. The breeding implication is that for open-pollinated (out-breeding) species, the frequency of the immigrant gene may be low, but its effect on the host gene and genotypes could be significant.

### Mutation

Natural mutations are generally rare. A unique mutation (non-recurrent mutation) would have little impact on gene frequencies. Mutations are generally recessive in gene action, but the dominant condition may also be observed. Recurrent mutation (occurs repeatedly at a constant frequency) may affect the gene frequency of the population. Natural mutations are of little importance to practical plant breeding. However, breeders may artificially induce mutation to generate new variability for plant breeding (see Chapter 12).

### Selection

Selection is the most important process for altering population gene frequencies by plant breeders (see Chapter 8). Its effect is to change the mean value of the progeny population from that of the parental population. This change may be greater or lesser than the population mean, depending on the trait of interest. For example, breeders aim for higher yield but may accept and select for less of a chemical factor in the plant that

may be toxic in addition to the high yield. For selection to succeed there must be:

- 1 Phenotypic variation for the trait to allow differences between genotypes to be observed.
- 2 The phenotypic variation must at least be partly genetic.

### Frequency-dependent selection

Selection basically concerns the differential rate of reproduction by different genotypes in a population. The concept of fitness describes the absolute or relative reproductive rate of genotypes. The contribution of genotypes to the next generation is called the **fitness** (or **adaptive value** or **selective value**). The relative fitness of genotypes in a population may depend on its frequency relative to others. Selection occurs at different levels in the plant – phenotype, genotype, zygote, and gamete – making it possible to distinguish between haploid and diploid selections. The **coefficient of selection** is designated  $s$ , and has values between 0 and 1. Generally, the contribution of a favorable genotype is given a score of 1, while a less favorable (less fit) genotype is scored  $1 - s$ .

If  $s = 0.1$ , it means that for every 100 zygotes produced with the favorable genotype, there will be 90 individuals with the unfavorable genotype. Fitness can exhibit complete dominance, partial dominance, no dominance, or overdominance. Consider a case of complete dominance of the  $A$  allele. The relative fitness of the genotypes will be:

Genotypes	$AA$	$Aa$	$aa$	Total
Initial frequency	$p^2$	$2pq$	$q^2$	1
Relative fitness	1	1	$1 - s$	
After selecting	$p^2$	$2pq$	$q^2(1 - s)$	$1 - sq^2$

The total after selection is given by:

$$\begin{aligned} & p^2 + 2pq + q^2(1 - s) \\ &= (1 - q)(1 - q) + 2(1 - q)q + q^2 - sq^2 \\ &= 1 - 2q + q^2 + 2q - 2q^2 + q^2 - sq^2 \end{aligned}$$

To obtain the gene frequency in the next generation, use:

$$\begin{aligned} Q &= (\frac{1}{2}H + Q)/N \\ &= [pq + q^2(1 - s)]/1 - sq^2 \end{aligned}$$

where  $p = 1 - q$ ; multiply  $(1 - s)$  by  $q^2$ :

$$\begin{aligned} q_1 &= [q(1 - q) + q^2 - sq^2]/1 - sq^2 \\ &= (q - q^2 + q^2 - sq^2)/1 - sq^2 \\ &= (q - sq^2)/1 - sq^2 \\ &= [q(1 - sq)]/1 - sq^2 \end{aligned}$$

The relationship between any two generations may be generalized as:

$$q(n+1) = [q_n(1 - sq_n)]/1 - sq_n^2$$

Similarly, the difference in gene frequency,  $\Delta q$ , between any two generations can be shown to be:

$$\begin{aligned} \Delta q &= q_1 - q \\ &= [sq^2(1 - q)]/1 - sq^2 \end{aligned}$$

Other scenarios of change in gene frequency are possible.

Plant breeders use artificial selection to impose new fitness values on genes that control traits of interest in a breeding program.

### Summary of key plant breeding applications

- 1 Selection is most effective at intermediate gene frequency ( $q = 0.5$ ) and least effective at very large or very small frequencies ( $q = 0.99$  or  $q = 0.01$ ). Further, selection for or against a rare allele is ineffective. This is so because a rare allele in a population will invariably occur in the heterozygote and be protected (heterozygote advantage).
- 2 Migration increases variation of a population. Variation of a population can be expanded in a breeding program through introductions (impact of germplasm). Migration also minimizes the effects of inbreeding.
- 3 In the absence of the other factors or processes, any one of the frequency-altering forces will eventually lead to the fixation of one allele or the other.
- 4 The forces that alter gene frequencies are usually balanced against each other (e.g., mutation to a deleterious allele is balanced by selection).
- 5 Gene frequencies attain stable values called equilibrium points.
- 6 In both natural and breeding populations, there appears to be a selective advantage for the heterozygote (hybrid). Alleles with low selection pressure may persist in the population in a heterozygote state for many generations.
- 7 As population size decreases, the effect of random drift increases. This effect is of importance in

germplasm collection and maintenance. The original collection can be genetically changed if a small sample is taken for growing to maintain the accession.

### Modes of selection

There are three basic forms of selection – **stabilizing**, **disruptive**, and **directional** – the last form being the one of most concern to plant breeders. These forms of selection operate to varying degrees under both natural and artificial selection. A key difference lies in the goal. In natural selection, the goal is to increase the fitness of the species, whereas in plant breeding, breeders impose artificial selection usually to direct the population toward a specific goal (not necessarily the fittest).

#### Stabilizing selection

Selection as a process is ongoing in nature. Regarding characters that directly affect the fitness of a plant (e.g., viability, fertility), selection will always be directionally toward the optimal phenotype for a given habitat. However, for other characters, once optimal phenotype has been attained, selection will act to perpetuate it as long as the habitat remains stable. Selection will be for the population mean and against extreme expressions of the phenotype. This mode of selection is called **stabilizing selection** (also called balancing or optimum selection). Taking flowering for example, stabilizing selection will favor neither early flowering nor late flowering. In terms of genetic architecture, dominance will be low or absent or ambidirectional, whereas epistasis will not generally be present. Stabilizing selection promotes additive variation.

#### Disruptive selection

Natural habitats are generally not homogeneous but consist of a number of “ecological niches” that are distinguishable in time (seasonal or long-term cycles), space (microniches), or function. These diverse ecological conditions favor diverse phenotypic optima in form and function. Disruptive selection is a mode of selection in which extreme variants have higher adaptive value than those around the average mean value. Hence, it promotes diversity (polymorphism). The question then is how the different optima relate (dependently or independently) for maintenance and functioning. Also, at what rate does gene exchange occur between the differentially selected genotypes? These two factors (functional

relationship and rate of gene exchange) determine the effect of a population’s genetic structure. In humans, for example, a polymorphism that occurs is sex (female and male). The two sexes are 100% interdependent in reproduction (gene exchange is 100%). In plants, self-incompatibility is an example of such genetically controlled polymorphism. The rarer the self-incompatibility allele at a locus, the higher the chance of compatible mating (and vice versa). Such frequency-dependent selection is capable of building up a large number of self-incompatibility alleles in a population. As previously indicated, several hundreds of alleles have been found in some species.

#### Directional selection

Plant breeders, as previously stated, impose **directional selection** to change existing populations or varieties (or other genotypes) in a predetermined way. Artificial selection is imposed on the targeted character(s) to achieve maximal or optimal expression. To achieve this, the breeder employs techniques (crossing) to reorganize the genes from the parents in a new genetic matrix (by recombination), assembling “coadapted” gene complexes to produce a fully balanced phenotype, which is then protected from further change by genetic linkage. The breeding system will determine whether the newly constituted gene combinations will be maintained. Whereas inbreeding (e.g., selfing) would produce a homozygous population that will resist further change (until crossed), outbreeding tends to produce heterozygous combinations. In heterozygous populations, alleles that exhibit dominance in the direction of the expression targeted by the breeder will be favored over other alleles. Hence, directional selection leads to the establishment of dominance and/or genic interaction (epistasis).

#### Effect of mating system on selection

Four mating systems are generally recognized. They may be grouped into two broad categories as **random mating** and **non-random mating** (comprising genetic assortative mating, phenotypic assortative mating, and disassortative mating).

#### Random mating

In plants, random mating occurs when each female gamete has an equal chance of being fertilized by

any male gamete of the same plant, or with any other plant of the population, and further, there is an equal chance for seed production. As can be seen from the previous statement, it is not possible to achieve true random mating in plant breeding since selection is involved. Consequently, it is more realistic to describe the system of mating as random mating with selection. Whereas true random mating does not change gene frequencies, existing variability in the population, or genetic correlation between close relatives, random mating with selection changes gene frequencies and the mean of the population, with little or no effect on homozygosity, population variance, or genetic correlation between close relatives in a large population. Small populations are prone to random fluctuation in gene frequency (genetic drift) and inbreeding, factors that reduce heterozygosity in a population. Random mating does not fix genes, with or without selection. If the goal of the breeder is to preserve desirable alleles (e.g., in germplasm composites), random mating will be an effective method of breeding.

### Non-random mating

Non-random mating has two basic forms: (i) mating occurs between individuals that are related to each other by ancestral descent (promotes an increase in homozygosity at all loci); and (ii) individuals mate preferentially with respect to their genotypes at any particular locus of interest. If mating occurs such that the mating pair has the same phenotype more often than would occur by chance, it is said to be **assortative mating**. The reverse is true in **disassortative mating**, which occurs in species with self-incompatibility or sterility problems, promoting heterozygosity.

### Genetic assortative mating

Genetic assortative mating or inbreeding entails mating individuals that are closely related by ancestry, the closest being selfing (self-fertilization). A genetic consequence of genetic assortative mating is the exposure of cryptic genetic variability that was inaccessible to selection and was being protected by heterozygosity (i.e., heterozygous advantage). Also, repeated selfing results in homozygosity and brings about fixation of types. This mating system is effective if the goal of the breeder is to develop homozygous lines (e.g., developing inbred lines for hybrid seed breeding or the development of synthetics).

### *Phenotypic assortative mating*

Mating may also be done on the basis of phenotypic resemblance. Called phenotypic assortative mating, the breeder selects and mates individuals on the basis of their resemblance to each other compared to the rest of the population. The effect of this action is the development of two extreme phenotypes. A breeder may choose this mating system if the goal is to develop an extreme phenotype.

### *Disassortative mating*

Disassortative mating may also be genetic or phenotypic. Genetic disassortative mating entails mating individuals that are less closely related than they would be under random mating. A breeder may use this system to cross different strains. In phenotypic disassortative mating, the breeder may select individuals with contrasting phenotypes for mating. Phenotypic disassortative mating is a conservative mating system that may be used to maintain genetic diversity in the germplasm, from which the breeder may obtain desirable genes for breeding as needed. It maintains heterozygosity in the population and reduces genetic correlation between relatives.

## Concept of inbreeding

As previously indicated, plant breeding is a special case of evolution, whereby a mixture of natural and, especially, artificial selection operates rather than natural selection alone. The Hardy–Weinberg equilibrium is not satisfied in plant breeding because of factors including non-random mating. Outcrossing promotes random mating, but breeding methods impose certain mating schemes that encourage non-random mating, especially inbreeding. Inbreeding is measured by the **coefficient of inbreeding** ( $F$ ), which is the probability of identity of alleles by descent. The range of  $F$  is 0 (no inbreeding; random mating) to 1 (prolonged selfing). It can be shown mathematically that:

$$[p^2(1 - F) + Fp] : [2pq(1 - F)] : [q^2(1 - F) + Fq]$$

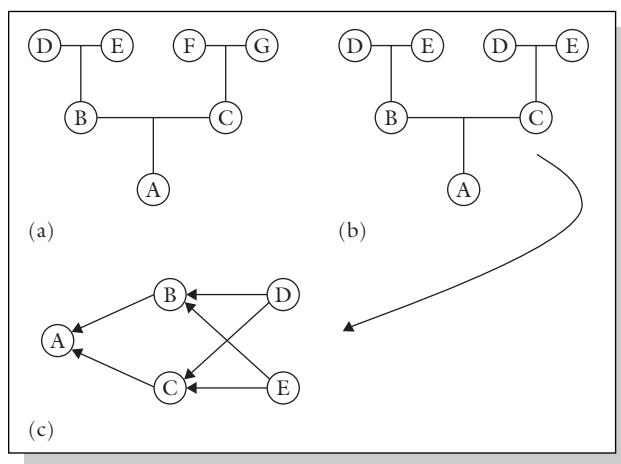
If  $F = 0$ , then the equation reduces to the familiar  $p^2 + 2pq + q^2$ . However, if  $F = 1$ , it becomes  $p : 0 : q$ . The results show that any inbreeding leads to **homozygosity** (all or nearly all loci homozygous), and extreme inbreeding leads to a complete absence of **heterozygosity** (all or nearly all loci heterozygous).

Differential fitness is a factor that mitigates against the realization of the Hardy–Weinberg equilibrium. According to Darwin, the more progeny left, on average, by a genotype in relation to the progeny left by other genotypes, the fitter it is. It can be shown that the persistence of alleles in the population depends on whether they are dominant, intermediate, or recessive in gene action. An unfit (deleterious) recessive allele is fairly quickly reduced in frequency but declines slowly thereafter. On the other hand, an unfit dominant allele is rapidly eliminated from the population, while an intermediate allele is reduced more rapidly than a recessive allele because the former is open to selection in the heterozygote. The consequence of these outcomes is that unfit dominant or intermediate alleles are rare in cross-breeding populations, while unfit recessive alleles persist because they are protected by their recessiveness. The point that will be made later but is worth noting here is that inbreeding exposes unfit recessive alleles (they become homozygous and are expressed) to selection and potential elimination from the population. It follows that inbreeding will expose any unfit allele, dominant or recessive. Consequently, species that are inbreeding would have an opportunity to purge out unfit alleles and hence carry less genetic unfitness load (i.e., have more allele fitness) than outcrossing species. Furthermore, inbreeders (self-pollinated species) are more tolerant of inbreeding whereas outcrossing species are intolerant of inbreeding.

Whereas outcrossing species have more heterozygous loci and carry more unfitness load, there are cases in which the heterozygote is fitter than either homozygote. Called **overdominance**, this phenomenon is exploited in hybrid breeding (see Chapter 18).

### Inbreeding and its implications in plant breeding

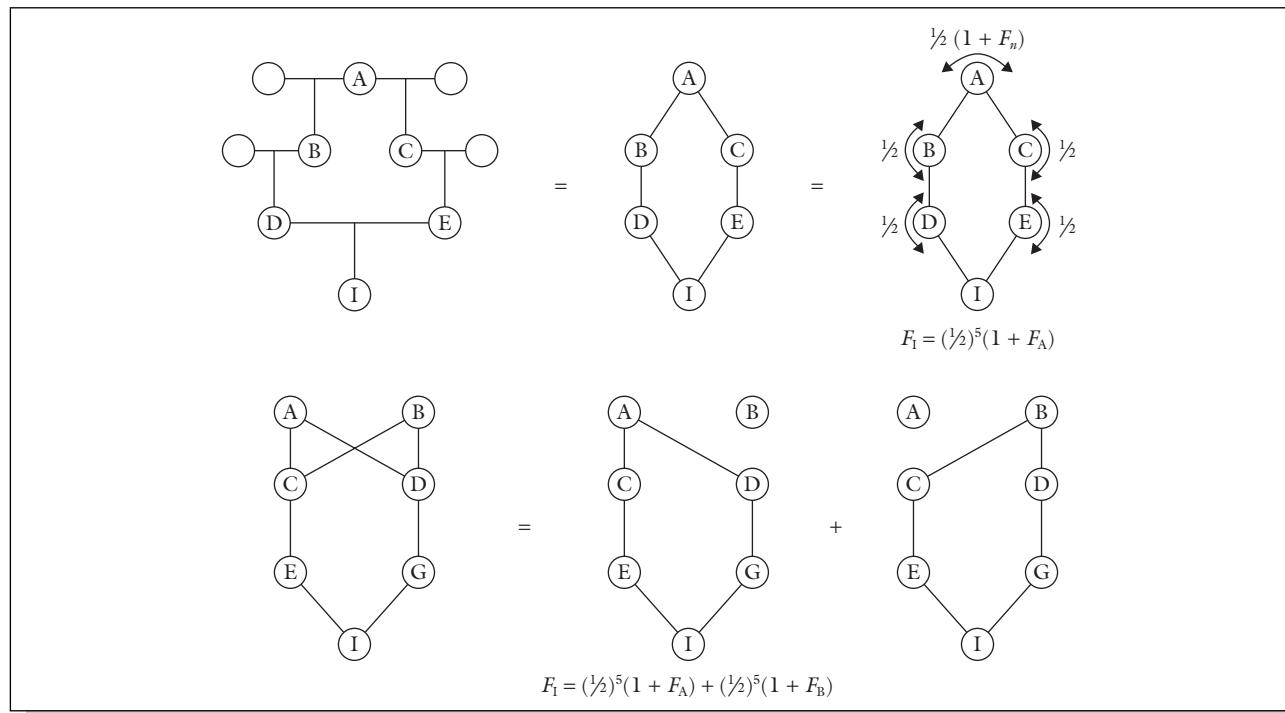
The point has already been made that the methods used by plant breeders depend on the natural means of reproduction of the species. This is because each method of reproduction has certain genetic consequences. In Figure 7.3a, there is no inbreeding because there is no common ancestral pathway to the individual A (i.e., all parents are different). However, in Figure 7.3b inbreeding exists because B and C have common parents (D and E), that is, they are full sibs. To calculate the amount of inbreeding, the standard pedigree is converted to an arrow diagram (Figure 7.3c). Each individual contributes one-half of its genotype to its offspring. The



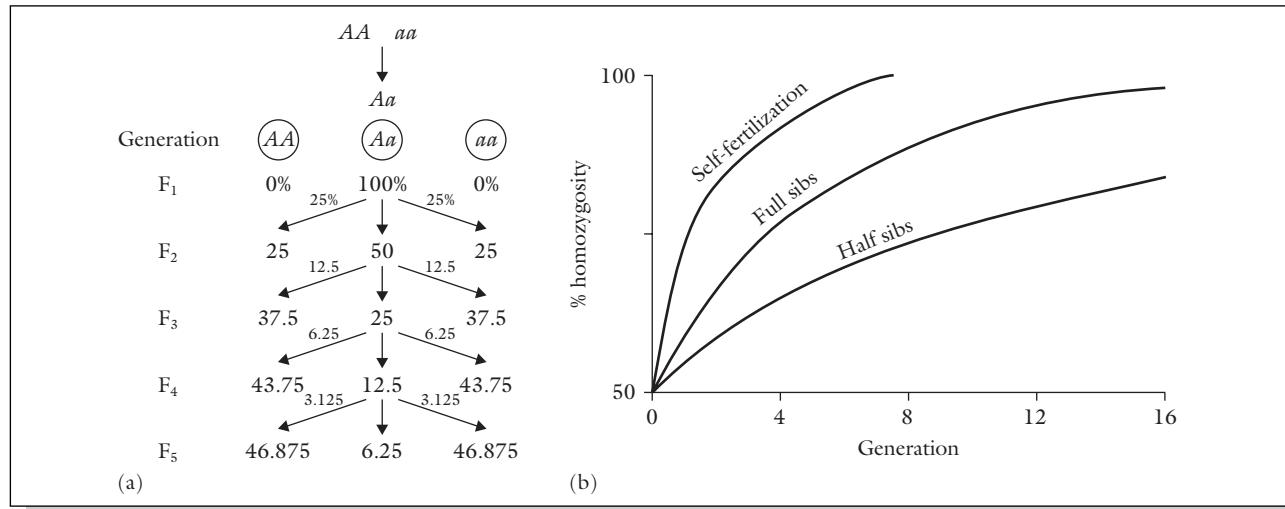
**Figure 7.3** Pedigree diagrams can be drawn in the standard form (a, b) or converted to into an arrow diagram (c).

coefficient of relationship ( $R$ ) is calculated by summing up all the pathways between two individuals through a common ancestor as:  $R_{BC} = \sum(\frac{1}{2})^s$ , where  $s$  is the number of steps (arrows) from B to the common ancestor and back to C. For example, B and C probably inherited  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$  of their genes in common through ancestor D. Similarly, B and C probably inherited one-quarter of their genes in common through ancestor E. The coefficient of relationship between B and C, as a result of common ancestry, is hence  $R_{BC} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} = 50\%$ . Other more complex pedigrees are shown in Figure 7.4.

As previously indicated, prolonged selfing is the most extreme form of inbreeding. With each selfing, the percent heterozygosity decreases by 50%, whereas the percent homozygosity increases by 50% from the previous generation. The approach to homozygosity depends on the intensity of inbreeding as illustrated in Figure 7.5. The more distant the relationship between parents, the slower is the approach to homozygosity. The coefficient of inbreeding ( $F$ ), previously discussed, measures the probability of identity of alleles by descent. This can be measured at both the individual level as well as at the population level. At the individual level,  $F$  measures the probability that any two alleles at any locus are identical by descent (i.e., they are both products of a gene present in a common ancestor). At the population level,  $F$  measures the percentage of all loci that were heterozygous in the base population but have now probably become homozygous due to the effects of inbreeding. There are several methods used for calculating  $F$ . The coefficient of inbreeding ( $F_x$ ) of an individual may be obtained by counting the number of arrows ( $n$ ) that connect the



**Figure 7.4** The inbreeding coefficient ( $F$ ) may be calculated by counting the number of arrows that connect the individual through one parent back to the common ancestor and back again to the other parent and applying the formula in the figure.



**Figure 7.5** Increase in percentage of homozygosity under various systems of inbreeding. (a) Selfing reduces heterozygosity by 50% of what existed at the previous generation. (b) The approach to homozygosity is most rapid under self-fertilization.

individual through one parent back to the common ancestor and back again to the other parent, using the mathematical expression:

$$F_x = \sum \left( \frac{1}{2} \right)^n (1 + F_A)$$

### Consequences

The genetic consequences of inbreeding were alluded to above. The tendency towards homozygosity with inbreeding provides an opportunity for recessive alleles to be homozygous and hence expressed. Whereas inbreeding generally has little or no adverse effect in inbred species, crossbred species suffer adverse consequences when the recessive alleles are less favorable than the dominant alleles. Called **inbreeding depression**, it is manifested as a reduction in performance, because of the expression of less fit or deleterious alleles. The severity of inbreeding depression varies among species, being extreme in species such as alfalfa in which inbreeding produces homozygous plants that fail to survive. Further, the effect of inbreeding is most significant in the first 5–8 generations, and negligible after the eighth generation in many cases.

### Applications

Inbreeding is desirable in some breeding programs. Inbred cultivars of self-pollinated species retain their genotype through years of production. In cross-pollinated species, inbred lines are deliberately developed for use as parents in hybrid seed production. Similarly, partially inbred lines are used as parents in the breeding of synthetic cultivars and vegetatively propagated species by reducing the genetic load. Another advantage of inbreeding is that it increases the genetic diversity among individuals in a population, thereby facilitating the selection process in a breeding program.

### Mating systems that promote inbreeding

Mating is a way by which plant breeders impact the gene frequencies in a population. Four **mating systems** are commonly used to affect inbreeding – self-fertilization, full-sib mating, half-sib mating, and backcrossing (see Section 6). Self-fertilization is the union of male and female gametes; full-sib mating involves the crossing of pairs of plants from a population. In half-sib mating, the pollen source is random from the population, but the female plants are identifiable. In a backcross, the  $F_1$  is repeatedly crossed to one of the parents. Self-fertilization

and backcrossing are the most extreme forms of inbreeding, attaining a coefficient of inbreeding ( $F$ ) of 15/16 after four generations of mating. Autopolyploids have multiple alleles and hence can accumulate more deleterious alleles that remain masked. Inbreeding depression is usually more severe in autopolyploids than diploid species. However the progression to homozygosity is much slower in autopolyploids than in diploids.

## Concept of population improvement

The general goal of improving open- or cross-pollinated species is to change the gene frequencies in the population towards fixation of favorable alleles while maintaining a high degree of heterozygosity. Unlike self-pollinated species in which individuals are the focus and homozygosity and homogeneity are desired outcomes of breeding, population improvement focuses on the whole group, not individual plants. Consequently, open-pollinated populations are not homogeneous.

### Types

The population can be changed by one of two general strategies (i.e., there are two basic types of open-pollinated populations in plant breeding) – by **population improvement** and by the development of **synthetic cultivars**. To develop cultivars by population improvement entails changing the population *en masse* by implementing a specific selection tactic. A cultivar developed this way is sustainable in a sense, maintaining its identity indefinitely through random mating within itself in isolation. The terminology “synthetic” is used to denote an open-pollinated cultivar developed from combining inbred or clonal parental lines. However, the cultivar is not sustainable and must be reconstituted from parental stock. Other usage of the term occurs in the literature.

### Methods of population improvement

Some form of evaluation precedes selection. A breeding material is selected after evaluating the variability available. Similarly, advancing plants from one generation to the next is preceded by an evaluation to determine individuals to select. In self-pollinated species, individuals are homozygous and when used in a cross their genotype is precisely reproduced in their progeny. A **progeny test** is hence adequate for evaluating an individual’s performance. However, open-pollinated species are heterozygous plants and are further pollinated by other

heterozygous plants growing with them in the field. Progeny testing is hence not adequately evaluative of the performance of individual plants of such species. A more accurate evaluation of performance may be achieved by using pollen (preferably from a homozygous source – inbred line) to pollinate the plants. As previously described, the method of evaluating the performance of different mother plants in a comparative way using a common pollen source (tester line) is called a **testcross**. The objective of such a test is to evaluate

the performance of a parent in a cross, a concept called combining ability (discussed in Chapter 8).

The methods used by plant breeders in population improvement may be categorized into two groups, based on the process for evaluating performance. One group of methods is based solely on phenotypic selection and the other on progeny testing (genotypic selection). The specific methods include mass selection, half-sib, full-sib, and recurrent selection, and synthetics (see Section 6).

## References and suggested reading

- Ayala, F.J., and C.A. Campbell. 1974. Frequency-dependent selection. *Ann. Ecol. Systemat.* 5:115–138.
- Cornelius, P.L., and J.W. Dudley. 1974. Effects of inbreeding by selfing and full-sib mating in a maize population. *Crop Sci.* 14:815–819.
- Crow, J.F., and M. Kimura. 1970. An introduction to population genetics theory. Harper & Row, New York.
- Hayward, M.D., and E.L. Breese. 1993. Population structure and variability. In: *Plant breeding: Principles and practices* (Hayward, M.D., N.O. Bosemark, and I. Ramagosa, eds). Chapman & Hall, London.
- Li, C.C. 1976. A first course in population genetics. Boxwood Press, Pacific Grove, CA.

## Outcomes assessment

### Part A

Please answer these questions true or false:

- 1 Inbreeding promotes heterozygosity.
- 2 Naturally cross-breeding species are more susceptible to inbreeding than naturally self-pollinated species.
- 3 In Hardy–Weinberg equilibrium gene frequencies add up to unity.
- 4 Open-pollinated species can be improved by mass selection.

### Part B

Please answer the following questions:

- 1 Define the terms (a) population, and (b) gene pool.
- 2 Give three major factors that influence the genetic structure of a population during the processes of transmission of genes from one generation to another.
- 3 Explain the phenomenon of inbreeding depression.
- 4 Distinguish between assortative and disassortative matings.
- 5 Discuss the main types of mating systems used by plant breeders to affect inbreeding.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the Hardy–Weinberg equilibrium and its importance in breeding cross-pollinated species.
- 2 Discuss the consequences of inbreeding.
- 3 Discuss the concept of combining ability.
- 4 Discuss the application of inbreeding in plant breeding.



# 8

# Introduction to quantitative genetics

---

## Purpose and expected outcomes

*Most of the traits that plant breeders are interested in are quantitatively inherited. It is important to understand the genetics that underlie the behavior of these traits in order to develop effective approaches for manipulating them. After studying this chapter, the student should be able to:*

- 1 Define quantitative genetics and distinguish it from population genetics.
  - 2 Distinguish between qualitative traits and quantitative traits.
  - 3 Discuss polygenic inheritance.
  - 4 Discuss gene action.
  - 5 Discuss the variance components of quantitative traits.
  - 6 Discuss the concept of heritability of traits.
  - 7 Discuss selection and define the breeders' equation.
  - 8 Discuss the concept of general worth of a plant.
  - 9 Discuss the concept of combining ability.
- 

## What is quantitative genetics?

Genetics has several branches, including population genetics, quantitative genetics, biometric genetics, and molecular genetics. Population genetics is an extension of Mendelian genetics applied at the population level. Population genetics does not assign a genotypic or numerical value to each of the individuals (genotypes) in the population (except in the case of coefficients of selection). **Quantitative genetics**, on the other hand, is a branch of genetics in which individual genotypes are unidentified, and the traits of individuals are measured. Genotypic values are assigned to genotypes in the population. Quantitative genetics emphasizes the role of selection in controlled populations of known ancestry. Some topics of population genetics are often discussed in quantitative genetics books, partly because population genetics is basic to quantitative genetics.

A quantitative geneticist observes the phenotype, a product of the genotype and the environment. The genotypic array depends on mating systems and genetic linkage relationships, as well as on allelic frequencies, which in turn are impacted by mutation, migration, random drift, and selection (see Chapter 7). To make effective observations about phenotypes, the quantitative geneticist has to make assumptions about the mating system, allelic frequency altering forces, and the environment.

Common assumptions of quantitative genetic analysis are as follow:

- 1 **Reference population defined.** Allelic and genotypic frequencies can only be defined with respect to a specified population. The researcher should define a base reference population. All inferences made about the estimates should depend upon the composition of this reference population.

- 2 Absence of linkage.** It is assumed that the trait (phenotype) observed is not affected by autosomal linkage genes.
- 3 Presence of diploid Mendelian inheritance.** The plants are assumed to be diploid in which genes segregate and assort independently. Analysis of polyploids is possible, but is involved and handled differently.
- 4 Absence of selection during the formation of inbred lines.** In order for the estimates of genetic variances to pertain to the base reference population, it is required that no selection occur when inbred lines are crossed.
- 5 No breeding of the reference population.** It is assumed that the inbreeding coefficient of the reference population is zero. The analysis becomes more complex when inbreeding is coupled with more than two loci and includes the presence of epistasis.

### Quantitative traits

The topic of quantitative traits was first discussed in Chapter 5. Most traits encountered in plant breeding are quantitatively inherited. Many genes control such traits, each contributing a small effect to the overall phenotypic expression of a trait. Variation in quantitative trait expression is without natural discontinuities (i.e., the variation is continuous). The traits that exhibit continuous variations are also called **metric traits**. Any attempt to classify such traits into distinct groups is only arbitrary. For example, height is a quantitative trait. If plants are grouped into tall versus short plants, one could find relatively tall plants in the short group and, similarly, short plants in the tall group.

#### Qualitative genetics versus quantitative genetics

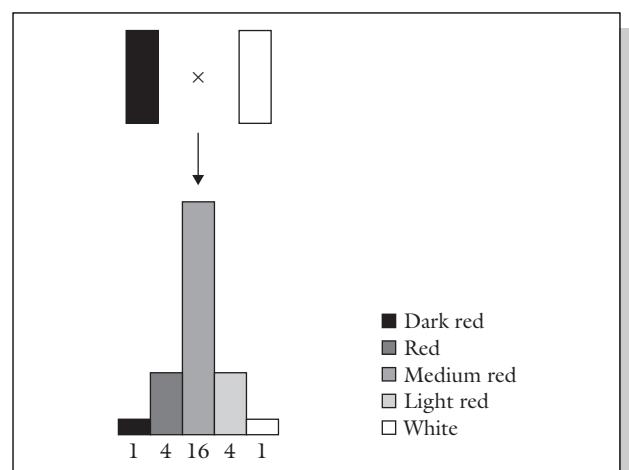
The major ways in which qualitative genetics and quantitative genetics differ may be summarized as:

- 1 Nature of traits.** Qualitative genetics is concerned with traits that have Mendelian inheritance and can be described according to kind and, as previously discussed, can be unambiguously categorized. Quantitative genetics traits are described in terms of the degree of expression of the trait, rather than the kind.
- 2 Scale of variability.** Qualitative genetic traits provide discrete (discontinuous) phenotypic variation, whereas quantitative genetic traits produce phenotypic variation that spans the full spectrum (continuous).

- 3 Number of genes.** In qualitative genetics, the effects of single genes are readily detectable, while in quantitative genetics, single gene effects are not discernible. Rather, traits are under polygenic control (genes with small indistinguishable effects).
- 4 Mating pattern.** Qualitative genetics is concerned with individual matings and their progenies. Quantitative genetics is concerned with a population of individuals that may comprise a diversity of mating kinds.
- 5 Statistical analysis.** Qualitative genetic analysis is quite straightforward, and is based on counts and ratios. On the other hand, quantitative analysis provides estimates of population parameters (attributes of the population from which the sample was obtained).

#### The environment and quantitative variation

All genes are expressed in an environment (phenotype = genotype + environmental effect). However, quantitative traits tend to be influenced to a greater degree than qualitative traits. It should be pointed out that, under significantly large environmental effects, qualitative traits (controlled by one or a few major genes) can exhibit a quantitative trait inheritance pattern (Figure 8.1). A strong environmental influence causes the otherwise distinct classes to overlap.



**Figure 8.1** Nilsson-Ehle's classic work involving wheat color provided the first formal evidence of genes with cumulative effect.

## Polygenes and polygenic inheritance

Quantitative traits are controlled by multiple genes or polygenes.

### What are polygenes?

**Polygenes** are genes with effects that are too small to be individually distinguished. They are sometimes called **minor genes**. In polygenic inheritance, segregation occurs at a large number of loci affecting a trait. The phenotypic expression of polygenic traits is susceptible to significant modification by the variation in environmental factors to which plants in the population are subjected. Polygenic variation cannot be classified into discrete groups (i.e., variation is continuous). This is because of the large number of segregating loci, each with effects so small that it is not possible to identify individual gene effects in the segregating population or to meaningfully describe individual genotypes. Instead, biometrics is used to describe the population in terms of means and variances. Continuous variation is caused by environmental variation and genetic variation due to the simultaneous segregation of many genes affecting the trait. These effects convert the intrinsically discrete variation to a continuous one. Biometric genetics is used to distinguish between the two factors that cause continuous variability to occur.

Another aspect of polygenic inheritance is that different combinations of polygenes can produce a particular phenotypic expression. Furthermore, it is difficult to measure the role of the environment on trait expression because it is very difficult to measure the environmental effect on the plant basis. Consequently, a breeder attempting to breed a polygenic trait should evaluate the cultivar in an environment that is similar to that prevailing in the production region. It is beneficial to plant breeding if a tight linkage of polygenes (called **Polygenic block** or **linkage block**) that has favorable effects on traits of interest to the breeder is discovered.

In 1910, a Swedish geneticist, Nilsson-Ehle provided a classic demonstration of polygenic inheritance and in the process helped to bridge the gap between our understanding of the essence of quantitative and qualitative traits. Polygenic inheritance may be explained by making three basic assumptions:

- 1 Many genes determine the quantitative trait.
- 2 These genes lack dominance.
- 3 The action of the genes are additive.

**Table 8.1** Transgressive segregation.

P <sub>1</sub>	$R_1R_1R_2R_2$ (dark red)	×	$r_1r_1r_2r_2$ (white)
F <sub>1</sub>	$R_1r_1R_2r_2$		
F <sub>2</sub>	1/16	=	$R_1R_1R_2R_2$
	4/16	=	$R_1R_1R_2r_2, R_1r_1R_2R_2$
	6/16	=	$R_1R_1r_2r_2, R_1r_1R_2r_2, r_1r_1R_2R_2$
	4/16	=	$R_1r_1r_2r_2, r_1r_1R_2r_2$
	1/16	=	$r_1r_1r_2r_2$

Nilsson-Ehle crossed two varieties of wheat, one with deep red grain of genotype  $R_1R_1R_2R_2$ , and the other white grain of genotype  $r_1r_1r_2r_2$ . The results are summarized in Table 8.1. He observed that all the seed of the F<sub>1</sub> was medium red. The F<sub>2</sub> showed about 1/16 dark red and 1/16 white seed, the remainder being intermediate. The intermediates could be classified into 6/16 medium red (like the F<sub>1</sub>), 4/16 red, and 4/16 light red. The F<sub>2</sub> distribution of phenotypes may be obtained as an expansion of the binomial  $(a + b)^4$ , where  $a = b = \frac{1}{2}$ .

His interpretation was that the two genes each had a pair of alleles that exhibited cumulative effects. In other words, the genes lacked dominance and their action was additive. Each  $R_1$  or  $R_2$  allele added some red to the phenotype so that the genotypes of white contained neither of these alleles, while the dark red genotype contained only  $R_1$  and  $R_2$ . The phenotypic frequency ratio resulting from the F<sub>2</sub> was 1 : 4 : 6 : 4 : 1 (i.e., 16 genotypes and five classes) (see Figure 8.1).

The study involved only two loci. However, most polygenic traits are conditioned by genes at many loci. The number of genotypes that may be observed in the F<sub>2</sub> is calculated as  $3^n$ , where  $n$  is the number of loci (each with two alleles). Hence, for three loci, the number of genotypes is 27, and for 10 loci, it will be  $3^{10} = 59,049$ . Many different genotypes can have the same phenotype, consequently, there is no strict one-to-one relationship between genotypes (Table 8.2). For  $n$  genes, there are  $3^n$  genotypes and  $2n + 1$  phenotypes. Many complex traits such as yield may have dozens and conceivably even hundreds of loci.

Other difficulties associated with studying the genetics of quantitative traits are dominance, environmental variation, and epistasis. Not only can dominance obscure the true genotype, but both the amount and direction can vary from one gene to another. For example, allele  $A$  may be dominant to  $a$ , but  $b$  may be

**Table 8.2** As the number of genes controlling a trait increases, the phenotypic classes become increasingly indistinguishable. Given  $n$  genes, the number of possible phenotypes in the  $F_2$  is given by  $2^n + 1$ .

Number of gene loci	1	2	3.....	$n$
Ratio of $F_2$ individuals expressing either extreme phenotype (parental)	1/4	1/16	1/64.....	$(1/4)^n$

dominant to  $B$ . It has previously been mentioned that environmental effects can significantly obscure genetic effects. Non-allelic interaction is a clear possibility when many genes are acting together.

### Number of genes controlling a quantitative trait

Polygenic inheritance is characterized by segregation at a large number of loci affecting a trait as previously discussed. Biometric procedures have been proposed to estimate the number of genes involved in a quantitative trait expression. However, such estimates, apart from not being reliable, have limited practical use. Genes may differ in the magnitude of their effects on traits, not to mention the possibility of modifying gene effects on certain genes.

### Modifying genes

One gene may have a major effect on one trait, and a minor effect on another. There are many genes in plants without any known effects besides the fact that they modify the expression of a major gene by either enhancing or diminishing it. The effect of modifier genes may be subtle, such as slight variations in traits like the shape and shades of color of flowers, or, in fruits, variation in aroma and taste. Those trait modifications are of concern to plant breeders as they conduct breeding programs to improve quantitative traits involving many major traits of interest.

### Decision-making in breeding based on biometric genetics

**Biometric genetics** is concerned with the inheritance of quantitative traits. As previously stated, most of the genes of interest to plant breeders are controlled by many

genes. In order to effectively manipulate quantitative traits, the breeder needs to understand the nature and extent of their genetic and environmental control. M. J. Kearsey summarized the salient questions that need to be answered by a breeder who is focusing on improving quantitative (and also qualitative) traits, into four:

- 1 Is the character inherited?
- 2 How much variation in the germplasm is genetic?
- 3 What is the nature of the genetic variation?
- 4 How is the genetic variation organized?

By having answers to these basic genetic questions, the breeder will be in a position to apply the knowledge to address certain fundamental questions in plant breeding.

### What is the best cultivar to breed?

As will be discussed later in the book, there are several distinct types of cultivars that plant breeders develop – pure lines, hybrids, synthetics, multilines, composites, etc. The type of cultivar is closely related to the breeding system of the species (self- or cross-pollinated), but more importantly on the genetic control of the traits targeted for manipulation. As breeders have more understanding of and control over plant reproduction, the traditional grouping between types of cultivars to breed and the methods used along the lines of the breeding system have diminished. The fact is that the breeding system can be artificially altered (e.g., self-pollinated species can be forced to outbreed, and vice versa). However, the genetic control of the trait of interest cannot be changed. The action and interaction of polygenes are difficult to alter. As Kearsey notes, breeders should make decisions about the type of cultivar to breed based on the genetic architecture of the trait, especially the nature and extent of dominance and gene interaction, more so than the breeding system of the species.

Generally, where additive variance and additive  $\times$  additive interaction predominate, pure lines and inbred cultivars are appropriate to develop. However, where dominance variance and dominance  $\times$  dominance interaction suggest overdominance predominates, hybrids would be successful cultivars. Open-pollinated cultivars are suitable where a mixture of the above genetic architectures occur.

### What selection method would be most effective for improvement of the trait?

The kinds of selection methods used in plant breeding are discussed in Chapters 16 and 17. The genetic control

of the trait of interest determines the most effective selection method to use. The breeder should pay attention to the relative contribution of the components of genetic variance (additive, dominance, epistasis) and environmental variance in choosing the best selection method. Additive genetic variance can be exploited for long-term genetic gains by concentrating desirable genes in the homozygous state in a genotype. The breeder can make rapid progress where heritability is high by using selection methods that are dependent solely on phenotype (e.g., mass selection). However, where heritability is low, the method of selection based on families and progeny testing are more effective and efficient. When overdominance predominates, the breeder can exploit short-term genetic gain very quickly by developing hybrid cultivars for the crop.

It should be pointed out that as self-fertilizing species attain homozygosity following a cross, they become less responsive to selection. However, additive genetic variance can be exploited for a longer time in open-pollinated populations because relatively more genetic variation is regularly being generated through the ongoing intermating.

#### *Should selection be on single traits or multiple traits?*

Plant breeders are often interested in more than one trait in a breeding program, which they seek to improve simultaneously. The breeder is not interested in achieving disease resistance only, but in addition, high yield and other agronomic traits. The problem with simultaneous trait selection is that the traits could be correlated such that modifying one affects the other. The concept of correlated traits is discussed next. Biometric procedures have been developed to provide a statistical tool for the breeder to use. These tools are also discussed in this section.

### Gene action

There are four types of gene action: **additive**, **dominance**, **epistatic**, and **overdominance**. Because gene effects do not always fall into clear-cut categories, and quantitative traits are governed by genes with small individual effects, they are often described by their gene action rather than by the number of genes by which they are encoded. It should be pointed out that gene action is conceptually the same for major genes as well as minor genes, the essential difference being that the action of a

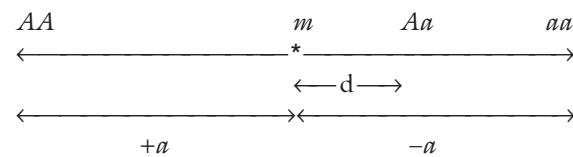
minor gene is small and significantly influenced by the environment.

#### Additive gene action

The effect of a gene is said to be additive when each additional gene enhances the expression of the trait by equal increments. Consequently, if one gene adds one unit to a trait, the effect of  $aabb = 0$ ,  $Aabb = 1$ ,  $AABb = 3$ , and  $AABB = 4$ . For a single locus ( $A, a$ ) the heterozygote would be exactly intermediate between the parents (i.e.,  $AA = 2$ ,  $Aa = 1$ ,  $aa = 0$ ). That is, the performance of an allele is the same irrespective of other alleles at the same locus. This means that the phenotype reflects the genotype in additive action, assuming the absence of environmental effect. Additive effects apply to the allelic relationship at the same locus. Furthermore, a superior phenotype will breed true in the next generation, making selection for the trait more effective to conduct. Selection is most effective for additive variance; it can be fixed in plant breeding (i.e., develop a cultivar that is homozygous).

#### Additive effect

Consider a gene with two alleles ( $A, a$ ). Whenever  $A$  replaces  $a$ , it adds a constant value to the genotype:



Replacing  $a$  by  $A$  in the genotype  $aa$  causes a change of  $a$  units. When both  $aa$  are replaced, the genotype is  $2a$  units away from  $aa$ . The midparent value (the average score) between the two homozygous parents is given by  $m$  (representing a combined effect of both genes for which the parents have similar alleles and environmental factors). This also serves as the reference point for measuring deviations of genotypes. Consequently,  $AA = m + a_A$ ,  $aa = m - a$ , and  $Aa = m + d_A$ , where  $a_A$  is the **additive effect** of allele  $A$ , and  $d$  is the dominance effect. This effect remains the same regardless of the allele with which it is combined.

#### Average effect

In a random mating population, the term **average effect** of alleles is used because there are no homozygous lines. Instead, alleles of one plant combine with alleles from

pollen from a random mating source in the population through hybridization to generate progenies. In effect the allele of interest replaces its alternative form in a number of randomly selected individuals in the population. The change in the population as a result of this replacement constitutes the average effect of the allele. In other words, the average effect of a gene is the mean deviation from the population mean of individuals that received a gene from one parent, the gene from the other parent having come at random from the population.

#### Breeding value

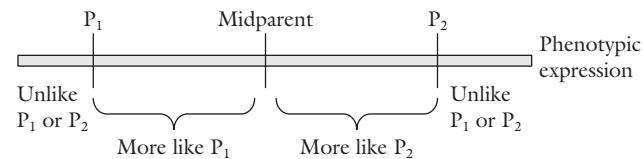
The average effects of genes of the parents determine the mean genotypic value of the progeny. Further, the value of an individual judged by the mean value of its progeny is called the **breeding value** of the individual. This is the value that is transferred from an individual to its progeny. This is a measurable effect, unlike the average effect of a gene. However, the breeding value must always be with reference to the population to which an individual is to be mated. From a practical breeding point of view, the additive gene effect is of most interest to breeders because its exploitation is predictable, producing improvements that increase linearly with the number of favorable alleles in the population.

#### Dominance gene action

Dominance action describes the relationship of alleles at the same locus. Dominance variance has two components – variance due to homozygous alleles (which is additive) and variance due to heterozygous genotypic values. Dominance effects are deviations from additivity that make the heterozygote resemble one parent more than the other. When dominance is complete, the heterozygote is equal to the homozygote in effects (i.e.,  $Aa = AA$ ). The breeding implication is that the breeder cannot distinguish between the heterozygous and homozygous phenotypes. Consequently, both kinds of plants will be selected, the homozygotes breeding true while the heterozygotes will not breed true in the next generation (i.e., fixing superior genes will be less effective with dominance gene action).

#### Dominance effect

Using the previous figure for additive effect, the extent of dominance ( $d_A$ ) is calculated as the deviation of the heterozygote,  $Aa$ , from the mean of the two homozygotes ( $AA$ ,  $aa$ ). Also,  $d_A = 0$  when there is



**Figure 8.2** An illustration of overdominance gene action. The heterozygote,  $Aa$ , is more valuable than either homozygote.

no dominance while  $d$  is positive if  $A$  is dominant, and negative if  $a_A$  is dominant. Further, if dominance is complete  $d_A = a_A$ , whereas  $d_A < a_A$  for incomplete (partial) dominance, and  $d_A > a_A$  for overdominance. For a single locus,  $m = \frac{1}{2}(AA + aa)$  and  $a_A = \frac{1}{2}(AA - aa)$ , while  $d_A = Aa - \frac{1}{2}(AA + aa)$ .

#### Overdominance gene action

**Overdominance gene action** exists when each allele at a locus produces a separate effect on the phenotype, and their combined effect exceeds the independent effect of the alleles (i.e.,  $aa = 1$ ,  $AA = 1$ ,  $Aa = 2$ ) (Figure 8.2). From the breeding standpoint, the breeder can fix overdominance effects only in the first generation (i.e.,  $F_1$  hybrid cultivars) through apomixis, or through chromosome doubling of the product of a wide cross.

#### Epistatic gene action

Epistatic effects in qualitative traits are often described as the masking of the expression of a gene by one at another locus. In quantitative inheritance, **epistasis** is described as non-allelic gene interaction. When two genes interact, an effect can be produced where there was none (e.g.,  $Aabb = 0$ ,  $aaBB = 0$ , but  $A-B- = 4$ ).

The estimation of gene action or genetic variance requires the use of large populations and a mating design. The effect of the environment on polygenes makes estimations more challenging. As N. W. Simmonds observed, at the end of the day, what qualitative genetic analysis allows the breeder to conclude from partitioning variance in an experiment is to say that a portion of the variance behaves *as though* it could be attributed to additive gene action or dominance effect, and so forth.

#### Variance components of a quantitative trait

The genetics of a quantitative trait centers on the study of its variation. As D. S. Falconer stated, it is in terms of

variation that the primary genetic questions are formulated. Further, the researcher is interested in partitioning variance into its components that are attributed to different causes or sources. The genetic properties of a population are determined by the relative magnitudes of the components of variance. In addition, by knowing the components of variance, one may estimate the relative importance of the various determinants of phenotype.

K. Mather expressed the phenotypic value of quantitative traits in this commonly used expression:

$$P(\text{phenotype}) = G(\text{genotype}) + E(\text{environment})$$

Individuals differ in phenotypic value. When the phenotypes of a quantitative trait are measured, the observed value represents the phenotypic value of the individual. The phenotypic value is variable because it depends on genetic differences among individuals, as well as environmental factors and the interaction between genotypes and the environment (called  $G \times E$  interaction).

Total variance of a quantitative trait may be mathematically expressed as follows:

$$V_p = V_G + V_E + V_{GE}$$

where  $V_p$  = total **phenotypic variance** of the segregating population,  $V_G$  = **genetic variance**,  $V_E$  = environmental variance, and  $V_{GE}$  = variance associated with the genetic and environmental interaction.

The genetic component of variance may be further partitioned into three components as follows:

$$V_G = V_A + V_D + V_I$$

where  $V_A$  = **additive variance** (variance from additive gene effects),  $V_D$  = **dominance variance** (variance from dominance gene action), and  $V_I$  = **interaction** (variance from interaction between genes). Additive genetic variance (or simply additive variance) is the variance of breeding values and is the primary cause of resemblance between relatives. Hence  $V_A$  is the primary determinant of the observable genetic properties of the population, and of the response of the population to selection. Further,  $V_A$  is the only component that the researcher can most readily estimate from observations made on the population. Consequently, it is common to partition genetic variance into two – additive versus all other kinds of variance. This ratio,  $V_A/V_p$ , gives what is called the **heritability** of a trait, an estimate that is of practical importance in plant breeding (see next).

The total phenotypic variance may then be rewritten as:

$$V_p = V_A + V_D + V_I + V_E + V_{GE}$$

To estimate these variance components, the researcher uses carefully designed experiments and analytical methods. To obtain environmental variance, individuals from the same genotype are used.

An inbred line (essentially homozygous) consists of individuals with the same genotype. An  $F_1$  generation from a cross of two inbred lines will be heterozygous but genetically uniform. The variance from the parents and the  $F_1$  may be used as a measure of environmental variance ( $V_E$ ). K. Mather provided procedures for obtaining genotypic variance from  $F_2$  and backcross data. In sum, variances from additive, dominant, and environmental effects may be obtained as follows:

$$\begin{aligned} V_{P1} &= E; V_{P2} = E; V_{F1} = E \\ V_{F2} &= \frac{1}{2}A + \frac{1}{4}D + E \\ V_{B1} &= \frac{1}{4}A + \frac{1}{4}D + E \\ V_{B2} &= \frac{1}{4}A + \frac{1}{4}D + E \\ V_{B1} + V_{B2} &= \frac{1}{2}A + \frac{1}{2}D + 2E \end{aligned}$$

This represents the most basic procedure for obtaining components of genetic variance since it omits the variances due to epistasis, which are common with quantitative traits. More rigorous biometric procedures are needed to consider the effects of interlocular interaction.

It should be pointed out that additive variance and dominance variance are statistical abstractions rather than genetic estimates of these effects. Consequently, the concept of additive variance does not connote perfect additivity of dominance or epistasis. To exclude the presence of dominance or epistasis, all the genotypic variance must be additive.

### Concept of heritability

Genes are not expressed in a vacuum but in an environment. A phenotype observed is an interaction between the genes that encode it and the environment in which the genes are being expressed. Plant breeders typically select plants based on the phenotype of the desired trait, according to the breeding objective. Sometimes, a genetically inferior plant may appear superior to other plants only because it is located in a more favorable region of the soil. This may mislead the breeder. In other words, the selected phenotype will not give rise to the same progeny. If the genetic variance is high and the

environmental variance is low, the progeny will be like the selected phenotype. The converse is also true. If such a plant is selected for advancing the breeding program, the expected genetic gain will not materialize. Quantitative traits are more difficult to select in a breeding program because they are influenced to a greater degree by the environment than are qualitative traits. If two plants are selected randomly from a mixed population, the observed difference in a specific trait may be due to the average effects of genes (hereditary differences), or differences in the environments in which the plants grew up, or both. The average effects of genes is what determines the degree of resemblance between relatives (parents and offspring), and hence is what is transmitted to the progenies of the selected plants.

### Definition

The concept of the reliability of the phenotypic value of a plant as a guide to the breeding value (additive genotype) is called the **heritability** of the metric trait. As previously indicated, plant breeders are able to measure phenotypic values directly, but it is the breeding value of individuals that determines their influence on the progeny. Heritability is the proportion of the observed variation in a progeny that is inherited. The bottom line is that if a plant breeder selects plants on the basis of phenotypic values to be used as parents in a cross, the success of such an action in changing the characteristics in a desired direction is predictable only by knowing the degree of correspondence (genetic determination) between phenotypic values and breeding values. Heritability measures this degree of correspondence. It does not measure genetic control, but rather how this control can vary.

**Genetic determination** is a matter of what causes a characteristic or trait; heritability, by contrast, is a scientific concept of what causes differences in a characteristic or trait. Heritability is, therefore, defined as a fraction: it is the **ratio of genetically caused variation to total variation** (including both environmental and genetic variation). Genetic determination, by contrast, is an informal and intuitive notion. It lacks quantitative definition, and depends on the idea of a normal environment. A trait may be described as genetically determined if it is coded in and caused by the genes, and bound to develop in a normal environment. It makes sense to talk about genetic determination in a single individual, but heritability makes sense only relative to a population in which individuals differ from one another.

### Types of heritability

Heritability is a property of the trait, the population, and the environment. Changing any of these factors will result in a different estimate of heritability. There are two different estimates of heritability.

- 1 **Broad sense heritability.** Heritability estimated using the total genetic variance ( $V_G$ ) is called broad sense heritability. It is expressed mathematically as:

$$H = V_G/V_P$$

It tends to yield a high value (Table 8.3). Some use the symbol  $H^2$  instead of  $H$ .

- 2 **Narrow sense heritability.** Because the additive component of genetic variance determines the response to selection, the narrow sense heritability estimate is more useful to plant breeders than the broad sense estimate. It is estimated as:

$$h^2 = V_A/V_P$$

However, when breeding clonally propagated species (e.g., sugarcane, banana), in which both additive and non-additive gene actions are fixed and transferred from parent to offspring, broad sense heritability is also useful. The magnitude of narrow sense heritability cannot exceed, and is usually less than, the corresponding broad sense heritability estimate.

Heritabilities are seldom precise estimates because of large standard errors. Characters that are closely related to reproductive fitness tend to have low heritability estimates. The estimates are expressed as a fraction, but

**Table 8.3** Heritability estimates of some plant architectural traits in dry bean.

Trait	Heritability
Plant height	45
Hypocotyl diameter	38
Number of branches/plant	56
Nodes in lower third	36
Nodes in mid section	45
Nodes in upper third	46
Pods in lower third	62
Pods in mid section	85
Pods in upper third	80
Pod width	81
Pod length	67
Seed number per pod	30
100 seed weight	77

may also be reported as a percentage by multiplying by 100. A heritability estimate may be unity (1) or less.

### Factors affecting heritability estimates

The magnitude of heritability estimates depends on the genetic population used, the sample size, and the method of estimation.

#### *Genetic population*

When heritability is defined as  $h^2 = V_A/V_P$  (i.e., in the narrow sense), the variances are those of individuals in the population. However, in plant breeding, certain traits such as yield are usually measured on a plot basis (not on individual plants). The amount of genotypic variance present for a trait in a population influences estimates of heritability. Parents are responsible for the genetic structure of the populations they produce. More divergent parents yield a population that is more genetically variable. Inbreeding tends to increase the magnitude of genetic variance among individuals in the population. This means that estimates derived from  $F_2$  will differ from, say, those from  $F_6$ .

#### *Sample size*

Because it is impractical to measure all individuals in a large population, heritabilities are estimated from sample data. To obtain the true genetic variance for a valid estimate of the true heritability of the trait, the sampling should be random. A weakness in heritability estimates stems from bias and lack of statistical precision.

#### *Method of computation*

Heritabilities are estimated by several methods that use different genetic populations and produce estimates that may vary. Common methods include the **variance component method** and **parent–offspring regression**. Mating schemes are carefully designed to enable the total genetic variance to be partitioned.

### Methods of computation

The different methods of estimating heritabilities have both strengths and weaknesses.

#### *Variance component method*

The variance component method of estimating heritability uses the statistical procedure of **analysis of**

**variance** (ANOVA, see Chapter 9). Variance estimates depend on the types of populations in the experiment. Estimating genetic components suffers from certain statistical weaknesses. Variances are less accurately estimated than means. Also, variances are unrobust and sensitive to departure from normality. An example of a heritability estimate using  $F_2$  and backcross populations is as follows:

$$\begin{aligned}V_{F_2} &= V_A + V_D + V_E \\V_{B1} + V_{B2} &= V_A + 2V_D + 2V_E \\V_E &= V_{P1} + V_{P2} + V_{F1} \\H &= (V_A + V_D)/(V_A + V_D + V_E) = V_G/V_P \\h^2 &= (V_A)/(V_A + V_D + V_E) = V_A/V_P\end{aligned}$$

**Example** For example, using the data in the table below:

	P <sub>1</sub>	P <sub>2</sub>	F <sub>1</sub>	F <sub>2</sub>	BC <sub>1</sub>	BC <sub>2</sub>
Mean	20.5	40.2	28.9	32.1	25.2	35.4
Variance	10.1	13.2	7.0	52.3	35.1	56.5

$$\begin{aligned}V_E &= [V_{P1} + V_{P2} + V_{F1}]/3 \\&= [10.1 + 13.2 + 7]/3 \\&= 30.3/3 \\&= 10.1\end{aligned}$$

$$\begin{aligned}V_A &= 2V_{F_2} - (V_{B1} + V_{B2}) \\&= 2(52.3) - (35.1 + 56.5) \\&= 104.6 - 91.6 \\&= 13.0\end{aligned}$$

$$\begin{aligned}V_D &= [(V_{B1} + V_{B2}) - F_2 - (V_{P1} + V_{P2} + F_1)]/3 \\&= [(35.1 + 56.5) - 52.3 - (10.1 + 13.2 + 7.0)]/3 \\&= [91.6 - 52.3 - 30.3]/3 \\&= 3.0\end{aligned}$$

#### *Broad sense heritability*

$$\begin{aligned}H &= [13.0 + 3.0]/[13.0 + 3.0 + 10.1] \\&= 16/26.1 \\&= 0.6130 \\&= 61.30\%\end{aligned}$$

#### *Narrow sense heritability*

$$\begin{aligned}h^2 &= 13.0[13.0 + 3.0 + 10.1] \\&= 13.0/26.1 \\&= 0.4980 \\&= 49.80\%\end{aligned}$$

#### *Other methods of estimation*

$$\begin{aligned}H &= [V_{F_2} - \frac{1}{2}(V_{P1} + V_{P2})]/F_2 \\&= [52.3 - \frac{1}{2}(10.1 + 13.2)]/52.3 \\&= 40.65/52.3 \\&= 0.7772 \\&= 77.72\%\end{aligned}$$

This estimate is fairly close to that obtained by using the previous formula.

#### *Parent–offspring regression*

The type of offspring determines if the estimate would be broad sense or narrow sense. This method is based on several assumptions: the trait of interest has diploid Mendelian inheritance; the population from which the parents originated is randomly mated; the population is in linkage equilibrium (or no linkage among loci controlling the trait); parents are non-inbred; and there is no environmental correlation between the performance of parents and offspring.

The parent–offspring method of heritability is relatively straightforward. First, the parent and offspring means are obtained. Cross products of the paired values are used to compute the covariance. A regression of offspring on midparent value is then calculated. Heritability in the narrow sense is obtained as follows:

$$h^2 = b_{op} = V_A/V_p$$

where  $b_{op}$  is the regression of offspring on midparent value, and  $V_A$  and  $V_p$  are the additive variance and phenotypic variance, respectively.

However, if only one parent is known or relevant (e.g., a polycross):

$$b = \frac{1}{2}(V_A/V_p)$$

and

$$h^2 = 2b_{op}$$

#### **Applications of heritability**

Heritability estimates are useful for breeding quantitative traits. The major applications of heritability are:

- 1 To determine whether a trait would benefit from breeding. If, in particular, the narrow sense heritability for a trait is high, it indicates that the use of plant breeding methods will likely be successful in improving the trait of interest.
- 2 To determine the most effective selection strategy to use in a breeding program. Breeding methods that use selection based on phenotype are effective when heritability is high for the trait of interest.
- 3 To predict gain from selection. Response to selection depends on heritability. A high heritability would

likely result in high response to selection to advance the population in the desired direction of change.

#### *Evaluating parental germplasm*

A useful application of heritability is in evaluating the germplasm assembled for a breeding project to determine if there is sufficient genetic variation for successful improvement to be pursued. A replicated trial of the available germplasm is conducted and analyzed by ANOVA as follows:

Source	Degrees of freedom (df)	Error mean sum of squares (EMS)
Replication	$r - 1$	
Genotypes	$g - 1$	$\sigma^2 + r\sigma_g^2$
Error	$(r-1)(g-1)$	$\sigma^2$

From the analysis, heritability may be calculated as:

$$H/b^2 = [\sigma_g^2]/[\sigma_g^2 + \sigma_e^2]$$

It should be pointed out that whether the estimate is heritability in the narrow or broad sense depends on the nature of the genotypes. Pure lines or inbred lines would yield additive type of variance, making the estimate narrow sense. Segregating population would make the estimate broad sense.

#### **Response to selection in breeding**

Selection was discussed in Chapter 7. The focus of this section is on the **response to selection (genetic gain or genetic advance)**. After generating variability, the next task for the breeder is the critical one of advancing the population through selection.

Selection, in essence, entails discriminating among genetic variation (heterogeneous population) to identify and choose a number of individuals to establish the next generation. The consequence of this is differential reproduction of genotypes in the population such that gene frequencies are altered, and, subsequently, the genotypic and phenotypic values of the targeted traits. Even though artificial selection is essentially directional, the concept of “complete” or “pure” artificial selection is an abstraction because, invariably, before the breeder gets a chance to select plants of interest, some amount of natural selection has already been imposed.

The breeder hopes, by selecting from a mixed population, that superior individuals (with high genetic potential)

will be advanced, and will consequently change the population mean of the trait in a positive way in the next generation. The breeder needs to have a clear objective. The trait to be improved needs to be clearly defined. Characters controlled by major genes are usually easy to select. However, polygenic characters, being genetically and biologically complex, present a considerable challenge to the breeder.

The response to selection ( $R$ ) is the difference between the mean phenotypic value of the offspring of the selected parents and the whole of the parental generation before selection. The response to selection is simply the change of population mean between generations following selection. Similarly, the **selection differential** ( $S$ ) is the mean phenotypic value of the individuals selected as parents expressed as a deviation from the population mean (i.e., from the mean phenotypic value of all the individuals in the parental generation before selection). Response to selection is related to heritability by the following equation:

$$R = h^2 S$$

### Prediction of response in one generation: genetic advance due to selection

The genetic advance achieved through selection depends on three factors:

- 1 The total variation (phenotypic) in the population in which selection will be conducted.
- 2 Heritability of the target character.
- 3 The selection pressure to be imposed by the plant breeder (i.e., the proportion of the population that is selected for the next generation).

A large phenotypic variance would provide the breeder with a wide range of variability from which to select. Even when the heritability of the trait of interest is very high, genetic advance would be small without a large amount of phenotypic variation (Figure 8.3). When the heritability is high, selecting and advancing only the top few performers is likely to produce a greater genetic advance than selecting many moderate performers. However, such a high selection pressure would occur at the expense of a rapid loss in variation. When heritability is low, the breeder should impose a lower selection pressure in order to advance as many high-potential genotypes as possible.

In principle, the prediction of response is valid for only one generation of selection. This is so because a

response to selection depends on the heritability of the trait estimated in the generation from which parents are selected. To predict the response in subsequent generations, heritabilities must be determined in each generation. Heritabilities are expected to change from one generation to the next because, if there is a response, it must be accompanied by a change in gene frequencies on which heritability depends. Also, selection of parents reduces the variance and the heritability, especially in the early generations. It should be pointed out that heritability changes are not usually large.

If heritability is unity ( $V_A = V_p$ ; no environmental variance), progress in a breeding program would be perfect, and the mean of the offspring would equal the mean of the selected parents. On the other hand, if heritability is zero, there would be no progress at all ( $R = 0$ ).

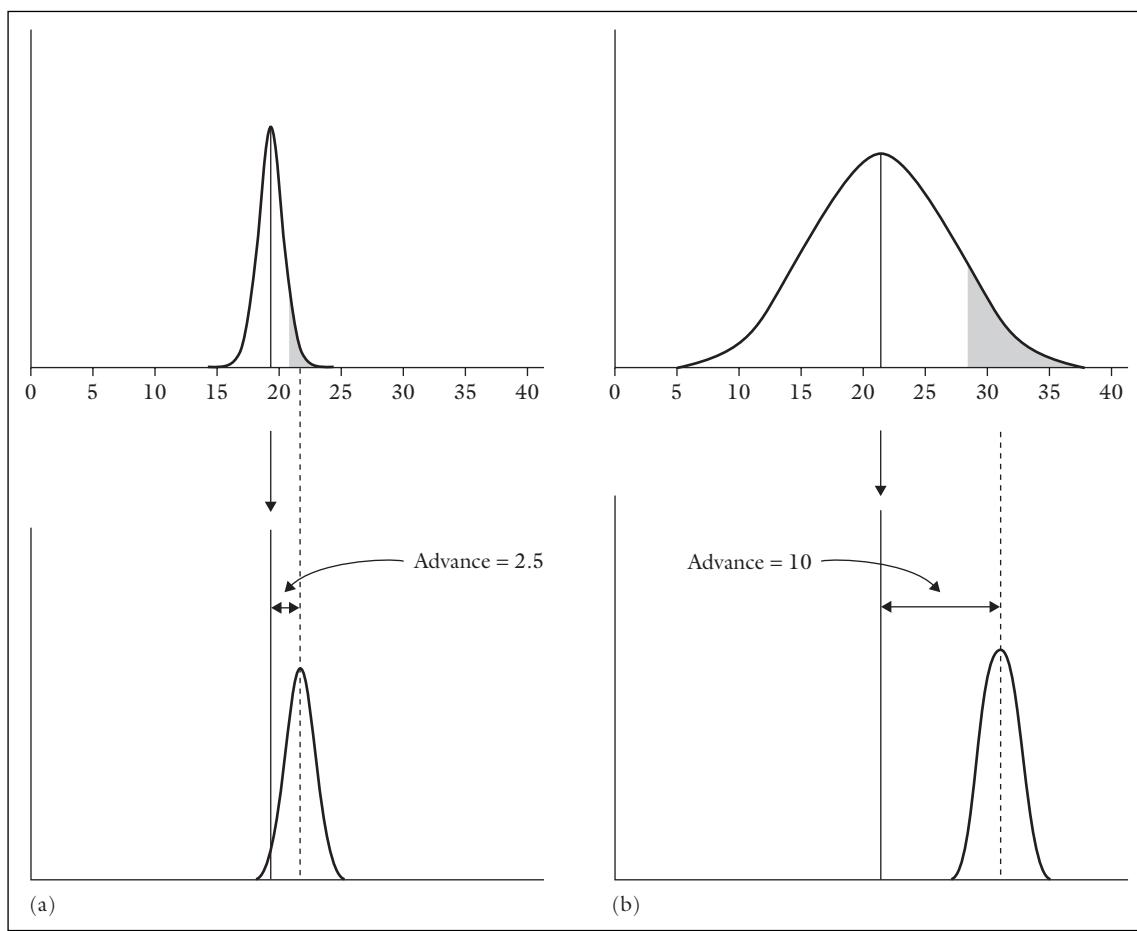
The response in one generation may be mathematically expressed as:

$$\bar{X}_o - \bar{X}_p = R = ih^2\sigma \text{ (or } \Delta G = ih^2\sigma_p)$$

where  $\bar{X}_o$  = mean phenotype of the offspring of selected parents,  $\bar{X}_p$  = mean phenotype of the whole parental generation,  $R$  = advance in one generation of selection,  $h^2$  = heritability,  $\sigma_p$  = phenotypic standard deviation of the parental population,  $i$  = intensity of selection, and  $\Delta G$  = genetic gain or genetic advance.

This equation has been suggested by some to be one of the fundamental equations of plant breeding, which must be understood by all breeders, and hence is called the **breeders' equation**. The equation is graphically illustrated in Figure 8.4. The factor "i", the intensity of selection, is a statistical factor that depends on the fraction of the current population retained to be used as parents for the next generation. The breeder may consult statistical tables for specific values (e.g., at 1%  $i = 2.668$ ; at 5%  $i = 2.06$ ; at 10%  $i = 1.755$ ). The breeder must decide the selection intensity to achieve a desired objective. The selection differential can be predicted if the phenotypic values of the trait of interest are normally distributed, and the selection is by truncation (i.e., the individuals are selected solely in order of merit according to their phenotypic value – no individual being selected is less good than any of those rejected).

The response equation is effective in predicting response to selection, provided the heritability estimate ( $h^2$ ) is fairly accurate. In terms of practical breeding, the parameters for the response equation are seldom available and hence not widely used. Over the long haul, repeated selection tends to fix favorable genes, resulting



**Figure 8.3** The effect of phenotypic variance on genetic advance. (a) If the phenotypic variance is too small, the genetic variability from which to select will be limited, resulting in a smaller genetic gain. (b) The reverse is true when the phenotypic variance is large.

in a decline in both heritability and phenotypic standard deviation. Once genes have been fixed, there will be no further response to selection.

**Example** For example:

	$X$	$\sigma_p$	$V_p$	$V_A$	$V_E$
Parents	15	2	6	4	3
Offspring	20.2	15	4.3	2.5	3

$$R = ih^2\sigma_p$$

*Parents*

$$\begin{aligned} h^2 &= V_A/V_p \\ &= 4/6 \\ &= 0.67 \end{aligned}$$

for  $i$  at  $P = 10\% = 1.755$  (read from tables and assuming a very large population).

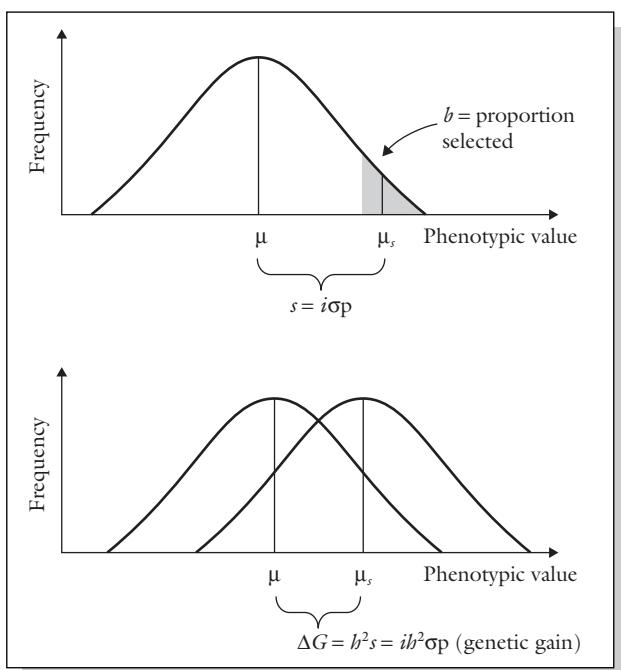
$$\begin{aligned} R &= 1.755 \times 0.67 \times 2 \\ &= 2.35 \end{aligned}$$

*Offspring*

$$\begin{aligned} h^2 &= V_A/V_p \\ &= 2.5/4.3 \\ &= 0.58 \end{aligned}$$

$$\begin{aligned} R &= 1.755 \times 0.58 \times 1.5 \\ &= 1.53 \end{aligned}$$

Generally, as selection advances to higher generations, genetic variance and heritability decline. Similarly, the advance from one generation to the next declines, while the mean value of the trait being improved increases.



**Figure 8.4** Genetic gain or genetic advance from selection indicates the progress plant breeders make from one generation to another based on the selection decisions they make.

### Concept of correlated response

Correlation is a measure of the degree of association between traits as previously discussed. This association may be on the basis of genetics or may be non-genetic. In terms of response to selection, genetic correlation is what is useful. When it exists, selection for one trait will cause a corresponding change in other traits that are correlated. This response to change by genetic association is called **correlated response**. Correlated response may be caused by pleiotropism or linkage disequilibrium. Pleiotropism is the multiple effect of a single gene (i.e., a single gene simultaneously affects several physiological pathways). In a random mating population, the role of linkage disequilibrium in correlated response is only important if the traits of interest are closely linked.

In calculating correlated response, genetic correlations should be used. However, the breeder often has access to phenotypic correlation and can use them if they were estimated from values averaged over several environments. Such data tend to be in agreement with genetic correlation. In a breeding program the breeder, even while selecting simultaneously for multiple traits,

has a primary trait of interest and secondary traits. The correlated response (CRy) to selection in the primary trait (y) for a secondary trait (x) is given by:

$$\text{CRy} = i_x b_x b_y \rho_g \sqrt{V_{Py}}$$

where  $b_x$  and  $b_y$  are square roots of the heritabilities of the two respective traits, and  $\rho_g$  is the genetic correlation between traits. This relationship may be reduced to:

$$\text{CRy} = i_x \rho_g b_x \sqrt{V_{Gy}}$$

$$\text{since } b_y = \sqrt{(V_{Gy}/V_{Py})}$$

It is clear that the effectiveness of indirect selection depends on the magnitude of genetic correlation and the heritability of the secondary traits being selected. Further, given the same selection intensity and a high genetic correlation between the traits, indirect selection for the primary trait will be more effective than directional selection, if heritability of the secondary trait is high ( $\rho_g b_x > b_y$ ). Such a scenario would occur when the secondary trait is less sensitive to environmental change (or can be measured under controlled conditions). Also, when the secondary trait is easier and more economic to measure, the breeder may apply a higher selection pressure to it.

Correlated response has wider breeding application in homozygous, self-fertilizing species and apomicts. Additive genetic correlation is important in selection in plant breeding. As previously discussed, the additive breeding value is what is transferred to offspring and can be changed by selection. Hence, where traits are additively genetically correlated, selection for one trait will produce a correlated response in another.

### Selection for multiple traits

Plant breeders may use one of three basic strategies to simultaneously select multiple traits: **tandem selection**, **independent curling**, and **selection index**. Plant breeders often handle very large numbers of plants in a segregating population using limited resources (time, space, labor, money, etc.). Along with the large number of individuals are the many breeding characters often considered in a breeding program. The sooner they can reduce the numbers of plants to the barest minimum, but more importantly, to the most desirable and promising individuals, the better. Highly heritable and readily scorable traits are easier to select for in the initial stages of a breeding program.

### Tandem selection

In this mode of selection, the breeder focuses on one trait at a time (serial improvement). One trait is selected for several generations, then another trait is focused on for the next period. The question of how long each trait is selected for before a switch and at what selection intensity, are major considerations for the breeder. It is effective when genetic correlation does not exist between the traits of interest, or when the relative importance of each trait changes throughout the years.

### Independent curling

Also called truncation selection, independent curling entails selecting for multiple traits in one generation. For example, for three traits, A, B, and C, the breeder may select 50% plants per family for A on phenotypic basis, and from that group select 40% plants per family based on trait B; finally, from that subset, 50% plants per family are selected for trait C, giving a total of 10% selection intensity ( $0.5 \times 0.4 \times 0.5$ ).

### Selection index

A breeder has a specific objective for conducting a breeding project. However, selection is seldom made on the basis of one trait alone. For example, if the breeding project is for disease resistance, the objective will be to select a genotype that combines disease resistance with the qualities of the elite adapted cultivar. Invariably, breeders usually practice selection on several traits, simultaneously. The problem with this approach is that as more traits are selected for, the less the selection pressure that can be exerted on any one trait. Therefore, the breeder should select on the basis of two or three traits of the highest economic value. It is conceivable that a trait of high merit may be associated with other traits of less economic value. Hence, using the concept of selection on total merit, the breeder would make certain compromises, selecting individuals that may not have been selected if the choice was based on a single trait.

In selecting on a multivariate phenotype, the breeder explicitly or implicitly assigns a weighting scheme to each trait, resulting in the creation of a univariate trait (an **index**) that is then selected. The index is the best linear prediction of an individual's breeding value. It takes the form of a multiple regression of breeding values on all the sources of information available for the population.

The methods used for constructing an index usually include heritability estimates, the relative economic

importance of each trait, and genetic and phenotypic correlation between the traits. The most common index is a linear combination that is mathematically expressed as follows:

$$I = \sum_{i=1}^m b_i z_i = b^T z$$

where  $z$  = vector of phenotypic values in an individual, and  $b$  = vector of weights. For three traits, the form may be:

$$I = aA^1 + bB^1 + cC^1$$

where  $a$ ,  $b$ , and  $c$  are coefficients correcting for relative heritability and the relative economic importance of traits A, B, and C, respectively, and  $A^1$ ,  $B^1$ , and  $C^1$  are the numerical values of traits A, B, and C expressed in standardized form. A standardized variable ( $X^1$ ) is calculated as:

$$X^1 = (X - \bar{X}) / \sigma_x$$

where  $X$  = record of performance made by an individual,  $\bar{X}$  = average performance of the population, and  $\sigma_x$  = standard deviation of the trait.

The classic selection index has the following form:

$$I = b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_m x_n$$

where  $x_1$ ,  $x_2$ ,  $x_3$ , to  $x_n$  are the phenotypic performance of the traits of interest, and  $b_1$ ,  $b_2$ , and  $b_3$  are the relative weights attached to the respective traits. The weights could be simply the respective relative economic importance of each trait, with the resulting index called the **basic index**, and may be used in cultivar assessment in official registration trials.

An index by itself is meaningless, unless it is used in comparing several individuals on a relative basis. Further, in comparing different traits, the breeder is faced with the fact that the mean and variability of each trait is different, and frequently, the traits are measured in different units. Standardization of variables resolves this problem.

### Concept of intuitive index

Plant breeding was described in Chapter 2 as both a science and an art. Experience (with the crop, the methods of breeding, breeding issues concerning the crop) is advantageous in having success in solving plant breeding problems. Plant breeders, as previously indicated, often



## Industry highlights

### Recurrent selection with soybean

Joe W. Burton

USDA Plant Science Building, 3127 Ligon Street, Raleigh, NC 27607, USA

#### Selection using a restricted index

Two commodities, protein meal and oil, are produced from soybean (*Glycine max* (L.) Merr.) and give the crop its value. Soybean seeds are crushed, oil is extracted, and protein meal is what remains. On a dry weight basis, soybeans are approximately 20% oil and 40% protein. Concentration of protein in the meal is dependant on protein concentration in soybean seeds. Protein meal is traded either as 44% protein or 48% protein. The 48% protein meal is more valuable, so increasing or maintaining protein concentration in soybean seeds has been a breeding objective. Protein is negatively associated with oil in seeds and in many breeding populations it is negatively associated with seed yield (Brim & Burton 1979).

The negative association between yield and protein could be due to genetic linkage as well as physiological processes (Carter et al. 1982). Thus a breeding strategy is needed that permits simultaneous selection of both protein and yield. Increased genetic recombination should also be helpful in breaking unfavorable linkages between genes that contribute to the negative yield and protein relation. We devised a recurrent  $S_1$  family selection program to satisfy the second objective and applied a restricted index to family performance to achieve the first objective.

#### Selection procedure

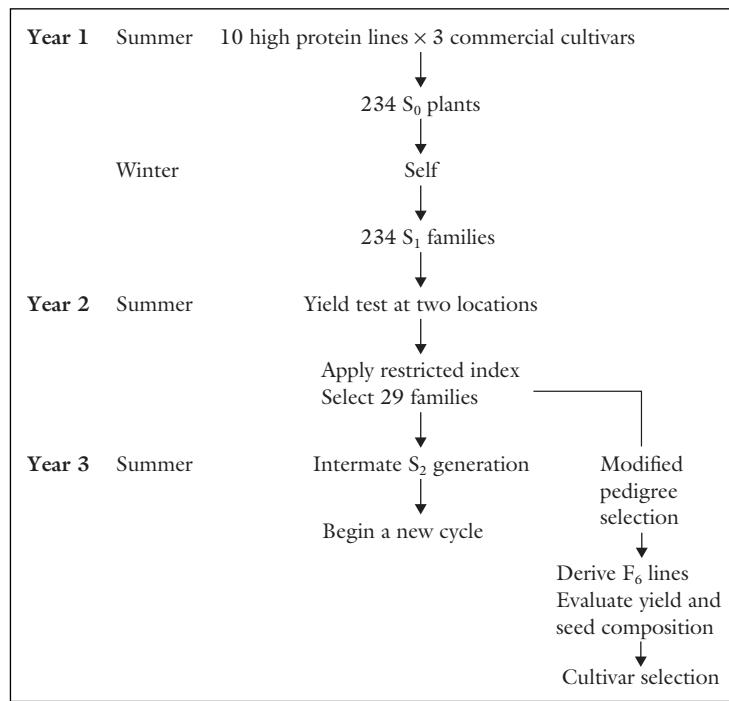
A population designated RS4 was developed using both high-yielding and high protein parents. The high-yielding parents were the cultivars, "Bragg", "Ransom", and "Davis". The high protein parents were 10  $F_3$  lines from cycle 7 of another recurrent selection population designated IA (Brim & Burton 1979). In that population, selection had been solely for protein. Average protein concentration of the 10 parental  $F_3$  lines was 48.0%. The base or  $C_0$  population was developed by making seven or eight matings between each high protein line and the three cultivars, resulting in 234 hybrids (Figure 1).

The  $S_0$  generation was advanced at the US Department of Agriculture (USDA) winter soybean nursery in Puerto Rico resulting in 234  $S_1$  families. These were tested in two replications at two locations. Both seed yield and protein concentration were determined for each family. Average protein concentration of the initial population was 45.6%. As this was an acceptable increase in protein, a restricted selection index was applied aimed at increasing yield and holding protein constant. This index was:

$$I = \text{yield} - (\sigma_{Gyp}/\sigma_{Gp}^2) \times \text{protein}$$

where  $\sigma_{Gyp}$  = estimated genetic covariance between yield and protein, and  $\sigma_{Gp}^2$  = estimated genetic variance of protein (Holbrook et al. 1989). Using this index, 29 families were selected.

The following summer, these 29 families (now in the  $S_2$  generation) were randomly intermated. To do this, we used the following procedure. Each day of the week, flowers for pollen were collected from 24 of the families and used to pollinate the remaining five families. A different set of 24 and five families were used as males and females, respectively, each



**Figure 1** Recurrent  $S_1$  family selection for yield and seed protein concentration using a restricted index.

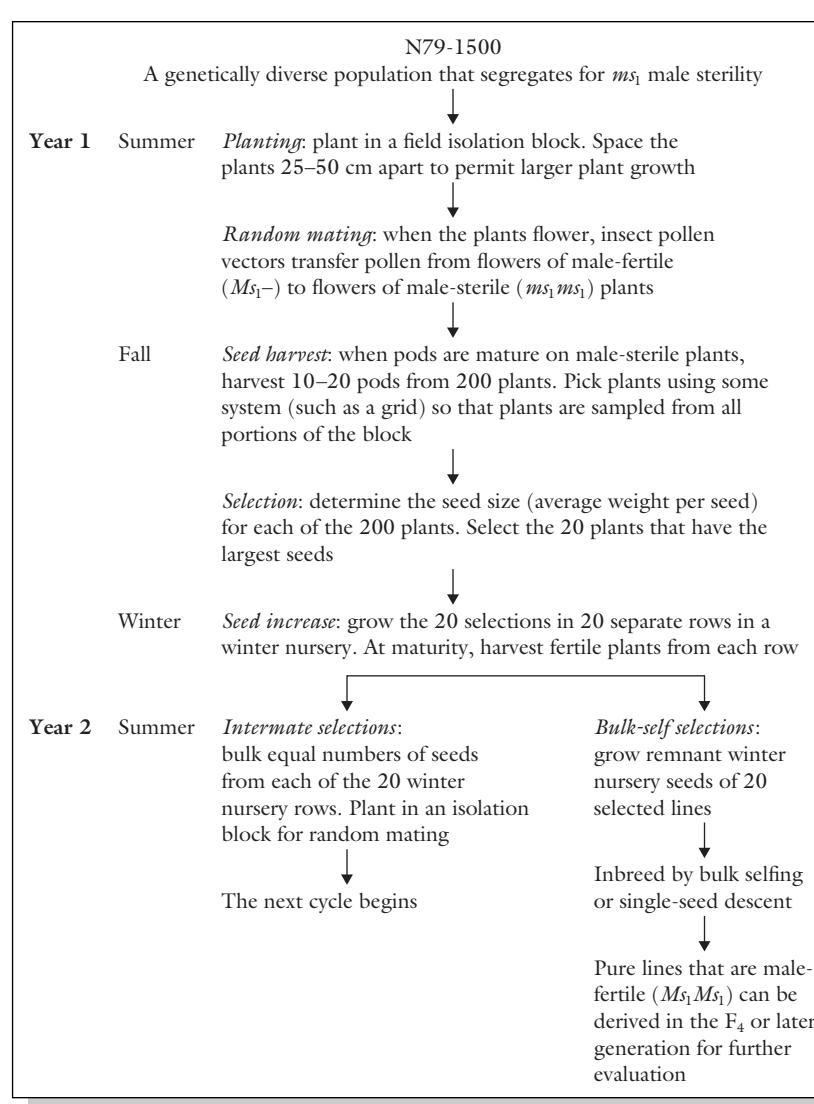
day. This process was followed until each family had at least seven successful pollinations on seven different plants within each family. These were advanced in the winter nursery to generate the  $S_1$  families for the next cycle of selection.

### **Development of “Prolina” soybean**

Modified pedigree selection was applied to the  $S_1$  families chosen in the first restricted index selection cycle.  $F_6$  lines were tested in replicated yield tests. One of those lines, N87-984, had good yielding ability and 45% seed protein concentration. Because of heterogeneity for plant height within the line,  $F_9$  lines were derived from N87-984 using single-seed descent. These were yield tested in multiple North Carolina locations. The two lines most desirable in terms of uniformity, protein concentration, and seed yield, were bulked for further testing and eventual release as the cultivar “Prolina” (Burton et al. 1999). At its release, “Prolina” had 45% protein compared with 42.7% for the check cultivar, “Centennial”, and similar yielding ability.

### **Recurrent selection using male sterility**

In the previous example, intermating the selections was done using hand pollinations. Hand pollination with soybean is time-consuming and difficult. The average success rate in our program during the August pollinating season has been 35%. Thus, a more efficient method for recombination would be helpful in a recurrent selection program that depends on good random mating among selected progeny for genetic recombination and reselection.



Genetic (nuclear) male sterility has been used for this purpose. Several nuclear male-sterile alleles have been identified (Palmer et al. 2004). The first male-sterile allele to be discovered ( $ms_1$ ) is completely recessive (Brim & Young 1971) to the male-fertility allele ( $Ms_1$ ). Brim and Stuber (1973) described ways that it could be used in recurrent selection programs. Plants that are homozygous for the  $ms_1$  allele are completely male-sterile. All seeds produced on male-sterile plants result from pollen contributed by a male-fertile plant ( $Ms_1Ms_1$  or  $Ms_1ms_1$ ) via an insect pollen vector. The  $ms_1ms_1$  male-sterile plants are also partially female-sterile, so that seed set on male-sterile plants is low in number, averaging about 35 seeds per plant. In addition, most pods have only one seed and that seed is larger (30–40% larger) than seeds that would develop on a fertile plant with a similar genetic background. The  $ms_1$  allele is maintained in a line that is 50%  $ms_1ms_1$  and 50%  $Ms_1ms_1$ . This line is planted in an isolation block. One-half of the pollen from male-fertile plants carries the  $Ms_1$  fertile allele and one-half carries the  $ms_1$  sterile allele. Male-sterile plants are pollinated by insect vectors, usually various bee species. At maturity, only seeds of male-sterile plants are harvested. These occur in the expected genotypic ratio of  $\frac{1}{2}Ms_1ms_1 : \frac{1}{2}ms_1ms_1$ .

**Figure 2** Recurrent mass selection for seed size in soybean using nuclear male sterility to intermate selections.

One of the phenotypic consequences of *ms<sub>1</sub>*, male sterility and low seed set is incomplete senescence. At maturity, soybeans normally turn yellow, leaves abscise, and the pods and seeds dry. Seed and pods on male-sterile plants mature and dry normally, but the plants remain green and leaves do not abscise. Thus, they are easily distinguished from male-fertile plants.

To use nuclear male sterility in a recurrent selection experiment, a population is developed for improvement that segregates for one of the recessive male-sterile alleles. This can be accomplished in a number of ways depending on breeding objectives. Usually, a group of parents with desirable genes are mated to male-sterile genotypes. This can be followed by one or more back-crosses. Eventually, an F<sub>2</sub> generation that segregates for male sterility is allowed to randomly intermate. Seeds are harvested from male-sterile plants. Then several different selection units are possible. These include the male-sterile plant itself (Tinius et al. 1991); the seeds (plants) from a single male-sterile plant (a half-sib family) (Burton & Carver 1993); and selfed seeds (plants) of an individual from a male-sterile plant (S<sub>1</sub> family) (Burton et al. 1990). Selection can be among and/or within the families (Carver et al. 1986). If appropriate markers are employed, half-sib selection using a tester is also possible (Feng et al. 2004). As with all recurrent selection schemes, selected individuals are intermated. These can be either remnant seed of the selection unit or progeny of the selection unit. The male-sterile allele segregate in both because both were derived in some manner from a single male-sterile plant.

### **Recurrent mass selection for seed size**

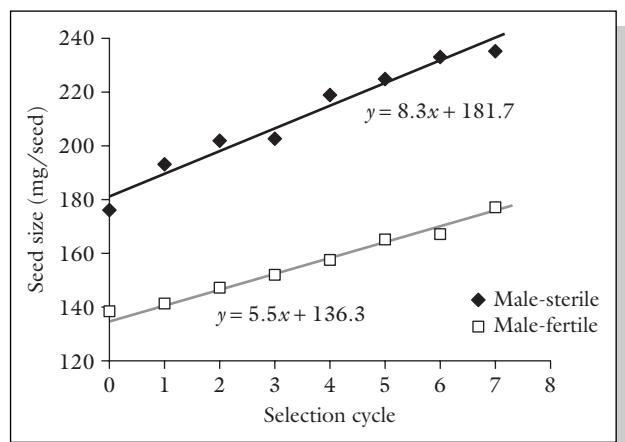
Because seed set on male-sterile plants is generally low in number, we hypothesized that size of the seed was not limited by source (photosynthate) inputs. Thus selecting male-sterile plants with the largest seeds would be selecting plants with the most genetic potential for producing large seeds. If so, this would mean that male-fertile plants derived from those selections would also produce larger seeds and perhaps have increased potential for overall seed yield.

To test this hypothesis, we conducted recurrent mass selection for seed size (mg/seed) in a population, N80-1500, that segregated for the *ms<sub>1</sub>* male-sterile allele and had been derived from adapted high-yielding cultivar and breeding lines (Burton & Brim 1981). The population was planted in an isolation block. Intermating between male-sterile and male-fertile plants occurred at random. In North Carolina there are numerous wild insect pollen vectors so introduction of domestic bees was not needed. If needed, bee hives can be placed in or near the isolation block. At maturity, seeds were harvested from approximately 200 male-sterile plants. To make sure that the entire population was sampled, the block was divided into sections, and equal numbers of plants were sampled from each section. Seeds from each plant were counted and weighed. The 20 plants with the largest seeds (greatest mass) were selected. These 20 selections were grown in a winter nursery and bulk-selfed to increase seed numbers. Equal numbers of seeds from the 20 selfed selections were combined and planted in another isolation block the following summer to begin another selection cycle (Figure 2).

With this method, one cycle of selection is completed each year. This is mass selection where only the female parent is selected. Additionally the female parents all have an inbreeding coefficient of 0.5 because of the selfing seed increase during the winter. Thus the expected genetic gain ( $\Delta_G$ ) for this selection scheme is:

$$\Delta_G = S(0.75)\sigma_A^2(\sigma_p^2)^{-1}$$

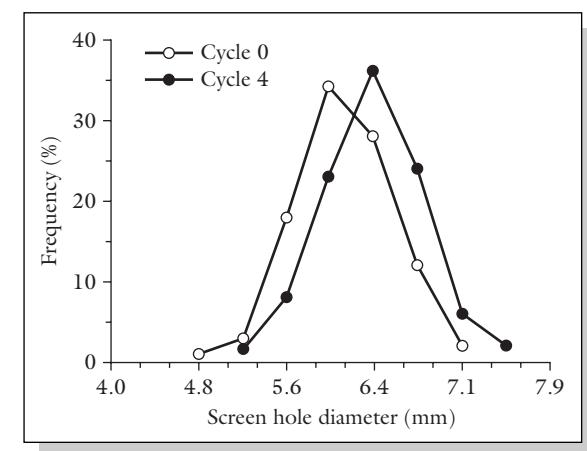
where  $S$  = selection differential,  $\sigma_A^2$  = additive genetic variance, and  $\sigma_p^2$  = phenotypic variance. This method was also used to increase oleic acid concentration in seed lipids (Carver et al. 1986).



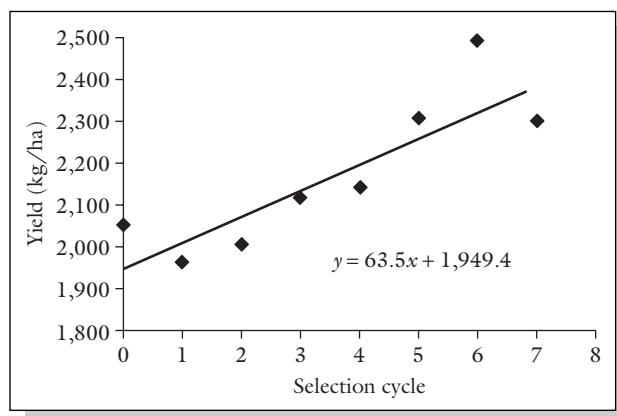
**Figure 3** Seed size changes with each selection for male-sterile and male-fertile soybeans.

At the end of cycle 4 and cycle 7, selected materials from each cycle were evaluated in replicated field trials. Results of those trials showed that the method had successfully increased both seed size and yield in the population. In seven cycles of selection, seed size of the male-sterile plants increased linearly from 182 to 235 mg/seed. Male-fertile seed size also increased linearly from 138 to 177 mg/seed (Figure 3). Not only the mass, but the physical size of the seeds increased. The range in seed diameter initially was 4.8 to 7.1 mm. After four cycles of selection, the diameter range had shifted and was 5.2 to 7.5 mm (Figure 4). Yield increased at an average rate of 63.5 kg/ha each cycle (Figure 5) or about 15% overall. There was some indication that after cycle 5 changes in yield were leveling off as yields of selections from cycle 5 and cycle 7 were very similar.

This method is relatively inexpensive. Little field space is required, and only a balance is needed to determine which individual should be selected. The ability to complete one cycle each year also makes it efficient. The largest expense is probably that needed to increase the seeds from selected male-sterile plants in a winter greenhouse or nursery. The



**Figure 4** Distribution of seed diameters initially, and after four cycles of selection, for larger seeds.



**Figure 5** Correlated changes in seed yield with selection for increased seed size.

method may be quite useful for introgressing unadapted germplasm into an adapted breeding population, followed by rapid improvement of productivity. The population could be sampled in any cycle using single-seed descent. Pure lines developed from these populations would be handled exactly as those developed from single crosses in typical modified pedigree selection programs.

### References

- Brim, C.A., and J.W. Burton. 1979. Recurrent selection in soybeans: II. Selection for increased protein in seeds. *Crop Sci.* 19:494–498.
- Brim, C.A., and C.W. Stuber. 1973. Application of genetic male sterility to recurrent selection schemes in soybeans. *Crop Sci.* 13:528–530.
- Brim, C.A., and M.F. Young. 1971. Inheritance of a male-sterile character in soybeans. *Crop Sci.* 11:564–566.
- Burton, J.W., and C.A. Brim. 1981. Registration of two soybean germplasm populations. *Crop Sci.* 21:801.
- Burton, J.W., T.E. Carter Jr., and R.F. Wilson. 1999. Registration of "Prolina" soybean. *Crop Sci.* 39:294–295.
- Burton, J.W., and B.F. Carver. 1993. Selection among S1 families vs. selfed half-sib and full-sib families in autogamous crops. *Crop Sci.* 33:21–28.
- Burton, J.W., E.M.K. Koinange, and C.A. Brim. 1990. Recurrent selfed progeny selection for yield in soybean using genetic male sterility. *Crop Sci.* 30:1222–1226.
- Carter, T.E., Jr., J.W. Burton, and C.A. Brim. 1982. Recurrent selection for percent protein in soybean seed – Indirect effects on plant N accumulation and distribution. *Crop Sci.* 22:513–519.
- Carver, B.F., J.W. Burton, T.E. Carter Jr., and R.F. Wilson. 1986. Cumulative response to various recurrent selection schemes in soybean oil quality and correlate agronomic traits. *Crop Sci.* 26:853–858.
- Feng, L., J.W. Burton, T.E. Carter Jr., and V.R. Pantalone. 2004. Recurrent half-sib selection with test cross evaluation for increased oil content in soybean. *Crop Sci.* 44:63–69.
- Holbrook, C.C., J.W. Burton, and T.E. Carter Jr. 1989. Evaluation of recurrent restricted index selection for increasing yield while holding seed protein constant in soybean. *Crop Sci.* 29:324–329.
- Palmer, R.G., T.W. Pfeiffer, G.R. Buss, and T.C. Kilen. 2004. Quantitative genetics. In: *Soybeans, improvement, production, and uses*, 3rd edn (H.R. Boerma, and J.E. Specht, eds), pp. 137–233. Agronomy Monograph No. 16. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Madison, WI.
- Tinius, C.N., J.W. Burton, and T.E. Carter Jr. 1991. Recurrent selection for seed size in soybeans. I. Response to selection in replicate populations. *Crop Sci.* 31:1137–1141.

must evaluate many plant characters in a breeding program. Whereas one or a few would be identified as key characters and focused on in a breeding program, breeders are concerned about the overall performance of the

cultivar. During selection, breeders formulate a mental picture of the product desired from the project, and balance good qualities against moderate defects as they make final judgments in the selection process.

Explicit indices are laborious, requiring the breeder to commit to extensive record-keeping and statistical analysis. Most breeders use a combination of truncation selection and intuitive selection index in their programs.

### **Concept of general worth**

For each crop, there are a number of characters, which considered together, define the overall desirability of the cultivar from the combined perspectives of the producer and the consumer. These characters may range between about a dozen to several dozens, depending on the crop, and constitute the primary pool of characters that the breeder may target for improvement. These characters differ in importance (economic and agronomic) as well as ease with which they can be manipulated through breeding. Plant breeders typically target one or few of these traits for direct improvement in a breeding program. That is, the breeder draws up a working list of characters to address the needs embodied in the stated objectives. Yield of the economic product is almost universally the top priority in a plant breeding program. Disease resistance is more of a local issue, since what may be economically important in one region may not be important in another area. Even though a plant breeder may focus on one or a few traits at a time, the ultimate objective is the improvement of the totality of the key traits that impact the overall desirability or general worth of the crop. In other words, breeders ultimately have a holistic approach to selection in a breeding program. The final judgments are made on a balanced view of the essential traits of the crop.

### **Nature of breeding characteristics and their levels of expression**

Apart from relative importance, the traits the plant breeder targets vary in other ways. Some are readily evaluated by visual examination (e.g., shape, color, size), whereas others require a laboratory assay (e.g., oil content) or mechanical measurement (e.g., fiber characteristics of cotton). Special provisions (e.g., greenhouse, isolation block) may be required in disease breeding, whereas yield evaluations are most reliable when conducted over seasons and locations in the field.

In addition to choosing the target traits, the breeder will have to decide on the level of expression of each one, below which a plant material would be declared worthless. The acceptability level of expression of a trait

may be narrowly defined (stringent selection) or broadly defined (loose selection). In industrial crops (e.g., cotton), the product quality may be strictly defined (e.g., a certain specific gravity, optimum length). In disease-resistance breeding, there may not be a significant advantage of selecting for extreme resistance over selecting for less than complete resistance. On the other hand, in breeding nutritional quality, there may be legal guidelines as to threshold expression for toxic substances.

### **Early generation testing**

Early generation testing is a selection procedure in which the breeder initiates testing of genetically heterogeneous lines or families in an earlier than normal generation. In Chapter 17, recurrent selection with testers was used to evaluate materials in early generations. A major consideration of the breeder in selecting a particular breeding method is to maximize genetic gain per year. Testing early, if effective, helps to identify and select potential cultivars from superior families in the early phase of the breeding program. The early generation selection method has been favorably compared with other methods such as pedigree selection, single-seed descent, and bulk breeding. The question of how early the test is implemented often arises. Should it be in the  $F_1$ -,  $F_2$ - or  $F_3$ -derived families? Factors to consider in deciding on the generation in which selection is done include the trait being improved, and the availability of off-season nurseries to use in producing additional generations per year (in lieu of selecting early).

### **Concept of combining ability**

Over the years, plant breeders have sought ways of facilitating plant breeding through the efficient selection of parents for a cross, effective and efficient selection within a segregating population, and prediction of response to selection, among other needs. Quantitative assessment of the role of genetics in plant breeding entails the use of statistical genetics approaches to estimate variances and to partition them into components, as previously discussed. Because variance estimates are neither robust nor accurate, the direct benefits of statistical genetics to the breeder have been limited.

In 1942, Sprague and Tatum proposed a method of evaluation of inbred lines to be used in corn hybrid production that was free of the genetic assumptions that accompany variance estimates. Called **combining**

**ability**, the procedure entails the evaluation of a set of crosses among selected parents to ascertain the extent to which variances among crosses are attributable to statistically additive characteristics of the parents, and what could be considered the effect of residual interactions. Crossing each line with several other lines produces an additional measure in the mean performance of each line in all crosses. This mean performance of a line, when expressed as a deviation from the mean of all crosses, gives what Sprague and Tatum called the **general combining ability (GCA)** of the lines.

The GCA is calculated as the average of all  $F_1$ s having this particular line as one parent, the value being expressed as a deviation from the overall mean of crosses. Each cross has an expected value (the sum of GCAs of its two parental lines). However, each cross may deviate from the expected value to a greater or lesser extent, the deviation being the **specific combining ability (SCA)** of the two lines in combination. The differences of GCA are due to the additive and additive  $\times$  additive interactions in the base population. The differences in SCA are attributable to non-additive genetic variance. Further, the SCA is expected to increase in variance more rapidly as inbreeding in the population reaches high levels. The GCA is the average performance of a plant in a cross with different tester lines, while the SCA measures the performance of a plant in a specific combination in comparison with other cross combinations.

The mathematical representation of this relationship for each cross is:

$$X_{\text{AP}} = \bar{X} + G_A + G_P + S_{\text{AP}}$$

where  $\bar{X}$  is the general mean and  $G_A$  and  $G_B$  are the general combining ability estimates of the parents, and  $S_{AB}$  is the statistically unaccounted for residual or specific combining ability. The types of interactions that can be obtained depend upon the mating scheme used to produce the crosses, the most common being the diallel mating design (full or partial diallel).

Plant breeders may use a variety of methods for estimating combining abilities, including the polycross and topcrossing methods. However, the diallel cross (each line is mated with every other line) developed by B. Griffing in 1956 is perhaps the most commonly used method. The GCA of each line is calculated as follows:

$$Gx = [Tx/(n-2)] - [\sum T/n(n-2)]$$

where  $x$  represents a specific line. Using the data in Table 8.4,  $G_A$  can be calculated as:

$$\begin{aligned}
 G_A &= [T_A/(n-2)] - [\sum T/n(n-2)] \\
 &= [226/(10-2)] - [2,024/10(10-2)] \\
 &= 28.25 - 25.3 \\
 &= 2.95
 \end{aligned}$$

The others may be calculated as for line A. The next step is to calculate the expected value of each cross. Using the cross CD as an example, the expected value is calculated as:

$$E(X_{CP}) = -4.18 + 5.33 + 22.49 = 23.64$$

The SCA is calculated as follows:

$$\text{SCA}_{\text{CD}} = 26 - 23.64 \\ = 2.36$$

**Table 8.4** Calculating general and specific combining abilities.

This is done for each combination and a plot of observed values versus expected values plotted. Because the values of SCA are subject to sampling error, the points on the plot do not lie on the diagonal. The distance from each point to the diagonal represents the SCA plus sampling error of the cross. Additional error would occur if the lines used in the cross are not highly inbred (error due to the sampling of genotypes from the lines).

Combining ability calculations are statistically robust, being based on first-degree statistics (totals, means). No genetic assumptions are made about individuals. The concept is applicable to both self-pollinated and cross-pollinated species, for identifying desirable cross combinations of inbred lines to include in a hybrid program or for developing synthetic cultivars. It is used to predict the performance of hybrid populations of cross-pollinated species, usually via a testcross or poly-cross. It should be pointed out that combining ability calculations are properly applied only in the context in which they were calculated. This is because GCA values are relative and depend upon the mean of the chosen parent materials in the crosses.

A typical ANOVA for combining ability analysis is as follows:

Source	df	Sum of squares	Mean sum of squares	EMS
		(SS)	(MS)	
GCA	$p - 1$	$S_G$	$M_G$	$\sigma_E^2 + \sigma_{GCA}^2 + \sigma_{GCA}^2$
SCA	$p(p - 1)/2$	$S_S$	$M_S$	$\sigma_E^2 + \sigma_{SCA}^2$
Error	$m$	$S_E$	$M_E$	$\sigma_E^2$

The method used for a combining ability analysis depends on the available data:

- 1 Parents +  $F_1$  or  $F_2$  and reciprocal crosses (i.e.,  $p^2$  combinations).
- 2 Parents +  $F_1$  or  $F_2$ , without reciprocals (i.e.,  $\frac{1}{2}p(p+1)$  combinations).
- 3  $F_1 + F_2 +$  reciprocals, without parents and reciprocals (i.e.,  $\frac{1}{2}p(p-1)$  combinations).
- 4 Only  $F_1$  generations, without parents and reciprocals (i.e.,  $\frac{1}{2}p(p-1)$  combinations).

## Mating designs

Artificial crossing or mating is a common activity in plant breeding programs for generating various levels of relatedness among the progenies that are produced. Mating in breeding has two primary purposes:

- 1 To generate information for the breeder to understand the genetic control or behavior of the trait of interest.
- 2 To generate a base population to initiate a breeding program.

The breeder influences the outcome of a mating by the choice of parents, the control over the frequency with which each parent is involved in mating, and the number of offspring per mating, among other ways. A mating may be as simple as a cross between two parents, to the more complex diallel mating.

### Hybrid crosses

These are reviewed here to give the student a basis for comparison with the random mating schemes to be presented.

- 1 Single cross =  $A \times B \rightarrow F_1$  (AB)
- 2 Three-way cross =  $(A \times B) \rightarrow F_1 \times C \rightarrow (ABC)$
- 3 Backcross =  $(A \times B) \rightarrow F_1 \times A \rightarrow (BC_1)$
- 4 Double cross =  $(A \times B) \rightarrow F_{AB}; (C \times D) \rightarrow F_{CD}; F_{AB} \times F_{CD} \rightarrow (ABCD)$

These crosses are relatively easy to genetically analyze. The breeder exercises significant control over the mating structure.

### Mating designs for random mating populations

The term mating design is usually applied to schemes used by breeders and geneticists to impose random mating for a specific purpose. To use these designs, certain assumptions are made by the breeder:

- 1 The materials in the population have diploid behavior. However, polyploids that can exhibit disomic inheritance (allopoloids) can be studied.
- 2 The genes controlling the trait of interest are independently distributed among the parents (i.e., uncorrelated gene distribution).
- 3 The absence of: non-allelic interactions, reciprocal differences, multiple alleles at the loci controlling the trait, and  $G \times E$  interactions.

### Biparental mating (or paired crosses)

In this design, the breeder selects a large number of plants ( $n$ ) at random and crosses them in pairs to produce  $\frac{1}{2}n$  full-sib families. The biparental (BIP) is the simplest of the mating designs. If  $r$  plants per progeny

family are evaluated, the variation within ( $w$ ) and between ( $b$ ) families may be analyzed as follows:

Source	df	MS	EMS
Between families	$(\frac{1}{2}n) - 1$	$MS_1$	$\sigma^2 w + r\sigma^2 b$
Within families	$\frac{1}{2}n(r-1)$	$MS_2$	$\sigma^2 w$

where  $\sigma^2 b$  is the covariance of full sibs ( $\sigma^2 b = \frac{1}{2}V_A + \frac{1}{4}V_D + V_{EC} = 1/r(MS_1 - MS_2)$  and  $\sigma^2 w = \frac{1}{2}V_A + \frac{3}{4}V_D + V_{EW} = MS_2$ ).

The limitation of this otherwise simple to implement design is its inability to provide the needed information to estimate all the parameters required by the model. The progeny from the design comprise full sibs or unrelated individuals. There is no further relatedness among individuals in the progeny. The breeder must make unjustifiable assumptions in order to estimate the genetic and environmental variance.

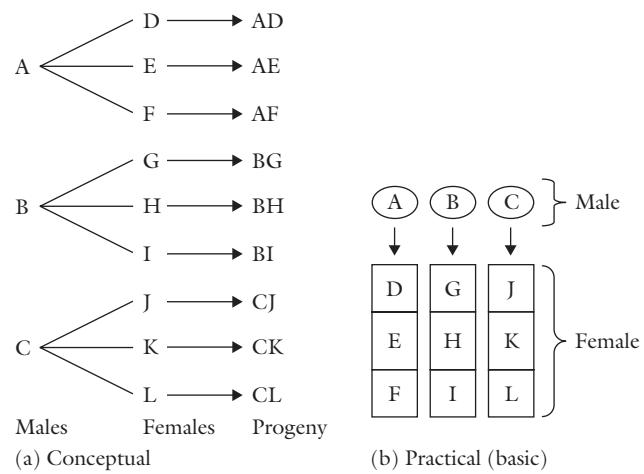
### Polycross

This design is for intermating a group of cultivars by natural crossing in an isolated block. It is most suited to species that are obligate cross-pollinators (e.g., forage grasses and legumes, sugarcane, sweet potato), but especially to those that can be vegetatively propagated. It provides an equal opportunity for each entry to be crossed with every other entry. It is critical that the entries be equally represented and randomly arranged in the crossing block. If 10 or less genotypes are involved, the Latin square design may be used. For a large number of entries, the completely randomized block design may be used. In both cases, about 20–30 replications are included in the crossing block. The ideal requirements are hard to meet in practice because of several problems, placing the system in jeopardy of deviating from random mating. If all the entries do not flower together, mating will not be random. To avoid this, the breeder may plant late flowering entries earlier.

Pollen may not be dispersed randomly, resulting in concentrations of common pollen in the crossing block. Half sibs are generated in a polycross because progeny from each entry has a common parent. The design is used in breeding to produce synthetic cultivars, recombining selected entries of families in recurrent selection breeding programs, or for evaluating the GCA of entries.

### North Carolina Design I

**Design I** is a very popular multipurpose design for both theoretical and practical plant breeding applications



**Figure 8.5** North Carolina Design I. (a) This design is a nested arrangement of genotypes for crossing in which no male is involved in more than one cross. (b) A practical layout in the field.

(Figure 8.5). It is commonly used to estimate additive and dominance variances as well as for the evaluation of full- and half-sib recurrent selection. It requires sufficient seed for replicated evaluation trials, and hence is not of practical application in breeding species that are not capable of producing large amounts of seed. It is applicable to both self- and cross-pollinated species that meet this criterion. As a nested design, each member of a group of parents used as males is mated to a different group of parents. NC Design I is a hierarchical design with non-common parents nested in common parents.

The total variance is partitioned as follows:

Source	df	MS	EMS
Males	$n - 1$	$MS_1$	$\sigma^2 w + r\sigma^2_{mf} + r\sigma^2_m$
Females	$n_1(n_2 - 1)$	$MS_2$	$\sigma^2 w + r\sigma^2_{mf}$
Within progenies	$n_1 n_2(r - 1)$	$MS_3$	$\sigma^2 w$

$$\sigma^2_m = [MS_1 - MS_2]/rn_2 = \frac{1}{4}V_A$$

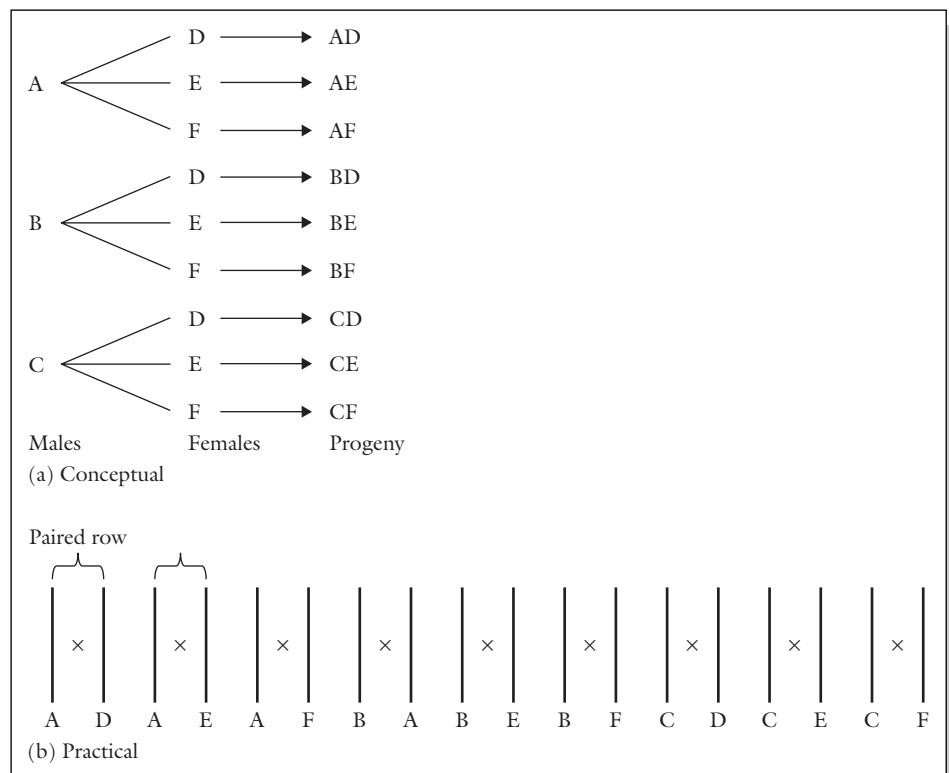
$$\sigma^2_{mf} = [MS_2 - MS_3]/r = \frac{1}{4}V_A + \frac{1}{4}V_D$$

$$\sigma^2 w = MS_3 = \frac{1}{2}V_A + \frac{3}{4}V_D + E$$

This design is most widely used in animal studies. In plants, it has been extensively used in maize breeding for estimating genetic variances.

### North Carolina Design II

In this design, each member of a group of parents used as males is mated to each member of another group of



**Figure 8.6** North Carolina Design II. (a) This is a factorial design. (b) Paired rows may be used in the nursery for factorial mating of plants.

parents used as females. **Design II** is a factorial mating scheme similar to Design I (Figure 8.6). It is used to evaluate inbred lines for combining ability. The design is most adapted to plants that have multiple flowers so that each plant can be used repeatedly as both male and female. Blocking is used in this design to allow all the mating involving a single group of males to a single group of females to be kept intact as a unit. The design is essentially a two-way ANOVA in which the variation may be partitioned into difference between males (m) and females (f) and their interaction. The ANOVA is as follows:

Source	df	MS	EMS
Males	$n_1 - 1$	$MS_1$	$\sigma^2_w + r\sigma_{mf}^2 + rn\sigma_m^2$
Females	$n_2 - 1$	$MS_2$	$\sigma^2_w + r\sigma_{mf}^2 + rn_1\sigma_f^2$
Males $\times$ females	$(n_1 - 1)(n_2 - 1)$	$MS_3$	$\sigma^2_w + r\sigma_{mf}^2$
Within progenies	$n_1 n_2(r - 1)$	$MS_4$	$\sigma^2_w$

$$\begin{aligned}\sigma_m^2 &= [MS_1 - MS_3]/rn_2 = \frac{1}{4}V_A \\ r\sigma_f^2 &= [MS_2 - MS_3]/rn_1 = \frac{1}{4}V_A \\ r\sigma_{mf}^2 &= [MS_3 - MS_4]/r = \frac{1}{4}V_D \\ \sigma^2_w &= MS_4 = \frac{1}{2}V_A + \frac{3}{4}V_D + E\end{aligned}$$

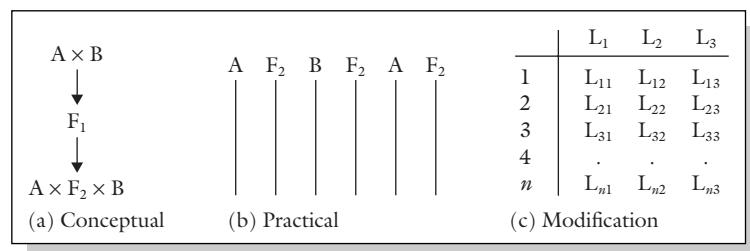
The design also allows the breeder to measure not only GCA but also SCA.

#### North Carolina Design III

In this design, a random sample of  $F_2$  plants is back-crossed to the two inbred lines from which the  $F_2$  was descended. It is considered the most powerful of all the three NC designs. However, it was made more powerful by the modifications made by Kearsey and Jinks that adds a third tester (not just the two inbreds) (Figure 8.7). The modification is called the **triple testcross** and is capable of testing for non-allelic (epistatic) interactions, which the other designs cannot, and also capable of estimating additive and dominance variance.

#### Diallel cross

A **complete diallel mating** design is one that allows the parents to be crossed in all possible combinations, including selfs and reciprocals. This is the kind of mating scheme required to achieve Hardy-Weinberg



**Figure 8.7** North Carolina Design III. The conventional form (a), the practical layout (b), and the modification (c) are shown.

equilibrium (see Chapter 7) in a population. However, in practice, a diallel with selfs and reciprocals is neither practical nor useful for several reasons. Selfing does not contribute to the recombination of genes between parents. Furthermore, recombination is achieved by crossing in one direction making reciprocals unnecessary. Because of the extensive mating patterns, the number of parents that can be mated this way is limited. For  $p$  entries, a complete diallel will generate  $p^2$  crosses. Without selfs and reciprocals, the number is  $p(p - 1)/2$  crosses.

When the number of entries is large, a **partial diallel mating** design, which allows all parents to be mated to some but not all other parents in the set, is used. A diallel design is most commonly used to estimate combining abilities (both general and specific). It is also widely used for developing breeding populations for recurrent selection.

Nursery arrangements for the application of complete and partial diallel are varied. Because a large number of crosses are made, diallel mating takes a large amount of space, seed, labor, and time to conduct. Because all possible pairs are contained in one half of a symmetric Latin square, this design may be used to address some of the space needs.

There are four basic methods developed by Griffing that vary in either the omission of parents or the

omission of reciprocals in the crosses. The number of progeny families (pf) for methods 1 through 4 are: pf =  $n^2$ , pf =  $\frac{1}{2}n(n + 1)$ , pf =  $n(n - 1)$ , and pf =  $\frac{1}{2}n(n - 1)$ , respectively. The ANOVA for method 4, for example, is as follows:

Source	df	EMS
GCA	$n_1 - 1$	$\sigma^2 e + r\sigma_g^2 + r(n - 2)\sigma^2$
SCA	$[n(n - 3)]/2$	$\sigma^2 e + r\sigma_g^2$
Reps × crosses	$(r - 1)[(n(n - 1)/2) - 1]$	$\sigma^2 e$

#### Comparative evaluation of mating designs

Hill, Becker, and Tigerstedt roughly summarized these mating designs in two ways:

- 1 In terms of coverage of the population: BIPs > NC I > polycross > NC III > NC II > diallel, in that order of decreasing effectiveness.
- 2 In terms of amount of information: diallel > NC II > NC III > NC I > BIPs.

The diallel mating design is the most important for GCA and SCA. These researchers emphasized that it is not the mating design *per se*, but rather the breeder who breeds a new cultivar. The implication is that the proper choice and use of a mating design will provide the most valuable information for breeding.

#### References and suggested reading

- Ali, A., and D.L. Johnson. 2000. Heritability estimates for winter hardiness in lentil under natural and controlled conditions. *Plant Breed.* 119:283–285.
- Bhatnagar, S., F.J. Betran, and L.W. Rooney. 2004. Combining abilities of quality protein maize inbreds. *Crop Sci.* 44:1997–2005.
- Bohren, B.B., H.E. McKean, and Y. Yamada. 1961. Relative efficiencies of heritability estimates based on regression of offspring on parent. *Biometrics* 17:481–491.
- Comstock, R.E., H.F. Robinson, and P.H. Harvey. 1949. A breeding procedure designed to make maximum use of both general and specific combining ability. *J. Am. Soc. Agron.* 41:360–367.
- Edwards, J.W., and K.R. Lamkey. 2002. Quantitative genetics of inbreeding in a synthetic maize population. *Crop Sci.* 42:1094–1104.
- Falconer, D.S. 1981. *Introduction to quantitative genetics*. Longman Group, New York.

- Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to quantitative genetics, 4th edn. Longman, Harlow, UK.
- Gardner, C.O. 1977. Quantitative genetic studies and population improvement in maize and sorghum. In: Proceedings of the International Conference on Quantitative Genetics (Pollak, E., O. Kempthorne, and T.B. Bailey, eds). Iowa State University, Ames, IA.
- Glover, M.A., D.B. Willmot, L.L. Darrah, B.E. Hibbard, and X. Zhu. 2005. Diallel analysis of agronomic traits using Chinese and US maize germplasm. *Crop Sci.* 45:1096–1102.
- Griffing, B. 1956a. A generalized treatment of the use of diallel crosses in quantitative inheritance. *Heredity* 10:31–50.
- Griffing, B. 1956b. Concept of general and specific combining ability in relation to a diallel crossing system. *Aust. J. Biol. Sci.* 9:463–493.
- Henderson, C.R. 1963. Selection index and expected genetic advance. In: Statistical genetics and plant breeding (Hanson, W.D., and H.F. Robinson, eds). National Academy of Sciences and National Research Council Publication No. 982. National Academy of Sciences and National Research Council, Washington, DC.
- Hill, J., H.C. Becker, and P.M.A. Tigerstedt. 1998. Quantitative and ecological aspects of plant breeding. Chapman & Hall, London.
- Holland, J.B. 2001. Epistasis and plant breeding. *Plant Breed. Rev.* 21:27–92.
- Lin, C.Y. 1978. Index selection for genetic improvement of quantitative characters. *Theor. Appl. Genet.* 52:49–56.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Quantitative traits are more influenced by the environment than qualitative traits.
- 2 Quantitative traits are controlled by polygenes.
- 3 Heritability is a population phenomenon.
- 4 The specific combining ability of a trait depends on additive gene action.
- 5 Polygenes have distinct and distinguishable effects.
- 6 Quantitative variation deals with discrete phenotypic variation.
- 7 Quantitative traits are also called metric traits.

### Part B

Please answer the following questions:

- 1 What is quantitative genetics, and how does it differ from qualitative genetics?
- 2 Give two specific assumptions of quantitative genetic analysis.
- 3 Describe additive gene action.
- 4 What is the heritability of a trait?
- 5 What is the breeders' equation?

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the role of the environment in quantitative trait expression.
- 2 Discuss the concept of general worth of a plant.
- 3 Discuss the concept of intuitive selection.
- 4 Discuss the application of combining ability analysis in plant breeding.
- 5 Discuss a method of estimating heritability of a trait.



# 9

# Common statistical methods in plant breeding

---

## Purpose and expected outcomes

*Statistics is indispensable in plant breeding. Breeders conduct the bulk of their work in the field under variable environmental conditions that tend to mask real effects. Further, plant breeders often handle large amounts of data that need to be summarized in order to facilitate sound decision-making. Computer software of all kinds is available for use in plant breeding. The critical first thing is to know what statistical procedure to use to address a specific problem. After studying this chapter, the student should be able to:*

- 1 Describe the role of statistics in plant breeding.
  - 2 Discuss the measures of central tendency.
  - 3 Discuss the measures of dispersion.
  - 4 Discuss the measures of association.
  - 5 Discuss the method of analysis of variance.
  - 6 Discuss multivariate analyses in plant breeding.
  - 7 Discuss the concept of path analysis.
- 

## Role of statistics in plant breeding

The development of **statistics** arose out of a need to assist researchers in those areas where the laws of cause and effect are not apparent to the observer, and where an objective approach is needed. Plant breeders use statistics to design studies, analyze results, and draw sound conclusions. The role of statistics in plant breeding may be summarized in three key applications as follows:

- 1 **To obtain a descriptive summary of the sample.** Research data are often large, requiring some mathematical reduction to expose hidden trends for easy interpretation. The values obtained from such descriptive analysis are sometimes called **summary statistics**.

- 2 **To provide a means of statistical inference.** The key purpose of collecting data in research is to enable the researcher to draw some kind of inference about a certain characteristic of the population from which the data were drawn. To do this, the values obtained about the sample are used.
- 3 **Comparison.** Often, the researcher has multiple sets of experimental data and needs to know whether they represent significantly different populations of measurement. Another way of putting this is that the objectives of statistics are the estimation of population parameters and the testing of hypotheses about the parameters.

Statistical methods used in plant breeding can range from the simple and straightforward such as arithmetic

averages, to the more complex multivariate analysis. Computers are required for complex analyses, but sometimes, the breeder may have a small amount of data and might want to use a handheld calculator for quick results. Hence, there is the need to know the computational basis of the commonly used statistical methods.

### **Population versus sample**

A statistical **population** is the totality of the units (individuals) of interest to the researcher. It follows then that, depending on the researcher's objectives, a population may be small or infinitely large. A small population can be measured in its entirety. Plant breeders often handle large populations, and obtaining measurement from the entire population is often impractical. Instead, researchers obtain measurements from a subset of the population, called a **sample**. The scores from the sample are used to infer or estimate the scores we would expect to find if it were possible to measure the entire population.

In order to draw accurate conclusions about the population, the sample must be representative of the population. To obtain a representative sample, the statistical technique of **random sampling** (in which all possible scores in the population have an equal chance of being selected for a sample) is used. There are other methods of drawing samples from a population for a variety of purposes. These include quota, convenience, and stratified sampling methods. A number that describes a characteristic of a population is called a **sample statistic** or simply, a **statistic**. A number that describes a population is called a population parameter, or simply, a **parameter**.

### **Issue of causality**

Scientific conclusions are drawn from the preponderance of the evidence obtained from properly conducted research. Cause and effect is implicit in the logic of researchers. However, it is difficult to definitely prove that variable *X* causes variable *Y*. There is always the possibility that some unknown variable is actually responsible for the effect observed (change in scores). No statistical procedure will prove that one variable causes another variable to change (i.e., statistics does not prove anything!). An experiment provides evidence to argue for a certain point of view, not prove it.

### **Statistical hypothesis**

A **hypothesis** is an informed conjecture (educated guess) about a phenomenon. It is arrived at after taking into account pertinent scientific knowledge and personal experience. Researchers often have preconceived ideas about the phenomenon that they seek to investigate. However, they should be willing to approach an investigation with an open mind. A hypothesis declares the prediction of the researcher concerning the relationship between two or more variables associated with the study. An experiment is designed to test this relationship.

In plant breeding, a breeder ends up with about a dozen promising genotypes from which one would eventually be selected for release to farmers for use in cultivation. The breeder conducts field tests or trials (over locations and years) to help in the decision-making process. He or she suspects or predicts differences among these genotypes. The predicted difference represents a true phenomenon. To avoid any biases, the hypothesis is formulated in the opposite direction to the predicted or suspected outcome. That is, the researcher would state that no real differences exist among the genotypes (i.e., any differences are due to chance). This is the **null hypothesis** ( $H_0$ ) or the hypothesis of no difference. The **alternative hypothesis** ( $H_1$ ) would indicate a real difference exists. There is a standard way of mathematically stating a hypothesis. If four genotypes were being evaluated, a hypothesis could be formulated as follows:

Null hypothesis,  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$   
 (i.e., all genotype means are equal)  
 Alternate hypothesis,  $H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$   
 (i.e., all genotype means are not equal)

The  $H_0$  is accepted (i.e., automatically reject  $H_1$ ), or rejected (accept  $H_0$ ), at a chosen level of **statistical significance**,  $\alpha$  (e.g.,  $\alpha = 0.01$  or  $0.05$ ; acknowledging that 1% or 5% of the time you could be mistaken in your conclusion). In other words, the research does not prove anything outright, as previously pointed out. Rejecting  $H_0$  when it is true (i.e., you are saying a difference exists when in fact none really does) is called a **Type I error**. On the other hand, failure to find a true difference when it exists (accepting a false  $H_0$ ) is called a **Type II error**.

The goal of a plant breeder is to conduct research in such a way that true differences, when they exist, are observed. This depends on the adoption of sound experimental procedures, often called field plot techniques, and is discussed later in this chapter.

## Concept of statistical error

Error in statistics does not imply a mistake. As previously stated, statistics is not used to prove anything. Experimental conditions are seldom, if ever, perfect. If five samples of a uniform cultivar (e.g., pure line) are planted under identical conditions, it would be expected that a measurement of a trait (e.g., height) would be identical for all samples. In practice, differences, albeit minor, would be observed. This variation that cannot be accounted for is called **experimental error** (or simply **error**). No effort can completely eliminate experimental error. However, efforts can be made to reduce it such that true differences in a study are not obscured. Laboratory or controlled environment research often allows the researcher more effective control over variation in the experimental environment. Field studies are subject to significant variation from the soil as well as the above-ground environment. Other sources of undesirable variation are competition among plants and operator (human) error. Plant breeders need to understand the principles of experimental design. A large error would not permit small real differences in the experiment to be detected.

Errors may be random or systematic, the former being responsible for inflated error estimates. Practical ways of reducing error include the use of proper plot size and shape. Within limits, rectangular plots and larger plots tend to reduce variation per plot. Also, the use of experimental designs that include local control of variation (e.g., randomized complete block design) helps to reduce error.

## Principles of experimental design

This subject is treated in detail in Chapter 23. It is introduced here only to further explain the concept of error. The unit to which a treatment is applied is called the **experimental unit**. In plant breeding common examples of treatment are genotypes (to be evaluated), locations (where genotypes will be evaluated), years, and seasons (over which evaluations are conducted). An experimental unit may be a plant or groups of plants (in a pot).

Experimental designs are statistical procedures for arranging experimental units (experimental design) such that experimental error is minimized. Three tactics or techniques are used in experimental designs for this purpose. These are **replication**, **randomization**, and **local control**.

### *Replication*

Replication is the number of times a treatment is repeated in a study. It is important in experimental design for several reasons, two key ones being:

- 1 **Estimation of statistical error.** To establish that experimental units treated alike vary in their response requires at least two of the same units that have been treated alike.
- 2 **To reduce the size of statistical error.** A measure of the consistency in a data set (standard error) will be presented later in the book. Calculated as  $\sigma/\sqrt{(\text{number of replications})}$ , it is obvious that the larger the number of replications, the smaller the error ( $\sigma$  = standard deviation).

Another pertinent question is the number of replications to use in a study. It should be noted that the more replications used, the more expensive the experiment will be to conduct. In plant breeding, breeders commonly use fewer replications (e.g., two) for preliminary field trials, which often contain hundreds of lines, and more replications (e.g., four) for advanced trials that contain about 10–20 entries.

### *Randomization*

This is the principle of equal opportunity whereby treatment allocation to experimental units is made without bias. To make the statistical test of significance valid, errors should be independent of treatment effect. Randomization may be completely random or may have restrictions to accommodate a specific factor in the experiment.

### *Local control*

This is an additional tactic used by researchers to “contain” variation within an experiment through grouping of experimental units on the basis of homogeneity. Variation within groups is kept to a minimum, while enhancing variation between groups. Statistical procedures are then used to extract this group-based variation from the error estimate. Blocking is recommended if a distinct variation occurs in the experimental field. For example, where a field has a slope, there will be a fertility gradient. Completely random allocation of treatments may place all the replications of one treatment in one fertility zone. Use of the blocking techniques will allow one replication of each treatment to be represented in each distinct fertility zone by placing a restriction on

randomization. The blocks should be laid across the fertility gradient.

## Probability

Statistical probability is a procedure for predicting the outcome of events. Probabilities range from 0 (an event is certain not to occur) to 1.0 (an event is certain to occur). There are two basic laws of probability – **product** and **sum laws**. The probability of two or more outcomes occurring simultaneously is equal to the product of their individual probabilities. Two events are said to be independent of one another if the outcome of each one does not affect the outcome of the other. Genetic ratios may be expressed as probabilities. Consider a heterozygous plant (*Rr*). The probability that a gamete will carry the *R* allele is one-half. In a cross, *Rr* × *Rr* (selfing), the probability of a homozygous recessive (*rr* offspring) is  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ . The probability that one or another of several mutually exclusive outcomes will occur is the sum of their individual probabilities. Using the cross *Rr* × *Rr*, the *F*<sub>2</sub> will produce *RR* : *Rr* : *rr* in the ratio  $\frac{1}{4} : \frac{1}{2} : \frac{1}{4}$ . The probability that a progeny will exhibit a dominant phenotype (*RR*, *Rr*) =  $\frac{1}{4} + \frac{1}{2} = \frac{3}{4}$ . Other examples were discussed in Chapter 3.

In using probabilities for prediction, it is important to note that a large population size is needed for accurate prediction. For example, in a dihybrid cross, the *F*<sub>2</sub> progeny will have a 9 : 3 : 3 : 1 phenotypic ratio, indicating 9/16 will have the dominant phenotype. However, in a sample of exactly 16 plants, it is unlikely that exactly nine plants will have the dominant phenotype. A larger sample is needed.

## Measures of central tendency

The distribution of a set of phenotypic values tends to cluster around a central value. The most common measure of this clustering is the **arithmetic mean**. Plant breeders use this statistical procedure very frequently in their work. The formulae for calculating means are:

$$\text{Sample mean, } \bar{X} = \sum X / n$$

$$\text{Population mean, } \mu = \sum X / N$$

where *X* = measured value of the item,  $\bar{X}$  = sample mean,  $\mu$  = population mean, *n* = sample size, and *N* = population sample.

**Table 9.1** Ungrouped data for distribution of plant seedling height.

	Distribution of seedling height (cm)													
	5	6	7	8	9	10	11	12	13	14	15	16	17	Total
<i>F</i> <sub>1</sub>					5	14	16	12	3					50
<i>F</i> <sub>2</sub>	4	10	13	17	20	28	25	18	17	13	11	10	7	193

**Table 9.2** Grouped data for frequency calculation.

<i>f</i> <sub>i</sub>	<i>x</i> <sub>i</sub>	<i>f</i> <sub>i</sub>	<i>f</i> <sub>i</sub> <i>x</i> <sub>i</sub>	<i>f</i> <sub>i</sub> <i>x</i> <sub>i</sub> <sup>2</sup>	<i>f</i> <sub>i</sub> <i>x</i> <sub>i</sub>	<i>f</i> <sub>i</sub> <i>x</i> <sub>i</sub> <sup>2</sup>	
	5	4	20	100			
	6	10	60	360			
	7	13	91	637			
5	8	17	136	1,088	40	320	
14	9	20	180	1,620	126	1,134	
16	10	28	280	2,800	160	1,600	
12	11	25	275	3,029	132	1,452	
3	12	18	216	2,592	36	432	
	13	17	221	2,873	494	4,938	
	14	13	182	2,548			
	15	11	165	2,475			
	16	10	160	2,560			
	17	7	119	2,023			
<i>n</i>			$\sum f_i x_i$	$\sum f_i x_i^2$			
193			2,105	24,705			

The sample mean is calculated as:

$$\bar{X} = \sum_{i=1}^n X_i / n \text{ (for ungrouped data; Table 9.1)}$$

or:

$$\bar{X} = \sum X_i f_i / n \text{ (for grouped data; Table 9.2)}$$

where *X*<sub>i</sub> = value of the *i*th unit included in the sample, *f*<sub>i</sub> = frequency of the *i*th class, and *n* =  $\sum f_i$ .

The sample mean of seed size of navy beans is:

$$\begin{aligned} \bar{X} &= \sum_{i=1}^n X_i / n \\ &= (17.2 + 18.1 + \dots + 19.7) / 10 \\ &= 190.9 / 10 \\ &= 19.01 \text{ g per 100 seed} \end{aligned}$$

Using data in Table 9.2, the mean of the  $F_2$  can be obtained as:

$$\bar{X} = \sum X_i f_i / n \\ = 2,105 / 193 \\ = 10.91 \text{ cm}$$

### Measures of dispersion

Measures of dispersion or variability concerns the degree to which values of a data set differ from their computed mean. The most commonly used measure of dispersion is the mean square deviation or **variance**. The **population variance** is given by:

$$\sigma^2 = [\sum (X_1 - \mu)^2] / N$$

where  $\sigma^2$  = population variance,  $X_1$  = value of observations in the population,  $\mu$  = mean of the population, and  $N$  = total number of observations in the population.

The sample variance is given by:

$$s^2 = [\sum (X - \bar{X})^2] / (n - 1)$$

where  $s^2$  = sample variance,  $X$  = value of the observation in the sample,  $\bar{X}$  = mean of the sample, and  $n$  = total number of observations in the sample.

The computational formula is:

$$s^2 = [\sum X^2 - (\sum \bar{X})^2 / n] / (n - 1)$$

Using the data below for number of leaves per plant:

Number of leaves	Total
$X$ 7 6 7 8 10 7 9 8 7 10 79 = $\sum X$	
$X^2$ 49 36 49 64 100 49 81 64 49 100 641 = $\sum X^2$	

$$(\sum X)^2 / n = 79^2 / 10 = 6,241 / 10 = 624.1$$

$$s^2 = (641 - 624.1) / 9 = 16.9 / 9 = 1.88$$

Variance may also be calculated from grouped data. Using the data in Table 9.2, variance may be calculated as follows:

$$s^2 = [n \sum fx^2 - (\sum fx)^2] / n(n - 1) \\ = [193(24,705) - (2,105)^2] / 193(193 - 1) \\ = (4,768,065 - 4,431,025) / 37,056 \\ = 337,040 / 37,056 \\ = 9.10 \text{ (for } F_2, \text{ the most variable generation following a cross)}$$

Variance for the  $F_1$  is 1.67.

### Standard deviation

The **standard deviation (SD)** measures the variability that indicates by how much the value in a distribution typically deviates from the mean. It is the positive square root of the population variance. The larger the value of the standard deviation, the more the observations (data) are spread about the mean, and vice versa.

The standard deviation of the sample is simply:

$$s = \sqrt{s^2}$$

For the number of leaves per plant example:

$$s = \sqrt{1.88} \\ = 1.37$$

Similarly, for the seedling height data:

$$\text{SD of the } F_2 = \sqrt{9.1} = 3.02 \\ \text{SD of the } F_1 = \sqrt{1.67} = 1.29$$

### Normal distribution

One of the most important examples of continuous probability distribution is the **normal distribution** or the **normal curve**. It is important because it approximates many kinds of natural phenomena. If the population is normally distributed, the mean = 0.0 and the variance = 1.0. Further, a range of  $\pm 1$  SD from the mean will include 68.26% of the observations, whereas a range of  $\pm 2$  SD from the mean will capture most of the observations (95.45%) (Figure 9.1). The shape of the curve varies depending on the nature of the population.

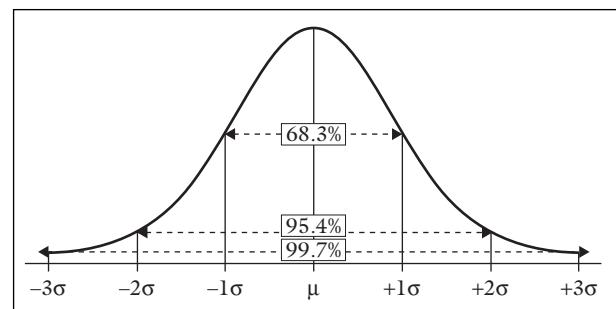


Figure 9.1 Normal distribution curve.

## Coefficient of variation

The **coefficient of variation** is a measure of the relative variability of given populations. Variance estimates have units attached to them. Consequently, it is not possible to compare population measurements of different units (i.e., comparing apples with oranges). For example, one population may be measured in kilograms (e.g., yield), while another is measured in centimeters or feet (e.g., plant height). A common application of variance is the test to find out if one biological sample is more variable for one trait than for another (e.g., is plant height in soybean more variable than the number of pods per plant?). Larger organisms usually vary more than smaller ones. Similarly, traits with larger means tend to vary more than those with smaller means. For example, grain yield per hectare of a cultivar (in kg/ha) will vary more than its 100 seed weight (in grams). For these and other enquiries, the coefficient of variation facilitates the comparison because it is unit free.

The coefficient of variation (**CV**) is calculated as:

$$CV = (s/\bar{X}) \times 100$$

For the number of leaves per plant example:

$$\begin{aligned} CV &= 1.37/7.9 = 0.173 \\ &= 17.3\% \end{aligned}$$

A CV of 10% or less is generally desirable in biological experiments.

## Standard error of the mean

The **standard error** measures the amount of variability among individual units in a population. If several samples are taken from one population, the individuals will vary within samples as well as among samples. The **standard error of the mean (SE)** measures the variability among different sample means taken from a population.

It is computed as:

$$s_x = s/\sqrt{n}$$

For the number of leaves per plant example:

$$\begin{aligned} s_x &= 1.37/\sqrt{10} \\ &= 0.433 \end{aligned}$$

The standard error of the mean indicates how precisely the population parameter has been estimated. It

may be attached to the mean in the presentation of results in a publication (e.g., for the leaves per plant example, it will be  $7.9 \pm 0.43$ ).

## Simple linear correlation

Plant breeders are not only interested in variability as regards a single characteristic of a population, but often they are interested in how multiple characteristics of the units of a population associate. If there is no association, covariance will be zero or close to zero. The magnitude of covariance is often related to the size of the variables themselves, and also depends on the scale of measurement.

The **simple linear correlation** measures the linear relationship between two variables. It measures a joint property of two variables. Plant breeding is facilitated when desirable genes are strongly associated on the chromosome. The relationship of interest in correlation is not based on cause and effect. The degree (closeness or strength) of linear association between variables is measured by the **correlation coefficient**. The correlation coefficient is free of scale and measurement, and has values that lie between +1 and -1 (i.e., correlation can be positive or negative). If there is no linear association between variables, the correlation is zero. However, a lack of significant linear correlation does not mean there is no association (the association could be non-linear or curvilinear).

The population correlation coefficient ( $\rho$ ) is given by:

$$\rho = \sigma_{XY}^2 / \sqrt{(\sigma_X^2 \times \sigma_Y^2)}$$

where  $\sigma_X^2$  = variance of  $X$ ,  $\sigma_Y^2$  = variance of  $Y$ , and  $\sigma_{XY}^2$  = covariance of  $X$  and  $Y$ . The sample covariance is called the **Pearson correlation coefficient ( $r$ )** and is calculated as:

$$r = s_{XY}^2 / \sqrt{(s_X^2 \times s_Y^2)}$$

where  $s_{XY}^2$  = sample covariance of  $X$  and  $Y$ ,  $s_X^2$  = sample variance of  $X$ , and  $s_Y^2$  = sample variance of  $Y$ .

The computational formula is:

$$r = [N(\sum XY) - (\sum X)(\sum Y)] / \sqrt{[N(\sum X^2) - (\sum X)^2][N(\sum Y^2) - (\sum Y)^2]}$$

The data in Table 9.3 shows the seed oil and protein content of 10 soybean cultivars. The calculation yields the following results:

Covariance = -2.45  
 Correlation = -0.757  
 Intercept = 59.34  
 Slope = -1.087  
 Standard error = 0.332  
 Student's  $t$ -value = 3.273; probability = 0.010

The results indicate a significant negative association between seed oil and protein content. The breeding implication is that as one selects for high seed oil, seed protein will decrease.

**Table 9.3** Data for oil and protein content of soybean seed.

Oil content (%)	Protein content (%)
20.1	35.7
21.2	35.1
19.5	33.2
18.3	40.6
19.0	37.5
21.3	36.1
19.8	39.5
22.6	34.8
17.5	39.1
19.9	40.2
Mean	19.92
	37.68

### Simple linear regression

Unlike simple linear correlation, **simple linear regression** is a relationship between two variables that involves cause and effect. There is a **dependent variable ( $Y$ )** and an **independent variable ( $X$ )**. For example, grain yield depends on seed size, number of seeds per pod, etc. The changes in the dependent variable (effect) are brought about by the changes in the independent variable (cause). Another way of looking at it is that regression is a study of the relationship between variables with the objective of identifying, estimating, and validating the relationship.

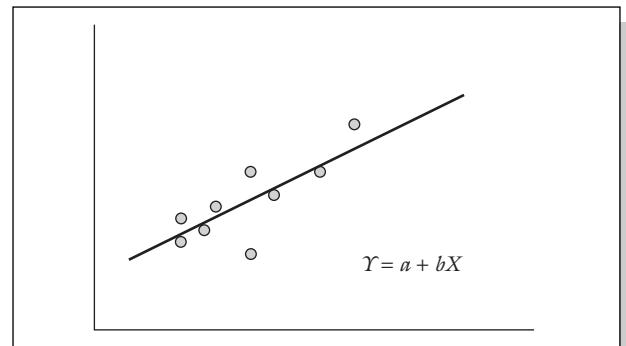
Simple linear regression has the mathematical form of the equation of a straight line:

$$Y = a + bX$$

where  $Y$  = dependent variable,  $X$  = independent variable,  $b$  = slope of the regression line, and  $a$  = intercept on the  $y$  axis.

**Table 9.4** Data on plant yield and maturity of soybean.

	Yield (bushels)	Days to maturity
44	138	
40	136	
38	125	
35	118	
33	115	
32	111	
30	110	
28	109	
24	98	
18	93	
Mean	32.2	115.3



**Figure 9.2** The linear regression line.

The regression coefficient is calculated as:

$$b = [N(\sum XY) - (\sum X)(\sum Y)] / N(\sum X^2) - (\sum X)^2$$

The data in Table 9.4 represent the yield of soybean corresponding to various days to maturity of the crop. The results of a regression analysis are as follows:

Covariance = 110.04  
 Correlation = 0.976  
 Intercept = -27.03  
 Slope ( $b$ ) = 0.514  
 Standard error = 0.040  
 Student's  $t$ -value = 12.794; probability = 0.000

The prediction equation is hence (Figure 9.2):

$$\hat{Y}_i = 127.03 + (0.514)X_i$$

By plugging values for  $X_i$ , corresponding  $Y$  values can be predicted. The results indicate that the regression line will be a good predictor of an unknown value of the independent variable.

### Chi-square test

The **chi-square** ( $\chi^2$ ) test is used by plant breeders to test hypotheses related to categorical data such as would be collected from inheritance studies. The statistic measures the deviations of the observed frequencies of each class from that of expected frequencies. Its values can be zero or positive but not negative. As the number of degrees of freedom increases, the chi-square distribution approaches a normal distribution. It is defined mathematically as:

$$\chi^2 = \sum [(f_o - f_e)^2] / f_e$$

where  $f_o$  = observed sample frequency, and  $f_e$  = expected frequency of the null hypothesis ( $H_0$ ), the hypothesis to be “disproved”.

#### Chi-square test of goodness-of-fit

Suppose a breeder is studying the inheritance of a trait. A cross is made and the following outcomes are recorded:

Character	Observed frequency	Expected frequency
Green cotyledon	78	75
Yellow cotyledon	22	25

If we assume that the trait is controlled by a single gene pair exhibiting dominance, we expect to find a phenotypic ratio of 3 : 1 in the  $F_2$  (or 1 : 1 ratio in a testcross). This is the null hypothesis ( $H_0$ ). The expected frequencies based on the 3 : 1 ratio are also given. The chi-square value is calculated as follows:

$$\begin{aligned} \chi^2 &= \sum [(f_o - f_e)^2] / f_e \\ &= (78 - 75)^2 / 75 + (22 - 25)^2 / 25 \\ &= 0.12 + 0.36 \\ &= 0.48 \end{aligned}$$

The degrees of freedom ( $df$ ) = 2 – 1 = 1; the tabulated  $t$ -value = 3.81 at probability = 0.05. The calculated chi-square value is less than the tabulated value; therefore, the discrepancy observed above is purely a chance event.

The null hypothesis is hence accepted, and the cotyledon color is declared to be controlled by a single gene pair with complete dominance.

#### Chi-square test of independence

Also called a **contingency chi test**, the chi-square test of independence may be applied to different situations. For example, it is applicable where a breeder has made one set of observations obtained under a particular set of conditions, and wishes to compare it with a similar set of observations under a different set of conditions. The question being asked in contingency chi square is whether the experimental results are dependent (contingent upon) or independent of the conditions under which they were observed. In general, whenever two or more systems of classification are used, one can check for independence of the system.

There is a cross-classification when one individual is classified in multiple ways. For example, a cultivar may be classified according to species and also according to resistance to a disease. The question then is whether the classification of one individual according to one system is independent of its classification by the other system. More specifically, if there is independence in this species–infection classification, then the breeder would interpret the results to mean that there is no difference in infection rate between species.

The short cut method for solving contingency chi-square problems is as follows:

	Categories of observation		Total
	I	II	
A	a	b	a + b
B	c	d	c + d
Total	a + c	b + d	a + b + c + d

$$\chi^2 = [(ad - bc)^2] n / [(a + b)(c + d)(a + c)(b + d)]$$

where a, b, c, and d are the observed frequencies. This is called a  $2 \times 2$  contingency chi-square, but can be extended to more complex problems ( $2 \times 4$ ,  $4 \times 6$ , etc.).

### t-test

The **t-test** is used to make inferences about population means. A breeder may wish to compare the yields of two cultivars, for example. Assuming the sample observations are drawn at random, the two population variances are equal, and the populations from which the samples

were drawn follow the normal distributions (these are assumptions made in order to use this test), the hypothesis to be tested is:

$$H_0: \mu_1 = \mu_2 \text{ (no difference between the two means)}$$

The alternative hypothesis to the null is:

$$H_1: \mu_1 \neq \mu_2 \text{ (the two populations are not equal)}$$

This may be tested as follows (for a small sample size):

$$t = [\bar{X}_1 - \bar{X}_2] / s_p \sqrt{[1/n_1 + 1/n_2]}$$

where:

$$s_p = \sqrt{\{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\} / (n_1 + n_2 - 2)}$$

= pooled variance

and  $\bar{X}_1$  and  $\bar{X}_2$  are the means of samples 1 and 2, respectively.

**Example** A plant breeder wishes to compare the seed size of two navy bean cultivars, A and B. Samples are drawn and the 100 seed weight obtained. The following data were compiled:

$$\begin{aligned} H_0: \mu_1 &= \mu_2 \\ H_1: \mu_1 &\neq \mu_2 \end{aligned}$$

	Cultivar	
	A	B
n	10	8
$\bar{X}$	21.2	19.5 (g/100 seeds)
s	1.3	1.1 (g/100 seeds)

where:

$$\begin{aligned} t &= [\bar{X}_1 - \bar{X}_2] / s_p \sqrt{[1/n_1 + 1/n_2]} \\ s_p &= [(10 - 1)(1.3)^2 + (8 - 1)(1.1)^2] / (10 + 8 - 2) \\ &= [9(1.69) + 7(1.21)] / 16 \\ &= [15.21 + 8.47] / 16 \\ &= 23.68 / 16 \\ &= 1.48 \\ t &= 10 - 8 \\ &= 2 / (1.48 \times 0.47) \\ &= 0.70 \\ &= 2 / 0.70 \\ t(\text{calculated}) &= 2.857 \end{aligned}$$

at  $\alpha_{0.05}$ :

$$\begin{aligned} df &= 10 + 8 - 2 = 16 \\ t(\text{tabulated}) &= 1.746 \end{aligned}$$

Since calculated  $t$  exceeds tabulated  $t$ , we declare a significant difference between the two cultivars for seed size (measured as 100 seed weight).

## Analysis of variance

Frequently, the breeder needs to compare more than two cultivars. In yield trials, several advanced genotypes are evaluated at different locations and in different years. The  $t$ -test is not applicable in this circumstance but its extension, the **analysis of variance (ANOVA)**, is used instead. ANOVA allows the breeder to analyze measurements that depend on several kinds of effects, and which operate simultaneously, in order to decide which kinds of effects are important, and to estimate these effects. As a statistical technique, ANOVA is used to obtain and partition the total variation in a data set according to the sources of variation present and then to determine which ones are important. The test of significance of an effect is accomplished by the  $F$ -test. The results of an analysis of variance are presented in the ANOVA table, the simplest form being as follows:

Source of variation		df	SS	MS	F
Treatment	$k - 1$	$SS_{Tr}$	$MS_{Tr} = SS_{Tr} / k - 1$	$MS_{Tr} / MS_E$	
Error	$N - k$	$SS_E$	$MS_E = SS_E / N - k$		
Total	$N - 1$	$SS_T$			

“Treatment” is the most important source of variance caused by the applied treatments (e.g., different cultivars, locations, years, etc.). The error is unaccounted variation.

In more detailed analysis, interaction effects between treatments are accounted for in the analysis. Examples of ANOVA for  $G \times E$  interaction analysis are provided in Chapter 23. ANOVA is usually done on the computer using software such as MSTAT and SAS. As previously stated, ANOVA permits the breeder to handle more than two genotypes (or variables) in one analysis. The  $t$ -test is not efficient for comparing more than two means. Commonly used tests to separate means under such conditions include the least significant difference (LSD) and Duncan's multiple range test (DMRT).

## Multivariate statistics in plant breeding

**Multivariate analysis** is the branch of statistics concerned with analyzing multiple measurements that have been made on one or several samples of individuals. Because these variables are interdependent among themselves, they are best considered together. Unfortunately, handling data with multicollinearity can be unwieldy and hence some meaningful summarization is needed.

The multivariate techniques in use may be divided into two groups:

- 1 **Interdependence models** – e.g., principal components analysis, factor analysis.
- 2 **Dependence models** – e.g., multivariate analysis of variance, classification functions, discriminant function analysis, cluster analysis, multiple correlation, canonical correlation.

W. W. Cooley and P. R. Jones further classified multivariate procedures into four categories according to the number of populations and the number of variables as follows:

- 1 One set of variables, one population – e.g., principal components analysis, factor analysis.
- 2 One set of variables, two or more populations – e.g., multivariate analysis of variance, discriminant functions, classification functions.

3 Two or more sets of variables, one population – e.g., polynomials fit, multiple correlation, canonical correlation, multiple partial correlation.

4 Two or more sets of variables, two or more sets of populations – e.g., multivariate covariance.

Multivariate analyses are done on computers because of their complexity. An overview of the common procedures is discussed next.

### Factor analysis

A variable can be explained to the extent that its variance can be attributed to an identifiable source. **Factor analysis** may be used to find ways of identifying fundamental and meaningful dimensions of a multivariate domain. It is a decision-making model for extracting subsets of covarying variables. To do this, natural or observed intercorrelated variables are reformulated into a new set (usually fewer in number) of independent variables, such that the latter set has certain desired properties specified by the analyst. Naming factors is only a mnemonic convenience. It should be done thoughtfully so as to convey information to both the analyst and the audience. For example, a large set of morphological traits may be reduced to several conceptual factors such as “architectural factor” (loaded by variables such as internode length, number of internodes, etc.), whereas a “seed size factor” may be loaded by traits such as seed length and seed width.



## Industry highlights

### *Multivariate analyses procedures: applications in plant breeding, genetics, and agronomy*

A. A. Jaradat

USDA-ARS, Morris, 56267 MN, USA

#### *Introduction*

Plant breeders, geneticists, and agronomists are increasingly faced with theoretical and practical questions of multivariate nature. With increases in germplasm sizes, the number of plant and crop variables, and evaluation and characterization data on molecular, biochemical, morphological, and agronomic traits, multivariate statistical analysis (MVA) methods are receiving increasing interest and assuming considerable significance. Some MVAs (e.g., multivariate analysis of variance, MANOVA, and covariance, MANCOVA) are extensions of uni- and bivariate statistical methods appropriate for significance tests of statistical hypotheses. However, most MVAs are used for data exploration, the extraction of fundamental components of large data sets, the discovery of latent structural relationships, and the visualization and description of biological patterns. This review focuses on the salient features and applications of MVAs in multivariate data analyses of plant breeding, genetics, and agronomy data. These include MANOVA, MANCOVA, data reduction methods (factor, principal components, principal coordinates, perceptual mapping, and correspondence analyses), and data classification methods (discriminant analysis, clustering and additive trees).

Crop improvement programs – through breeding, selection, and agronomic evaluation – rely on available genetic diversity for specific trait(s) in the primary and, if needed, in the secondary gene pool of a particular crop species. Classic univariate analysis

**Table 1** Summary of the significant effect ( $P < 0.05$ ) for leaf area index (LAI) and dry weight of stems per plant in MANOVA and determination of the smallest set of variables.

Variable	Significant source of variation	Wilks lambda	F approximation	Final set
LAI	Year	0.182	40.45***	$y_{\max}$
	Genotype	0.175	4.73*	$x_{\max}$
	Growth habit	0.479	26.85***	
Dry weight of stems	Year	0.552	7.30*	$x_{\inf}$
	Genotype	0.067	13.99**	$x_{\inf}$
	Growth habit	0.480	70.17***	
	Within winter types	0.105	13.12***	

$y_{\max}$ , maximum value of the variable;  $x_{\max}$ , time in growing-degree days from sowing to  $y_{\max}$ ;  $x_{\inf}$ , time in growing degree-days from sowing to reach maximum rate of growth. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ .

procedures, limited to estimation and hypotheses testing, are not capable of detecting patterns and exploring multivariate data structures in genetic resources, breeding lines, or cultivars. Therefore, MVA methods to classify and order large numbers of breeding material, trait combinations, and genetic variation are gaining considerable importance and assuming considerable significance.

### MANOVA and MANCOVA

MANOVA and MANCOVA perform a multivariate analysis of variance or covariance when multiple dependent variables are specified. MANOVA tests whether mean differences among groups for a combination of dependent variables are likely to have occurred by chance. A new dependent variable that maximizes group differences is created from the set of dependent variables. The new dependent variable is a linear combination of measured dependent variables, combined so as to separate the groups as much as possible. ANOVA is then performed on the newly created dependent variable. MANCOVA asks if there are statistically reliable mean differences among groups after adjusting the newly created dependent variable for differences on one or more covariates. In this case, variance associated with the covariate(s) is removed from error variance; smaller error variance provides a more powerful test of mean differences among groups.

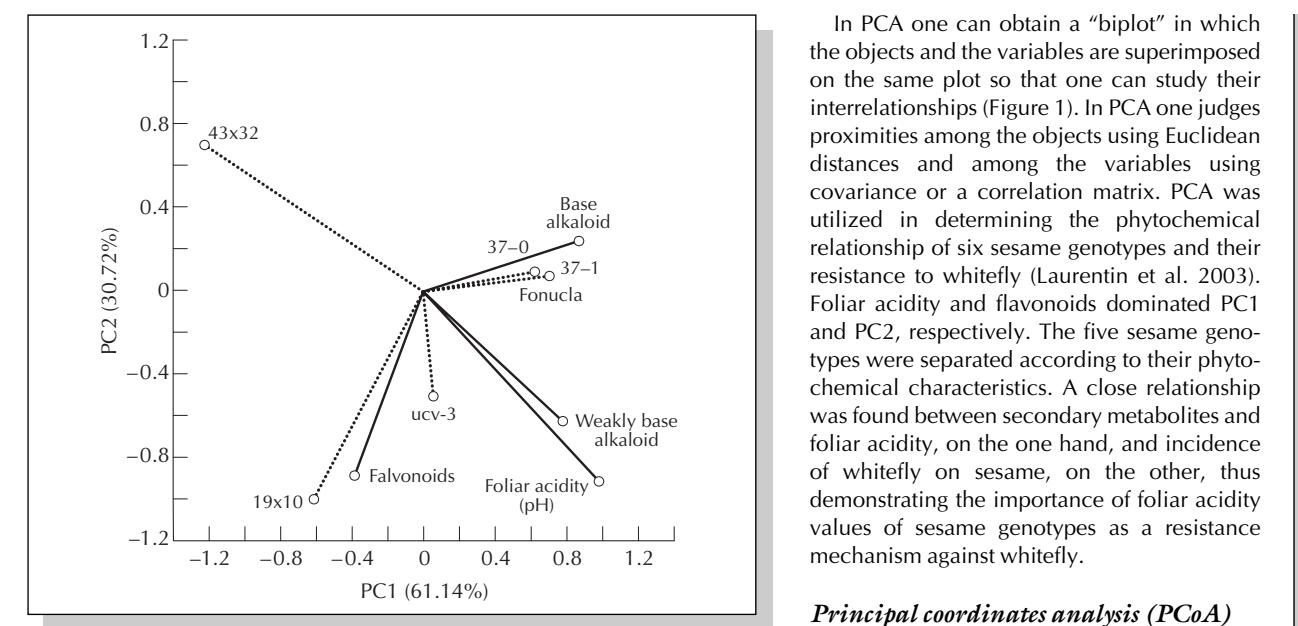
MANOVA was used in the analysis of growth patterns and biomass partitioning of crop plants as a prerequisite for interpreting results of field experiments and in developing crop simulation models. Royo and Blanco (1999) utilized MANOVA to compare non-linear regression growth curves in spring and winter triticale and identified variables responsible for the differences between these curves. Results of these studies are partially presented in Table 1, along with the smallest set of variables required to characterize the growth curves. Wilk's lambda is the criteria for statistical inference and is estimated as the pooled ratio of error variance to effect variance plus error variance. In this example, all Wilk's lambda and F-approximation estimates are significant. For example, the differences within each growth habit (Table 1) were non-significant but differences between growth habits were significant. Thermal time needed to reach the maximum leaf area index was the variable responsible for these differences.

### Variance components analysis (VCA)

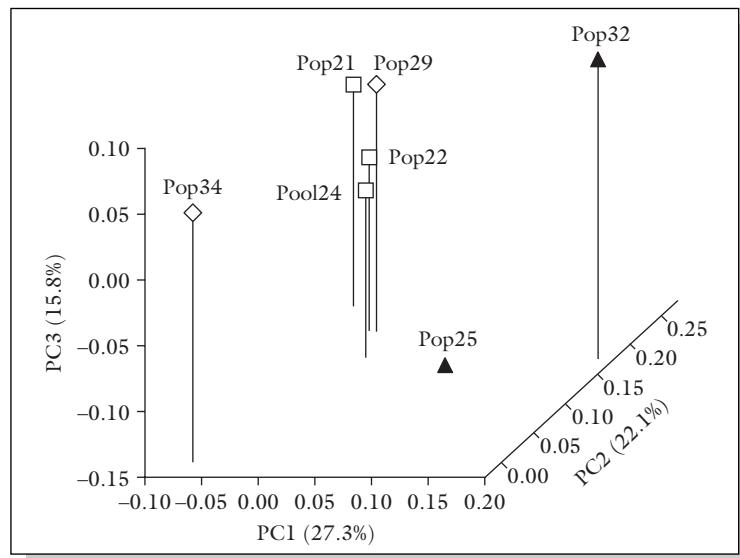
Experimentation is sometimes mistakenly thought to involve only the manipulation of levels of the independent variables and the observation of subsequent responses on the dependent variables. Independent variables whose levels are determined or set by the experimenter are said to have fixed effects. A second class of effects, random effects, are classification effects where the levels of the effects are assumed to be randomly selected from an infinite population of possible levels. Many independent variables of research interest are not fully amenable to experimental manipulation, but nevertheless can be studied by considering them to have random effects.

### Factor analysis (FA) and principal components analysis (PCA)

The primary purpose of FA and PCA is to define the underlying structure in a data matrix. As data reduction or exploratory methods, these procedures are used to reduce the number of variables and to detect structure in the relationships between these variables. FA reproduces the correlation matrix among variables with a few orthogonal factors; however, contrary to PCA, most forms of FA are not unique. PCA is a procedure for finding hypothetical variables (components) that account for as much of the variance in multidimensional data as possible. PCA is a unique mathematical solution; it performs simple reduction of the data set to a few components, for plotting and clustering purposes, and can be used to hypothesize that the most important components are correlated with some other underlying variables.



**Figure 1** A graph based on PCA of five sesame genotypes as operational taxonomic units (dotted lines), and three secondary metabolites in leaves and foliar acidity as variables (solid lines).



**Figure 2** PCoA plot of seven tropical maize populations based on modified Roger's distance. PC1, PC2, and PC3 are the first, second, and third principal coordinates, respectively. Heterotic group A (Pop21, Pop22, and Pool24), heterotic group B (Pop25, Pop32), and populations not yet assigned to heterotic groups (Pop29, Pop34) are shown.

differentiation of the genotypes, the similarity and differences among the genotypes in their response to locations, and the nature and magnitude of the interaction between any genotype and any location.

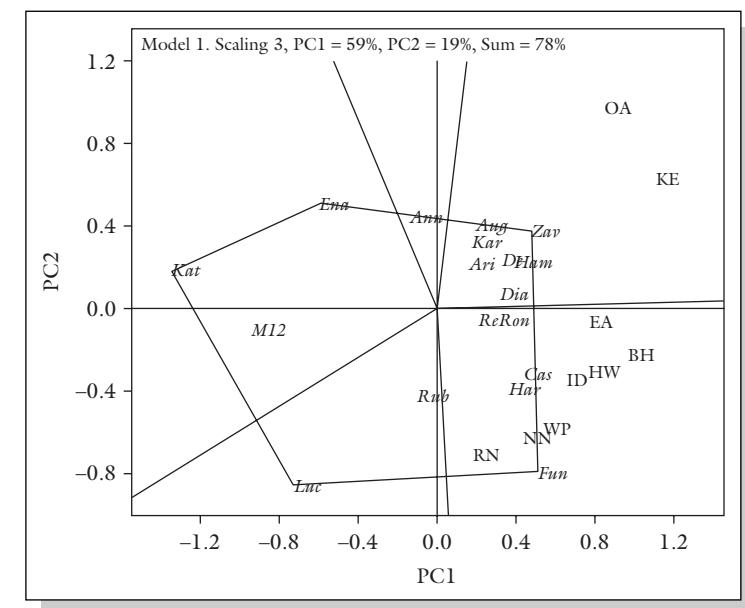
In PCA one can obtain a “biplot” in which the objects and the variables are superimposed on the same plot so that one can study their interrelationships (Figure 1). In PCA one judges proximities among the objects using Euclidean distances and among the variables using covariance or a correlation matrix. PCA was utilized in determining the phytochemical relationship of six sesame genotypes and their resistance to whitefly (Laurentin et al. 2003). Foliar acidity and flavonoids dominated PC1 and PC2, respectively. The five sesame genotypes were separated according to their phytochemical characteristics. A close relationship was found between secondary metabolites and foliar acidity, on the one hand, and incidence of whitefly on sesame, on the other, thus demonstrating the importance of foliar acidity values of sesame genotypes as a resistance mechanism against whitefly.

#### Principal coordinates analysis (PCoA)

PCoA focuses on samples rather than variables and is based on a matrix containing the distances between all data points. A typical usage of PCoA is the reduction and interpretation of large multivariate data sets with some underlying linear structure. PCoA was instrumental in delineating relationships among tropical maize populations based on simple sequence repeats for breeding purposes (Reif et al. 2003). PCoA revealed very clear association among populations within certain heterotic groups (Figure 2). Reif et al. (2003) succeeded in identifying genetically similar germplasm based on molecular markers, and concluded that PCoA provides a more economic and solid approach for making important breeding decisions early in the breeding program.

#### Perceptual mapping (biplot and GE)

Success in evaluating germplasm, breeding lines, and cultivars in multiple environments and for complex traits to identify superior genotypes with specific or wide adaptation can be achieved if the genotypic (*G*) and environmental (*E*) effects and their interaction (*GE*) are precisely estimated (Yan et al. 2000). The *GE* biplot procedure has been used by breeders and agronomists for dissecting *GE* interactions and is being used to analyze data from genotype  $\times$  trait, genotype  $\times$  marker, environment  $\times$  QTL, and diallel cross data. The biplot allows a readily visualized display of similarity and differences among environments in their differentiation of the genotypes, the similarity and differences among the genotypes in their response to locations, and the nature and magnitude of the interaction between any genotype and any location.



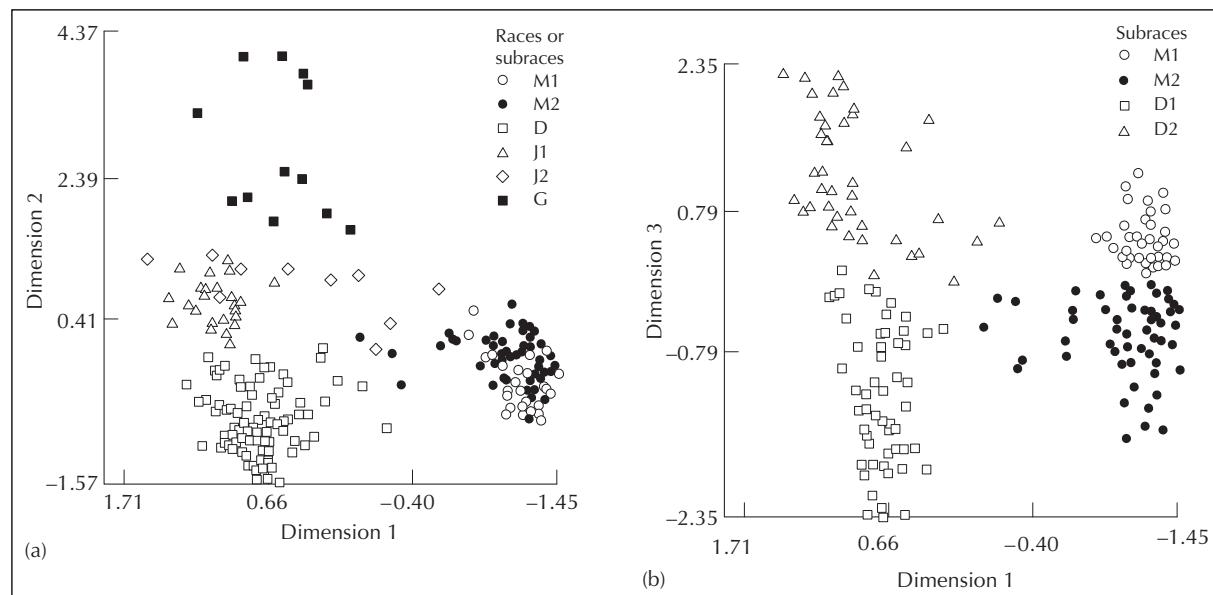
**Figure 3** Biplot showing performance of different wheat cultivars (in italics) in different environments (in capital letters) as a selection method to identify superior cultivars for a target environment.

(random amplified polymorphic DNA), among common bean landraces of Middle American origin for breeding purposes. MCA results (Figure 4) indicated that the Middle American bean germplasm is more complex than previously thought with certain regions holding important genetic diversity that has yet to be properly explored for breeding purposes. The first dimension

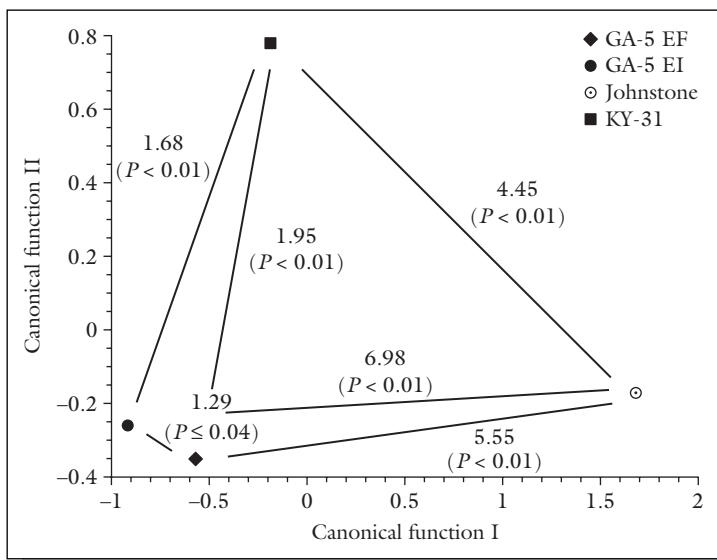
Biplot was used to compare the performance of wheat cultivars under several environments in the Ontario wheat performance trials (Figure 3) and to estimate relative variance components and their level of significance. Results of biplot analysis have several implications for future breeding and cultivar evaluation. A test for optimal adaptation can be achieved through the deployment of different cultivars for mega-environments, and the unpredictable genotype  $\times$  location interaction can be avoided or minimized through cultivar evaluation and selection focusing on the main effects of genotype.

#### Multiple correspondence analysis (MCA)

MCA is a recently developed interdependence MVA procedure that facilitates both dimensional reduction of object ratings on a set of attributes and the perceptual mapping of objects relative to these attributes. MCA helps researchers quantify the qualitative data found in nominal variables and has the ability to accommodate both non-metric data and non-linear relationships. In order to facilitate the use of common bean landraces in genetic improvement, Beebe et al. (2000) used MCA to study the structure of genetic diversity, based on RAPD



**Figure 4** Plot of (a) 250 Middle American bean genotypes in dimension 1 and 2, and (b) 206 genotypes of two races in dimensions 1 and 3 of MCA based on RAPD data.



**Figure 5** Scatterplot of centroid values of four tall fescue cultivars on two canonical discriminant functions. Mahalanobis distances and their probability values, in parentheses, measure the extent of genetic diversity between the four cultivars.

uses Mahalanobis distance to differentiate between cultivars or populations. The higher the canonical loadings (measures of the simple linear correlation between an original independent variable and the canonical variate) of traits of particular significance, the higher the genetic variation as compared with traits having low canonical loadings. Plant breeders can use this information to focus on particular trait(s) for genetic improvement of a particular crop. Vaylay and van Santen (2002) employed CDA in the assessment of genetic variation in tall fescue (Figure 5). They found that the genetic composition of four tall fescue cultivars differ mainly, in decreasing order, in maturity, cell wall content, flag leaf length, tiller number, and dry matter yield. Therefore, tall fescue breeders can concentrate on the most important traits of this perennial pasture crop knowing that the genetic composition of its cultivars changes with time.

#### Cluster analysis (CA)

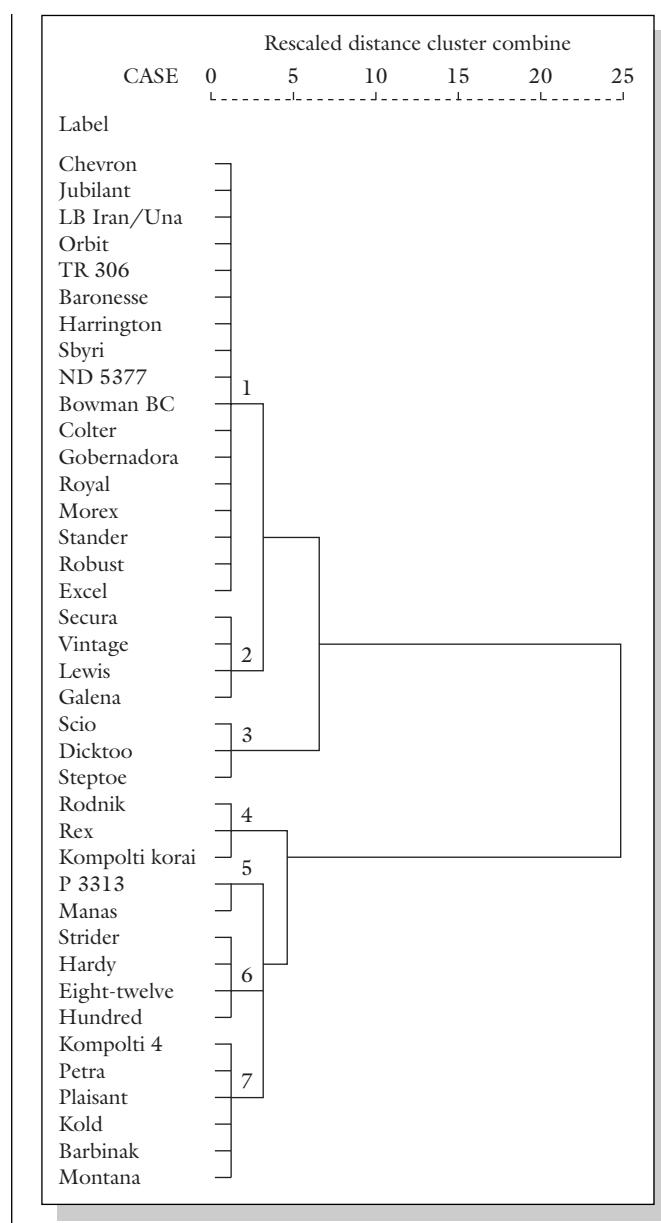
CA is an analytical MVA procedure for developing meaningful subgroups of objects. It classifies a sample of objects into a small number of mutually exclusive groups based on the similarities among the objects. Stepwise clustering involves a combination or division of objects into clusters. Hierarchical CA starts with each case in a separate cluster and then combines the clusters sequentially, reducing the number of clusters at each step until only one cluster is left. The divisive clustering method begins with all the objects in a single cluster, which is then divided at each step into two clusters that contain the most dissimilar objects. Additive trees, as an extension of clustering, are based on a dissimilarity distance matrix among all possible pairs of objects in order to retain the original distances among all pairs of these objects. Unlike other clustering algorithms that are based on the rigorous ultrametric relationships between objects, the additive tree precisely reflects distances among the objects.

Cluster analysis was used as a tool to optimize and accelerate barley breeding. Karsai et al. (2000) evaluated barley cultivars for five physiological and agronomic traits that have significant effects on heading date and winter hardiness. CA helped identify groups of cultivars representing different adaptational types. The wide level of diversity identified in the germplasm set was valuable in studying the genetics of adaptation to certain environments. It was possible to identify (numbered 1 through 7 in Figure 6) winter and spring groups, groups of cultivars with no vernalization response that had the lowest earliness *per se*, and other group of cultivars least sensitive to changes in photoperiod but with a strong vernalization response. A breeding scheme was designed on the basis of the clustering results (Figure 6) and was aimed at developing new cultivars better adapted to a given environment.

(Figure 4a) discriminated between lowland and highland races. The second dimension discriminated among highland races, whereas the third dimension (Figure 4b) divided the highland races according to their growth habit, geographic, distribution, and seed type. Results of MCA can be used to orient plant breeders in their search for distinct genes that can be recombined, thus contributing to higher genetic gain.

#### Canonical discriminant analysis (CDA)

CDA is used to study the variation among two or more groups (samples) of crop cultivars relative to the average variation found within the groups. Linear combinations of the original variables that account for as much as possible of total variation in the data set are constructed using PCA, then canonical correlation is used to determine a linear association between predictor variables identified in PCA and criterion measures. In CDA more distinct differentiation of cultivars is achieved as compared with univariate analysis, since all independent variables (e.g., traits) are considered simultaneously in the process. CDA can separate "among population" effects from "within population" effects thus maximizing the overall heritability estimates of canonical variates by placing very large weight on traits with low levels of environmental variability. CDA



### References

- Beebe, S., P.W. Skroch, J. Tohme, M.C. Duque, F. Pedraza, and J. Nienhuis. 2000. Structure of genetic diversity among common bean landraces of Mesoamerican origin based on correspondence analysis of RAPD. *Crop Sci.* 40:264–273.
- Karsai, I., K. Meszaros, L. Lang, P.M. Hayes, and Z. Bedo. 2000. Multivariate analysis of traits determining adaptation in cultivated barley. *Plant Breed.* 120:21–22.
- Laurentin, H., C. Pereira, and M. Sanabria. 2003. Phytochemical characterization of six sesame (*Sesamum indicum* L.) genotypes and their relationships with resistance against the sweetpotato whitefly *Bemisia tabaci* Gennadius. *Agron. J.* 95:1577–1582.
- Reif, J.C., A.E. Melchinger, X.C. Xia, et al. 2003. Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. *Crop Sci.* 43:1275–1282.
- Royo, C., and R. Blanco. 1999. Growth analysis of five spring and five winter triticale genotypes. *Agron. J.* 91:305–311.
- Vaylay, R., and E. van Santen. 2002. Application of canonical discriminant analysis for the assessment of genetic variation in tall fescue. *Crop Sci.* 42:534–539.
- Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavincs. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40:597–605.

**Figure 6** Cluster analysis of 39 barley cultivars based on a matrix of vernalization response, photoperiod sensitivity, earliness *per se*, frost tolerance at –10 and –13°C, and heading dates under different photoperiod regimes. The dendrogram was created using the Ward minimum variance method. Groups (1–7) were characterized by having specific levels of one or more traits.

### Principal components analysis

**Principal components analysis** (PCA) reduces the dimensions of multivariate data by removing inter-correlations among the traits being studied and thereby enabling multidimensional relationships to be plotted on two or three principal axes. PCA reduces the number of variables to be used for prediction and description.

By examining a set of 15 quality traits, researchers at Michigan State University Bean Breeding Program were able to ascertain that certain quality traits (dry characteristics, soaking characteristics, cooking characteristics) of dry beans were independent. This prompted the researchers to suggest a tandem selection procedure to be followed by the construction of selection indices for their breeding program.

### Discriminant function analysis

**Discriminant function analysis** assumes a population is made up of subpopulations, and that it is possible to find a linear function of certain measures and attributes of the population that will allow the researcher to discriminate between the subpopulations. Consequently, discriminant procedures are not designed for seeking population groupings (that is what cluster analysis does) because the population has already been grouped. Discriminant analysis may be used in conjunction with the  $D^2$  statistic (Mahalanobis  $D^2$ ) to indicate the biological distance between separated groups.

### Cluster analysis

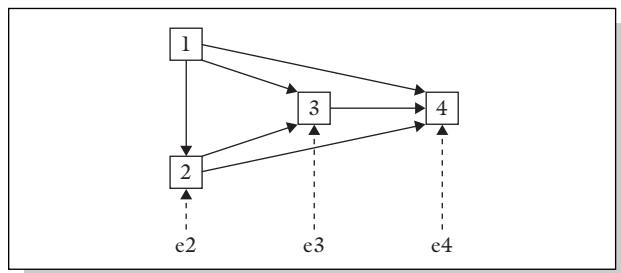
Genetic assessment of germplasm is commonly undertaken by plant breeders to understand genetic variation in the germplasm and to discover patterns of genetic diversity. **Cluster analysis**, unlike discriminant function analysis, groups genetically similar genotypes. Clustering can be done on a morphological or molecular basis (e.g., using DNA markers). Analysis of genetic diversity levels in germplasm helps plant breeders to make proper choices of parents to use in breeding programs.

### Canonical correlation analysis

The **canonical correlation** analysis is a generalization of the multiple correlation procedure. The technique is used to analyze the relationship between two sets of variables drawn from the same subjects. An assumption is made that there are unobserved variables dependent on a known set of variables  $X$ , and determining another known set,  $Y$ . The intermediately unobserved variables are used to canalize the influence of set  $X$  on set  $Y$ .

### Path analysis

Path analysis is a technique for decomposing correla-



**Figure 9.3** The basic concept of path analysis.

tions into different pieces for the interpretation of effects. The procedure is closely related to multiple regression analysis. Path analysis allows the researcher to test theoretical propositions about cause and effect without manipulating variables. Variables may be assumed to be causally related and propositions about them tested. However, it should be cautioned that, should such propositions be supported by the test, one cannot conclude that the causal assumptions are necessarily correct. A breeder may want to understand the relative contributions of yield components and morphophenological traits to grain yield.

The general display of a path analysis is shown in Figure 9.3. Arrows are used to indicate assumed causal relations. A single-headed arrow points from the assumed cause to its effect. If an arrow is double-headed, only correlation is present (no causal relations are assumed). Variables to which arrows are pointed are called endogenous variables or dependent variables ( $Y$ ). Exogenous variables have no arrows pointing to them; they are independent variables ( $X$ ). The direct effect of a variable assumed to be a cause on another variable assumed to be an effect, is called a **path coefficient**. Path coefficients are standardized partial regression coefficients.

### References and suggested reading

- Akroda, M.O. 1983. Principal components analysis and metroglyph of variation among Nigerian yellow yams. *Euphytica* 32:565–573.
- Cooley, W.W., and P.R. Lohnes. 1971. Multivariate data analysis. John Wiley & Sons, Inc., New York.
- Denis, J.C., and M.W. Adams. 1978. A factor analysis of plant variables related to yield in dry beans. I. Morphological traits. *Crops Sci.* 18:74–78.
- Kendall, M.G. 1965. A course in multivariate analysis. Charles Griffin & Co., London.
- Snedecor, G.W., and W.G. Cochran. 1967. Statistical methods, 6th edn. Iowa State University, Ames, IA.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Statistics do not prove anything.
- 2 Lack of significant linear regression does not mean a lack of relationship.
- 3 A *t*-test is used for separating more than three means.
- 4 Discriminant analysis is used for seeking population groupings.
- 5 Chi-square analysis is used for testing a hypothesis involving continuous data.

### Part B

Please answer the following questions:

- 1 What is a statistic?
- 2 Distinguish between standard error and standard error of the mean.
- 3 Distinguish between simple linear correlation and linear regression.
- 4 What is contingency chi-square analysis?

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the role of statistics in plant breeding.
- 2 Discuss the issue of causality in research.
- 3 Discuss the application of multivariate statistics in plant breeding.
- 4 Discuss the procedure of path analysis and its application in breeding.

# **Section 5**

## **Tools in plant breeding**

- Chapter 10 Sexual hybridization and wide crosses in plant breeding**
  - Chapter 11 Tissue culture and the breeding of clonally propagated plants**
  - Chapter 12 Mutagenesis in plant breeding**
  - Chapter 13 Polyploidy in plant breeding**
  - Chapter 14 Biotechnology in plant breeding**
  - Chapter 15 Issues in the application of biotechnology in plant breeding**
- 

This section discusses the application of the principles and concepts of plant genetics and breeding to the development of new cultivars. Methods of plant genetic manipulation have changed over the years as new knowledge and technology have become available. The first generation technology for plant manipulation was non-invasive, requiring artificial crossing to accomplish gene mixing to create new genotypes. This method is restrictive in that only sexually compatible plants can be improved this way. Second generation technologies of tissue culture, mutagenesis, and chromosome manipulation allowed more intrusive plant genetic manipulation, targeting cells or the genetic material relatively more directly. The first and second generation technologies are generally classified as conventional plant breeding tools. In this section, the student will learn the scientific basis and application of hybridization, tissue culture, induced mutations, and ploidy in the genetic manipulation of plants. In addition, the cutting edge technology of genetic engineering that enables researchers to directly manipulate plants at the molecular level will be discussed. Because this powerful technology allows scientists to transfer DNA across natural biological barriers, its developments and applications are often mired in controversy. Consequently, the issues of risk and regulation of biotechnology are also discussed in this section.

---



# 10

# Sexual hybridization and wide crosses in plant breeding

---

## Purpose and expected outcomes

*One of the principal techniques of plant breeding is artificial mating (crossing) of selected parents to produce new individuals that combine the desirable characteristics of the parents. This technology is restricted to sexually reproducing species that are compatible. However, in the quest for new desirable genes, modern plant breeders sometimes attempt to mate individuals that are biologically distant in relationship. It is important for the breeder to understand the problems associated with making a cross, and how barriers to crossing, where they exist, can be overcome. After studying this chapter, the student should be able to:*

- 1 Define sexual hybridization and discuss its genetic consequences.
  - 2 Define a wide cross and discuss its objectives and consequences.
  - 3 Discuss the challenges to wide crosses and techniques for overcoming them.
- 

### Concept of gene transfer

Crop improvement entails genetically manipulating plants in a predetermined way, which often involves the transfer of genes from one source or genetic background to another. When a plant breeder has determined the direction in which a crop is to be improved, the next crucial step is to find a source of the appropriate gene(s) for making the desired change(s). Once an appropriate source (germplasm) has been found, the next step is to transfer the gene(s) to the parent to be improved. In flowering species, the conventional method of gene transfer is by crossing or **sexual hybridization**. This procedure causes genes from the two parents to be assembled into a new genetic matrix. It follows that if parents are not genetically compatible, gene transfer by sexual means cannot occur at all, or at best may be fraught with complications. The product of hybridization is called a **hybrid**.

Sexual hybridization can occur naturally through agents of pollination (see Chapter 4). Even though self-pollinating species may be casually viewed as “self-hybridizing”, the term hybridization is reserved for crossing between unidentical parents (the degree of divergence is variable). Artificial sexual hybridization is the most common conventional method of generating a segregating population for selection in breeding flowering species. In some breeding programs, the hybrid ( $F_1$ ) is the final product of plant breeding (see Chapter 18). However, in most situations, the  $F_1$  is selfed (to give an  $F_2$ ) to generate recombinants (as a result of recombination of the parental genomes) or a segregating population, in which selection is practiced.

The tools of modern biotechnology now enable the breeder to transfer genes by circumventing the sexual process (i.e., without crossing). More significantly, gene transfer can transcend natural reproductive or genetic barriers. Transfers can occur between unrelated plants and even between plants and animals (see Chapter 14).

## Applications of crossing in plant breeding

Sometimes, crossing is done for specific purposes, within the general framework of generating variability. Hybridization precedes certain methods of selection in plant breeding to generate general variability.

- 1 Gene transfer.** Sometimes, only a specific gene (or a few) needs to be incorporated into an adapted cultivar. Crossing is used for the gene transfer process, followed by additional strategic crossing to retrieve the desirable genes of the adapted cultivar (see backcrossing in Chapter 16).
- 2 Recombination.** Genetically diverse parents may be crossed in order to recombine their desirable traits. The goal of recombination, which is a key basis of plant breeding, is to forge desirable linkage blocks.
- 3 Break undesirable linkages.** Whereas forging desirable linkage blocks is a primary goal of plant breeding, sometimes crossing is applied to provide opportunities for undesirable linkages to be broken (see recurrent selection in Chapter 17).
- 4 For heterosis.** Hybrid vigor (heterosis) is the basis of hybrid seed development. Specially developed parents are crossed in a predetermined fashion to capitalize on the phenomenon of heterosis for cultivar development.
- 5 For maintenance of parental lines.** In hybrid seed development programs, crossing is needed to maintain special parents used in the breeding program (e.g., cytoplasmic male-sterile (CMS) lines, maintainer lines).
- 6 For maintenance of diversity in a gene pool.** Plant breeders may use a strategy of introgression (crossing and backcrossing selected entries with desired traits into adapted stocks) and incorporation to develop dynamic gene pools from which they can draw materials for crop improvement.
- 7 For evaluation of parental lines.** Inbred lines for hybrid seed development are evaluated by conducting planned crosses to estimate combining abilities, in order to select appropriate parents for use in hybrid seed development.
- 8 For genetic analysis.** Geneticists make planned crosses to study the inheritance and genetic behavior of traits of interest.

## Artificial hybridization

**Artificial hybridization** is the deliberate crossing of selected parents (controlled pollination). There are specific methods for crossing different species, which differ according to factors including floral morphology,

floral biology, genetic barriers, and environmental factors. Methods for selected species are described in Part II of this book. However, there are certain basic factors to consider in preparation for hybridization:

- 1** The parents should be unidentical but reproductively compatible. Generally, corn is crossed with corn, and tomato with tomato. Further, parents to be crossed are usually obtained from the same species.
- 2** The parents together should supply the critical genes needed to accomplish the breeding objective.
- 3** One parent is usually designated as female. Whereas some breeding methods may not require this designation, breeders usually select one parent to be a female and the other a male (pollen source). This is especially so when hybridizing self-pollinated species. In some cases, selected parents of cross-pollinated species may be isolated and allowed to randomly cross-pollinate each other.
- 4** The female parent usually needs some special preparation. In complete flowers (with both male and female organs), the flowers of the parent selected to be female are prepared for hybridization by removing the anthers, a tedious procedure called **emasculatio**n (discussed below). Emasculation is eliminated in some crossing programs by taking advantage of male sterility (renders pollen sterile) when it occurs in the species (see Chapter 4).

Pollen is often physically or manually transferred. Artificial hybridization often includes artificial pollination, whereby the breeder physically deposits pollen from the male parent onto the female stigma. However, when hybridization is conducted on a large scale (e.g., commercial hybrid seed development), hand pollination is not a feasible option in nearly all cases.

## Flower and flowering issues in hybridization

The flower has a central role in hybridization. The success of a crossing program depends on the condition of the flower regarding its overall health, readiness or receptiveness to pollination, maturity, and other factors. The actual technique of crossing depends on floral biology (time of pollen shedding, complete or incomplete flower, self- or cross-pollinated, flower size and shape).

### Flower health and induction

It is important that plants in a crossing block (or that are to be crossed) be in excellent health and be properly

developed. This is especially so when flowers are to be manually emasculated. Once successfully crossed, adequate amount of seed should be obtained for planting the first generation. The parents to be mated should receive proper lighting, moisture supply, temperature, nutrition, and protection from pests.

Plants growing in the greenhouse should be provided with the proper intensity and duration of light. If the species is photoperiod sensitive, the lighting should be adjusted accordingly. A suitable temperature is required for proper plant growth and development. In some species, a special temperature treatment (vernalization) is required for flower induction. Furthermore, temperature affects the pollen shed in flowers. Consequently, extreme temperatures may cause inadequate amounts of pollen to be shed for successful artificial pollination. Pollen quantity and quality are influenced by the relative humidity of the growing environment. Extreme moisture conditions should be avoided. Parents should be fertilized with the proper amounts of nitrogen, phosphorus, and potassium for vigorous plant growth to develop an adequate number of healthy flowers.

### Synchronization of flowering

In artificial pollination, the breeder should be familiar enough with the species to know its flowering habits regarding time from planting to flowering, duration of flowering, mechanisms of natural pollen dehiscence and fertilization, and time of peak pollen production, in order to take advantage of the window of opportunity of anthesis (pollen shed) for best crossing outcomes. To ensure that parents in a crossing program will have flowers at the same time, the practice of staggered planting is used by some breeders to plant sets of parents at different times. This way, flowering will occur over a longer period of time. When depending on natural pollination, interspersed planting on different dates will favor even pollen distribution.

Photoperiod may be manipulated in photoperiod-sensitive species to delay or advance flowering as appropriate, in order to synchronize flowering of the parents in a cross. Other techniques that have been used in specific cases include manipulation of temperature and planting density, removal of older flowers to induce a new flush of flowers, and pinching (e.g., removal of plant apex to induce tillering or branching for additional flowers). In corn, the silk of an early-flowering inbred parent may be cut back to delay the time to readiness for pollination.

### Selecting female parents and suitable flowers

After selecting lines to be parents in a cross, it is necessary in artificial crosses to designate one parent as female (as previously stated), as well as identify which type of flowers on the parent would be most desirable to cross. In crossing programs in which the CMS system is being used, it is critical to know which plants to use as females (these would be the male-sterile genotypes, or A- and B-lines). Because the pollen or male gamete is practically without cytoplasm, and because certain genes occur in the extranuclear genome (such as CMS), it is critical that parents to be treated as female plants are selected judiciously.

Markers are important to plant breeding as was previously discussed. Some markers may be used to distinguish between selfed and hybrid seed on the female plant. For example, in sorghum, waxy endosperm is conditioned by a recessive allele while normal endosperm is under the control of the dominant allele. If a waxy female is crossed with a normal male, all  $F_1$  seed with waxy endosperm would be products of selfing (undesirable) while normal seed would indicate a successful hybrid. In terms of flower characteristics, bigger flowers are easier to handle than tiny ones. Whenever possible, the parent with bigger flowers should be used as female.

Another critical aspect of flower physiology is the age of the flower when it is most receptive to pollination. The breeder usually determines the optimal stage of flower maturity by examining its physical appearance. Tell-tale signs are variable among species. Usually, fully opened flowers would have already been pollinated by undesirable pollen. Soybean flowers are emasculated in the bud stage just as the petals begin to show through the bud. Rice is ready in the boot stage, whereas wheat is best emasculated when florets are light green with well-developed anthers and feathery stigmas that extend about a quarter of the length of the florets. Furthermore, flowers in the same inflorescence have different maturity levels. In species such as the broad bean, the first inflorescence is more suitable for crossing than later ones; also, flowers at the base and middle of the inflorescence give better results than those at the top.

### Emasculation

The process of making a bisexual flower female by removing the male parts or incapacitating them is called **emasculati**n. It should be pointed out right away that

emasculcation is not a universal requirement for artificial crossing of plants. Species with fertility-regulating mechanisms (e.g., male sterility, self-incompatibility, protogyny, monoecy, dioecy) may be crossed without the often tedious and time-consuming process of emasculation.

### Factors to consider for success

Some of these factors were discussed above. Apart from picking the right flowers, it is critical to know the duration of stigma receptivity and pollen viability. The maximum time between emasculation and pollination that can be tolerated varies among species. Sometimes, it is convenient to emasculate flowers and pollinate at a later time, either during the same day or even later. The caution to observe is that prolonged delay between the two operations increases the chance of contamination from undesirable pollen. To reduce this risk, emasculated flowers may be covered with bags (e.g., glassine bag, cloth bag).

Pollen quality and quantity varies with the weather and time of day. For example, in chickpea, some breeders prefer to emasculate in the evening and pollinate in the morning. Because emasculation is done before anthers are mature in species such as wheat and barley, pollination is done 2–3 days later, when the stigma is receptive. In extreme cases, such as in sugar beet, pollination may immediately follow emasculation or be delayed for up to 12 days.

### Methods of emasculation

There are several techniques of emasculation used by plant breeders that include the use of instruments or chemicals. A pair of forceps or tweezers is one of the most widely used instruments in the emasculation of flowers. Different shapes and sizes are used according to the size and structure of the flower. The methods of emasculation may be classified as direct or indirect.

#### *Direct anther emasculation*

The technique of removing anthers from selected flowers is the most common procedure for the emasculation of flowers (usually using a pair of forceps). When handling plants with inflorescences, it is important to first thin out the bunch by removing immature flowers as well as old ones. This will improve the survival of the emasculated flowers. Breeders of various crops have developed convenient ways of removing the anthers.

Sometimes, the sepals are first removed, followed by the petals, before access is gained to the anthers. In soybean and sesame, a skilled person may be able to remove the petals and anthers in one attempt. In flowers such as soybean, the pedicel is easily broken as a result of physical handling of the delicate flower during emasculation. In wheat and barley, the florets are clipped with scissors. Specific techniques for specific crops are discussed in Part II of this book.

#### *Indirect anther emasculation*

In these methods, the anthers are incapacitated without being removed from the flower. Incapacitation is achieved in several ways.

- 1 **Thermal inactivation.** The inflorescence is first thinned out to leave only flowers at the proper stage for emasculation. It is then immersed in hot water (e.g., held in a thermos bottle) to kill the pollen without injuring the pistil. The temperature and time of emasculation is variable (e.g., 43°C for 5 minutes in rice; 47–48°C for 10 minutes in sorghum). The inflorescence is allowed to dry before pollinating in about 30–60 minutes' time.
- 2 **Alcohol emasculation.** In species such as alfalfa, the raceme is immersed in 57% ethanol for 10 seconds and then rinsed in water for a few seconds.
- 3 **Commercial gametocides.** These are chemicals designed to kill the anthers (e.g., sodium methyl arsenate).

If pollination will not immediately follow emasculation, the flowers should be covered and tagged with an appropriate label, to exclude contaminants.

## Pollination

### Collection and storage

In some species (e.g., soybean) pollination immediately follows emasculation. In this case, there is no need for storage. Fresh pollen gives the best success of crossing. Good pollen flowers may be picked and placed in a Petri dish or some suitable container for use. In some species, mechanical vibrations may be used to collect pollen. Pollen is most copious at peak anthesis. Generally, pollen loses viability quickly. However, in some species, pollen may be stored at a cool temperature and appropriate humidity for the species for an extended period of time.

### Application of pollen

Commonly, pollen is applied directly to the stigma by using a fine brush or dusting off the pollen onto the stigma directly from the flower of the pollen source (e.g., the staminal column may be used as brush). Sometimes, a tooth pick or pointed object is used to deposit pollen on the stigma. In some flowers, pollen deposition is made without direct contact with the stigma. Instead, pollen may be injected into a sack covering the emasculated inflorescence and agitated to distribute the pollen over the inflorescence. A key precaution against contamination during pollination is for the operators to disinfect their hands and tools between pollinations, when different varieties are involved. It is critical to tag the pollinated flower for identification at the time of harvesting.

### Number of F<sub>1</sub> crosses to make

There are practical factors to consider in deciding on the number of crosses to make for a breeding project. These include the ease of making the crosses from the stand-point of floral biology, and the constraints of resources (labor, equipment, facilities, funds). It is easier to make more crosses in species in which emasculation is not needed (e.g., monoecious and dioecious species) than in bisexual species. Some breeders make a small number of carefully planned crosses, while others may make thousands of cross combinations.

Generally, a few hundred cross combinations per crop per year would be adequate for most purposes for species in which the F<sub>1</sub> is not the commercial product. More crosses may be needed for species in which hybrids are commonly produced, in order to discover heterotic combinations. As will be discussed next, breeding programs that go beyond the F<sub>1</sub> usually require very large F<sub>2</sub> populations. Regarding the number of flowers per cross combination, there is variation according to fecundity. Species such as tomato may need only one or two crosses, since each fruit contains over 100 seeds. Plants that tiller also produce large numbers of seed.

### Genetic issues in hybridization

#### Immediate effect

The immediate effect of hybridization is the assembly of two different genomes into a newly created individual. Several genetic consequences may result from such a

union of diverse genomes, some of which may be desirable, some not.

- 1 **Expression of recessive lethal gene.** Crossing may bring together recessive lethal genes into the expressible homozygous state. The resulting hybrid may die or lose vigor. By the same token, hybridization can also mask the expression of a recessive allele by creating a heterozygous locus. Individuals carry a certain genetic load (or genetic burden), representing the average number of recessive lethal genes carried in the heterozygous condition by an individual in a population. Selfing or inbreeding predisposes an individual to having deleterious recessive alleles, which were protected in the heterozygous state, expressed in the homozygous recessive form.
- 2 **Heterosis.** Genes in the newly constituted hybrid may complement each other to enhance the vigor of the hybrid. The phenomenon of hybrid vigor (heterosis) is exploited in hybrid seed development (see Chapter 18).
- 3 **Transgressive segregation.** Hybrids have features that may represent an average of the parental features, or a bias toward the features of one parent, or even new features that are unlike either parent (transgressive segregates). When the parents “nick” in a cross, transgressive segregates, with performances superseding either parent, are likely to occur in the segregating population.

#### Subsequent effect

The subsequent effect of hybridization, which is often the reason for hybridizing parents by breeders, occurs in the F<sub>2</sub> and later generations. By selfing the F<sub>1</sub> hybrid, the parental genes are reorganized into new genetic matrices in the offspring. This occurs through the process of meiosis, a nuclear division process that occurs in flowering plants. Contrasting alleles segregate and subsequently recombine in the next generation to generate new variability. Furthermore, the phenomenon of crossing over that leads to the physical exchange of parts of chromatids from homologous chromosomes provides an opportunity for recombination of linked genes, also leading to the generation of new variation.

#### Gene recombination in the F<sub>2</sub>

The goal of crossing for generating variability for selection is to produce a large number of gene recombinations from the parents used in the cross. In hybrid seed programs, the F<sub>1</sub> is the end product for commercial use.

However, in other crosses, the  $F_2$  and subsequent generations are evaluated to select genotypes that represent the most desirable recombination of parental genes. The  $F_2$  generation has the largest number of different gene combinations of any generation following a cross. The critical question in plant breeding is the size of  $F_2$  population to generate in order to have the chance of including the ideal homozygous recombinant for all the desirable genes in the parent. Three factors determine the number of gene recombinations that would be observed in an  $F_2$  population:

- 1 The number of gene loci for which the parents in a cross differ.
- 2 The number of alleles at each locus.
- 3 The linkage of the gene loci.

Plant breeders are often said to play the numbers game. Table 10.1 summarizes the challenges of breeding in terms of size of the  $F_2$  population to grow. If the parents differ by only one pair of allelic genes, the breeder needs to grow at least four plants in the  $F_2$  to have the chance to observe all the possible gene combinations (according to Mendel's laws). On the other hand, if the parents differ in 10 allelic pairs, the minimum  $F_2$  population size needed is 1,048,576 (obtained by the formula  $4^n$ , where  $n$  is the number of loci). The frequencies illustrate how daunting a task it is to select for quantitative traits.

The total possible genotypes in the  $F_2$  based on the number of alleles per locus is given by the relationship  $[k(k + 1)/2]^n$  where  $k$  is the number of alleles at each locus, and  $n$  is the number of heterozygous loci. With one heterozygote and two alleles, there will be only three kinds of genotypes in the  $F_2$ , while with one heterozygote and four alleles, there will be 10. The effect on gene recombination by linkage is more important than for the number of alleles. Linkage may be desirable

or undesirable. Linkage reduces the frequency of gene recombination (it increases parental types). The magnitude of reduction depends on the phase: the **coupling phase** (with both dominant gene loci in one parent, e.g.,  $AB/ab$ ) and the **repulsion phase** (with one dominant and one recessive locus in one parent, e.g.,  $Ab/ab$ ). The effect of linkage in the  $F_2$  may be calculated as  $\frac{1}{4}(1 - P)^2 \times 100$  for the coupling phase, and  $\frac{1}{4}P^2 \times 100$  for the repulsion phase, for the proportion of  $AB/AB$  or  $ab/ab$  genotypes in the  $F_2$  from a cross between  $AB/ab \times Ab/aB$ . Given, for example, a crossing over value ( $P$ ) of 0.10, the percentage of the homozygotes will be 20.25% in the coupling phase versus only 0.25% in the repulsion phase. If two genes were independent (crossing over value = 0.50), only 6.25% homozygotes would occur. The message here is that the  $F_2$  population should be as large as possible.

With every advance in generation, the heterozygosity in the segregating population decreases by 50%. The chance of finding a plant that combines all the desirable alleles decreases as the generations advance, making it practically impossible to find such a plant in advanced generations. Some calculations by J. Sneep will help clarify this point. Assuming 21 independent gene pairs in wheat, he calculated that the chance of having a plant with all the desirable alleles (either homozygous or heterozygous) are one in 421 in the  $F_2$ , one in 49,343 in the  $F_3$ , and one in 176,778 in the  $F_4$ , and so on. However, to be certain of finding such a plant, he recommended that the breeder grow four times as many plants.

Another genetic consequence of hybridization is the issue of **linkage drag**. As previously noted, genes that occur in the same chromosome constitute a linkage block. However, the phenomenon of crossing over provides an opportunity for linked genes to be separated and not inherited together. Sometimes, a number of

**Table 10.1** The variability in an  $F_2$  population as affected by the number of genes that are different between the two parents.

Number of heterozygous loci ( $n$ )	Number of heterozygous in $F_2$ ( $2^n$ )	Number of different genotypes in $F_2$ ( $3^n$ )	Minimum population size for a chance to include each genotype ( $4^n$ )
1	2	3	4
2	4	9	16
6	64	729	4,096
10	1,024	59,049	1,048,576
15	32,768	14,348,907	1,076,741,824

genes are so tightly linked they are resistant to the effect of recombination. Gene transfer by hybridization is subject to the phenomenon of linkage drag, the unplanned transfer of other genes associated with those targeted. If a desired gene is strongly linked with other undesirable genes, a cross to transfer the desired gene will invariably be accompanied by the linked undesirable genes.

### Types of populations generated through hybridization

A breeding program starts with an initial population that is obtained from previous programs, existing variable populations (e.g., landraces), or is created through a planned cross. Hybridization may be used to generate a wide variety of populations in plant breeding, ranging from the very basic two-parent cross (single cross) to very complex populations in which hundreds of parents could be involved. Simple crosses are the most widely used in breeding. Commercial hybrids are mostly produced by single crosses. Complex crosses are important in breeding programs where the goal is population improvement. Hybridization may be used to introgress new alleles from wild relatives into breeding lines. Because the initial population is critical to the success of the breeding program, it cannot be emphasized enough that it be generated with much planning and thoughtfulness.

Various mating designs and arrangements are used by breeders and geneticists to generate plant populations. These designs require some type of cross to be made. Factors that affect the choice of a mating design, as outlined by C. Stuber include: (i) the predominate type of pollination (self- or cross-pollinated); (ii) the type of crossing used (artificial or natural); (iii) the type of pollen dissemination (wind or insect); (iv) the presence of a male-sterility system; (v) the purpose of the project (for breeding or genetic studies); and (vi) the size of the population required. In addition, the breeder should be familiar with how to analyze and interpret or use the data to be generated from the mating.

The primary purpose of crossing is to expand genetic variability by combining genes from the parents involved in the cross to produce offspring that contain genes they never had before. Sometimes, multiple crosses are conducted to generate the variability in the base population to begin the selection process in the program. Based on how the crosses are made and their effects on the genetic structure of the plants or the population, methods of crossing may be described as either divergent or convergent.

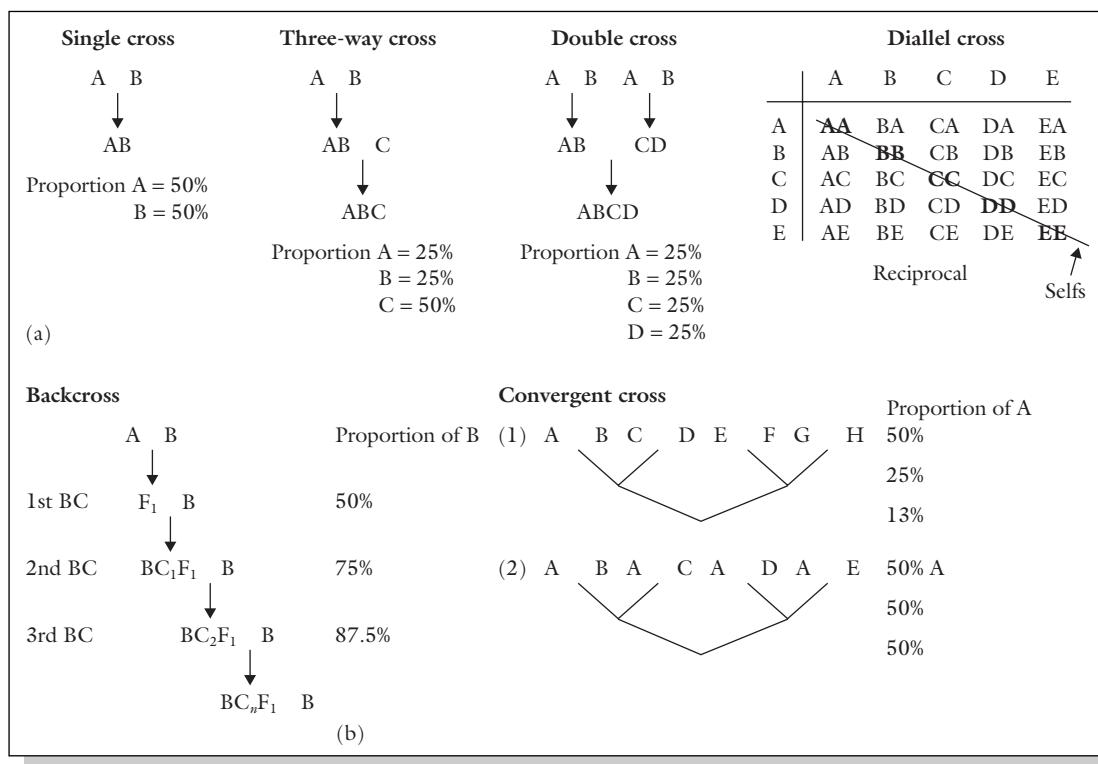
### Divergent crossing

Genetically divergent parents are crossed for recombination of their desirable genes. To optimize results, parents should be carefully selected to have the maximum number of positive traits and a minimum number of negative traits (i.e., elite  $\times$  elite cross). This way, recombinants that possess both sets of desirable traits will occur in significant numbers in the  $F_2$ . The  $F_1$  contains the maximum number of desirable genes from both parents. There are several ways to conduct divergent crosses (Figure 10.1a).

- 1 Single cross.** If two elite lines are available that together possess adequate traits, one cross [single cross ( $A \times B$ )] may be all that is needed in the breeding program.
- 2 Three-way cross.** Sometimes, desirable traits occur in several cultivars or elite germplasm. In this case, multiple crosses may be required in order to have the opportunity of obtaining recombinants that consist of all the desirable traits. The method of three-way crosses [ $(A \times B) \times C$ ] may be used. If a three-way cross product will be the cultivar, it is important that the third parent ( $C$ ) be adapted to the region of intended use.
- 3 Double cross.** A double cross is a cross of two single crosses [ $(A \times B) \times (C \times D)$ ]. The method of successive crosses is time-consuming. Further, complex crosses such as double crosses have a low frequency of yielding recombinants in the  $F_2$  that possess a significant number of desirable parental genes. When this method is selected, the targeted desirable traits should be small (about 10). The double-cross hybrid is more genetically broad-based than the single-cross hybrid but is more time-consuming to make.
- 4 Diallel cross.** A diallel cross is one in which each parent is crossed with every other parent in the set (complete diallel), yielding  $n - (n - 1)/2$  different combinations (where  $n$  is the number of entries). This method entails making a large number of crosses. Sometimes, a partial diallel is used in which only certain parent combinations are made. The method is tedious to apply to self-pollinated species. Generally, it is a crossing method for genetic studies.

### Convergent crossing

These are conservative methods of crossing plants. The primary goal of convergent crossing is to incorporate a specific trait into an existing cultivar without losing any



**Figure 10.1** The basic types of crosses used by plant breeders. Some crosses are divergent (a) while others are convergent (b).

of the existing desirable traits. Hence, one (or several) parent(s) serves as a donor of specific genes and is usually involved in the cross only once. Subsequent crosses entail crossing the desirable parent (recurrent parent) repeatedly to the F<sub>1</sub>, in order to retrieve all the desirable traits. A commonly used convergent cross is the **back-cross** (Figure 10.1b) (see Chapter 16).

### Issue of reproductive isolation barriers

Hybridization is often conducted routinely without any problems when individuals from the same species are involved, provided there are no fertility-regulating mechanisms operating. Even when such mechanisms exist, hybridization can be successfully conducted by providing appropriate pollen sources. Sometimes, plant breeders are compelled to introduce desired genes from distant relatives or other species. Crossing plants from two different species or sometimes even plants from different genera is more challenging and has limited success. Often, the breeder needs to use additional techniques (e.g., embryo rescue) to intervene at some point

in the process in order to obtain a mature hybrid plant. This kind of crossing involving parents from different species is called a **wide cross** and is described further below.

Reproductive isolation barriers may be classified into three categories as suggested by researchers such as G. L. Stebbins, T. Dobzhansky, and D. Zohary (Table 10.2).

**Table 10.2** A summary of the reproductive isolation barriers in plants as first described by G. L. Stebbins.

#### External barriers

**Spatial isolation mechanisms:** associated with geographic distances between two species

**Prefertilization reproductive barriers:** prevents union of gametes. Includes ecological isolation (e.g., spring and winter varieties), mechanical isolation (differences in floral structures), and gametic incompatibility

#### Internal barriers

**Postfertilization reproductive barriers:** leads to abnormalities following fertilization (hybrid inviability or weakness and sterility of plants)

The purpose of these barriers is to maintain the genetic integrity of the species by excluding gene transfer from outside species. Some barriers occur before fertilization, some after fertilization. These barriers vary in degree of difficulty to overcome through breeding manipulations.

#### *Spatial isolation*

**Spatial isolation** mechanisms are usually easy to overcome. Plants that have been geographically isolated may differ only in photoperiod response. In which case, the breeder can cross the plants in a controlled environment (e.g., greenhouse) by manipulating the growing environment to provide the proper duration of day length needed to induce flowering.

#### *Prefertilization reproductive barriers*

These barriers occur between parents in a cross. Crops such as wheat have different types that are **ecologically isolated** – there are spring wheat types and winter wheat types. Flowering can be synchronized between the two groups by, for example, vernalization (a cold temperature treatment that exposes plants to temperatures of about 3–4°C) of the winter wheat to induce flowering (normally accomplished by exposure to winter conditions). **Mechanical isolation** may take the form of differences in floral morphology that prohibit the same pollinating agent (e.g., insect) from pollinating different species. A more serious barrier to gene transfer is **gametic incompatibility** whereby fertilization is prevented. This mechanism is a kind of self-incompatibility (see Chapter 4). The mechanism is controlled by a complex of multiple allelic systems of *S* genes that prohibit gametic union. The breeder has no control over this barrier.

#### *Postfertilization reproductive barriers*

These barriers occur between hybrids. After fertilization, various hindrances to proper development of the embryo (hybrid) may arise, sometimes resulting in abortion of the embryo, or even formation of a haploid (rather than a diploid). The breeder may use embryo rescue techniques to remove the embryo and culture it to full plant. Should the embryo develop naturally, the resulting plant may be unusable as a parent in future breeding endeavors because of a condition called **hybrid weakness**. This condition is caused by factors such as disharmony between the united genomes. Some hybrid plants may fail to flower because of **hybrid**

**sterility** ( $F_1$  sterility) resulting from meiotic abnormalities. On some occasions, the hybrid weakness and infertility manifest in the  $F_2$  and later generations (called **hybrid breakdown**).

## **Wide crosses**

The first choice of parents for use in a breeding program are cultivars and experimental materials with desirable traits of interest. Most of the time, plant breeders make elite × elite crosses (they use adapted and improved materials). Even though genetic gains from such crosses may not always be dramatic, they are nonetheless significant enough to warrant the practice. After exhausting the variability in the elite germplasm as well as in the cultivated species, the breeder may look elsewhere, following the recommendation by Harlan and de Wet, as previously noted. These researchers proposed that the search for desired genes should start from among materials in the primary gene pool (related species), then proceed to the secondary gene pool, and if necessary, to the tertiary gene pool. Crosses involving materials outside the cultivated species are collectively described as **wide crosses**. When the wide cross involves another species, it is called an **interspecific cross** (e.g., kale). When it involves a plant from another genus, it is called an **intergeneric cross** (e.g., wheat).

#### **Objectives of wide crosses**

Wide crosses may be undertaken for practical and economic reasons, research purposes, or to satisfy curiosity. Specific reasons for wide crosses include the following:

- 1 **Economic crop improvement.** The primary purpose of wide crosses is to improve a species for economic production by transferring one or a few genes, or segment of chromosomes or whole chromosomes, across interspecific or intergeneric boundaries. The genes may condition a specific disease or pest resistance, or may be a product quality trait, amongst other traits. In some species such as sugarcane, cotton, sorghum, and potato, hybrid vigor is known to have accompanied certain crosses.
- 2 **New character expression.** Novelty is highly desirable in the ornamental industry. Combining genomes from diverse backgrounds may trigger a complementary gene action or even introduce a few genes that could produce previously unobserved phenotypes that may be superior to the parental expression of both qualitative and quantitative traits.

- 3 Creation of new allopoloids.** Wide crosses often produce sterile hybrids. The genome of such hybrids can be doubled to create a new fertile allopoloid species (a polyploid with the genomes of different species), such as triticale.
- 4 Scientific studies.** Cytogenetic studies following a wide cross may be used to understand the phylogenetic relationships between the parents of a cross.
- 5 Curiosity and aesthetic value.** Wide crosses may produce unique products of ornamental value, which can be useful to the horticultural industry. Sometimes just being curious is a good enough reason to try new things.

### Selected success with wide crosses

Developing commercial cultivars with genes introduced from the wild can be an expensive and long process (see prebreeding in Chapter 6). Some linkages with wild genes need to be broken. In tomato, it took 12 years to break the linkage between nematode resistance and undesirable fruit characteristics. Nonetheless, some significant successes have been accomplished through wide crosses.

#### Natural wide crosses

Natural wide crosses have been determined by scientists to be the origin of numerous modern-day plants of economic importance. Ornamentals such as irises, cannas, dahlias, roses, and violets, are among the list of such species. In tree crops, apples, cherries, and grapes are believed to have originated as natural wide crosses, and so are field crops such as wheat, tobacco, and cotton, as well as Irish and sweet potatoes. Most natural wide cross products of economic value to modern society are used as ornamentals and are usually propagated vegetatively. This led G. L. Stebbins to observe that wide crosses may be more valuable in vegetatively propagated species than seed-propagated species.

#### Synthetic (artificial) wide crosses

Apart from natural occurrences, plant breeders over the years have introgressed desirable genes into adapted cultivars from sources as close as wild progenitors to distant ones such as different genera. Practical applications of wide crosses may be grouped into three categories.

- 1 Gene transfer between species with the same chromosome number.** Wide crosses between two tomato species, *Lycopersicon pimpinellifolium* × *L.*

*esculentum*, have been conducted to transfer resistance genes to diseases such as leaf mold and *Fusarium* wilt. Gene transfers in which both parents have identical chromosome numbers is often without complications beyond minor ones (e.g., about 10% reduction in pollen fertility). It is estimated that nearly all commercially produced tomatoes anywhere in the world carry resistance to *Fusarium* that derived from a wild source.

- 2 Gene transfer between species with a different number of chromosomes.** Common wheat is a polyploid (an allohexaploid) with a genomic formula of *AABBDD*. It has 21 pairs of chromosomes. There is diploid wheat, einkorn (*Triticum monococcum*), with seven pairs of chromosomes and a genomic formula of *AA*. There are several tetraploid wheats (*AABB*) such as emmer wheat (*T. dicoccum*). Transfer of genes from species of lower ploidy to common wheat is possible (but not always the reverse). Stem rust resistance is one such gene transfer that was successful.
- 3 Gene transfer between two genera.** Common wheat comprises three genomes of which one (*DD*) is from the genus *Aegilops*. Consequently, gene transfers have been conducted between *Triticum* and *Aegilops* (e.g., for genes that confer resistance to leaf rust).

### Overcoming the challenges of reproductive barriers

The reproductive barriers previously discussed confront plant breeders who attempt gene transfer between distant genotypes via hybridization. The primary challenge of wide crosses is obtaining fertile F<sub>1</sub> hybrids, because of the mechanisms that promote, especially, gametic incompatibility. As previously indicated, this mechanism acts to prevent: (i) the pollen from reaching the stigma of the other species; (ii) germination of the pollen and inhibition of growth of the pollen tube down the style, or the union of the male gamete and the egg if the pollen tube reaches the ovary; and (iii) the development of the zygote into a seed and the seed into a mature plant. Gametic incompatibility ends when fertilization occurs. However, thereafter, there are additional obstacles to overcome. Gametic incompatibility and hybrid breakdown are considered to be barriers to hybridization that are outside the control of the breeder.

Several techniques have been developed to increase the chance of recovering viable seed and plants from a wide cross. These techniques are based on the nature of the barrier. All techniques are not applicable to all species.



## Industry highlights

### *The use of the wild potato species, Solanum etuberosum, in developing virus- and insect-resistant potato varieties*

Richard Novy

USDA-ARS, Aberdeen, ID 83210, USA

#### *Historical background*

Oftentimes referred to as the "Irish" potato, *Solanum tuberosum* subsp. *tuberousum* might more aptly be termed the "Inca" potato. The origin of the fourth most widely grown cultivated crop in the world is thought to be the Central Andes region of South America, with the possibility of independent domestication in Chile as well. Potato was an important food crop for the Incas, but it is unlikely that they were the civilization responsible for its domestication. Potato food remnants have been found in preceramic archeological sites in South America that date to over 5,000 years ago, indicating that the potato truly is an ancient food crop.

The Spanish and English are thought to have brought this New World crop back to Europe in the late 16th century. Adapted to form tubers under the short day conditions near the equator (approximately 12 hours), potatoes did not successfully produce tubers in most northern latitudes prior to being killed by freezing temperatures in the fall. The exceptions were the milder climates of Spain, Italy, southern France, and Ireland, where the potato was maintained as a botanical oddity in private and botanical gardens. Over the course of 150–200 years, potato clones (with the help of man) were identified and propagated that formed tubers under the longer day lengths of the northern latitudes. This environmental adaptation allowed for the expansion and adoption of potato as a food crop in Europe, and eventually throughout the world.

#### *Viruses of potato*

Accompanying the introduction of the potato to Europe were pathogens that had coevolved with the crop, most notable being potato viruses X and Y (PVX and PVY) and potato leafroll virus (PLRV). These viruses are transmitted from an infected plant to the tubers it produces. When virus-infected tubers are cut and used to establish the potato crop the following growing season (asexual propagation), plants developing from the tuber seed are infected with the virus as well. Symptoms of virus infection in potato include stunting, chlorosis/necrosis of leaf tissue, and, in the case of PLRV (as indicated by the name) rolling or cupping of leaves. Total yield in a growing season can be reduced by as much as 80% if virus-infected seed is used. Transmission of PVY and PLRV from infected to healthy plants is mediated by aphids, most notably green peach aphid; PVX is transmitted to healthy plants via mechanical contact with an infected plant or with PVX-contaminated field equipment. The detrimental impact of viruses on potato was termed "degeneration" or "running out" by early growers of potato who did not yet know of the existence of viruses.

Potato varieties with resistances to viruses can be effective in reducing crop losses. Cultivated potato is fortunate in having > 200 wild *Solanum* relatives, many of which have been identified as virus resistant. In the United States, potato species collected in Mexico and Central and South America are maintained at the Potato Genebank in Sturgeon Bay, Wisconsin. This species collection has been systematically screened for resistance to the major pests and diseases of potato. Many species have been identified with high levels of resistance to PVX and PVY, with a lesser number identified as having desirable levels of PLRV resistance.

Ideally, from the standpoint of a potato breeder, it is desirable to work with species that have a high level of resistance to all three viruses. A search of the Potato Genebank collection – consisting of 5,634 introductions representing 168 species – identified only one introduction (PI 245939) of the wild potato species, *S. etuberosum*, as having a high level of resistance to PVY, PVX, and PLRV. This accession also was identified as having resistance to green peach aphid – a primary insect vector of PVY and PLRV. The introgression of these multiple virus and insect vector resistances from *S. etuberosum* into cultivated potato will be the focus of the remainder of this box.

#### *Solanum etuberosum: its use in the genetic improvement of potato*

*Solanum etuberosum* is a wild potato species endemic to Chile. Its natural habitat is among rocks on slopes with seepage, or along streams. It is generally found in the open, or in the shade of trees and shrubs (Correll 1962). It is notable for its large, deep purple flowers (Figure 1). The attractiveness and abundance of its flowers and its striking foliage led a taxonomist in 1835 to propose that it be grown as a hardy perennial for ornamental purposes in England (Correll 1962). It also is notable among wild potato species in that it does not form tubers.

A diploid ( $2n = 2x = 24$ ) species, *S. etuberosum* does not cross readily with either tetraploid ( $2n = 4x = 48$ ) or dihaploid ( $2n = 2x = 24$ ) forms of cultivated potato. Bridging species and ploidy manipulations have been used by breeders in synthesizing *etuberosum*–*tuberousum* hybrids with limited success.



**Figure 1** Flower and foliage of *S. etuberosum*. Considered a weed in Chile, this wild relative of potato has desirable genetic resistances to viruses and insects that plague cultivated potato.

Backcrossing of somatic hybrids to potato cultivars was undertaken to improve tuberization and yield. Crosses using somatic hybrids as the male parent yielded few berries and no seeds. Stylar analyses showed blockage of somatic hybrid pollen tube growth generally occurred in the upper third of *Gp. tuberosum* styles. Pollen tube blockage of cultivated potato was not observed in the styles of somatic hybrids: 503 pollinations produced 99 berries containing 24 seeds. Five of the seed germinated to produce viable BC<sub>1</sub> progenies that were at or near the tetraploid level (48–49 chromosomes).

The five progenies obtained had much improved tuberization relative to the somatic hybrid parent, while still retaining 11–13 *S. etuberosum* chromosomes. One of the five progenies produced an average of six seeds per berry when crossed to cultivated potato. Viable BC<sub>2</sub> progenies were obtained from this seed. Tuberous BC<sub>2</sub>, now looking like those of cultivated potato, are shown in Figure 2.

#### *Virus and green peach aphid resistances of somatic hybrids and their progeny*

Novy and Helgeson (1994b) analyzed the fusion parents, their somatic hybrids, and the sexual progeny of the somatic hybrids for resistance to PVY following their mechanical inoculation in the greenhouse over a 2-year period. The *S. etuberosum* fusion parent was highly resistant to PVY infection whereas the *tuberosum-berthaultii* fusion parent was highly susceptible. Three somatic hybrids analyzed in this study did not show the high level of resistance found in the *S. etuberosum* parent; however they were significantly more resistant than the cultivars "Katahdin" (moderate field resistance to PVY) and "Atlantic" (PVY susceptible). Five progenies of the somatic hybrids also were analyzed in this study. Three displayed PVY resistance comparable to the somatic hybrid parents, whereas the remaining two were more susceptible with absorbance means comparable to the potato varieties included in the study.

*Solanum etuberosum* also had been identified as having resistance to PLRV and green peach aphid. Resistance to green peach aphid (*Myzus persicae*) can aid in decreasing the transmission of viruses by decreasing aphid population size and subsequent opportunities for virus transmission. However, green peach aphid resistance alone is not adequate to confer the necessary level of resistance needed by the industry. This is especially true in the case of PVY, which can be quickly transmitted by the stylar probings of many different aphid species – species that may not include potato as a primary host and therefore will not be adversely impacted by host plant resistance.

A combination of green peach aphid and PVY/PLRV resistances is the most effective means to reduce virus infection and spread. Novy et al. (2002) evaluated five BC<sub>2</sub> progenies of the *S. etuberosum* somatic hybrids (the recurrent parents being potato varieties) for green peach aphid, PLRV, and PVY resistance. Virus resistances were evaluated in both open field and field cage trials; aphid resistance was evaluated in the field and greenhouse.

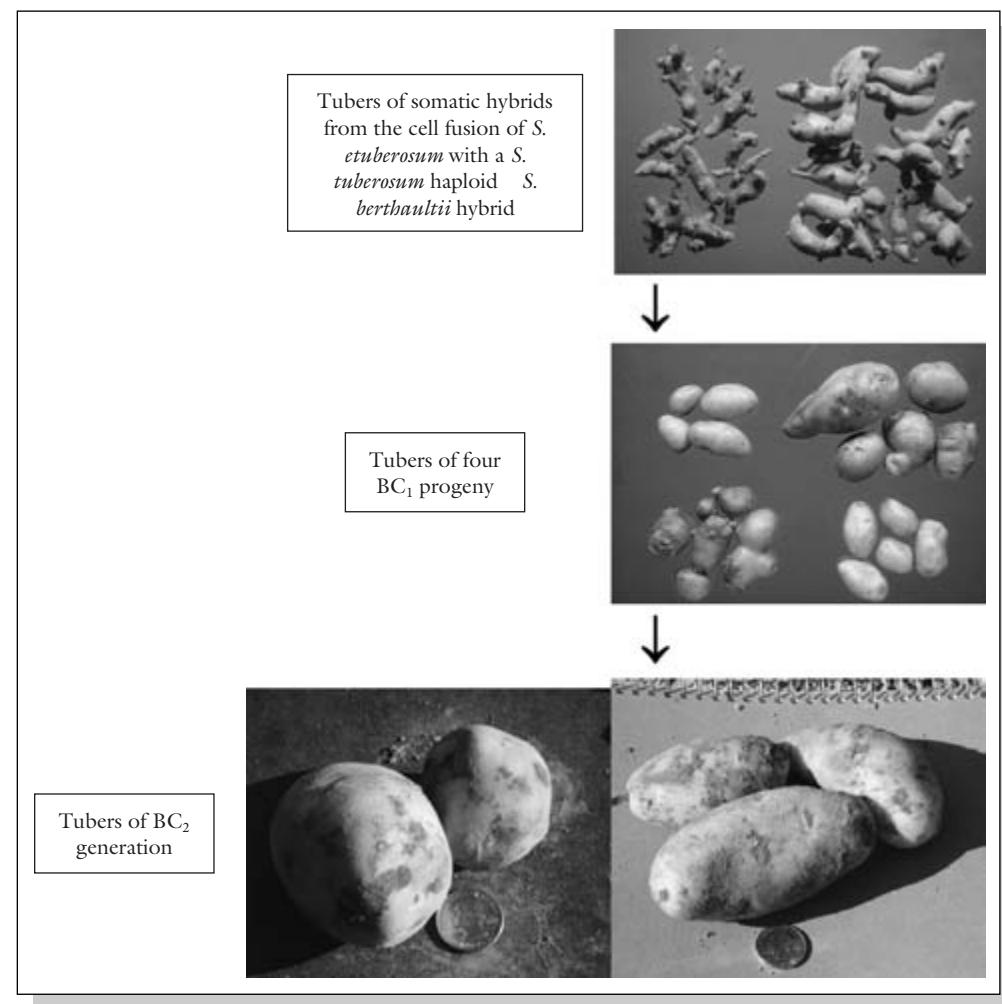
The authors identified resistance to green peach aphid in all *S. etuberosum*-derived BC<sub>2</sub> progeny. Resistance was characterized by reduced adult body size and fecundity. One BC<sub>2</sub> individual also exhibited reduced nymph survival. Prolonged development from nymph to adult also appeared to contribute to reduced aphid populations on the BC<sub>2</sub> relative to susceptible checks.

Another means of circumventing reproductive barriers is to bypass them completely by the use of somatic hybridization. This technique involves the isolation of potato cells from leaf tissue of parental clones, the enzymatic digestion of their cell walls to form protoplasts, and the fusion (using chemicals or electric currents) of parental protoplasts. Fused protoplasts are then cultured on medium whereby they re-form a cell wall and are allowed to divide to form undifferentiated tissue called callus. Calli are then placed on culture media that promotes cell differentiation and the formation of plants. These hybrid plants can then be excised from the calli, induced to form roots, and can then be grown as a normal plant in field or greenhouse environments.

Using somatic hybridization, Novy and Helgeson (1994a) successfully generated hybrids between a *S. etuberosum* clone from virus-resistant PI 245939 and a subsp. *tuberosum* dihaploid × *S. berthaultii* hybrid clone ( $2n = 2x = 24$ ). The trispecies hybrids, based on cytological and molecular analyses, were at or near the expected  $2n = 4x = 48$ . Somatic hybrids had very vigorous foliar growth in the field with limited tuberization (Figure 2); poor tuber type and yield was not unexpected in that half the genome of the somatic hybrids was from non-tuber-bearing *S. etuberosum*.

Backcrossing of somatic hybrids to potato cultivars was undertaken to improve tuberization and yield. Crosses using somatic hybrids as the male parent yielded few berries and no seeds. Stylar analyses showed blockage of somatic hybrid pollen tube growth generally occurred in the upper third of *Gp. tuberosum* styles. Pollen tube blockage of cultivated potato was not observed in the styles of somatic hybrids: 503 pollinations produced 99 berries containing 24 seeds. Five of the seed germinated to produce viable BC<sub>1</sub> progenies that were at or near the tetraploid level (48–49 chromosomes).

The five progenies obtained had much improved tuberization relative to the somatic hybrid parent, while still retaining 11–13 *S. etuberosum* chromosomes. One of the five progenies produced an average of six seeds per berry when crossed to cultivated potato. Viable BC<sub>2</sub> progenies were obtained from this seed. Tuberous BC<sub>2</sub>, now looking like those of cultivated potato, are shown in Figure 2.



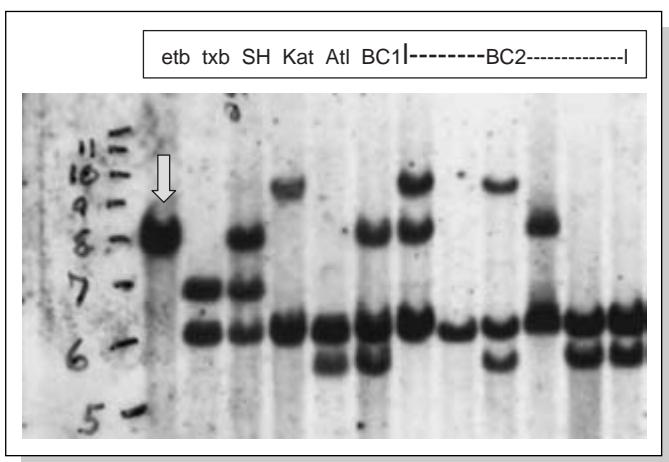
**Figure 2** The progression of tuber type in somatic hybrids of *S. etuberosum* and their backcross progenies in successive hybridizations with cultivated potato.

Analogous to observations in three of the  $BC_1$ , all  $BC_2$  (derived from a PVY-resistant  $BC_1$ ) exhibited statistically significant reduced PVY infection relative to the PVY-susceptible potato variety, "Russet Burbank".  $BC_2$  were found to segregate for resistance to PLRV, with two of five displaying resistance to infection on the basis of field and cage evaluations. Progeny of a PLRV-resistant  $BC_2$  clone also show high levels of resistance to PLRV in field evaluations in Idaho (author, unpublished data). *S. etuberosum*-derived PLRV resistance is highly heritable on the basis of its expression in third generation progeny of the somatic hybrids.

Backcross progenies are being analyzed with molecular markers to identify chromosomal regions from *S. etuberosum* associated with its observed PVY and PLRV resistances (Figure 3). Once identified, prospective regions will be further saturated with additional markers to identify those that are tightly linked to the virus resistances. Such markers can then be used for marker-assisted selection (MAS) in our breeding program. These correlated DNA markers can be used to assess whether an individual will express resistance to PLRV or PVY, thereby speeding the development of potato varieties with the virus resistances of *S. etuberosum*.

#### *Wireworm resistance of backcross progeny derived from somatic hybrids*

Wireworm, a larval stage of a beetle, is the most damaging soil-dwelling pest of potato. The larvae spend 3–4 years in the soil and their entry and feeding into potato tubers detrimentally impacts the marketability of the tuber for fresh consumption or processing.



**Figure 3** Segregation of a RFLP (restriction fragment length polymorphism) specific to *S. etuberosum*; the RFLP probe used was TG65. The arrow indicates the RFLP unique to *S. etuberosum* that also is present in the somatic hybrid (SH), BC<sub>1</sub>, and two of six BC<sub>2</sub>. This molecular marker is not present in the *S. tuberosum* × *S. berthaultii* fusion parent (txb) or the potato cultivars "Katahdin" (Kat) or "Atlantic" (Atl) that were used as parents in the generation of the backcross progeny. Numbers on the side are approximations of the RFLP fragment sizes.

resistant clones had levels  $\geq 47$  mg/100 g. However, the remaining three clones had levels  $\leq 13$  mg/100 g; all three, relative to susceptible "Russet Burbank", had reduced percentages of wireworm-damaged tubers. These data indicate that high total tuber glycoalkaloid levels are not necessary for conferring wireworm resistance – an important finding if wireworm-resistant potato cultivars with acceptable glycoalkaloid levels are to be developed.

#### References

- Correll, D.S. 1962. The potato and its wild relatives. Texas Research Foundation, Renner, TX, 606 pp.  
 Novy, R.G., and J.P. Helgeson. 1994a. Somatic hybrids between *Solanum etuberosum* and diploid, tuber bearing *Solanum* clones. *Theor. Appl. Genet.* 89:775–782.  
 Novy, R.G., and J.P. Helgeson. 1994b. Resistance to potato virus Y in somatic hybrid between *Solanum etuberosum* and *S. tuberosum* × *S. berthaultii* hybrid. *Theor. Appl. Genet.* 89:783–786.  
 Novy, R.G., A. Nasruddin, D.W. Ragsdale, and E.B. Radcliffe. 2002. Genetic resistances to potato leafroll virus, potato virus Y, and green peach aphid in progeny of *Solanum etuberosum*. *Am. J. Potato Res.* 79:9–18.

#### Overcoming barriers to fertilization

**1 Conduct reciprocal crosses.** Generally, it is recommended to use the parent with the larger chromosome number as the female in a wide cross for a higher success rate. This is because some crosses are successful only in one direction. Hence, where there is no previous information about crossing behavior, it is best to cross in both directions.

Insecticide use has been the primary means of controlling wireworm damage. However, insecticides currently used for wireworm control may lose their registration in the future. Genetic resistance to wireworm could be an important component of an integrated pest management (IPM) program, and it was decided to evaluate the progeny of *S. etuberosum* for wireworm resistance.

In collaboration with Dr Juan Alvarez, an entomologist with the University of Idaho, one BC<sub>1</sub> and four BC<sub>2</sub> clones derived from the *S. etuberosum* somatic hybrids were evaluated for wireworm resistance; comparisons of these clones were made relative to a susceptible cultivar, "Russet Burbank". An additional treatment also was included in the evaluation; susceptible "Russet Burbank" was treated with Genesis®, an insecticide commonly used for wireworm control. Four of the five backcross clones had a percentage of damaged tubers comparable to or lower than that observed with the use of Genesis®. Wireworm entry points or "holes" per tuber among the five clones were comparable to the numbers observed with the use of Genesis®.

The resistance of two of the five clones was attributable to high levels of certain chemical compounds naturally produced in the tuber called glycoalkaloids. Total tuber glycoalkaloid concentrations need to be less than 20 mg/100 g tuber fresh weight for safe consumption by humans – these two highly

resistant clones had acceptable total tuber glycoalkaloid levels of  $\leq 13$  mg/100 g; all three, relative to susceptible "Russet Burbank", had reduced percentages of wireworm-damaged tubers. These data indicate that high total tuber glycoalkaloid levels are not necessary for conferring wireworm resistance – an important finding if wireworm-resistant potato cultivars with acceptable glycoalkaloid levels are to be developed.

- 2 Shorten the length of the style.** The pollen tube of a short-styled species may not be able to grow through a long style to reach the ovary. Thus, shortening a long style may improve the chance of a short pollen tube reaching the ovary. This technique has been successfully tried in corn.  
**3 Apply growth regulators.** Chemical treatment of the pistil with growth-promoting substances (e.g., naphthalene acetic acid, gibberellic acid) tends to

promote rapid pollen tube growth or extend the period over which the pistil remains viable.

- 4 **Modify ploidy level.** A diploid species may be converted to a tetraploid to be crossed to another species. For example, narrow leaf trefoil (*Lotus tenuis*,  $2n = 12$ ) was successfully crossed with broadleaf bird's foot trefoil (*L. corniculatus*,  $2n = 24$ ).
- 5 **Use mixed pollen.** Mixing pollen from a compatible species with pollen from an incompatible parent makes it possible to avoid the unfavorable interaction associated with cross-incompatibility.
- 6 **Remove stigma.** In potato, wide crosses were accomplished by removing the stigma before pollination and by substituting it with a small block of agar fortified with sugar and gelatin.
- 7 **Grafting.** Grafting the female parent to the male species has been reported to promote pollen tube growth and subsequent fertilization.
- 8 **Protoplast fusion.** A protoplast is all the cellular component of a cell excluding the cell wall. Protoplasts may be isolated by either mechanical or enzymatic procedures. Mechanical isolation involves slicing or chopping of the plant tissue to allow the protoplast to slip out through a cut in the cell wall. This method yields low numbers of protoplasts. The preferred method is the use of hydrolytic enzymes to degrade the cell wall. A combination of three enzymes – cellulase, hemicellulase, and pectinase – is used in the hydrolysis. The tissue used should be from a source that would provide stable and metabolically active protoplasts. This calls for monitoring plant nutrition, humidity, day length, and other growth factors. Often, protoplasts are extracted from leaf mesophyll or plants grown in cell culture. The isolated protoplast is then purified, usually by the method of flotation. This method entails first centrifuging the mixture from hydrolysis at about  $50\times$  the force of gravity, and then resuspending the protoplasts in a high concentration of fructose. Clean, intact protoplasts float and can be retrieved by pipetting. Protoplasts can also be used to create hybrids *in vitro* (as opposed to crossing mature plants in conventional plant breeding).

#### *Overcoming the problem of inadequate hybrid seed development*

Abnormal embryo or endosperm development following a wide cross may be overcome by using proper parent selection and reciprocal crossing as previously described. In addition, the technique of embryo rescue is an effective and common technique. The embryo is aseptically extracted and nurtured into a full plant under tissue culture conditions (see Chapter 11).

#### *Overcoming lack of hybrid vigor*

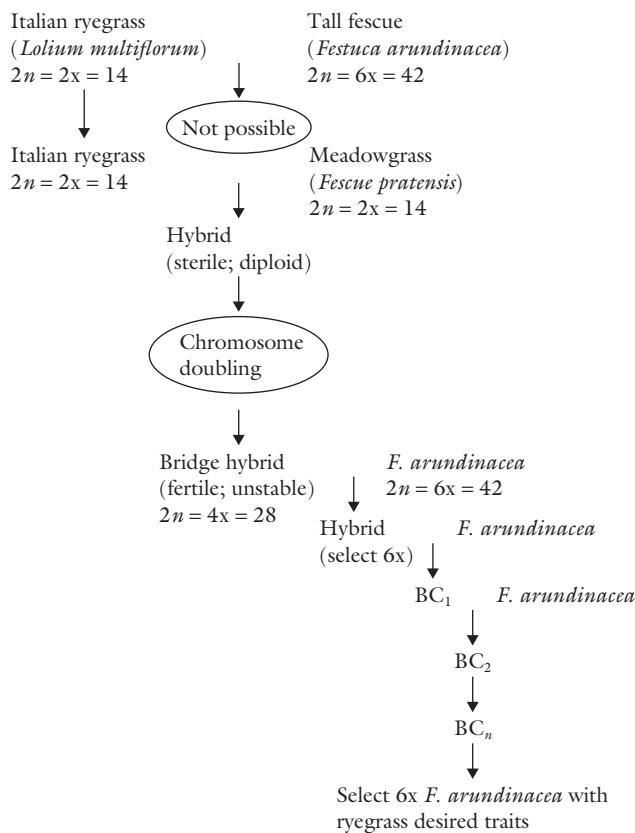
Hybrids may lack the vigor to grow properly to flower and produce seed. Techniques such as proper parent selection, reciprocal crossing, and grafting the hybrid onto one of the parents may help.

#### *Overcoming hybrid sterility*

Sterility in hybrids often stems from meiotic complications due to lack of appropriate pairing partners. Sterility may be overcome by doubling the chromosomes of the hybrid to create pairing mates for all chromosomes, and hence producing viable gametes.

### **Bridge crosses**

**Bridge crossing** is a technique of indirectly crossing two parents that differ in ploidy levels through a transitional or intermediate cross (Figure 10.2). For example



**Figure 10.2** An example of a bridge cross. In order to hybridize Italian ryegrass and tall fescue, the breeder may first make an intermediary cross with meadowgrass, followed by chromosome doubling.

R. C. Buckner and his colleagues succeeded in crossing the diploid Italian ryegrass (*Lolium multiflorum*,  $2n = 2x = 14$ ) with the hexaploid tall fescue (*Festuca arundinacea*,  $2n = 6x = 42$ ) via the bridge cross technique. The intermediate cross was between *L. multiflorum* and diploid meadow fescue (*Fescue pratensis*,  $2n = 2x = 14$ ). The resulting embryo was rescued and the chromosome number doubled to produce a fertile but genetically unstable tetraploid hybrid (ryegrass–meadow fescue), to be used as a genetic bridge. Using tall fescue as the recipient, the genetic bridge was repeatedly backcrossed to tall fescue. A 42-chromosome cultivar of tall fescue with certain Italian ryegrass traits was eventually recovered and stabilized.

## Developing new species via wide crossing

A species is defined as a population of individuals capable of interbreeding freely with one another but which, because of geographic, reproductive, or other barriers, do not in nature interbreed with members of other species. One of the long-term “collaborative” breeding efforts is the development of the triticale ( $\times$  *Triticosecale* Wittmack). The first successful cross, albeit sterile, is traced back to 1876; the first fertile triticale was produced in 1891. The development of this new species occurred over a century, during which numerous scientists tweaked the procedure to reach its current status where the crop is commercially viable. Triticale is a wide cross between *Triticum* (wheat) and *Secale* (rye), hence triticale (a contraction of the two names). It is a predominantly self-fertilizing crop, and the breeding of triticale is discussed in Chapter 13.

## References and suggested reading

- Acquaah, G. 2003. Understanding biotechnology: An integrated and cyberbased approach. Prentice Hall, Upper Saddle River, NJ.
- Acquaah, G. 2004. Horticulture: Principles and practices, 3rd edn. Prentice Hall, Upper Saddle River, NJ.
- Chandler, J.M., and B.H. Beard. 1983. Embryo culture of *Helianthus* hybrids. Crop Sci. 23:1004–1007.
- Forsberg, R.A. (ed.). 1985. Triticale. Crop Science of America Special Publication No. 9. American Society of Agronomy, Madison, WI.
- Harlan, J.R., and J.M.J. de Wet. 1971. Toward a rational classification of cultivated plants. Taxon 20:509–517.
- Morrison, L.A., O. Riera-Lizaraza, L. Cremieux, and C.A. Mallory-Smith. 2002. Jointed goatgrass (*Aegilops cylindrica* Host)  $\times$  wheat (*Triticum aestivum* L.) hybrids: Hybridization dynamics in Oregon wheat fields. Crop Sci. 42:1863–1872.
- Singh, A.K., J.P. Moss, and J. Smartt. 1990. Ploidy manipulations for interspecific gene transfer. Adv. Agron. 43:199–240.
- Stalker, H.T. 1980. Utilization of wild species for crop improvement. Adv. Agron. 33:111–147.
- Stoskopf, N.C. 1993. Plant breeding: Theory and practice. Westview Press, Boulder, CO.
- Zohary, D. 1973. Gene-pools for plant breeding. In: Agricultural genetics (R. Moav, ed.). John Wiley & Sons, New York.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 A hybrid is a product of unidentical parents.
- 2 Emasculation is undertaken to make a flower female.
- 3 An integeneric cross occurs between two species.
- 4 Wheat is a product of a wide cross.
- 5 Bridge crosses are used to facilitate crosses between two parents of identical ploidy levels.

### Part B

Please answer the following questions:

- 1 What is hybridization?
- 2 What are wide crosses?

- 3 Give three specific reasons why wide crosses may be undertaken.
- 4 Explain the phenomenon of linkage drag.
- 5 Give examples of major crops that arose by wide crosses.

**Part C**

Please write a brief essay on each of the following topics:

- 1 Discuss the basic steps in artificial hybridization.
- 2 Discuss the challenges of wide crosses.
- 3 Discuss the techniques used for overcoming the challenges to wide crosses.
- 4 Discuss the technique of bridge crossing.
- 5 Discuss the genetic issues in hybridization.



# 11

## Tissue culture and the breeding of clonally propagated plants

---

### Purpose and expected outcomes

*The cell is the fundamental unit of structure and function of a plant. Conventional plant breeding entails manipulating plants at the whole-plant level. However, modern technologies enable scientists to manipulate plants at the tissue and cellular levels. Tissues and even single cells can be nurtured into full plants. The technique of tissue culture may be used to assist plant breeders who conduct wide crosses to be able to nurture young embryos into full plants. In biotechnology, it is critical to be able to nurture a single cell into a full plant in order to apply some of the sophisticated techniques such as gene transfer or transformation. Plant germplasm of vegetatively propagated species may be maintained in germplasm banks or tissue culture systems. Breeding vegetatively or clonally propagated species often includes the use of tissue culture systems. After completing this chapter, the student should be able to:*

- 1 Describe the general properties of a tissue culture medium.
  - 2 Discuss how cells and tissues can be regenerated into full plants.
  - 3 Discuss micropropagation and its applications.
  - 4 Discuss the importance of cell and tissue culture in plant breeding.
  - 5 Discuss the method of *in vitro* selection for generating variability.
  - 6 Discuss the method of somatic hybridization
  - 7 Discuss the methods of asexual propagation.
  - 8 Describe the characteristics of asexual propagation that have breeding implications.
  - 9 Discuss the breeding of apomictic species.
  - 10 Compare the advantages and limitations of asexual propagation.
- 

### Concept of totipotency

Plants reproduce sexually or asexually. Pieces of plant parts (leaf, stem, roots) can be used to grow full plants in the soil. *In vitro* (growing plants under sterile conditions) plant culture was first proposed in the early 1900s. By the 1930s, cell culture had been accomplished. Each cell in a multicellular organism is **totipotent** (i.e., endowed with the full complement of genes to direct the development of the cell into a

full organism). In theory, a cell can be taken from a root, leaf, or stem, and cultured *in vitro* into a complete plant. However, some cells are unable to differentiate into all the kinds of cells in an adult organism and are said to be **multipotent** or **pluripotent**. *In vitro* culture of cells, tissues, organs, and protoplasts is used as the technique by plant breeders for propagation and also for manipulating the genetics of plants to produce new materials for breeding and for genetic studies.

## Overview of the tissue culture environment

The most critical aspect of *in vitro* culture is the provision of a sterile environment. A plant has certain natural defenses against pathogens and the abiotic environment in which it grows. Cells and tissues lack such protection once extracted from the parent plant. An indispensable piece of equipment in a tissue culture lab is an autoclave, which is used for sterilizing some materials. All glassware and other tools and plant materials are sterilized before use. Chemicals such as ethanol and Chlorox® (household bleach) are used in a tissue culture lab for sterilizing the working areas and materials. Another key piece of equipment for maintaining a sterile environment is the laminar flow hood, which blows air horizontally over the working area, towards the worker.

The growth environment for growing plants in the soil under natural conditions should provide adequate moisture, nutrients, light, temperature, and air. Plant performance can be enhanced by supplementing the growth environment (e.g., by fertilization, irrigation). In tissue and cell culture, plant materials are grown in a totally artificial environment in which the same growth factors, plus additional ones (e.g., growth regulators), are supplied. The cultural environment in tissue culture may be manipulated by the researcher to control the growth and development of the cultured material. For example, the researcher may manipulate the hormonal balance in the culture medium to favor only root or shoot development.

Tissue culture is conducted under controlled environmental conditions. The temperature is maintained at about 25°C (slightly lower at night: 23 or 24°C). The photoperiod is about 16 hours and is maintained by cool white fluorescent tube lighting with an intensity of about 3,000–5,000 lux.

### ***In vitro* culture medium**

Growing plants in the field requires a medium (e.g., soil) containing nutrients and other growth factors for success. The components of a tissue culture medium may be categorized into four groups: a physical support system, and mineral elements, organic compounds, and growth regulators.

#### *Support system*

*In vitro* culture occurs in either **liquid medium** or on **solid medium**, depending on the objectives of the project. In liquid media culture (or **suspension culture**

as it is called), tissues or cells are cultured in water containing nutrients and other growth factors. The liquid medium has to be frequently agitated for good aeration. Solid media are prepared by using gelling agents (e.g., **agar** and **agarose**). Agar is the most widely used gelling agent. It is easy to prepare and handle. It is usually prepared at a concentration of between 0.5% and 1.0%. Agar is resistant to enzymes and does not react with media components. Agar provides reduced contact of the explant with the medium and is an additional cost to the operation.

Agarose is a purer support material that is extracted from agar. It is preferred by some researchers for its lack of impurities (agarpectin and sulfate groups) found in agar, and higher gel strength (thus requires smaller amounts for preparing a solid medium). There are other gelling agents, such as gellan gums (e.g., Phytigel®), that provide clear gels (rather than translucent gels).

#### *Nutrients*

The basic components of a tissue culture medium are inorganic salts, organic salts, amino acids, sugar, and vitamins. A large variety of basic media have been developed for various uses, the more common and broad-use media including MS (after its developers, Murashige and Skoog), Gamborg, and White media. These ingredients supply both macronutrients and micronutrients. One of the most popular media is the MS (Table 11.1).

In addition to these basic components, growth regulators (auxin and cytokinins) are included in the tissue

**Table 11.1** Murashige and Skoog (MS) medium salts.

Nutrient	Source
Nitrate	$\text{NH}_4\text{NO}_3$ $\text{KNO}_3$
Sulfate	$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$
Halide	$\text{CaCl}_2 \cdot \text{H}_2\text{O}$ KI $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$
P, B, Mo	$\text{KH}_2\text{PO}_4$ $\text{H}_3\text{BO}_3$ $\text{Na}_2\text{MoO}_4$
NaFeEDTA	$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ NaEDTA

culture medium to manipulate growth and development. Auxins (e.g., naphthalene acetic acid, indole-3-butyric acid, and 2,4-dichlorophenoxyacetic acid (2,4-D)) are used to induce rooting, while cytokinins (e.g., kinetin, benzylaminopurine) are used to induce shoot formation. In actuality, it is the ratio of cytokinin to auxin that has the morphogenetic effect, a higher ratio promoting shoot formation, while a higher auxin to cytokinin ratio promotes rooting. Some plant materials have appreciable endogenous levels of hormones, needing only exogenous amounts of cytokinin for optimal shoot multiplication.

### Micropagation

Seed is the preferred propagule for use in the propagation and cultivation of most agronomic and forest species. This is so because they are easy to handle before and during the production of the plant. However, a number of major food crops and horticultural species are vegetatively propagated as a preferred method because of biological reasons (e.g., self-incompatibility) and the lack of uniformity in seed. **Micropagation** is the *in vitro* clonal propagation of plants. It is used for commercial propagation of ornamentals and other high-priced horticultural species, rather than for agronomic species. Micropagation can utilize pre-existing meristems or non-meristematic tissue. The method of micropagation commonly used may be divided into three categories: **axillary shoot production**, **adventitious shoot production**, and **somatic embryogenesis**.

Micropropagation can be summarized in five general steps:

- 1 Selection of explant.** The plant part (e.g., meristem, leaf, stem tissue, buds) to initiate tissue culture is called the **explant**. It must be in good physiological condition and be disease-free. Factors that affect the success of the explant include its location on the plant, age, or developmental phase. Explants that contain shoot primordia (e.g., meristems, node buds, shoot apices) are preferred. Also, explants from younger (juvenile) plants are more successfully used in micropropagation.
- 2 Initiation and aseptic culture establishment.** The explant is surface sterilized (e.g., with Chlorox®, alcohol) before being placed on the medium. Small amounts of plant growth regulators may be added to the medium for quick establishment of the explant.
- 3 Proliferation of axillary shoots.** Axillary shoot proliferation is induced by adding cytokinin to the shoot

culture medium. A cytokinin : auxin ratio of about 50 : 1 produces shoots with minimum callus formation. New shoots may be subcultured at an interval of about 4 weeks.

- 4 Rooting.** The addition of auxin to the medium induces root formation. Roots must be induced on the shoot to produce plantlets for transfer into the soil. It is possible to root the shoot directly in the soil.
- 5 Transfer to the natural environment.** Before transferring into the field, seedlings are gradually moved from ideal laboratory conditions to more natural conditions by reducing the relative humidity, and increasing the light intensity, a process called **hardening off**.

### Axillary shoot production

Pre-existing meristems are used to initiate **shoot culture** (or shoot tip culture). The size of the shoot tip ranges between 1 and 10 mm in length. Cytokinin is used to promote axillary shoot proliferation. Some species (e.g., sweet potato) do not respond well to this treatment. Instead, shoots consisting of single or multiple nodes per segment are used. These explants are placed horizontally on the medium and from them single unbranched shoots arise that may be induced to root to produce plantlets.

Shoot tips are easy to excise from the plant and are genetically stable. They contain preformed incipient shoot and are phenotypically homogeneous. These explants have high survival and growth rates. Axillary and terminal buds have the advantages of shoot tips, but they are more difficult to disinfect. On the other hand, meristem tips contain preformed meristems and are genetically stable and phenotypically homogeneous, but are more difficult to extract from the plant. Further, they have low survival rates.

### Adventitious shoot production

Adventitious shoots originate from adventitious meristems. Plant growth occurs at specific regions called meristems where cells are undifferentiated (no specific assigned roles or function). Non-meristematic tissue can be induced to form plant organs (e.g., embryos, flowers, leaves, shoots, roots). Differentiated plant cells (with specific functional roles) can be induced to dedifferentiate from their current structural and functional state, and then embark upon a new developmental path to produce new characteristics. This method of micropropagation also goes through the stages previously discussed. Adventitious shoot production through

organogenesis occurs by one of two pathways: **indirect** or **direct**.

### *Indirect organogenesis*

The indirect organogenetic pathway goes through a stage in which a mass of dedifferentiated cells (**callus**) forms (i.e., the explant forms a callus from which adventitious meristems are induced and from which plant regeneration is initiated). The callus consists of an aggregation of meristem-like cells that are developmentally plastic (can be manipulated to redirect the morphogenic end point). The negative side of this method is that the callus phase sometimes introduces mutations (some clonal variation, making this not always a 100% clonal procedure). The callus phase also makes it more technically challenging than shoot tip micropropagation.

### *Direct organogenesis*

Direct organogenesis bypasses a callus stage in forming plant organs. The cells in the explant act as direct precursors of a new primordium. This pathway is less common than the callus-mediated pathway.

### **Somatic embryogenesis**

As previously discussed, a zygote is formed after an egg has been fertilized by a sperm. The zygote then develops into an embryo (**zygotic embryo**). *In vitro* tissue culture techniques may be used to induce the formation of embryos from somatic tissue (**non-zygotic embryo** or **somatic embryogenesis**) using growth regulators. Somatic embryos arise from a single cell rather than budding from a cell mass as in zygotic embryos. This event is very important in biotechnology since transgenesis in plants may involve the manipulation of single somatic cells. However, without successful regeneration, plant transformation cannot be undertaken. Somatic embryogenesis has been extensively studied in Apiaceae, Fabaceae, and Solanaceae. Embryo development, zygotic or somatic, goes through certain stages: globular, scutellar, and coleoptilar stages (in monocots), and globular, heart, torpedo, and cotyledonary stages (in dicots). It is generally difficult to obtain plants from somatic embryos.

### **Other tissue culture applications**

There are other tissue culture-based applications besides micropropagations, such as the following.

### **Synthetic seed**

Somatic embryogenesis has potential commercial applications, one of which is in the **synthetic seed** technology (production of artificial seeds). A synthetic seed consists of somatic embryos enclosed in protective coating. There are two types currently being developed:

- 1 **Hydrated synthetic seed.** This kind of seed is encased in hydrated gel (e.g., calcium alginate).
- 2 **Desiccated synthetic seed.** This kind of seed is coated with water-soluble resin (e.g., polyoxethylene).

Synthetic seed technology is currently very expensive.

To develop synthetic seed, it is critical to achieve a quiescent phase, which is typically lacking in somatic embryogenesis (i.e., without quiescence there is continuous growth, germination, and eventually death). The application will depend on the crop. Lucerne (*Medicago sativa*) and orchardgrass (*Dactylis glomerata*) are among the species that have received significant attention in artificial seed development. Potential application of artificial seed is in species that are highly heterozygous and in which conventional breeding is time-consuming. Trees can be cloned more readily by this method. In some typical species that are seed propagated but have short duration of viability, artificial seed production could be economic, because of the high economic value of these crops (e.g., cacao, coconut, oil palm, coffee). Also, hybrid synthetic seed could be produced in species in which commercial hybrid production is problematic (e.g., cotton, soybean).

### **Production of virus-free plants**

Viral infections are systemic, being pervasive in the entire affected plant. Heat therapy is a procedure that is used for ridding infected plants of viral infections. After heat treatment, subsequent new growth may be free of viruses. More precisely, meristems dissected from leaf and shoot primordia are more often free of viruses even when the plant is infected. Tissue culture technology is used to nurture the excised meristematic tissue into full plants that are free from viruses.

The process starts with detection (e.g., by enzyme-linked immunosorbent assay, ELISA) of the presence of a viral infection in the plant. Once confirmed, the meristems on the shoots are aseptically removed and sterilized (dipped in 75–99% ethanol or 0.1–0.5% sodium hypochlorite or household bleach for a few seconds or minutes). The explant is submitted to tissue culture as

previously described. Sometimes, to increase the success of viral elimination, researchers may include chemicals (e.g., Ribavirin®, Virazole®) in the tissue culture medium. The plants produced must be tested to confirm virus-free status.

The virus-free plants are used to produce more materials (by micropropagation) for planting a virus-free crop. It should be pointed out that virus elimination from plants does not make them virus resistant. The producer should adopt appropriate measures to protect the crop from infection.

### Applications in wide crosses

#### Embryo rescue

Sometimes, the embryo formed after fertilization in wide crosses fails to develop any further. The breeder may intervene in the development process by dissecting the flower to remove the immature embryo. The embryo is then nurtured into a full plant by using the tissue culture technology. This technique is called **embryo rescue**. The fertilized ovary is excised within several days of fertilization to avoid an abortion (due to, for example, abnormal endosperm development). Normal embryogenesis ends at seed maturation. The development of the embryo goes through several stages with certain distinct features. The globular stage is undifferentiated, while the heart stage is differentiated and capable of independent growth. The torpedo stage and cotyledonary stage of embryo development follow these early stages. Prior to differentiation, the developing embryo is heterotrophic and dependent on the endosperm for nutrients. Excising the embryo prematurely gives it less of a chance of surviving the embryo rescue process. Just like all tissue culture work, embryo rescue is conducted aseptically and cultured on the medium appropriate for the species.

#### Somatic hybridization

The first step in somatic hybridization is to isolate intact protoplasts. Mesophyll protoplasts are preferred to protoplasts from other sources. Young tissues from healthy and well-watered and shaded plants are used. The cell wall is removed enzymatically using commercial enzyme preparations (e.g., pectinase, cellulose) to digest it. The excised leaves are sterilized prior to subjection to about a 16-hour digestion. Protoplasts may also be obtained from suspension culture. Protoplasts are uniformly negatively charged and hence repel each

other, a force that must be overcome for fusion to occur.

A **protoplast** is all the cellular component of a cell excluding the cell wall. Protoplasts from two different plants can be fused to create a hybrid. Protoplasts may be isolated by either mechanical or enzymatic procedures, as discussed in Chapter 10.

The most common methods of fusion are by chemical agents or electrical manipulation. Fusogenic agents include salt solutions (e.g., KCl, NaCl). However, the most commonly used agent is polyethylene glycol (PEG). The protoplasts are agglutinated by the application of PEG to facilitate the fusion. Addition of the compound called concanavalin A to PEG enhances the fusion. Protoplast fusion can also be accomplished by an electrical process (**electrofusion**). Protoplasts are agglutinated by the technique of dielectrophoresis, in which they are subjected to a non-uniform AC field of low intensity (500–1,000 V/cm) for a very short time. This is followed by an application of high voltage AC pulse to destabilize the cell membrane at specific sites to facilitate the fusion. Maintenance of proper osmotic potential is critical to the success of fusion. Chemicals (e.g., manitol, sorbitol) are added to the tissue culture medium for this purpose.

Fusion of protoplast does not necessarily guarantee fusion of nuclei. For a stable hybrid to form, the two nuclei must fuse within a single cell, followed by mitosis involving the two genomes. Somatic hybrids are difficult to identify. A selection system is used to verify hybridity since fusion is non-specific and therefore allows the formation of various products – multiple fusions, homokaryons (fusion of protoplasts from the same parent), heterokaryons (fusion of protoplasts from different parents), and unfused protoplasts. Some of the methods used to authenticate hybridity include genetic complementation of non-allelic mutants, use of selective media, isozyme analysis, and microisolation. The mechanical methods are most precise but tedious. A microscope is used to examine the products to identify fused products. After fusion, the tissue culture environment is modified to induce cell wall formation.

Sexual hybridization and somatic hybridization have some differences. Sexual hybridization involves fusing of two haploid nuclei and one maternal cytoplasm; somatic hybridization combines diploid nuclear genomes and two maternal cytoplasmic genomes (**symmetric hybrid**). Whereas sexual hybrids are uniform, somatic hybrids produce significant variability in the population, resulting from genetic instability, mitotic recombination, somaclonal variation, and cytoplasmic segregation.

Products of somatic hybridization may be true hybrids or parasexual hybrids with the complete genomes of two parents, partial genomes of the parents (called **asymmetric hybrids**), or **cybrids** (combination of the nuclear genome of one parent and the cytoplasm of another).

Protoplast has been isolated from numerous species (e.g., barley, carrot, cassava, cotton, pea, soybean). One of the earliest successes was the **pomato** (potato-tomato fusion product). Intergeneric protoplast fusion has been accomplished in carrot × petunia, maize × sorghum, and soybean × barley. Interspecific hybrids include *Dacus carota* × *D. capillifolius* and *Nicotiana tabacum* × *N. sylvestris*. When closely related species are used, polyploid somatic hybrids are formed. For example, a somatic hybrid of potato was produced that combined the genomes of *Solanum tuberosum* ( $2n=48$ ) and *S. brevidens* Phil. (a wild non-tuberous variety;  $2n=24$ ) to produce a somatic hybrid with 72 chromosomes. This hybrid has resistance to potato leafroll virus. Protoplast fusion has also been used for a one-step cross (rather than via a backcross) to transfer cytoplasmic male sterility (CMS) to different fertile varieties.

### Production of haploids

Haploids contain half the chromosome number of somatic cells. Anthers contain immature microspores or pollen grains with the haploid ( $n$ ) chromosome number. If successfully cultured (anther culture), the plantlets resulting will have a haploid genotype. Haploid plantlets may arise directly from embryos or indirectly via calli, as previously discussed. To have maximum genetic variability in the plantlets, breeders usually use anthers from  $F_1$  or  $F_2$  plants. Usually, the haploid plant is not the goal of anther culture. Rather, the plantlets are diploidized (to produce diploid plants) by using colchicine for chromosome doubling. This strategy yields a highly inbred line that is homozygous at all loci, after just one generation.

Methods used for breeding self-pollinated species generally aim to maintain their characteristic narrow genetic base through repeated selfing over several generations for homozygosity. The idea of using haploids to produce instant homozygotes by artificial doubling has received attention. Haploids may be produced by one of several methods:

- 1 Anther culture to induce androgenesis.
- 2 Ovary culture to induce gynogenesis.
- 3 Embryo rescue from wide crosses.

### Anther culture

Flower buds are picked from healthy plants. After surface sterilization, the anthers are excised from the buds and cultured onto an appropriate tissue culture medium. The pollen grains at this stage are in the uninucleate microspore stage. In rice the late uninucleate stage is preferred. Callus formation starts within 2–6 weeks, depending on the species, genotype, and physiological state of the parent source. A high nitrogen content of the donor plant and exposure to a low temperature at meiosis reduces albinos and enhances the chance of green plant regeneration. Pretreatment (e.g., storing buds at 4–10°C for 2–10 days) is needed in some species. This and other shock treatments promote embryogenic development. The culture medium is sometimes supplemented with plant extracts (e.g., coconut water, potato extract). To be useful for plant breeding, the haploid pollen plants are diploidized (by artificial doubling with 0.2% colchicine, or through somatic callus culture).

### Applications

**1 Development of new cultivars.** Through diploidization, haploids are used to generate instant homozygous true-breeding lines. It takes only two seasons to obtain doubled haploid plants, versus about seven crop seasons using conventional procedures to attain near homozygous lines. The genetic effect of doubling is that doubled haploid lines exhibit variation due primarily to additive gene effects and additive × additive epistasis, enabling fixation to occur in only one cycle of selection. Heritability is high because dominance is eliminated. Consequently, only a small number of doubled haploid plants in the  $F_1$  are needed, versus several thousands of  $F_2$  plants for selecting desirable genotypes.

**2 Selection of mutants.** Androgenic haploids have been used for selecting especially recessive mutants. In species such as tobacco, mutants have been selected that are resistant to the methionine analogue (methionine sulfoxide) of the toxin produced by *Pseudomonas tabaci*.

**3 Development of supermales in asparagus.** Haploids of *Asparagus officinalis* may be diploidized to produce homozygous males or females.

### Disadvantages

- 1 The full range of genetic segregation of interest to the plant breeder is observed because only a small

- fraction of androgenic grains develop into full sporophytes.
- 2 High rates of albinos occur in cereal haploids (of no agronomic value).
  - 3 Chromosomal aberrations often occur, resulting in plants with higher ploidy levels, requiring several cycles of screening to identify the haploids.
  - 4 The use of haploids for genetic studies is hampered by the high incidence of nuclear instability of haploid cells in culture.

### Ovule/ovary culture

Gynogenesis using ovules or ovaries has been achieved in species such as barley, wheat, rice, maize, tobacco, sugar beet, and onion. The method is less efficient than androgenesis because only one embryo sac exists per ovary as compared to thousands of microspores in each anther. Ovaries ranging in developmental stages from uninucleate to mature embryo sac stages are used. However, it is possible for calli and embryos to develop simultaneously from gametophytic and sporophytic cells, making it a challenge to distinguish haploids from those of somatic origin. Generally, gynogenesis is selected when androgenesis is problematic (as in sugar beet and onion).

### Haploids from wide crosses

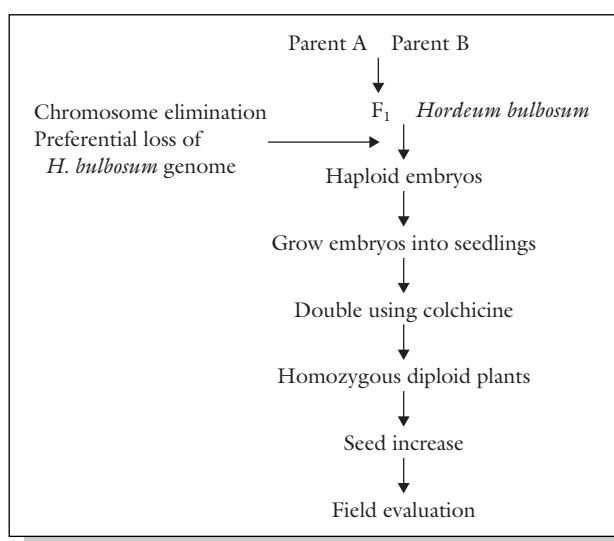
Certain specific crosses between cultivated and wild species are known to produce haploids. Well-established systems include the interspecific crosses between *Hordeum vulgare* ( $2n = 2x = 14$ , VV) and *H. bulbosum* ( $2n = 2x = 14$ , BB), commonly called the **bulbosum method**, and also in wheat × maize crosses. The bulbosum method is illustrated in Figure 11.1. The  $F_1$  zygote has  $2n = 2x = 14$  (7V + 7B). However, during tissue culture of the embryo, the *bulbosum* chromosomes are eliminated, leaving a haploid ( $2n = x = 7V$ ). This is then doubled by colchicine treatment to obtain  $2n = 2x = 14$  VV.

### Doubled haploids

Researchers exploit haploidy generally by doubling the chromosome number to create a cell with a double dose of each allele (homozygous).

#### Key features

Inbred lines are homozygous genotypes produced by repeated selfing with selection over several generations. The technique of **doubled haploids** may be used to



**Figure 11.1** Generating haploids in barley by the *bulbosum* method.

produce complete homozygous diploid lines in just 1 year (versus more than 4 years in conventional breeding) by doubling the chromosome complement of haploid cells. Such doubling may be accomplished *in vivo* naturally or through crossing of appropriate parents, or *in vitro* through the use of colchicine. The success of doubled haploids as a breeding technique depends on the availability of a reliable and efficient system for generating haploids and doubling them in the species.

#### Applications

Doubled haploids have been successfully used in breeding species in which efficient haploid generation and doubling systems have been developed. These include canola, barley, corn, and wheat. Additionally, doubled haploids are used to generate general genetic information that can be applied to facilitate the breeding process. Such information includes gene action and interaction, estimating the number of genetic variances, calculating combining abilities, and the detection of gene linkages, pleiotropy, and chromosome locations. Haploids are used in mutation studies (recessive mutants are observable instantly) and in selecting against undesirable recessive alleles.

#### Procedure

The first step in using doubled haploids in breeding is identifying the source of haploids.

**Procedure using natural sources** Haploids originate in nature through the phenomenon of **parthenogenesis** (gamete formation without fertilization). The haploids may be maternal or paternal in origin. It is estimated that haploids occur in corn at the rate of one in 1,000 diploids, 99% of which arise from parthenogenesis of maternal origin. Spontaneous doubling occurs in corn at the rate of 10% of haploids developed. The key is distinguishing between haploid and diploid plants. A marker system for this purpose was first developed by S. S. Chase based on seedling color, purple plants being encoded by the dominant gene (*P*) while normal green plants are recessive (*p*). A cross of  $F_1\ pp \times PP$  would yield 999 *Pp* (purple diploids) and one *pp* (green haploid). Another marker used is the purple aleurone color.

To use this marker system, the breeder should cross a heterozygous female to a male with marker genes. The seed from those with dominant endosperm marker of the male are saved and planted, discarding seedlings with the dominant male marker. Next, cytological evaluation of plants with the recessive female marker (by root tip squash) is conducted. The haploid plants are retained and grown in the greenhouse or field, and self-pollinated to produce diploids.

**Procedure using artificial sources** Haploid production through interspecific and intergeneric crosses is in use, one of the most well known being the barley system (discussed above). After doubling the chromosome, the diploid plants are grown to maturity. Seeds are harvested for planting in plant rows. Because diploids produced by this method are normally completely homozygous, there is no need to grow segregating generations as in conventional programs.

#### Advantages

- 1 Complete homozygosity is attainable in a shorter period.
- 2 Duration of the breeding program can be reduced by several (two or three) generations.
- 3 It is easier and more efficient to select among homogeneous progeny (versus heterogeneous progeny in conventional breeding).
- 4 The cultivar released is homogeneous.

#### Disadvantages

- 1 The procedure requires special skills and equipment in some cases.

- 2 Additional technology for doubling may increase the cost of a breeding program.
- 3 The frequency of haploids generated is not predictable.
- 4 There is a lack of opportunity to observe line performance in early generations prior to homozygosity.

#### Genetic issues

Unlike conventional methods of inbreeding, it is possible to achieve completely homozygous individuals. Using an  $F_1$  hybrid or a segregating population as the female parent in the production of maternally derived haploids increases genetic diversity in the doubled haploid line. It is advantageous if the female also has agronomically desirable traits.  $F_1$  hybrids are suitable because their female gametes will be segregating.

#### *In vitro* selection

*In vitro* selection is essentially selection of desirable genotypes under controlled environments in the Petri dish. Reports on the use of this technology have dwindled over the years, after an explosion of reports following the awareness of its potential as a source of biological variation. As a tool, it is applicable to both the sporophyte and the gametophyte.

#### Using whole plants or organs

Whole plants, seedlings, or embryos may be the units of *in vitro* selection. As previously indicated, an appropriate tissue culture system is needed for *in vitro* selection. The method has been used for screening for resistance to diseases such as *Fusarium culmorum* in wheat, and susceptibility to fungal spores in disease breeding. Tolerance (or resistance) to inorganic salts (e.g., salt tolerance in sugar beet) has been reported.

#### Using undifferentiated tissue

The capacity for regeneration by the callus makes the material attractive for use in selection. Plants can be multiplied using a callus system as well. Sometimes, spontaneous variability arises in callus or suspension culture, some of which may be heritable, while some is epigenetic, disappearing when plants are regenerated or reproduced sexually. Variability originating during tissue culture is called **somaclonal variation**.



## Industry highlights

### *Haploids and doubled haploids: their generation and application in plant breeding*

Sergey Chalyk

12 Goldfinch Court, Apt. 1007, Toronto M2R 2C4, Canada

#### ***Introduction***

Haploid plants are intensively utilized for the investigation and improvement of many agricultural crops. Haploids are unique plants and can provide researchers with genetic information that is not possible with normal diploid individuals. This box discusses methods of obtaining haploids and some advantages of using haploids in plant breeding.

#### ***What are haploids?***

The term haploid refers to a plant or an embryo that contains a gametic chromosome set. Spontaneous haploid plants have been found to occur in many crop species such as cotton, tomato, potato, soybean, tobacco, maize, barley, wheat, rice, rye, etc.

In general haploids can be divided into three types. The first type is called a maternal haploid. These haploids contain only nuclear material and cytoplasm from the maternal parent, and either result from the elimination of the chromosomes provided by the paternal parent during embryo development or by paternal sperm nuclei that are incapable of fertilization.

The second type is called an *in vitro* androgenic haploid. They can be obtained through anther or microspore culture and contain both the cytoplasm and nucleus of the developing microsporocyte.

The third type of haploid is called an *in vivo* androgenic haploid since they arise by *in vivo* embryogenesis. This class of haploid develops from an egg cell or any other cell of the embryo sac with the chromosomes of the maternal parent being lost during embryogenesis. Such haploids contain the cytoplasm of the maternal plant and the chromosomes of only the paternal parent.

#### ***Advantages of using haploids***

Haploid plants contain only one set of chromosomes. All their genes are hemizygous and each gene has only one allele. This particular feature of haploid plants allows them to be utilized in unique ways for breeding or genetic studies.

First, haploid plants can be utilized for the accelerated development of homozygous lines and pure cultivars. For this purpose it is essential to double the chromosome number after a haploid individual is generated. Following diploidization, two identical sets of chromosomes are present in the doubled haploid individual. In these instances, each gene is now represented in two exact copies or two identical alleles. By utilizing this approach, breeders can obtain homozygous lines and pure cultivars two to three times faster than by utilizing conventional methods of breeding.

Second, haploids can also be utilized for the selection of genotypes that contain favorable genes. Since haploids possess only a single dose of their respective genomes, there is no possibility for intra-allelic genetic interactions. Each gene is expressed in a single dose. This significantly facilitates the search and selection of favorable genes and the development of superior breeding genotypes. In addition, since haploids possess only a single dose of each chromosome, the possible number of gene segregation products is significantly reduced. This allows the breeder to identify a favorable combination of genes at a higher probability. This approach has special value for breeders or geneticists interested in developing an understanding of the inheritance and expression of quantitative traits. The enhanced probability of finding a favorable genotype is identified in the following example. Assume that the progeny of a hybrid plant is segregating for 10 genes. In a normal diploid, it would be necessary to grow 1,048,576 plants to obtain all possible combinations of these genes. For haploids, it would take only 1,024 haploid plants to generate all the possible combinations at least once. This example clearly shows that a desired combination of genes can be found with far fewer individuals when haploids are utilized.

Third, natural selection on haploids can be utilized as a genetic filter to identify or remove harmful mutant genes. Normally, a certain "genetic load" exists in any line, cultivar, or population as a result of spontaneous mutation over time. In diploid plants this is hidden by homologous alleles and can weaken the plants. By utilizing haploids, all genes are expressed in a single dose, including mutant or deleterious genes. Consequently, haploids containing harmful, lethal, and semilethal mutations either perish or are completely sterile. This approach leads to the natural cleansing of breeding material of genes that can reduce plant viability and productivity.

### ***Generation of haploids and doubled haploids***

#### **Chromosome elimination**

It was discovered that haploid plants of *Hordeum vulgare* could be obtained on a large scale following the hybridization of *H. vulgare* with *H. bulbosum* (Kasha & Kao 1970). When *H. vulgare* and *H. bulbosum* are crossed, a normal double fertilization event occurs. However, during seed development chromosomes of *H. bulbosum* are eliminated in both the embryo and endosperm. At approximately 10 days after fertilization, most dividing cells in the embryo are haploid. Seeds possessing a haploid embryo are removed from the spikes and placed on an embryo rescue nutrient agar culture. Approximately 50–60% of the cultured embryos develop into mature haploid plants. Colchicine, a mitotic inhibitor is applied to the haploid seedlings generating fertile spikelet/seed sectors with double the chromosome number. Haploids with these fertile sectors generate seed that have a normal diploid chromosome number.

Chromosome elimination is an alternative method for producing haploids commonly utilized in wheat. In this approach pollen from either *H. bulbosum* or maize pollen is applied to the silks of an emasculated wheat spike. The application of maize pollen has proven to be the most successful approach by providing the highest frequency of haploids.

#### ***In vitro* androgenesis**

*In vitro* androgenesis refers to the culturing of the male gamete either in the form of an anther or as isolated microspores onto an appropriate culture media. For most crop species appropriate *in vitro* androgenesis culture media have been developed. The most successful culture media will be useful for a wide range of genotypes, such as that developed for barley (Kasha et al. 2001). In their experiments more than 30 different barley genotypes have been examined and, in general, between 5,000 and 15,000 embryos are produced per plate. Regeneration ability of the embryos ranged from 36% to 97%. About 70% of plants obtained by the method of isolated microspore cultures double their chromosome number spontaneously, eliminating the necessity for the use of chromosome doubling procedures. This method can be used for mass production of haploids from any genotype of barley and, with minor modification, for genotypes of wheat.

#### ***In vivo* androgenesis**

Kermicle (1969) reported on the possibility of obtaining androgenic haploids in maize. He found that pollination of plants containing the homozygous gene *ig1* (indeterminate gametophyte 1) results in the development of 1–3% of seeds with an androgenic haploid embryo. Additional research has identified that the *ig1* gene causes an increased number of mitotic divisions during the formation of the megasporangium mother cell. The extra divisions of the egg cells lead to the loss of a normal fertilization event. Sperm usually penetrate the egg cell, but the sperm and egg cell fail to fuse. In this event, the developing embryo possesses only the chromosomes from the sperm nuclei.

Androgenic haploids are mostly used for the accelerated development of lines containing male-sterile cytoplasm. For this purpose a series of *ig1-ig1 B-3Ld-ig1* trisomics were developed that contain the following types of sterile cytoplasm: C, S, SD, Vg, ME, MY, CA, L, and Q (Kindiger & Hamann 1993; Kindiger 1994).

#### **Induction of maternal haploids in maize**

In maize, maternal haploids can occur spontaneously. Their rate is usually about one haploid per 1,000–2,000 normal diploid plants. Extensive research investigations by Chase (1969) suggested their use in inbred-line development; however, the low frequency of natural haploid generation prevents an efficient use of this approach in breeding programs. An alternative approach to obtain and investigate maternal haploids in maize was reported following the discovery of a line called "Stock 6" (Coe 1959). In this report, it was observed that pollen of "Stock 6" induced the generation of haploidy. The discovery of a maize pollen source that contains a haploid-inducing factor, simplified and facilitated the ability to obtain haploids from a wide array of different genotypes.

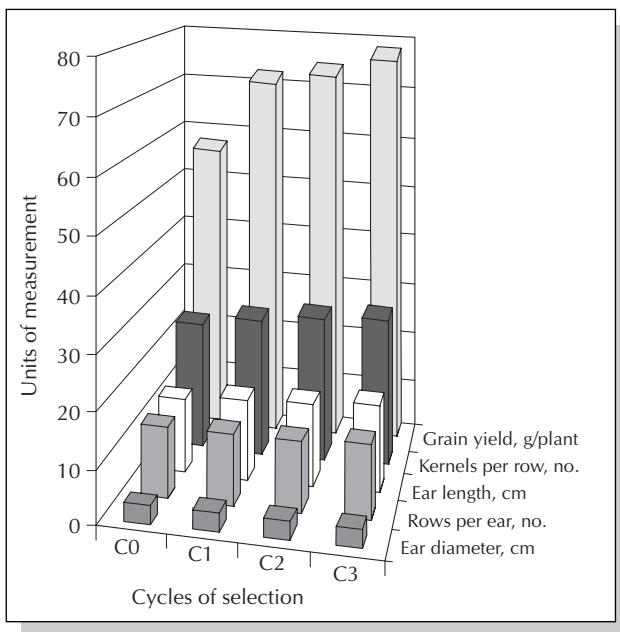
"Stock 6" has since been utilized for the development of many new haploid-inducing lines that possess dominant marker genes for the isolation of haploids. Typically, dominant marker genes conferring anthocyanin production are utilized. Such genes cause the development of a deep red or purple pigment in the seeds, seedlings, and/or plants. One such marker is called *R1-nj* (*R-navaajo*). This gene expresses anthocyanin in both the aleurone layer of the crown of the seed as well as the embryo. Following pollination of a breeding material with a haploid-inducing line that possesses the marker gene *R1-nj*, seeds that develop with pigment in the aleurone layer but with no pigment in the embryo will provide haploid plants (Figure 1). Seeds with pigment in both the aleurone and the embryo are hybrid.

#### **Generating doubled haploids**

Doubling the chromosome number in haploids is often conducted through the use of colchicine or other mitotic inhibitors, such as nitrous oxide gas and some herbicides. Following treatment of the apical meristem with a mitotic inhibitor, chromosome



**Figure 1** Ear developed following pollination by the haploid-inducer MHI that contains the marker gene *R1-nj*. The arrow shows the kernel containing the haploid embryo.



**Figure 2** Grain yield and ear traits of synthetic population SA after three cycles of haploid sib recurrent selection, averaged over 4 years of testing.

at the same cycle of selection. This step requires fertility in the haploid ears. Typically, following pollination, nearly every ear possesses seed with a normal fertilized diploid embryo. This unique tendency of maize haploid ears allows the breeder to move forward in the breeding process without doubling the chromosome number of the haploid individuals. This makes the utilization of haploids simple and inexpensive.

numbers are doubled in small sectors of the haploid plant, including some sectors of the spike, ear, or tassel. Normally the doubled sectors produce seeds. These seeds are doubled haploids – pure-line cultivars.

In our research, a method of colchicine treatment was evaluated on maize haploids. In this method, haploid seedlings were treated with colchicine at the stage when the length of the coleoptile was 1 cm or longer. Initially, the seeds were germinated at 26°C for 4–5 days. Then a small tip of the coleoptile was removed and the seedlings were placed into a 0.06% colchicine solution with 0.5% dimethyl sulfoxide (DMSO) for 12 hours. Thereafter the seedlings were washed and planted in the field. Following the treatment, approximately half of the haploids produced fertile pollen (49.4%). Most were selfed and 27.3% of the haploid individuals produced seeds. These seeds were doubled haploids since they contained normal diploid embryos. The seeds were quite viable and generated normal viable seedlings following planting in the field or greenhouse.

#### *Application of haploids and doubled haploids in plant breeding*

Haploids have two primary uses in plant breeding. The first is the accelerated production of homozygous lines and pure cultivars. As we have discussed earlier, the doubling of haploids is the most rapid route toward the development of pure cultivars in self-pollinated seed crops. It can be achieved in a single generation and can be performed at any generation in a breeding program. For cross-pollinated crops, haploids are used primarily for the production of homozygous lines, which are in themselves utilized in the production of hybrid seed. At the present time, more than 200 varieties have been developed by utilizing a doubled haploid approach (Thomas et al. 2003).

The second main use is that haploids provide a possibility of screening breeding material for the presence of advantageous genes. In both haploids and doubled haploids, all alleles are expressed. This facilitates the selection of genotypes that are important for breeders. Selected haploids can be used for the improvement of any breeding material, including increasing the frequency of favorable genes in populations (e.g., in recurrent selection). As one example of haploid utility in a breeding program, the method of "haploid sib recurrent selection" can be presented (Eder & Chalyk 2002). In this approach, the selection of favorable genotypes is performed on haploid plants. Every cycle of selection consists of two steps. The first step is to obtain haploids from a synthetic population. In our experiment, haploids were obtained in a space-isolated nursery following pollination with a haploid-inducing line. The second step is growing haploid plants, the selection of haploid plants, and pollination with a mixture of pollen collected from diploid plants of the same synthetic population

**Table 1** Estimated gain per cycle for grain yield, and ear and plant traits of SP and SA synthetic populations averaged for 4 years of testing.

Traits	Gain per cycle (%)									
	SP					SA				
	1998	1999	2000	2001	Average	1998	1999	2000	2001	Average
Grain yield	11.0	16.4	1.7	17.8	13.1	16.7	21.0	10.3	8.4	12.0
Ear length	9.2	6.6	4.4	7.5	7.8	4.5	9.2	4.4	5.8	6.2
Seeds per row	2.6	6.7	1.1	9.6	6.2	7.5	11.4	4.8	4.8	6.3
Rows number	7.3	6.2	3.0	5.2	2.9	1.4	4.7	3.0	4.4	6.1
Ear diameter	5.7	4.8	0.6	4.1	4.1	3.5	3.7	2.2	2.6	2.0
Plant height	9.3	9.6	10.5	7.7	10.0	12.1	3.7	7.3	4.8	4.8
Ear height	10.0	7.6	10.0	11.8	10.9	16.6	8.4	13.8	11.1	6.7
Leaf length	8.5	6.2	3.8	9.0	7.8	3.3	2.8	1.5	3.2	2.6

In our experiments, selection was carried out for ear size. Each season, 1,000–2,000 haploid plants from an improved synthetic population were planted in the field. Of these, 200–300 haploids were pollinated by diploid representatives of the same synthetic population. At harvest time about 20–30 haploids with the largest ears were selected for the next cycle of selection. Three cycles of haploid sib recurrent selection were completed for two synthetic populations, SP and SA. Initial and improved synthetics were evaluated in the field for 4 years. The performance results of synthetic SA are presented in the Figure 2. The data indicate that selection for ear size, utilizing haploids, can result in a significant increase of grain yield.

An important method that effectively estimates the efficiency of a recurrent selection program is the determination of gain per cycle. This parameter is used for comparing the efficiency of different recurrent selection schemes. Normally gain per cycle for grain yield in a recurrent selection scheme in maize approximates 2–5% and seldom exceeds 7% (Gardner 1977; Weyhrich et al. 1998). The results obtained by haploid sib recurrent selection are presented in Table 1. For synthetic population SA, the gain per cycle was 12.0%, and for synthetic SP it was 13.1%. Gain per cycle was distinctly higher than that which is observed when utilizing conventional recurrent selection methods. The conclusion drawn from the experiment is that the utilization of haploid plants for the selection of favorable genotypes greatly increases the efficiency of recurrent selection.

Overall, numerous experiments have indicated that haploid generation can be successfully applied to several species on a large scale. The methods and citations given above provide only a few examples of this useful and efficient method. The utilization of haploids and doubled haploids can simplify the identification of genotypes that can provide a significant improvement in a variety of agronomic traits. In addition, haploids and doubled haploids can accelerate the generation of homozygous lines and pure cultivars. Therefore, haploid-inducement technologies have a bright future in plant breeding.

## References

- Chase, S.S. 1969. Monoploids and monoploid-derivatives of maize (*Zea mays* L.). *Bot. Rev.* 35:117–167.
- Coe, E.H. 1959. A line of maize with high haploid frequency. *Am. Nat.* 93:381–382.
- Eder, J., and S. Chalyk. 2002. *In vivo* haploid induction in maize. *Theor. Appl. Genet.* 104:703–708.
- Gardner, C.O. 1977. Population improvement in maize. In: *Maize breeding and genetics* (J. Janick, ed.), pp. 207–228. Wiley, New York.
- Kasha, K.J., and K.N. Kao. 1970. High frequency haploid production in barley (*Hordeum vulgare* L.). *Nature* 225:874–876.
- Kasha, K.J., E. Simion, R. Oro, Q.A. Yao, T.C. Hu, and A.R. Carlson. 2001. An improved *in vitro* technique for isolated microspore culture of barley. *Euphytica* 120:379–385.
- Kermicle, J.L. 1969. Androgenesis conditioned by a mutation in maize. *Science* 116:1422–1424.
- Kindiger, B. 1993. Registration of 10 genetic stocks of maize for the transfer of cytoplasmic male sterility. *Crop Sci.* 34:321–322.
- Kindiger, B., and S. Hamann. 1994. Generation of haploids in maize: A modification of the *indeterminate gametophyte* (*ig*) system. *Crop Sci.* 33:342–344.
- Thomas, W.T.B., B.P. Forster, and B. Gertsson. 2003. Doubled haploids in breeding. In: *Doubled haploid production in crop plants* (Maluszynski, M., K. Kasha, B.P. Forster, and I. Szarejko, eds), pp. 337–350. Kluwer Academic, Dordrecht.
- Weyhrich, R.A., K.R. Lamkey, and A.R. Hallauer. 1998. Responses to seven methods of recurrent selection in the BS11 maize population. *Crop Sci.* 38:308–321.

### Somaclonal variation

A variety of mechanisms have been implicated in this phenomenon. Chromosomal changes, both polyploidy and aneuploidy, have been observed in potato, wheat, and ryegrass. Some research suggests mitotic crossovers to be involved whereas cytoplasmic factors (mitochondrial genes) have been implicated by others. Further, point mutation, transposable elements, DNA methylation, and gene amplification are other postulated mechanisms for this phenomenon.

As a breeding tool, breeders may deliberately plan and seek these variants by observing certain factors in tissue culture. Certain genotypes are more susceptible to genetic changes in tissue culture, polyploid genotypes generally being more so than diploid. Also, holding the callus in an undifferentiated state for prolonged periods of time enhances the chances of somaclonal variation occurring. Not unexpectedly, the tissue culture environment (medium components) can induce heritable changes in the callus. The inclusion of auxin 2,4-D enhances the chances of somaclonal variation.

Somaclonal variation from calli with disease resistance (e.g., *Helminthosporium sacchari* in sugarcane and *Fusarium* in *Apium graveolens*) has been found. Somaclones with resistance to various abiotic stresses have also been reported.

### Directed selection

Rather than allowing the variation to arise spontaneously, sometimes plant breeders apply selection pressure during the *in vitro* cultural process to influence the variability that might arise.

#### Selection for disease resistance

Various toxin metabolites have been included in tissue culture for use as the basis for selection, assuming that such metabolites have a role in pathogenesis. The culture filtrates from various fungi (*Fusarium*, *Helminthosporium maydis*) have been used to exert selection pressure for cells that are resistant to the pathogen. The main constraint to the use of directed selection in plant breeding for disease resistance is the inability of the *in vitro* system to be used to select for hypersensitivity, a major strategy in disease resistance.

#### Selection for herbicide tolerance

Mutants with about 10–100 times the level of resistance to herbicides (e.g., imidazolinone in sugar beet) have

been successfully isolated, characterized, and incorporated into commercial cultivar development. Many of the recorded successes with *in vitro* selection have been with herbicide tolerance.

#### Selection for tolerance to abiotic stresses

Selection for tolerance to salinity, metals (Zn, Al) and temperature (cold tolerance) has been attempted with varying degrees of success.

#### Single-cell selection system

Some researchers use single-cell tissue culture systems (suspension culture, protoplast culture) for *in vitro* selection. The advantages of this approach include a lack of chimerism, higher chances of isolation of true mutants, and an ability to more effectively apply microbial procedures to the large number of individual cells than can be screened in a small space. Selection for biotic stress resistance, herbicide tolerance (the author did this for chlorsulfuron), and aluminum tolerance, are among successful applications of direct selection using a single-cell selection system.

### Germplasm preservation

Germplasm preservation in tissue culture was discussed in Chapter 6. This method of germplasm storage is often used for vegetatively propagated species.

## Breeding vegetatively (clonally) propagated species

Because seed is produced via the process of meiosis, cultivars propagated by seed always have some heterozygosity, the degree of it varying with the method of breeding. Asexual or clonal propagation entails the use of parts of the plant other than the seed as the propagule. Clones have identical genotypes because they reproduce via mitosis not meiosis.

### Clones, inbred lines, and pure lines

As previously discussed, plants may be naturally sexually or asexually propagated. Further, sexually propagated species may be self-fertilized or cross-fertilized. These natural modes of reproduction have implications in the genetic structure and constitution of plants and breeding implications. Plant breeders are able to manipulate

the natural reproductive systems of species to develop plants that have an atypical genetic constitution. The terms **pure line**, **inbred line**, and **clone** are applied to materials developed by plant breeders to connote sameness of genetic constitution in some fashion. However, there are some significant distinctions among them.

- 1 Pure lines.** These genotypes are developed as cultivars of self-pollinated crops for direct use by farmers. As products of repeated selfing of single plants, pure lines are homogeneous and homozygous and can be naturally maintained by selfing.
- 2 Inbred lines.** These are genotypes that are developed to be used as parents in the production of hybrid cultivars and synthetic cultivars in the breeding of cross-pollinated species. They are not meant for direct release for use by farmers. They are homogenous and homozygous, just like pure lines. However, unlike pure lines, they need to be artificially maintained because they are produced by forced selfing (not natural selfing) of naturally cross-pollinated species.
- 3 Clones.** Clones are identical copies of a genotype. Together, they are phenotypically homogeneous. However, individually, they are highly heterozygous. Asexually or clonally propagated plants produce genetically identical progeny.

### Categories of asexually propagated species

In Chapter 4, asexually propagated species were grouped into two according to economic use as those cultivated for vegetative products and those cultivated for their fruits. For breeding purposes, vegetatively propagated crops may be grouped into four, based on flowering behavior.

- 1 Species with normal flowering and seed set.** Species in this category produce normal flowers and are capable of sexual reproduction (to varying extents) without artificial intervention (e.g., sugarcane). However, in crop production, the preference is to propagate them sexually. Such species enjoy the advantages of both sexual and asexual reproduction. Hybridization is used to generate recombinants (through meiosis) and introduce new genes into the adapted cultivar, while vegetative propagation is used to maintain, indefinitely, the advantages of the heterozygosity arising from hybridization.
- 2 Species with normal flowers and poor seed set.** Some plant species produce normal-looking flowers that have poor seed set, or set seed only under certain conditions but not under others. These restrictions

on reproduction make it unattractive to use seed as a means of propagation. However, the opportunity for hybridization may be exploited to transfer genes into adapted cultivars.

- 3 Species producing seed by apomixis.** The phenomenon of apomixis results in the production of seed without fertilization, as was first discussed in Chapter 4. Over 100 species of perennial grasses have this reproductive mechanism.
- 4 Non-flowering species.** These species may be described as “obligate asexually propagated species” because they have no other choice. Without flowers (or with sterile flowers) those species can only be improved by asexual means. Genetic diversity is not obtained via recombination but by other sources (e.g., mutation).

### Genetic issues in asexual breeding

#### Genetic makeup

All the progeny from an individual propagated asexually are genetically identical (clones) and uniform. Clones are products of mitosis. Any variation occurring among them is environmental in origin.

#### Heterozygosity and heterosis

Many species that are asexually propagated are highly heterozygous; they are highly heterotic. Consequently, they are susceptible to inbreeding depression. For those species that can be hybridized without problems, an advantage of asexual propagation is that heterosis, where it occurs, is fixed for as long as the cultivar is propagated asexually.

#### Ploidy

Many known species that are asexually propagated are interspecific hybrids or have high ploidy.

#### Chimerism

Clones are stable over many generations of multiplication. The only source of natural variation, albeit rare, is somatic mutation in the bud. Plant breeders may generate variability by the method of mutagenesis. Whether natural or artificial, somatic mutations are characterized by tissue mosaicism, a phenomenon called **chimerism**.

A chimera or chimeric change occurs when an individual consists of two or more genetically different types of cells. Though heritable changes, these mosaics can

only be maintained by vegetative propagation (not transferable to progenies by sexual means).

There are four basic types of chimeras:

- 1 Sectorial.** This chimera is observed in a growing shoot as two different tissues located side by side. The effect of this modification is that the stem develops with two distinct tissues on each half.
- 2 Periclinal.** This type of chimera consists of two thin layers of different genetic makeup, one over the other.
- 3 Mericlinal.** When an outer layer of different genetic tissue does not completely extend over the layer below, the chimera is mericlinal.
- 4 Graft chimeras.** Unlike the first three chimeras that have a genetic origin, a graft chimera is a non-heritable mixture of tissues that may occur after grafting is made.

Additional information on chimeras is found in Chapter 12. Whereas they are undesirable in crop plants, chimeras may be successfully exploited in horticulture.

### Breeding approaches used in asexual crops

Several breeding approaches are used in the breeding of asexually propagated species.

#### Planned introduction

Just like seed, vegetative material (whole plants or parts) may be introduced into a new production area for evaluation and adaptation to the new area. Seedlings or cuttings may be introduced. However, the technology of tissue culture allows a large variety of small samples to be introduced in sterile condition. These disease-free samples are easier to process through plant quarantine.

#### Clonal selection

Clonal selection has two primary goals – to maintain disease-free and genetically pure clones, and the development of new cultivars.

##### Purifying an infected cultivar

Clonal cultivars may become infected by pathogens, some of which may be systemic (e.g., viruses). Two general approaches may be used to purify a cultivar to restore it to its disease-free original genetic purity.

**1 Screening for disease-free material.** Plant materials may be visually inspected for the presence of pathogens. However, because some pathogens may be latent, a variety of serological and histological techniques are used to detect the presence of specific pathogens. Called **indexing**, these techniques can detect latent viruses (**viral indexing**) as well as other pathogens. A negative test may not always be proof of the absence of pathogens. It could be that the particular assay is not effective. The clean clonal material is then used as starting material for multiplication for propagation.

**2 Elimination of pathogens.** A positive test from indexing indicates the presence of a pathogen. Should this be the only source of planting material, the breeder has no choice but to eliminate the pathogen from the plant tissue by one of several methods:

- (a) **Tissue culture.** Even when the pathogen is systemic, it is known that tissue from the terminal growing points is often pathogen-free. Tissue from these points may be aseptically removed and cultured under tissue culture conditions to produce disease-free plantlets. Through micropropagation, numerous disease-free plants can be obtained.
- (b) **Heat treatment.** This may be of short or long duration. Short-duration heat treatment is administered to the plant material for about 30 minutes to 4 hours at 43–57°C. This could be in the form of hot air treatment or by soaking the material in hot water. This works well for fungal, bacterial, and nematode infections. For viruses, a longer treatment of several weeks (2–4 weeks) is used. Potted plants are held at 37°C in a controlled environment for the duration of the treatment. Cuttings from the treated plants may be used as scions in grafts, or rooted into a seedling.
- (c) **Chemical treatment.** This surface sterilization treatment is suitable for elimination of pathogens that are external to the plant material (e.g., in tubers).
- (d) **Use of apomictic seed.** Viral infections are generally not transmitted through seed in cultivars that are capable of apomixis (e.g., citrus).

##### Clonal selection for cultivar development

This procedure is effective if variability exists in the natural clonal population.

**Year 1** Assemble the clonal population. Plant and expose to diseases of interest. Select resistant

- clones with other superior traits and harvest individually.
- Year 2** Grow progenies of selected clones and evaluate as in year 1. Select superior clones.
- Year 3** Conduct preliminary yield trials. Select superior clones.
- Years 4–6** Conduct advanced yield trials at multilocations for cultivar release.

### Hybridization with clonal selection

This procedure is applicable to species that are capable of producing seed in appreciable quantities. Because heterosis can be fixed in clonal populations, the breeder may conduct a combining ability analysis to determine the best combiners to be used in hybridization.

- Year 1** Cross selected parents. Harvest F<sub>1</sub> seed.
- Year 2** Plant and evaluate the F<sub>1</sub>s. Select vigorous and healthy plants.
- Year 3** Space plant clonal progeny rows of selected plants. Select about 100–200 superior plant progenies.
- Year 4** Conduct preliminary yield trials.
- Years 5–7** Conduct advanced yield trials for cultivar release.

Other techniques that are applicable include backcrossing to transfer specific traits and wide crossing. The challenges with backcrossing are several. As previously indicated, clonal species are very heterozygous and prone to inbreeding depression. Backcrossing to one parent (the recurrent parent) provides an opportunity for homozygosity and consequently inbreeding depression. To prevent this, breeders may cross the backcross to another clone instead of the recurrent parent, followed by selection to identify superior plants. The process is repeated as needed.

### Mutation breeding

The subject is discussed in detail in Chapter 12. Inducing variability via mutagenesis is challenging for two key reasons. Being rare events, a large population of M<sub>1</sub>V<sub>2</sub> is needed to have a good chance of observing desired mutants. Obtaining a large number of vegetative propagules is difficult. Also, mutations occur in individual cells. Without the benefit of meiosis, the mutated clonal material develops chimeras. Using adventitious buds as starting material reduces the chance of chimeras. A mutation in the epidermal cell (usually there is one) would result in an adventitious shoot that originated

from a single mutant cell. This technique is not universally applicable.

### Breeding implications, advantages, and limitations of clonal propagation

The breeding implications of vegetative propagation were discussed in Chapter 4. There are several advantages and limitations of breeding vegetatively propagated species.

#### Advantages

- 1 Sterility is not a factor in clonal propagation because seed is not involved.
- 2 Because clonal plants are homogeneous, the commercial product is uniform.
- 3 Micropropagation can be used to rapidly multiply planting material.
- 4 Heterozygosity and heterosis are fixed in clonal populations.

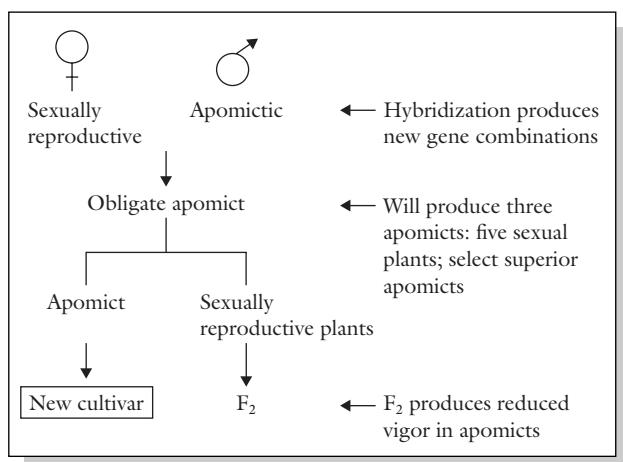
#### Disadvantages

- 1 Clonal propagules are often bulky to handle (e.g., stems, bulbs).
- 2 Clones are susceptible to devastation by an epidemic. Because all plants in the clonal population are identical, they are susceptible to the same strain of pathogen.
- 3 Clonal propagules are difficult to store for a long time because they are generally fresh and succulent materials.

### Breeding apomictic cultivars

**Apomixis**, the phenomenon of seed development without fertilization was discussed in Chapter 4, including its occurrence in nature, mechanisms, and benefits to the farmer and breeder.

Genetic control of apomixis has been demonstrated, implicating a few genes. Efforts using modern molecular genetic tools continue to be made to isolate these apomictic genes. Apomixis can be a two-edge sword – it can hinder breeding progress or it can be an effective breeding tool. To improve an apomict, there should be suitable materials, i.e., sexual or partially sexual plants for use as female plants for crossing. Generally, an obligate apomict cannot be used as a female parent in a hybridization program. However, most apomictic



**Figure 11.2** Steps in a method for breeding apomictic species.

plants produce adequate amounts of viable pollen to be usable as males in crossing. This leaves the identification of a suitable female the first critical step in an apomictic improvement program.

Once suitable parents have been selected, crossing can be conducted as for regular plants. A sexual female plant may be crossed with an apomictic male to produce  $F_1$  hybrids, some of which will be obligate and true breeding apomicts, while others will be asexual hybrids that will segregate in the  $F_2$  (Figure 11.2). Because apomicts are generally heterozygous, selfing of sexual clones will yield variability from which the breeder can practice selection. The use of markers and precautions in emasculation will help in distinguishing hybrids from other heterozygous plants. It should be pointed out that the breeder should aim to identify superior genotypes in the  $F_1$  where heterosis is at a maximum, rather than in later sexual generations.

As previously indicated, facultative apomicts are more challenging to improve, partly because the breeder cannot control variation in their progenies (they produce both sexual and apomictic plants). Furthermore, the stability of the reproductive process is influenced by environmental factors (especially photoperiod). Photoperiod has been observed to significantly affect the relative frequency of sexual versus apomictic embryo sacs in ovules of certain species.

## References and suggested reading

- Chaudhury, A.M., L. Ming, C. Miller, S. Craig, E.S. Dennis, and W.J. Peacock. 1997. Fertilization-independent seed development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 94:4223–4228.
- Hanna, W.W., and E.C. Bashaw. 1987. Apomixis: Its identification and use in plant breeding. Crop Sci. 27:1136–1139.
- Jain, M., K. Chengalrayan, M. Gallo-Meacher, and P. Misley. 2005. Embryogenic callus induction and regeneration in a pentaploid hybrid bermudagrass cv Tifton 85. Crop Sci. 45:1069–1072.
- Kamo, K. 1995. A cultivar comparison of plant regeneration from suspension cells, callus, and cormel slices of *Gladiolus*. In vitro Cell. Develop. Biol. 31:113–115.
- Kindinger, B., D. Bai, and V. Sokolov. 1996. Assignment of a gene(s) conferring apomixis in *Tripsacum* to a chromosome arm: Cytological and molecular evidence. Genome 39:1133–1141.
- Koltunow, A.M., R.A. Bicknell, and A.-M. Chaudhury. 1995. Apomixis: Molecular strategies for the generation of genetically identical seeds without fertilization. Plant Physiol. 108:1345–1352.
- Mohammadi, S.A., B.M. Prasanna, and N.N. Singh. 2003. Sequential path model for determining interrelationship among grain yield and related characters in maize. Crop Sci. 43:1690–1697.
- Murashige, T., and T. Skoog. 1962. A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol. Plantar. 15:473–497.
- Scityavathi, V.V., P.P. Janhar, E.M. Elias, and M.B. Rao. 2004. Effects of growth regulators on *in vitro* plant regeneration in durum wheat. Crop Sci. 44:1839–1846.
- Stefaniak, B. 1994. Somatic embryogenesis and plant regeneration of gladiolus. Plant Cell Rep. 13:386–389.
- Tae-Seok, K.O., R.L. Nelson, and S.S. Korban. 2004. Screening multiple soybean cultivars (MG 00 to MG VIII) for embryogenesis following *Agrobacterium*-mediated transformation of immature cotyledons. Crop Sci. 44:1825–1831.
- Trigiano, R.N., and D.J. Gray (eds). 1996. Plant tissue culture concepts and laboratory exercises. CRC Press, New York.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 MS medium was developed by Morris and Stevenson.
- 2 Agar is used as a gelling agent in tissue culture.
- 3 A protoplast is all the cellular component of a cell plus the cell wall.
- 4 An auxin : cytokinin ratio in favor of auxin promotes rooting.
- 5 Propagation by cuttings is a form of clonal propagation.
- 6 Diplospory is the most common mechanism of apomixis in higher plants.
- 7 Facultative apomicts reproduce exclusively by apomixis.
- 8 Pathogenesis is equivalent to haploidy.

### Part B

Please answer the following questions:

- 1 The part of the plant used to start tissue culture is called the .....
- 2 The MS tissue culture medium is named after ..... and .....
- 3 A mass of undifferentiated cells such as meristematic cells is called .....
- 4 What is clonal propagation?
- 5 What is apomixis?
- 6 Distinguish between apospory and diplospory as mechanisms of apomixis.
- 7 Species that reproduce exclusively (or nearly so) by apomixis are described as .....
- 8 Give a specific advantage of clonal propagation.

### Part C

Please write a brief essay on each of the following topics:

- 1 Why is it possible (at least theoretically) to raise a full plant, or for that matter any organism, from just one of its own cells?
- 2 All cells are not totipotent. Explain.
- 3 What is the importance of a callus phase in plant tissue culture research?
- 4 Discuss the rationale for the composition of a tissue culture medium.
- 5 Describe the *in vitro* production of hybrids.
- 6 Discuss practical applications of tissue culture in plant breeding.
- 7 Discuss the key breeding implications of clonal propagation.
- 8 Discuss the benefits of apomictic cultivars in crop production.
- 9 Apomixis can be a two-edged sword in plant breeding. Explain.
- 10 Discuss the occurrence of apomixis in nature.



# 12

# Mutagenesis in plant breeding

---

## Purpose and expected outcomes

*It was previously pointed out that mutation is the ultimate source of variation. Without adequate variation, plant breeding is impossible. To start a breeding program, the breeder must find the appropriate genotype (containing the desired genes) from existing variation, or create the variation if it is not found in nature. Mutagenesis is the process by which new alleles are created. The purpose of this chapter is to discuss mutagenesis as both a technique and a breeding method. The newly created mutants may be used as parents in future breeding programs, in which case mutagenesis is a breeding technique as a source of variation. However, an induced mutant can be systematically processed through conventional breeding steps to be released as a cultivar, hence making it a breeding method (mutation breeding). Mutations arise spontaneously in nature and are pivotal in natural evolution. After completing this chapter, the student should be able to:*

- 1 Define mutation and mutagenesis.
  - 2 Discuss mutagenic agents.
  - 3 Discuss the steps in a mutation breeding program.
  - 4 Discuss the limitations of mutation breeding.
- 

## Brief historical perspective

The discovery of the mutagenic effects of X-rays on the fruit fly (*Drosophila*) by H. Muller in the 1920s paved the way for researchers to experiment with its effects on various organisms. In 1928, H. Stubbe demonstrated the use of **mutagenesis** in producing mutants in tomato, soybean, and other crops. The first commercial mutant was produced in tobacco in 1934. Reports by B. Sigurbjornsson and A. Micke mentioned 77 cultivars that were developed via mutagenesis prior to 1995. In 1995, the number was 484. This number has since been significantly exceeded. They include food crops (e.g., corn, wheat, pea), ornamentals (e.g., chrysanthemum, poinsettia, dahlia), and fruit trees (e.g., citrus, apple, peach). Traits modified include agronomic ones such as plant maturity, winter hardiness, lodging resistance, and

product quality (e.g., protein and lysine content), and numerous ornamental mutants.

The perceived role of mutation in plant breeding was initially treated with skepticism by some, as demonstrated by L. J. Stadler, who is said to have advised his students against using mutation breeding for commercial crop improvement as a reaction to the overoptimism by protagonists who saw it as a revolutionary plant breeding method. Currently, induced mutations are used more often in a supplementary role as a source of new alleles. However, it is still important in breeding vegetatively propagated species, including field crops, ornamentals, and fruit and forest species. It is especially useful in ornamental plant breeding where novelty is often advantageous and can become commercially significant. Furthermore, with the advent of genetic engineering and its radical tools, which allow targeted

genetic alteration (versus the random genetic alteration produced by conventional mutagenesis), it appears that breeders are gravitating towards this truly revolutionary technology for creating new variability. However, no approach should be written off as every now and then some breeders find good reason to use a technique or technology that has been marginalized by advances in science and technology.

In conventional breeding of sexual plants, genetic variability is derived from recombination. Parents must not be identical, or else there would be no segregation in the  $F_2$  generation. Even when parents are dissimilar, they often have similar “housekeeping genes” that are common to both parents. Whereas segregation will not occur for these common genes, mutagenesis can create variability by altering them.

### Types of mutation

In terms of origin, mutations may be **spontaneous** (natural) or **induced** (artificial, with the aid of agents). Spontaneous mutations arise at the very low rate of about  $10^{-5}$  or  $10^{-6}$  per generation for most loci in most organisms. This translates to one in 100,000 or one in 1,000,000 gametes that may carry a newly mutated allele at any locus. They are caused by mistakes in molecular processes associated with the replication of DNA, recombination, and nuclear division. However, because mutagenic agents are common in the general environment, induced mutations, as a result of these agents (natural radiations), are hard to distinguish from spontaneously induced mutations due to cellular processes.

Mutations may also be classified according to the type of structural change produced:

- 1 Genomic mutation:** changes in chromosome number (gain or loss in complete sets of chromosomes or parts of a set).
- 2 Structural mutation:** changes in chromosome structure (e.g., duplications of segments, translocation of segments).
- 3 Gene mutation:** changes in the nucleotide constitution of DNA (by deletion or substitution).

Mutation may occur in the nuclear DNA or chromosomes, or in extranuclear (cytoplasmic) genetic systems. A good example of the practical application of mutations in plant breeding is the cytoplasmic-genetic male-sterility gene, which occurs in chloroplasts.

In terms of gene action, a mutation may be recessive or dominant:

**1 Recessive mutation:** change of a dominant allele to a recessive allele ( $A \rightarrow a$ ).

**2 Dominant mutation:** change of a recessive allele to a dominant allele ( $a \rightarrow A$ ).

Mutations that convert the **wild type** (the common phenotype) to the mutant form (the rare phenotype) are called **forward mutations**, while those that change a mutant phenotype to a wild phenotype are called **reverse mutations**. Forward mutations are more common than reverse mutations. Recessive mutations are the most common types of mutations. However, recessive alleles in a diploid are expressed only when they are in the homozygous state. Consequently, an organism may accumulate a genetic load without any consequence because of heterozygous advantage. As previously discussed, outcrossing species are susceptible to inbreeding depression (loss of vigor), because of the opportunities for expression of deleterious recessive alleles.

### Induced mutations versus spontaneous mutations

Spontaneous mutations produce novel alleles for the evolutionary process. Natural mutations have the benefit of being subjected to the evolutionary process whereby viable mutants become recombined with existing forms and become adapted under the guidance of natural selection. Mutagenesis can be used to create new alleles that can be incorporated into existing cultivars through recombination following hybridization and under the guidance of artificial selection. Modern crop production systems are capable of providing supplemental care to enable a mutant that would not have survived natural selection to become productive. As previously discussed, a significant number of commercial cultivars originated from mutation breeding techniques. Furthermore, the rate of spontaneous mutation is low ( $10^{-6}$  per locus). Artificial mutagenesis aims to increase mutation rates for desired traits.

### Cell type: gametic versus somatic mutations

Mutations may originate in the gametic or somatic cells. **Gametic mutations** are heritable from one generation to the next and are expressed in the entire plant. However, mutations in a somatic tissue will affect only that portion of the plant, resulting in a condition called **chimera**. In species that produce tillers, it is possible to have a tiller originate from a chimeric tissue, while others are normal. A chimera consists of two genetically distinct tissues and may produce two distinct flowers on

the same plant. However, the dual color is impossible to reproduce by either sexual or asexual propagation. Commercial use of chimera is not attractive because the vegetative propagules must, of necessity, comprise both kinds of tissues in order to reproduce the maternal features.

### Gene action: dominant versus recessive mutations

As previously indicated, mutations may cause a dominant allele to be changed into a recessive allele (**recessive mutation**), or a recessive allele to be changed into a dominant allele (**dominant mutation**). Open-pollinated species may accumulate a large amount of recessive mutant alleles without any adverse effects. However, upon selfing, the recessive alleles become homozygous and are expressed, leading to the phenomenon of inbreeding depression. Using recessive genes in breeding takes a longer time because it requires an additional step of selfing in order to identify and select the desired recombinants. On the other hand, dominant mutations manifest in the current generation, needing no additional regeneration to be observable.

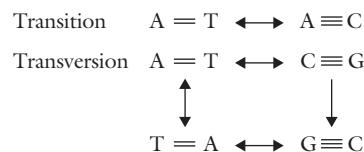
### Structural changes at the chromosomal level

Three types of structural changes in the chromosome can occur as a result of mutation.

#### Gene mutation

**Kind of mutation** Gene mutations entail a change in the nucleotide constitution of the DNA sequence, adding or deleting nucleotides.

**1 Transitions and transversions.** As previously described, the DNA consists of four bases – A, T, C, and G – that pair in a specific pattern, G–C and A–T.

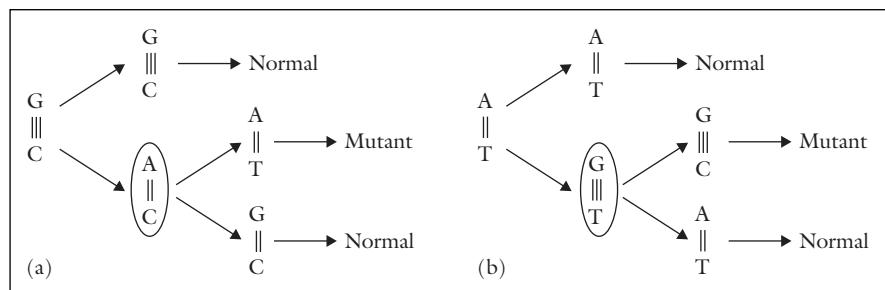


**Figure 12.1** Mutations may occur by transition or transversion.

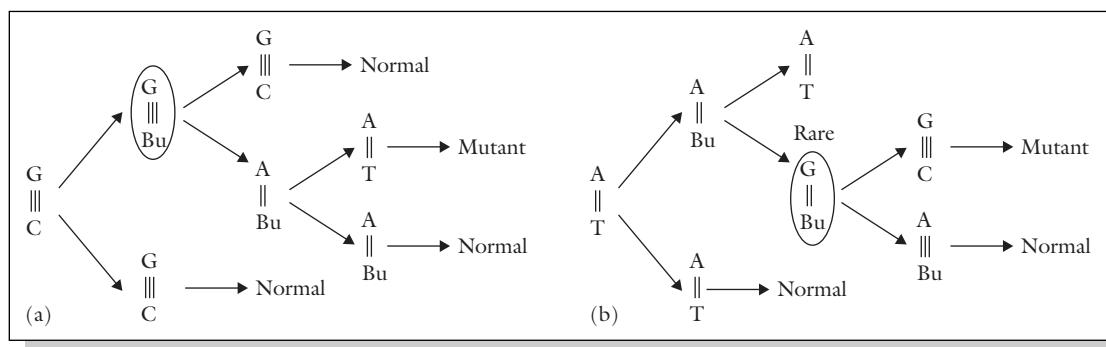
The structure of the DNA may be modified in two ways – transition or transversion of the bases (Figure 12.1). Mutation by transition entails the conversion of one purine base to another purine (or a pyrimidine to another pyrimidine). During replication, the second purine (a different purine), which has altered base-pairing properties, guides an incorrect base into position. Consequently, one normal base pair is converted to another pair that is genetically incorrect. An agent of mutation (a mutagen) such as nitrous acid has been known to cause deamination of adenine to hypoxanthine, cytosine to uracil, and guanine to xanthine, the net effect being a replacement of A–T with G–C in the DNA structure. A transversion involves the substitution of a purine by a pyrimidine and vice versa.

**2 Tautomeric shifts.** It is known that each of the bases of DNA can exist in rare states as a result of the redistribution of electrons and protons in the molecule, events called tautomeric shifts. When this occurs, the base sometimes is unable to hydrogen bond with its complementary base. Instead, some of these altered bases succeed in bonding with the wrong bases, resulting in mutations when, during replication, one purine (or pyrimidine) is substituted for the other (Figure 12.2).

**3 Effect of base analogues.** Certain analogues of the naturally occurring bases in the DNA molecule have



**Figure 12.2** Mutations may be caused by tautomeric shifts: (a) shift involving cytosine, and (b) shift involving thymine.



**Figure 12.3** Mutations may be caused by transition resulting from the substitution of base analogues for a natural base: (a) mistake in incorporation, and (b) mistake in replication.

been shown to have mutagenic effects. For example, the natural base thymine (T), a 5-methyluracil, has a structural analogue, 5-bromouracil (5-BU). The two bases are so similar that 5-BU can substitute for T during replication, leading to base pair transition (Figure 12.3).

**4 Single base deletions and additions.** A variety of alkylating agents (e.g., sulfur and nitrogen mustards) can act on the DNA molecule, reacting mainly with guanine (G) to alkylate and remove it from the DNA chain. The missing spot may be occupied by any of the four bases to create mutations, usually by transition. Acridine is also known to express its mutagenic effect through the addition of deletion of bases.

**Effect at the protein level** Four different effects are known to occur as a result of nucleotide substitution.

**1 Silent mutation.** Because the genetic code is degenerate (one amino acid can be coded by more than one triplet), a change from ACG → CGG has no effect as both triplets code for arginine.

**2 Neutral mutation.** This kind of mutation involves an altered triplet code that codes for a different but chemically equivalent amino acid. For example, CAC may change to CGC, altering histidine to a chemically equivalent amino acid, arginine. The change causes a change in the primary structure of the protein (amino acid sequence) but the formation of the resultant protein may be unchanged.

**3 Missense mutation.** Unlike neutral mutations, a missense mutation results when an altered triplet codes for a different amino acid that results in the protein being non-functional. For example, in hemoglobin of humans, a change of GAG (Glu) to GTG (Val) results in serious consequences.

**4 Nonsense mutation.** A nonsense mutation causes an existing amino acid to be changed to a stop codon (e.g., TAA, TAG), resulting in premature termination of protein synthesis.

**Frame shift mutation** Insertion–deletion mutations (indels) may cause significant changes in the amino acid composition of a protein and hence its function. For example, GAG-CCG-CAA-CTT-C (corresponding to Glu-Pro-Glu-Leu) may be altered by a deletion of G that shifts the reading frame to the right by one nucleotide to produce AGC-CGC-AAC-TTC (corresponding to Ser-Arg-Asp-Phe). Very simply, CAT-CAT-CAT-CAT + A (at the beginning) → ACA-TCA-TCA, etc.

#### Genomic mutation

Errors in cell division resulting from disorders in the spindle mechanism may result in improper distribution of chromosomes to daughter cells. Such errors may cause some cell division products to inherit more or less of the normal chromosome number. These errors, called **chromosomal mutations**, are of two main kinds: **euploidy** (cells inherit an additional complete set of the basic chromosome set,  $n$ ) and **aneuploidy** (certain chromosomes are missing from the basic set or added to the set in some cell division products). The subject is discussed in more detail in Chapter 13.

#### Structural chromosomal changes (aberrations)

Changes in chromosome structure begin with a physical break that may be caused by ionizing radiation (e.g., X-rays). After a break, several events may occur:

- 1 The ends of the segment may be disunited.
- 2 The break may be repaired to restore the chromosome to its original form (restitution).
- 3 One or both ends of a break may join to the ends produced by a different break event (non-restitutional union). These events may result in one of four types of rearrangement – **deletions, duplications, inversions, or translocations**. The resulting consequences are variable.

### Mutagenic agents

Agents of artificial mutations are called **mutagens**. They may be grouped into two broad categories – **physical mutagens** and **chemical mutagens**. The specific agents vary in ease of use, safety issues, and effectiveness in inducing certain genetic alterations, suitable tissue, and cost, among other factors.

#### Physical mutagens

The principal physical mutagens are ionizing radiations (Table 12.1). X-rays were the first to be used to induce

**Table 12.1** Examples of commonly used physical mutagens.

Mutagen	Characteristics
X-rays	Electromagnetic radiation; penetrate tissues from a few millimeters to many centimeters
Gamma rays	Electromagnetic radiation produced by radioisotopes and nuclear reactors; very penetrating into tissues; sources are Co <sup>60</sup> and Ce <sup>137</sup>
Neutrons	A variety exists (fast, slow, thermal); produced in nuclear reactors; uncharged particles; penetrate tissues to many centimeters; source is U <sup>235</sup>
Beta particles	Produced in particle accelerators or from radioisotopes; are electrons; ionize; shallowly penetrating; sources include P <sup>32</sup> and C <sup>14</sup>
Alpha particles	Derived from radioisotopes; a helium nucleus capable of heavy ionization; very shallowly penetrating
Protons	Produced in nuclear reactors and accelerators; derived from hydrogen nucleus; penetrate tissues up to several centimeters

mutations. Since then, various subatomic particles (neutrons, protons, beta particles, alpha particles) have been generated using nuclear reactors. Gamma radiation from radioactive cobalt (Co<sup>60</sup>) is widely used. It is very penetrating and dangerous. Neutrons are hazardous and less penetrating, but they are known to seriously damage chromosomes. They are best used for materials such as dry seed, whereas the gentler gamma radiation is suitable also for irradiating whole plants and delicate materials such as pollen grains. The relative biological effectiveness (RBE) of fast neutrons is higher than for gamma rays and X-rays. The breeder is interested in identifying and using treatments with high RBE for maximizing the number of mutants produced. Treatments with high RBE have high ionization density. Modifying the treatment environment (e.g., oxygen, moisture content of tissue) can increase the effectiveness of the treatment.

Ionizing radiations cause mutations by breaking chemical bonds in the DNA molecule, deleting a nucleotide, or substituting it with a new one. The radiation should be applied at the proper dose, a factor that depends on radiation intensity and duration of exposure. The dosage of radiation is commonly measured in roentgen (r or R) units. The exposure may be **chronic** (continuous low dose administered for a long period) or **acute** (high dose over a short period). The quality of mutation (proportion of useful mutations) is not necessarily positively correlated with dose rate. A high dose does not necessarily yield the best results. A key limitation of use of physical mutagens is the source. Special equipment or facilities are required for X-rays and nuclear-based radiations.

#### Chemical mutagens

Chemical mutagens are generally milder in their effect on plant material. They can be applied without complicated equipment or facilities. The ratio of mutational to undesirable modifications is generally higher for chemical mutagens than for physical mutagens. However, practical success with chemical mutagens lags behind achievements with physical mutagens. Usually, the material is soaked in a solution of the mutagen to induce mutations. Chemical mutagens are generally carcinogenic and must be used with great caution. One of the most effective chemical mutagenic groups is the group of alkylating agents (these react with the DNA by alkylating the phosphate groups as well as the purines and pyrimidines). Another group is the base analogues (they are closely related to the DNA bases and can be wrongly

**Table 12.2** Examples of commonly used chemical mutagens.

Mutagen group	Examples
Base analogues	5-bromouracil, 5-bromodeoxyuridine
Related compounds	Maleic hydrazide, 8-ethoxy caffeine
Antibiotics	Actinomycin D, mitomycin C, streptonigrin
Alkylating agents	
Sulfur mustards	Ethyl-2-chloroethyl sulfide
Nitrogen mustards	2-chloroethyl-dimethyl amine
Epoxides	Ethylene oxide
Ethyleneimines	Ethyleneimine
Sulfonates, etc.	Ethyl methane sulfonate (EMS), diethylsulfonate (DES)
Diazoalanes	Diazomethane
Nitroso compounds	<i>N</i> -ethyl- <i>N</i> -nitroso urea
Azide	Sodium azide
Hydroxylamine	Hydroxylamine
Nitrous acid	Nitrous acid
Acridines	Acridine orange

incorporated during replication); examples are 5-BU and maleic hydrazide (Table 12.2). Chemical mutagens commonly used are ethyl methane sulfonate (EMS) and diethylsulfonate (DES). The half-life (time taken for degradation of the initial amount of alkylating agent) for EMS in water is about 93 hours at 20°C but only 10 hours at 37°C. Consequently, chemical mutagens are best freshly prepared for each occasion. Plant breeders who undertake wide crosses often need to manipulate the chromosome number to have fertility. To do this, the most common chemical mutagen used is colchicine. Chemical mutagens produce more gene mutations and less chromosome damage.

### Types of materials used for mutagenesis

Whole plants and plant parts may be used for mutagenesis, depending on the mutagen to be used, the species, and the research objective, among other factors. Seeds are the most commonly used material for mutagenesis involving physical or chemical mutagens. Whole plants are suited to chronic exposure to physical mutagens. A major drawback to seeds and whole plants is that chimeras are commonly produced. Pollen grains can be used to avoid chimeras. The resulting plants from pollen mutagenesis are heterozygous for any genetic change.

However, sufficient quantities of pollen are difficult to obtain, and also they do not stay viable for long.

The plant breeder may also use cuttings and cell or tissue culture explants for mutagenesis. In tissue culture, cells or tissue are treated with the appropriate mutagens and regenerated into new whole plants.

### Factors affecting the success of mutagenesis

Mutations are random events even when scientists induce them. The plant breeder using the technique can increase the success rate by observing the following cautions.

- 1 **Clear objective.** A program established to select one specific trait is more focused and easier to conduct with a higher chance of success, than a program designed to select more than one trait.
- 2 **Efficient screening method.** Mutation breeding programs examine large segregating populations to increase the chance of finding the typically rare desirable mutational events. An efficient method of screening should be developed for a mutation breeding program.
- 3 **Proper choice of mutation and method of treatment.** Mutagens, as previously discussed, vary in various properties including source, ease of use, penetration of tissue, and safety. Some are suitable for soft tissues, whereas others are suited to hard tissue.
- 4 **Dose rate.** The breeder should decide on the appropriate and effective dose rate (dose and duration of application). The proper dose rate is determined by experimentation for each species and genotype. Plant materials differ in sensitivity to mutagenic treatment. It is difficult to find the precise dose (intensity), but careful experimentation can identify an optimum dose rate. Mutagenic treatments invariably kill some cells. Overdose kills too many cells and often produces crippled plants, whereas underdose tends to produce too few mutants. Not only are dose-response relationships rarely known, but they are influenced by the experimental conditions. It is recommended that three dose levels be used in a project – the optimal rate, plus one dose above and below this rate.
- 5 **Proper experimental conditions.** It is known that the oxygen level in the plant material affects the amount of damage caused by the mutagen – the higher the oxygen level, the greater the injury to the material tends to be. The change in the effect of a mutagen with oxygen supply is called the **oxygen enhancement ratio**. In studies in which enhanced

mutagen frequency is desired, the experimental conditions may be supplemented with oxygen (e.g., bubbling oxygen through the mutagen solution in the case of chemical mutagenesis). Where such enhancement is undesirable, an oxygen-free environment should be used. The effect of oxygen on mutagenesis is dependent upon the moisture content of the tissue. The higher the tissue moisture, the lower the tissue oxygen supply. Mutagens vary in the importance of moisture in their effectiveness in inducing mutation, X-rays being more affected than gamma rays, for example. Dry seeds are better to use when enhanced mutation frequency due to oxygen status of the research environment is not desired. In chemical mutagenesis, temperature has an effect on the half-life of the mutagen, high temperature accelerating the reaction.

Another experimental factor affecting the success of mutagenesis is the pH of the environment in chemical mutagenesis. For example, EMS is most effective at pH 7.0, whereas sodium azide is most effective at pH 3.0. Sometimes, dry seeds need to be presoaked to prepare the cells to initiate metabolic activity. It is important to mention that it is easier to modify the experimental condition when seeds are being used for the mutation program than when other materials are used.

## Mutation breeding of seed-bearing plants

### Objectives

The breeder should have clear objectives regarding the trait to be induced. Induced mutations are neutral mutants. They are equivalent to natural mutations, and hence once observed the normal procedure for plant breeding is applicable. It should be borne in mind that mutagenic treatments have both primary and secondary effects. In addition to the specific desired alteration (primary effect), mutagens tend to alter the background genotype (secondary effect). Mutagenesis may cause significant variation in quantitative traits.

### Genotypes and source of material

Both self- and cross-pollinated species can be improved with induced mutations. Seeds from self-pollinated species are homozygous and homogeneous, making it easier to identify mutants in the field. Outcrossed species are heterozygous and heterogeneous, making it more challenging to identify mutants. Dormant seeds are easier to handle than vegetative material.

Mutation breeding is often used to correct a specific deficiency in an adapted and high-yielding genotype. This otherwise desirable genotype may be susceptible to a disease or may need modification in plant architecture. A mutant is easier to spot if the parent (source genotype) is genetically pure (as opposed to being a mixture). Also, mutants are easier to identify if the mutant trait is distinctly different from the parental trait (e.g., it is easier to spot a true dwarf mutant if a tall parent is used for the project). Some suggest using an  $F_1$  in some instances because it contains two genomes, which may increase gene recombination and thereby produce a greater amount of diversity in gene mutations.

### Treatment

The goal in seed treatment is to treat enough seed to eventually produce a large enough segregating population in the second generation ( $M_2$ ). This means the number of treated seed should generate sufficient first generation plants ( $M_1$ ) that have sufficient fertility to produce the size of  $M_2$  population needed. The appropriate number is determined through preliminary experimentation. One of the side effects of mutagen treatment is sterility, which should be factored into the decision process to determine the number of seeds to treat. The  $M_2$  population may consist of 20,000–50,000 plants.

The seed is multicellular and hence a mutation in a single cell will give rise to a chimeric plant. Thus, an  $M_1$  plant is subject to both **diplobiotic** (competition between mutated and normal tissue during the vegetative phase of development) and **haplobiotic** (competition between mutated and normal pollen during fertilization) selection in order to be included in the tissues forming seed on the plants. Treating a seed with physical agents is done by placing the seed in an appropriate container, positioning it at a proper distance from the source of radiation, and exposing it at a predetermined dose rate. Gamma irradiation doses that have been used may be as low as 0.5 krad (e.g., in corn) or as high as 25 krad (e.g., in dry seed of wheat), whereas the dose rate for X-rays ranges between 10 and 25 krad. When using chemicals, the proper concentration of mutagen is prepared at the desirable pH and temperature. Seeds are soaked in the mutagenic solution for the appropriate duration. EMS concentrations that have been used range from less than 1% (e.g., in tomato) to about 4% (e.g., in wheat).

The common physiological injuries caused by the mutagenic treatment are reduction in seedling height (most frequently used identification of injury in the  $M_1$ ),

cytological effects (chromosomal aberrations), and sterility (evaluated by counting the number of inflorescences per plant, etc.).

### Field planting and evaluation

Treated seed ( $M_1$ ) may be planted in a small plot to produce  $M_2$  seed that is harvested for planting an  $M_2$  spaced population. It is advisable to plant plots of the untreated genotype ( $M_0$ ) used in the project for comparison, to aid in readily identifying mutants. Dominant mutants are identifiable in the  $M_1$  generation; recessive mutants are observable after selfing to produce the  $M_2$  generation.

Furthermore, in species that produce tillers, it is possible that only one of several tillers produced will carry an induced mutation. To produce the mutation throughout the plant, gametes (pollen grain) should be treated. Using seed also frequently produces chimeras, since mutations are induced in single cells that divide and differentiate into parts of the plant. Consequently, the stem may be a mutant tissue while the leaf is normal.

It should be pointed out that  $M_1$  and  $F_1$  generation plants could be genetically different.  $F_1$  plants produced from inbred lines are genetically identical. However  $M_1$  plants in a population may have different mutations. Hence, it is logical and proper to handle an  $M_1$  population as a segregating population such as an  $F_2$ .

To use mutagenesis as a breeding method for producing new cultivars, various cultivar selection strategies may be used, including the following.

**1 Bulk selection.** First, the breeder grows  $M_1$  plants, and then harvests and bulks all the seed. A sample of seed is planted in the next season, harvested, and the seed bulked. Alternatively, individual  $M_2$  plants may be harvested and bulked for progeny testing in the next season. Seed from progeny rows showing the desired mutant phenotype are identified and harvested. If rows are segregating, the  $M_3$  plants may be harvested and advanced individually. The breeder conducts replicated tests, evaluating on the basis of the desired mutant and other desirable agronomic traits. A weakness in the bulk method is that mutant plants often have low productivity. Consequently, by planting only a sample of the  $M_2$ , it is likely to exclude the desired mutant.

**2 Single-seed descent.** In this method, one or a few  $M_2$  seeds are selected from each plant and bulked for planting  $M_2$  plants. Desirable  $M_2$  plants are harvested and progeny rowed; alternatively, seeds from

desirable plants may be harvested and bulked. This method also has the potential of excluding desirable mutants, even though it allows a larger number of  $M_1$  plants to be sampled.

**3 Pedigree method.** Each  $M_1$  plant is harvested separately.  $M_2$  progenies are grown from  $M_1$  plants. Desirable  $M_2$  plants are harvested and progeny rowed. Desirable rows are harvested and bulked separately. The  $M_{2:4}$  lines are evaluated in replicated tests.

### Mutation breeding of clonally propagated species

Species that reproduce vegetatively or by apomixis may also be improved through mutation breeding. Vegetatively propagated species tend to be highly heterozygous. Consequently, selfing is accompanied by inbreeding depression. Mutation breeding offers a method of crop improvement whereby the genetic structure is largely unperturbed. Physical mutagens are more frequently used in such species than chemical mutagens. The plant parts targeted for mutagen treatment are those that can produce a bud from which a plant can develop. These parts include modified parts (e.g., tubers, rhizomes, shoot apex, cuttings, bulbs). The exposure to the mutagen must occur as early as possible in the development of the bud and target the meristematic cell. A mutation in the meristematic cell is critical in avoiding chimeras, which can delay selection because the breeder may not be able to identify the mutant without propagating all the material a number of times.

Chimeras are desirable in the breeding of certain ornamental species such as African violet (*Saintpaulia ionantha*). Induced mutations in clones are of necessity dominant mutations, unless the starting material was heterozygous and hence could yield recessive mutants. Furthermore, clonally induced mutants are primarily chimerical and start as **sectorial** mutants, later becoming **periclinal**. The growing point commonly has two layers (some have three), the outer layer generating the epidermis as well as some leaf mesophyll, while the inner layer generates the remaining parts of the plant. The innermost layer (LI) has one cell per layer, while LIII has multiple layers. Mutations, being one-celled events, result in part of whole layers exhibiting chimeras. A periclinal chimera is one that involves the entire layer (the whole layer is mutated), whereas a **mericinal** chimera involves only part of one layer. Periclinal mutations are stable. The number of initial cells in the layers in the



## Industry highlights

### *Current apple breeding programs to release apple scab-resistant scion cultivars*

F. Laurens

UMR Génétique et Horticulture (GenHort) (INRA/INH/UA), INRA Centre d'Angers, 49070 Beaucouzé, France

#### **Introduction**

The apple is the fourth largest fruit crop grown in the world (59 million tons) after citrus (108 million tons), banana (70 million tons), and grape (65 million tons) (FAO 2004). It is widely grown throughout the world in various pedoclimatic conditions, from the European Nordic countries to subtropical regions in Brazil or South Africa, yet it is mostly widespread in temperate regions. Apples can be consumed as fresh fruits or after processing.

More than 10,000 cultivars have been recorded (Way et al. 1990), yet only a few are grown on a commercial scale. We can distinguish four main categories according to the ways and times they have been released:

- 1 Old and local cultivars grown specifically in some regions or countries, e.g., "McIntosh" in Canada, "Belle de Boskoop" in the Netherlands, "Reinette du Canada" and "Reine des reinettes" in France, "Cox's Orange Pippin" and "Worcester Pearmain" in England, and "Rome Beauty" in the USA.
- 2 Cultivars selected from chance seedlings in the 18th or 19th centuries. The best examples are "Golden Delicious", "Red Delicious", and "Granny Smith" found in the USA. They are still among the most grown cultivars in the world.
- 3 Most of the cultivars released in the last 20 years, which are derived from controlled crossing programs from all over the world, e.g., "Gala" (New Zealand), "Elstar" (the Netherlands), "Jonagold" (USA), "Fuji" (Japan), and more recently "Pink Lady" (Australia). These cultivars derived from a very narrow genetic basis usually involving only a very few genitors, e.g., "Golden Delicious", "Jonathan", and "Red Delicious". They show significant improvement over older cultivars; in general, they fit the main current commercial requirements: high and regular cropping, fruit attractiveness (nice coloring, few cosmetic defaults), good fruit size, good fruit taste, adaptation to transport, and long storage ability. These cultivars have been propagated by grafting on a very large scale with many millions of trees planted. During the propagation process some rare mutation events appear affecting some characters of the tree or fruit. The most interesting and stable sports are then selected and propagated by grafting to get new cultivars that differ slightly from the original one. New sports can also be produced by inducing mutation, mostly by irradiation (X-rays or gamma rays). Mutations have had a major economic impact on the apple industry – in 1995, CPVR (community plant variety rights) applications for apple mutation cultivars were as important as those for apple seedling cultivars (Semon 2004). They affect two major traits: (i) the shape and volume of the tree, with the spur (compact) types allowing a higher orchard density (mainly for "Red Delicious"); and (ii) the coloring of the fruit. Nearly all the most propagated bicolored apples are not from the original clone but from sports with more intense or higher amounts of skin coloring, which make the fruits more attractive for the consumer and more profitable for the growers.
- 4 The latest generation of cultivars, from breeding programs that aim to combine fruit quality, good cropping, and scab resistance.

As with most fruit trees, the apple is grown as a composite tree that combines a rootstock and a fruiting scion. The focus in this box will be on the improvement of the dessert cultivar scion. First, I will present a short review of the main apple breeding programs in the world and their main objectives, and then describe the specific constraints of fruit trees that limit the development of breeding. Finally, I will describe the cycle of selection developed at the Institut National de la Recherche Agronomique (INRA), France, to release new apple cultivars resistant to scab.

#### **Review of the main apple breeding programs and their objectives**

Europe is the region where the highest number of dessert apple scion breeding programs are recorded: 32 have been reviewed by Y. Lespinasse (personal communication); 26 are developed by public institutes, but the new trend is an increasing amount of privately funded ones. On average, there is at least one apple breeding program in each European country; in Italy, up to five have been registered. In other continents there are fewer programs and they are generally present in countries or states where apple production is important (Laurens 1999).

Fruit quality and disease resistance are the most important breeding targets. Adaptation to climatic conditions is also of prime interest for the countries located in marginal areas, either in Nordic or subtropical regions. Many apple breeders also try to improve the tree habit to obtain productive and regular cropping trees.

### Fruit quality

Releasing apple cultivars with high fruit quality is obviously the major aim for each breeder. All around the world, the same criteria are taken into account to assess fruit quality. For fruit appearance, the first aim is to select big size and nicely colored fruits. Selections with cosmetic defaults (russetting) are discarded. Many traits are involved to build the fruit taste. The most important are flesh texture, firmness, juiciness, and sugar and acidity content. Flavor is also very important but more difficult to assess. Assessment of fruit quality is very subjective, it varies between people and countries. Breeding programs in China, Japan, Brazil, and India focus on sweet-tasting fruits with "Fuji", "Gala", or "Red Delicious" as references. People from northern countries prefer more acidic fruits such as those from "Elstar", "Jonagold", or "Braeburn".

Many apple breeders aim to improve storage ability but there are very few detailed studies on this topic.

The ongoing European project HiDRAS (High Quality Disease Resistance Apples Project) (2003–2006) is a collaborative effort between 11 European groups aimed at the identification of genetic factors controlling fruit quality based on a multidisciplinary approach, including apple breeding, genetics, molecular biology, statistics, and bioinformatics (Gianfranceschi & Soglio 2004). Breeders will be provided with new tools, such as molecular markers linked to fruit quality and pathogen resistance, to develop marker-assisted selection (MAS).

### Disease and pest resistance

Scab, caused by *Venturia inaequalis*, and powdery mildew, caused by *Podosphaera leucotricha*, are the two main diseases in apple orchards. About 15–20 specific treatments per year are required to control them.

**Scab** For the last 60 years, apple breeders all over the world have aimed to release new apple cultivars resistant to scab. The *V<sub>f</sub>* gene has been widely used in all scab-resistance programs with more than 80% of the scab-resistant cultivars released today carrying the *V<sub>f</sub>* gene. All over the world, crossing designs followed the same strategy: a series of modified backcrosses between scab-resistant hybrids and commercial scab-susceptible cultivars. At each step of selection, scab-resistant hybrids with the highest fruit quality were selected and crossed with other commercial cultivars. But, in the 1990s, the first strains of *V. inaequalis* able to overcome the *V<sub>f</sub>* gene were found in Europe (Parisi et al. 1993; Roberts & Crute 1993) and have been progressively spreading in Northern Europe. These events have led apple breeders to change crossing strategies: the aim is now to combine several types of resistance in one genotype. Molecular markers are necessary to detect in progenies those seedlings that carry more than one gene of resistance.

The European project DARE (Durable Apple Resistance in Europe) (1998–2002) was partly set up to control this issue. Decisive results were obtained characterizing the pathogenicity and diversity of *V. inaequalis* (Parisi et al. 2004) and in determining the resistance status of a large range of apple cultivars carrying various levels of durable resistance (Laurens et al. 2004). The central part of the project studied the genetic bases of apple scab resistance (Calenge et al. 2004; Tartarini et al. 2004) through the development of genetic maps on various progenies. Major resistant genes and quantitative trait loci (QTLs) were localized and tested for their durability against the new virulent strains of *V. inaequalis*. New molecular markers usable by the breeders were proposed for MAS (Kellerhals et al. 2004).

**Powdery mildew** Powdery mildew is also an important disease worldwide. Resistance to it is one of the major aims in most breeding programs. Up to now, all the cultivars released with improved mildew resistance have been based on partially resistant parents. But some major genes originating from wild *Malus* species have now been combined with scab-resistant ones. The most used are *Pl<sub>2</sub>* from *M. zumi* (Knight & Alston 1968), *Pl<sub>1</sub>* from *M. robusta*, and *Pl<sub>w</sub>* from "White Angel", an ornamental crab apple from northern America.

**Other disease and pest resistance** Fire blight (*Erwinia amylovora*) is one of the worst plagues in most of the apple-producing countries. No major resistance has yet been detected. Genetic studies are ongoing to localize QTLs of partial resistance (F. Calenge, personal communication). Some markers should soon be available to further develop MAS strategies.

Other fungi that can cause severe damage in some regions (*Nectria galligena*, *Valsa ceratosperma*, *Gymnosporangium juniperi-virginianae*, etc.) are also studied but only at a small scale (Laurens 1999).

Some breeding works are also being developed for aphids – the rosy apple aphid (*Dispahis plantaginea*) and woolly apple aphid (*Eriosoma lanigerum*).

### Climatic adaptation

The main programs are based in the Nordic countries (Drudze 2004; Nybom 2004) and in Canada (Khanizadeh et al. 2004). The objectives are to release winter-hardy cultivars adapted to a short and cool growing season. Specific local cultivars are used as progenitors.

Programs are also being conducted in the Republic of South Africa and Brazil, releasing cultivars with low chilling requirements adapted to the apple areas of these regions (Labuschagné 2004).

### Tree habit

Almost all breeders concentrate their tree habit programs on compact trees. "Wijick", a sport of "McIntosh" with columnar habit, is often used but it is very susceptible to biennial bearing. At INRA, the aim for tree habit is to create weeping and not too vigorous trees that give good and regular cropping and are easy to prune and harvest. Parents that transmit the trait "one fruit per cluster" are also included (Laurens et al. 2000).

### *Selection process to release new scab-resistant apple cultivars*

Compared to the main annual crops, breeding for fruit trees is particularly costly and long: the whole process to select a new apple cultivar is likely to last between 15 and 20 years. Some specific characteristics of fruit trees explain this.

The selection process in all fruit species is linked to genetic, morphological, physiological, and agronomical constraints. The length of the juvenile period is certainly the most restrictive. For the apple tree, it lasts from 3 to 6 years according to the genotype and tree management. This period is characterized by a strong vegetative growth and no flowering. During these first years of growing, no assessment can be performed on the fruits or on the tree, selection is limited to early resistance tests on seedlings. Another big constraint for the breeder is the volume of each individual tree, which limits considerably the number of progenies that can be studied in one site compared with annual crops.

Few genetic studies have been published on genetic parameters of agronomic traits in the apple; most of them have been dedicated to resistance. In fact, work on fruit quality faces many problems: the length of the juvenile period, establishing objective criteria, the close interaction between vegetative growth of the tree and fruit-quality traits, the complex genetic nature of the traits, etc. More generally, the highly heterozygous status of the apple cultivars, due to self-incompatibility, makes the estimation of genetic parameters very difficult as well as the breeding process.

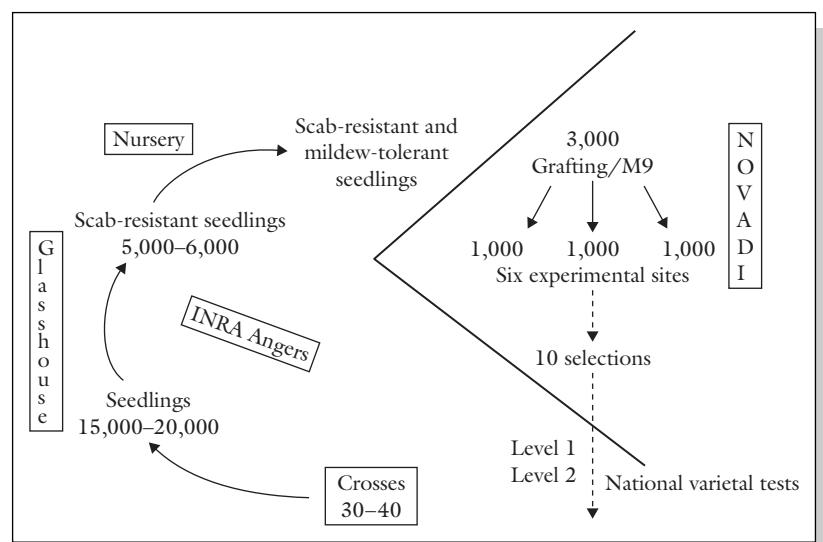
Fortunately, the breeder can take advantage of the apple tree's ability to be propagated by grafting. Once one initial tree is found promising, it is easy to propagate it quickly and in large numbers to be tested in different sites.

### Cycle of selection

All apple breeders roughly follow the same process of selection; only some practical points may change from one team to another. As an illustration, this box will present the cycle of selection performed in France through a partnership started in 1996

between the INRA, a public research institute, and NOVADI Ltd Co., which employs 16 nurserymen involved in the development of new, resistant apple cultivars. Work and responsibilities are shared between INRA and NOVADI as illustrated in Figure 1 (Laurens & Pitiot 2003).

Each year, cross-pollinations are performed. Flowers are hermaphrodite, so each genitor can be used as the male or female parent. In general, the earliest flowering parent is used as the male. The collection of the pollen and the different crossing techniques are described by Janick et al. (1996). The first step of selection occurs 1 year later on young seedlings that are tested in the greenhouse for scab resistance (Figure 1). All the seedlings are sprayed with a scab mixture, mainly race 1 (L. Parisi, personal communication), obtained from scabbed leaves collected in the INRA orchards. After two



**Figure 1** Cycle of selection in the collaborative INRA-NOVADI apple breeding program.

**Table 1** Main descriptors used to characterize each individual in each experimental site of the collaborative INRA–NOVADI apple breeding program.

	Minimum descriptors	Additional descriptors
Tree	Cropping, picking time	Number of fruits/cluster
Fruit aspect	Size, shape, attractiveness, opening calyx	Ground and overcolor, % <sup>1</sup> and type <sup>2</sup> of overcoloring, amount of russet <sup>1</sup> , cracking
Fruit taste	Firmness, flesh quality, flavor	Juiciness, sugar : acid ratio
Biotic and abiotic resistance		Powdery mildew, bitter pit

<sup>1</sup> 0–1/4, 1/4–1/2, 1/2–3/4 and 3/4–1.

<sup>2</sup> Stripes, blush, blush and stripes, or stripes and blush.

inoculations, the susceptible plants are discarded. The remaining resistant trees are planted in nursery conditions at the INRA. After 18 months, a mildew assessment is performed, and individuals that show a good level of resistance are selected. Budwoods of scab- and mildew-resistant individuals are then sent to the nurserymen who graft them on the rootstock M9, which confers a low vigour to the scion. The trees are planted at six nursery sites representing various French apple-cropping regions: Val de Loire (northwest), Auvergne (central), Aquitaine (southwest), and Languedoc (southeast).

#### Fruit evaluation

Fruit evaluation is managed at each site at the optimum date of maturity. The description is performed by means of 20 descriptors (Table 1): nine “minimum” descriptors are assessed for all individuals at all the sites: picking time, tree cropping, fruit size, attractiveness, shape, opening calyx, flesh firmness, flesh quality, and flavor. Eleven additional (optional) descriptors are also recorded (Table 1). Most of the descriptors are assessed on a 1 (low/weak/bad) to 5 (high/very good) scale, except, for example, for the shape of the fruit (1–9), which is actually a merging assessment of the height and conicity of the fruit. All the taste traits are sensorially assessed. Each individual is evaluated for at least 2 years in each site.

Figure 1 also illustrates the very high selection pressure applied: of the 15,000–20,000 seedlings grown in the greenhouses, only 10 elite individuals will be finally selected. It takes about 6–7 years to reach this first level of selection. The 10 selected individuals are then tested together with all the best worldwide selections in a national experimental network involving INRA, Ctifl (Centre Technique Interprofessionnel des Fruits et Légumes), and regional experimental stations in various sites across France. This long (10 years), multisite, experimental process is essential to test agronomic behavior, climatic adaptation, and consumer acceptance of the elite selections, thus giving a guarantee to the growers before investing in a new cultivar.

#### Conclusion

A major problem of the commercial apple cultivars currently available is their susceptibility to the main pests and diseases, which requires a lot of phytosanitary treatment in the orchards. After long processes of crossing and selection, apple breeders are now releasing a new generation of cultivars. Some are already being planted significantly, for example “Goldrush” (Crosby et al. 1994), “Topaz”, and “Ariane” (Laurens et al. 2005), which combine very high standards of fruit quality with resistance to the common strains of scab. The aim of the breeder is now to release new cultivars carrying durable resistance to scab and other diseases and pests. For this purpose, they are developing new methodologies and tools (molecular markers) to shorten and make more efficient the cycles of selection.

#### References

- Calenge, F., A. Faure, M. Goerre, C. Gebhardt, W.E. Van de Weg, L. Parisi, and C.E. Durel. 2004. Quantitative trait loci (QTL) analysis reveals both broad-spectrum and isolate-specific QTL for scab resistance in an apple progeny challenged with eight isolates of *Venturia inaequalis*. *Phytopathology* 94(4):370–379.
- Crosby, J.A., J. Janick, and P.C. Pecknold. 1994. Goldrush apple. *Hortscience* 29(7):827–828.
- Drudze, U. 2004. New apple and pear selections from hybrid material of ‘ledzeni’ in Latvia. *Acta Hort.* 663:895–898.
- FAO. 2004. FAOSTAT database results 2004. <http://faostat.fao.org>.
- Gianfranceschi, L., and V. Soglio. 2004. The European project HiDRAS: innovative multidisciplinary approaches to breeding high quality disease resistant apples. *Acta Hort.* 663:321–326.
- Janick, J., J.N. Cummins, S.K. Brown, and M. Hemmat. 1996. Apples. In: *Fruit breeding*, Vol. I: Tree and tropical fruits (J. Janick, and J.N. Moore, eds), pp. 1–77. Wiley & Sons, New York.
- Kellerhals, M., C. Sauer, B. Guggenbuehl, et al. 2004. Apple breeding for high quality and durable disease resistance. *Acta Hort.* 663:751–756.

- Khanizadeh, S., F. Laurens, Y. Lespinasse, Y. Groleau, O. Carisse, and J.R. DeEll. 2004. 'Reinette Russet®' and 'Galarina®': two new winter hardy disease resistant apple cultivars released through joint collaboration of INRA (France) and AAFC (Canada). *Acta Hort.* 663:887–889.
- Knight, R.L., and F.H. Alston. 1968. Sources of field immunity to mildew (*Podosphaera leucotricha*). *Can. J. Genet. Cytol.* 10:294–298.
- Labuschagné, I.F. 2004. Budbreak number as selection criterion for breeding apples adapted to mild winter climatic conditions: a review. *Acta Hort.* 663:775–777.
- Laurens, F. 1999. Review of the current apple breeding programmes in the world: objectives for scion cultivar improvement. *Acta Hort.* 484:163–170.
- Laurens, F., J.M. Audergon, J. Claverie, et al. 2000. Integration of architectural types in French programmes of ligneous fruit species genetic improvement. *Fruits* 55:141–152.
- Laurens, F., M. Chevalier, E. Dolega, et al. 2004. Local European cultivars as sources of durable scab resistance in apple. *Acta Hort.* 663:115–121.
- Laurens, F., Y. Lespinasse, and A. Fouillet. 2005. A new scab-resistant apple: 'Ariane'. *Hortscience* 40(2):484–485.
- Laurens, F., and C. Pitiot. 2003. French apple breeding program: a new partnership between INRA and the nurserymen of NOVADI. *Acta Hort.* 622:575–582.
- Nybom, H. 2004. 'Frida' and 'Fredrik', the first scab resistant apple cultivars developed in Sweden. *Acta Hort.* 663:871–873.
- Parisi, L., V. Fouillet, H.J. Schouten, et al. 2004. Variability of the pathogenicity of *Venturia inaequalis* in Europe. *Acta Hort.* 663:107–113.
- Parisi, L., Y. Lespinasse, J. Guillaumès, and J. Krüger. 1993. A new race of *Venturia inaequalis* virulent to apples with resistance due to Vf gene. *Phytopathology* 83(5):533–537.
- Roberts, T., and I. Crute. 1993. Apple scab resistance from *Malus floribunda* 821 (V<sub>f</sub>) is rendered ineffective by isolates of *Venturia inaequalis* from *Malus floribunda*. Abstracts, Third Workshop on Integrated Control of Pome Fruit Diseases. Lofthus (Norway), 43.
- Semon, S.F.A., 2004. Developments in breeding and techniques within the community plant variety rights fruit sector. *Acta Hort.* 663:707–711.
- Tartarini, S., F. Gennari, D. Pratesi, et al. 2004. Characterization and genetic mapping of a major scab resistance gene from the old Italian apple cultivar 'Durello di Forlì'. *Acta Hort.* 663:129–133.
- Way, R.D., H.S. Aldwinckle, R.C. Lamb, et al. 1990. Genetic resources of temperate fruit and nut. *Acta Hort.* 290:1–62.

growing tip determines the sectorial patterns. Sectorial patterns are mostly unstable, eventually producing periclinal structures. Because of this behavior, the breeder needs to grow several clonal generations successively, and select for the desired mutant as well as vegetative stability. Vegetative stability is attained when uniformly periclinal lines of subclones have been isolated. In a vegetative system, the treated plant meristem is called the M<sub>1</sub>V<sub>1</sub>, the next vegetative plant being M<sub>1</sub>V<sub>2</sub>, etc.

### Mutations from tissue culture systems

Useful mutations have arisen spontaneously or are deliberately induced in culture conditions. Such processes, **somaclonal variation** and **somatic selection**, were discussed in Chapter 11.

### Using induced mutants

A plant may be selected for a specific mutant trait. However, it does not mean that this plant contains only one genetic change. It is recommended to backcross this mutant to its original parents to reselect a pure version

of the mutant. Mutants may be directly used in cultivar development. Like the backcross method, an adapted line can be improved through mutation breeding to correct a shortcoming. In sexually reproducing species, the target mutant traits can be fixed and isolated in the M<sub>2</sub> or M<sub>3</sub> generation (compared to F<sub>6</sub> or F<sub>7</sub> generations in conventional breeding). The mutant lines may be used in genetic studies as markers, among other uses.

### Limitations of mutagenesis as a plant breeding technique

There are a number of disadvantages of mutagenesis when used in plant breeding.

**1 Associated side effects.** Mutations induced by mutagens are very diverse in nature. Invariably, the mutagen kills some cells outright while surviving plants display a wide range of deformities. Even those plants with the desirable mutations always inherit some undesirable side effects (akin to linkage drag, or is it a "mutation drag"?). Therefore, it is often necessary to transfer a new mutant into a stable genetic background, free of some of the associated

undesirable side effects such as sterility. To do that, the new mutant must be successively propagated for several generations and sometimes even crossed with other genotypes.

- 2 Large numbers of segregating populations are needed.** Another weakness is the need to produce and sort out a large number of segregating populations in order to have a good chance of finding a desirable mutant. Most mutations are deleterious or undesirable. Mutation breeding may be likened to finding a needle in a haystack. To sort through all the garbage, the plant breeder should have an easy method of screening the enormous variation. Morphological changes (e.g., shape, color) are easy to screen. However subtle changes require more definitive tests to evaluate and hence are more expensive to undertake.
- 3 Recessivity of mutants.** Most mutations are recessive and hence are observed only when the homozygous genotype occurs. This condition is readily satisfied in species that are naturally inbreeding. The situation practically excludes species that are polyploid, propagated clonally, or have a fertility-regulating mechanism (e.g., self-incompatibility) from being amenable to mutation breeding (except when dominant mutations are targeted).
- 4 Limited pre-existing genome.** The researcher cannot change what does not exist. Consequently, mutations can only be induced in existing genes.
- 5 Mutations are generally random events.** A modern biotechnology technique, site-directed mutagenesis, has been used to induce targeted mutations. However, conventional mutagenesis remains unpredictable and can not be directed to specific genes.

### Selected significant successes of mutation breeding

Mutation breeding has been used to improve a number of important crops (Table 12.3). The first commercial

**Table 12.3** Selected general areas of achievement in mutation breeding.

<b>Disease resistance:</b> e.g., <i>Verticillium</i> wilt resistance in peppermint, victorial blight resistance in barley, downy mildew resistance in pearl millet
<b>Modification of plant structure:</b> e.g., bush habit in dry bean, dwarf mutants in wheat and other cereals
<b>Nutritional quality augmentation:</b> e.g., opaque and floury endosperm mutants in maize
<b>Chemical composition alteration:</b> e.g., low erucic acid mutants of rape seed
<b>Male sterility:</b> for use in hybrid breeding in various crops
<b>Horticultural variants:</b> development of various floral mutants
<b>Breeding of asexually propagated species:</b> numerous species and traits
<b>Development of genetic stock:</b> various lines for breeding and research
<b>Development of earliness:</b> achieved in many species

cultivar derived from mutagenesis was “Chlorina”, a tobacco cultivar. It was developed with X-ray radiation in 1930. Since then, several hundreds of commercial cultivars or ornamentals, field crops, fruit crops, and other plant kinds have been released for production. These include barley cultivars “Pallas” and “Mari” that were developed by Swedish scientists. These genotypes, and other cultivars developed through hybridizations involving the two mutants, significantly impacted barley production in Denmark and Sweden. In the USA, “Pemrad” and “Luther” cultivars of barley were significant in the production of the crop in the 1970s. Dwarf cultivars of cereals have been developed by mutagenesis. Classic examples are “Norin 10” of the Green Revolution fame, and “Calrose 76” rice that strongly impacted California rice production.

### References and suggested reading

- Anon. 1991. Plant mutation breeding for crop improvement. Proceeding of the FAO/IAEA Symposium, Vienna, 1990. International Atomic Energy Agency, Vienna.
- Broertjes, C., and A.M. Van Harten. 1988. Applied mutation breeding for vegetatively propagated crops. Elsevier, Amsterdam.
- Micke, A. 1992. Fifty years induced mutations for improving disease resistance of crop plants. Mutation Breed. Newsl. 39:2–4. IAEA, Vienna.
- Peloquin, S.J. 1982. Meiotic mutants in potato breeding. Stadler Symp. 14:99–109.
- Pozniak, C.J., and P.J. Huel. 2004. Genetic analysis of imidazoline resistance in mutation-derived lines of common wheat. Crop Sci. 44:23–30.
- Rasmusson, D.C., and R.L. Phillips. 1997. Plant breeding progress and genetic diversity from de novo variation and elevated epistasis. Crop Sci. 37:303–310.

## Outcomes assessment

### Part A

Please answer the following question true or false:

- 1 Gamma rays are chemical mutagens.
- 2 A forward mutation converts a wild type to a mutant phenotype.
- 3 An acute application of a mutagen occurs over a short period of time.
- 4 EMS is a chemical mutagen.
- 5 Most mutations are recessive.

### Part B

Please answer the following questions:

- 1 Mutations may be spontaneous or .....
- 2 Mutagens may be chemical or .....
- 3 Applications of mutagens may be acute or .....
- 4 Mutations may be forward or .....
- 5 Give three examples of common mutagens.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss mutations according to the structural change produced.
- 2 Discuss the factors that impact the success of mutagenesis.
- 3 Discuss the importance of mutation in plant breeding.



# 13

# Polyplody in plant breeding

---

## Purpose and expected outcomes

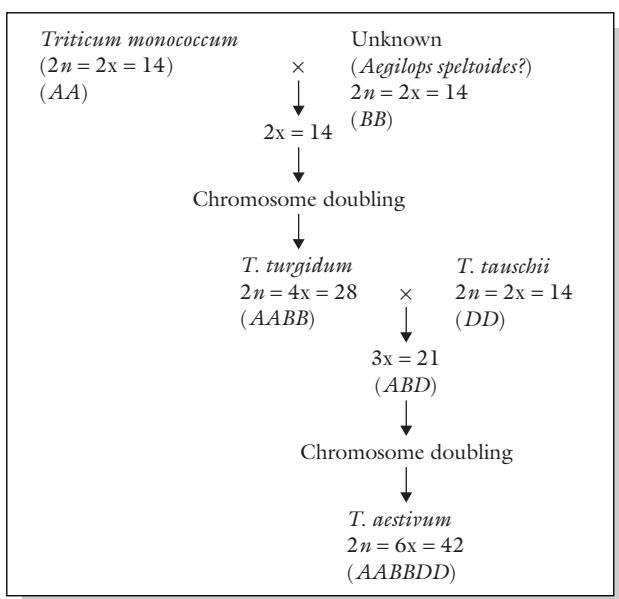
*Hybridization, as previously discussed, is a means of reorganizing genes from the parents involved in the cross in a new genetic matrix. Whereas the contents of the chromosomes may change because of the phenomenon of genetic recombination, normal hybridization does not alter the chromosome number of the species. However, certain natural processes can result in altered chromosome numbers. Similarly, the breeder may develop new variability by altering the number of chromosomes in the species through various processes. Furthermore, a number of the major crop species contain altered chromosome numbers. After studying this chapter, the student should be able to:*

- 1 Define the term polyplody.
  - 2 Discuss the variations in chromosome number in plants.
  - 3 Discuss the effects of polyploidy on plants.
  - 4 Discuss the importance of autopolyploidy in crop production.
  - 5 Discuss the genetics of autopolyploidy.
  - 6 Discuss the implications of autopolyploidy in plant breeding.
  - 7 Discuss the occurrence of allopolyploidy in nature.
  - 8 Discuss the genetics and breeding of allopolyploidy in plant breeding.
  - 9 Discuss the applications of aneuploidy.
- 

## Terminology

**Ploidy** refers to the number of copies of the entire chromosome set in a cell of an individual. The complete chromosome set is characteristic of, or basic to, a species (see Table 3.2). A set of chromosomes (the genome) is designated by “x”. Furthermore, the basic set is called the **monoploid** set. The **haploid number** ( $n$ ) is the number of chromosomes that occurs in gametes. This represents half the chromosome number in somatic cells, which is designated  $2n$ . A diploid species, such as corn, has  $n=10$  and  $2n=20$ . Also, a diploid species has  $2n=2x$  in its somatic cells, and  $n=x$  in its gametes. Some species have a higher ploidy, for example an autotetraploid (four basic sets of chromosomes) which has somatic cells

with  $2n=4x$  and gametes with  $n=2x$ . For corn, for example,  $2n=2x=20$ , while for wheat, a hexaploid with 42 chromosomes and a basic set of seven,  $2n=6x=42$ . Sometimes species that have more than two genomes comprise sets from different origins. To distinguish the source, each genome is designated by a different letter. For example, wheat has chromosome sets from three different origins and hence has a genetic designation (**genomic formula**) of *AABBDD* (Figure 13.1). To indicate the number of haploids derived from individuals of different ploidy levels for a single genome, a prefix is added to the term “haploid” to denote the number of sets ( $x$ ) of the basic genome present. For example a **monohaploid** ( $n=1x$ ) is derived from a diploid, while a **dihaploid** ( $n=2x$ ) is from a tetraploid, and so on.



**Figure 13.1** The proposed origin of common wheat *Triticum aestivum*.

In some species of higher plants, a pattern of ploidy emerges whereby the gametic (haploid) and somatic (diploid) chromosome numbers increase in an arithmetic progression, as illustrated by oat and wheat

(Table 13.1). The set of species displaying this pattern constitute a **polyploid series**.

### Variation in chromosome number

In nature, there exist two types of variation in chromosome number. In one type, called **euploidy**, the individuals contain multiples of the complete set of chromosomes that is characteristic of the species (the basic number,  $x$ ). In another, called **aneuploidy**, individuals contain incomplete sets of chromosomes that may be equivalent to the euploid number plus or minus one or more specific chromosomes (Table 13.2). The state of having multiples of the basic set in the somatic cell in excess of the diploid number is called **polyploidy**, and the individuals with such cells, **polyploids**. Polyploids are euploids. When euploids comprise multiples of the genome (i.e., duplicates of the genome from the same species) they are called **autopolyploids** and the condition **autopolyploidy** (or **autopolyploidy**). However, when a combination of genomes from different species are involved, the term **allopolyploid** (and similarly, **allopolyploidy** or **allopolyploidy**) is used. Alternatively, the term **amphiploid** or **amphidiploid** (and similarly, **amphiploidy** or **amphidiploidy**) is also used to describe polyploids with

**Table 13.1** Polyploid series of selected species.

Ploidy	Oat ( <i>Avena</i> spp.)	Wheat ( <i>Triticum</i> spp.)
Diploid ( $2n = 2x = 14$ )	<i>A. brevis</i> (short oat) <i>A. strigosa</i> (sand oat)	<i>T. monococcum</i> (einkorn) <i>T. tauschii</i> (wild oat)
Tetraploid ( $2n = 4x = 28$ )	<i>A. barbata</i> (slender oat) <i>A. abyssinica</i> (Abyssinia oat)	<i>T. timopheevii</i> (wild) <i>T. turgidum</i> (emmer)
Hexaploid ( $2n = 6x = 42$ )	<i>A. sativa</i> (common oat) <i>A. byzantina</i> (red oat)	<i>T. aestivum</i> (common bread wheat)

**Table 13.2** Classification of polyploidy.

Ploidy	Genome	Description
Diploidy	<i>AA BB</i>	Contains two of a basic chromosome set
Euploidy		
Autopolyploidy	<i>AAA BBBB</i>	Multiples of a basic set ( $n$ ) of one specific genome
Allotriploid	<i>AAB AABB</i>	Multiples of the basic number but of different genomes
Segmental allotriploid	<i>AA'B AB'B'</i>	Multiples of the basic number but the genomes have similar parts
Aneuploidy	<i>AA</i>	$2n \pm 1, 2, \dots, k$

**Table 13.3** Naming of polyploids.

	<b>Genome formula</b>	<b>General name</b>	<b>Specific name</b>
$n$	$A$	Haploid (monoploid)	
$2n$	$AA$	Diploid	
$3n$	$AAA$	Triploid	Autotriploid
	$AAB$	Triploid	Allotriploid
$4n$	$AAAA$	Tetraploid	Autotetraploid
	$AABB$	Tetraploid	Allotetraploid
$6n$	$AAAAAA$	Hexaploid	Autohexaploid
	$AABBDD$	Hexaploid	Allohexaploid

different genomes. It should be pointed out that autopoloidy and allopoloidy are extreme forms of polyploidy. Intermediates occur between them on a continuum of genomic relationships. C. L. Stebbins called the intermediates **segmental allopoloids**. Polyploids are named such that the prefix to the standard suffix (ploid) refers to the basic chromosome set (Table 13.3). For example “triploid” refers to a cell with three genomes (3x) while “hexaploid” refers to a cell with six genomes (6x).

### General effects of polyploidy of plants

In terms of general morphology, an autopoloid would resemble the original parent whereas an allotopoloid would tend to exhibit a phenotype that is intermediate between its parental species. Autopoloidy increases cell size, especially in meristematic tissues. Autopoloids usually have thicker, broader, and shorter leaves. Other plant organs may increase in size compared to their corresponding parts in diploids, an effect called **gigas** features. The **gigas luxuriance** contributes more to moisture content of the plant parts than to biomass. The plants tend to be determinate in growth.

The growth rate of polyploids is less than that of diploids. This may be due to their lower auxin content than that of their diploid counterparts, as found for tetraploids. Polyploids tend to flower later and over a longer period of time. In grasses, autopoloidy tends to reduce branching or tillering.

Polyploidy also affects the chemical composition of plant parts. For example, vitamin A activity in tetraploid corn is about 40% more than in diploid species. Similarly, the vitamin C content of vegetables and fruits has been known to increase following chromosome doubling. The nicotine content of tetraploid tobacco is about 18–33% higher than in diploid species. Autopoloids, generally, have fertility problems, and have

poor pollen production. In some cases, reduction in fertility as compared to their diploid counterparts may be as high as 80–95%. This reduction in fertility is attributed to genetic imbalance following chromosome doubling that leads to disharmonies in development (e.g., abnormal pollen sac, failure of fertilization). Some changes in ecological requirements such as photoperiod and heat requirements have been reported in some species following chromosome doubling.

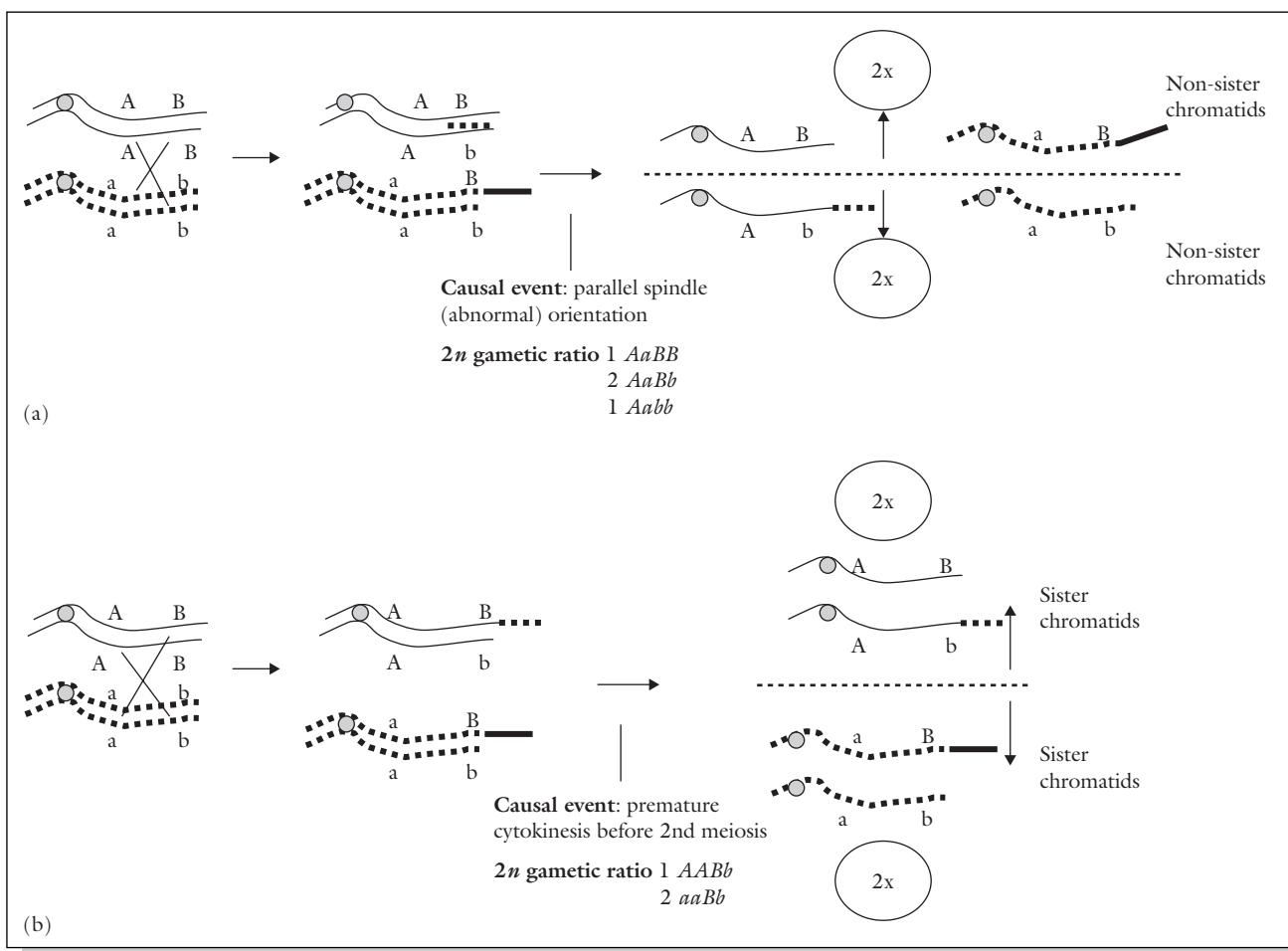
### Origin of polyploids

The breeding strategies employed in the breeding of polyploids are determined primarily on their origin. J. R. Harlan and J. M. de Wet (1975) concluded from an extensive review of the literature that nearly all polyploids arise by the path of unreduced gametes. They pointed out that the most common factor leading to polyploidy is the fusion of  $2n$  and  $n$  gametes to form a triploid, followed by either backcrossing or selfing to produce a tetraploid. Further, they observed that the occurrence of unreduced gametes is variable and pervasive in the plant kingdom.

The unreduced ( $2n$ ) gametes arise by one of two mechanisms – **first division restitution (FDR)** or **second division restitution (SDR)** – during meiosis (Figure 13.2). Each mechanism has a different genetic consequence. In FDR, the  $2n$  gametes result from parallel spindle formation after the normal first division of meiosis. The cleavage furrows occur across the plane of the parallel spindles, producing dyads and  $2 \times 2n$  pollen. The genetic consequence of the mechanism is that most of the heterozygosity of the diploid hybrid is conserved in the  $2n$  gametes. In the SDR mechanism, the first meiotic division is followed by cytokinesis, but the second division is absent. This results in a dyad with  $2 \times 2n$  gametes. However, in terms of genetic consequence, SDR results in significantly reduced heterozygosity in the  $2n$  gametes. Researchers such as T. Bingham have proposed, in breeding potatoes, the fusion of two FDR  $2n$  gametes to harness the heterosis that results. This heterosis can be fixed; the elite lines produced will then be clonally propagated. Seed-propagated species (e.g., alfalfa) cannot benefit from this strategy.

### Autopoloidy

As previously defined, autopoloids comprise duplicates of the same genome. Autopoloids are useful in making allopoloids and wide crosses.



**Figure 13.2** The origin of polyploidy by (a) first division restitution (FDR) and (b) second division restitution (SDR). Dyads occur in telophase II. FDR is caused by the presence of parallel or fused spindles, while SDR is caused by the presence of a cell plate before anaphase II.  $2n$  pollen tend to be bigger in size than  $1n$  pollen.

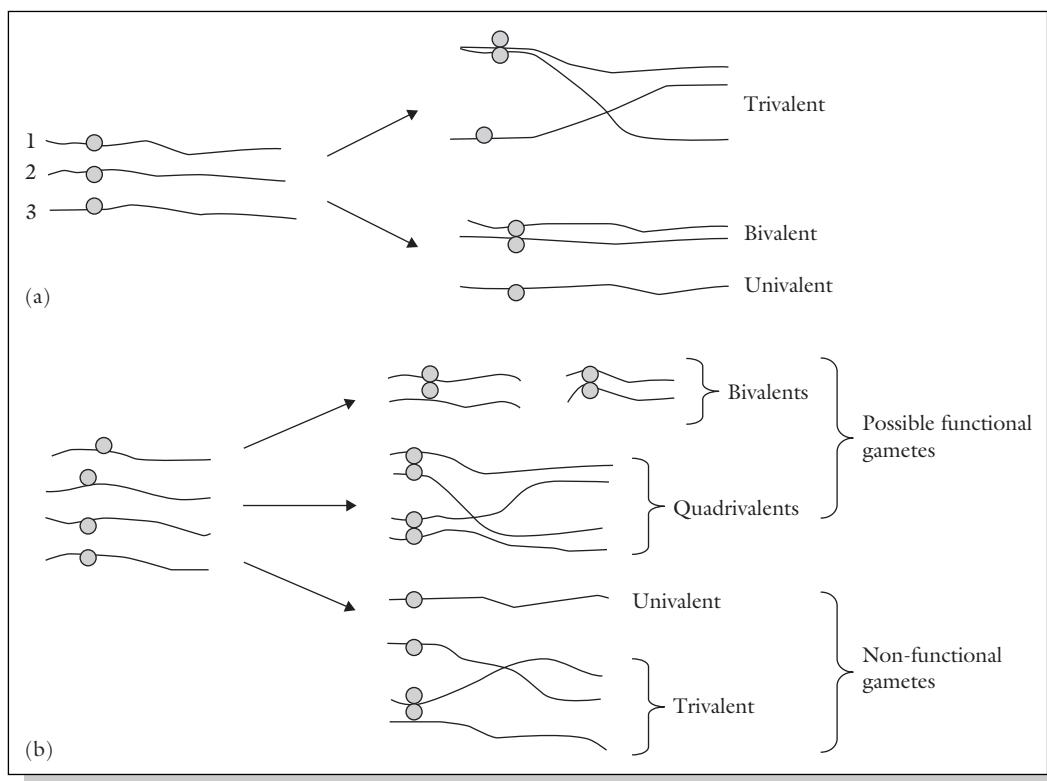
### Natural autopoloids of commercial importance

Autopoloidy is not known to have profoundly impacted the evolution of species. Having increased sets of chromosomes does not necessarily increase performance. Autopoloids of commercial value include banana, a triploid, which is seedless (diploid bananas have hard seeds not desirable in production for food). Other important autopoloids are tetraploid crops such as alfalfa, peanut, potato, and coffee. Spontaneous autopoloids are very important in the horticultural industry where the gigas feature has produced superior varieties of flowering ornamentals of narcissus, tulip, hyacinth, gladiolus, and dahlia among others. Autopoloid red clovers and ryegrasses with lusher and larger leaves,

taking advantage of the gigas feature of polyploidy, have been bred for commercial use as palatable and digestible livestock forage. It should be mentioned that there is no overwhelming evidence to suggest that autotetraploids are productively superior to their diploid counterparts.

### Cytology of autopoloids

Autopoloids contain more than two homologous chromosomes. Consequently, instead of forming bivalents during meiosis as in diploids, there are also multivalents (Figure 13.3). For example, autopoloids have mostly trivalents but some bivalents and univalents are also present. Tetraploids have quadrivalents or bivalents as well



**Figure 13.3** Cytology of polyploids: (a) triploidy and (b) autotetraploidy. Bivalents and quadrivalents usually produce functional gametes, while univalents and trivalents produce sterile gametes.

as some trivalents and univalents. These meiotic abnormalities are implicated in sterility to some extent, more so in triploids. The microspores and megasporangia with  $x$  or  $2x$  genomes are usually viable.

The amount and nature of chromosome pairing directly impacts the breeding behavior of autopolyploids. Autopolyploids are induced artificially by chromosome doubling using colchicine. Doubling a hybrid between two diploid cultivars would produce a tetraploid in which there may be a tendency for the doubled set of chromosomes from one parent to pair independently of the doubled set of chromosomes of the other parent. This propensity is called **preferential or selective pairing**, a phenomenon with genetic consequences. If preferential pairing is complete, there would be no new genetic recombination and hence the progeny would look like the doubled  $F_1$ . Furthermore, the bivalent pairing would contribute to sterility originating from meiotic disorders, while preserving heterosis indefinitely, should there be any produced by the original cross. The concept of preferential pairing is applied in the modern breeding of polyploids whereby allopolyploids are stabilized

and made reliable as diploids, a process called **diploidization**.

#### Genetics of autopolyploids

The ploidy level may also be defined as the number of different alleles that an individual can possess for a single locus on a chromosome. A diploid can have two alleles per locus, whereas an autotetraploid can have four different alleles. The genetics of autopolyploids is complicated by multiallelism and multivalent association of chromosomes during meiosis. Consider the segregation of alleles of a single locus ( $A$ ,  $a$ ). In a diploid species, there would be three possible genotypes  $AA$ ,  $Aa$ , and  $aa$ . However, in an autotetraploid there would be five genotypes ranging from nulliplex ( $aaaa$ ) to quadruplex ( $AAAA$ ) (Table 13.4). The proportion of dominant ( $A$ ) to recessive ( $a$ ) genes is different in two of the five genotypes ( $AAAA$  and  $Aaaa$ ) in autotetraploids from that which obtains in diploids. The number of phenotypes observed depends on the dominance relationship of  $A$  and  $a$ . If allele  $A$  is completely dominant

**Table 13.4** Genetics of autopolyploids.

Diploid	Polyplody	Name
Cross $Aa \times Aa$	$AAaa \times AAaa$	
Products		
1/4 $AA$	1/36 $AAAA$	Quadruplex
2/4 $Aa$	8/36 $AAAa$	Triplex
1/4 $aa$	18/36 $AAaa$ 8/36 $Aaaa$ 1/36 $aaaa$	Duplex Simplex Nulliplex

**Table 13.5** Genetic frequencies following chromosome segregation of an autotetraploid.

Genotype	Gametic frequency		
	$AA$	$Aa$	$aa$
$AAAA$	1	0	0
$AAAa$	1/2	1/2	0
$AAaa$	1/6	4/6	1/6
$Aaaa$	0	1/2	1/2
$aaaa$	0	0	1

Note: chromatid segregation occurs less frequently than chromosomes segregation and produces alternative types of segregation. For example the simplex ( $Aaaa$ ) can produce gametes that are homozygous ( $AA$ ) by the process called double reduction.

Assuming complete dominance and chromosome segregation the following phenotypic ratios are observed. Certain segregation ratios are sometimes indicative of the nature of autotetraploid inheritance.

Cross	Progeny (dominant : recessive)
$AAAA \times AAAA$	1 : 0
$AAAa \times AAAa$	1 : 0
$AAaa \times AAaa$	35 : 1
$AAaa \times Aaaa$	11 : 1
$AAaa \times aaaa$	5 : 1
$Aaaa \times Aaaa$	3 : 1
$Aaaa \times aaaa$	1 : 1
$aaaa \times aaaa$	0 : 1

to allele  $a$ , there would be only two phenotypes. If dominance is incomplete or the effect of allele  $A$  is cumulative, there could be up to five phenotypes. Upon selfing, a dominant phenotype in a diploid ( $AA$ ,  $Aa$ ) would produce a progeny that is all dominant, or

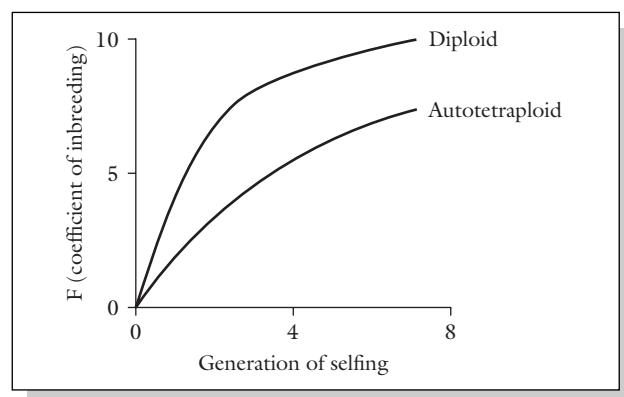
**Table 13.6** Multiple alleleism in autotetraploids.

Tetrasomic condition	1st	2nd	3rd	Total
$a_1a_2a_3a_4$	6	4	1	11
$a_1a_2a_2a_3$	3	1	0	4
$a_1a_1a_2a_2$	1	0	0	1
$a_1a_1a_1a_2$	1	0	0	1
$a_1a_1a_1a_1$	0	0	0	0

segregate in the 3 : 1 ratio. Selfing each of the five categories would produce many different outcomes in autotetraploids, assuming random chromosome segregation (Table 13.5).

An autopolyploid individual can have up to four alleles ( $abcd$ ) per locus. Five different genotype categories are similarly possible except that there may be only four nulliplex genotypes ( $aaaa$ ,  $bbbb$ ,  $cccc$ ,  $dddd$ ) and only one tetra-genic genotype ( $abcd$ ), but numerous combinations for the intermediates (Table 13.6). The possible gametic array is shown for each genotype. Interallelic and intra-allelic interactions may occur for as many as four alleles per locus in an autotetraploid. The degree to which intra-allelic interaction occurs determines the expression of heterosis and inbreeding depression in an autotetraploid. Because four identical alleles are required to achieve homozygosity in an autotetraploid compared with only two in a diploid, homozygosity is achieved at a less rapid rate in autotetraploids (Figure 13.4).

Another aspect of autopolyploid genetics with a plant breeding implication, is the difficulty of distinguishing between a triplex and a quadruplex on the basis of a progeny test (assuming random chromosome segregation). Both genotypes ( $AAAA$  and  $AAAa$ ) will breed true for the dominant allele. To identify a triplex plant, the breeder would have to advance the progeny one more generation to identify the duplex plants of the  $S_1$ . Achieving genetic purity in autotetraploid stocks is



**Figure 13.4** The effect of ploidy on the inbreeding coefficient as demonstrated by diploids and autotetraploids.

difficult, not only because it is challenging to identify triplex plants, but also because deleterious genes may persist in an autotetraploid, manifesting themselves only rarely in the homozygous genotype. The breeder would need an additional two generations in order to identify the homozygous-dominant genotype unequivocally.

### Induction of autopoloids

Plant breeders were initially attracted to induce polyploidy primarily because of the gigas effects, which increased cell size (but also reduced fertility). These pros and cons of the gigas effects make the induction of autopoloids more suited to crops whose economic part is vegetative. The primary technique for inducing autopoloids is the use of colchicine ( $C_{22}H_{25}O_6$ ), an alkaloid from the autumn crocus (*Colchicum autumnale*). This chemical compound works by disrupting the spindle mechanism in mitosis, thereby preventing the migration of duplicate chromosomes to opposite poles at anaphase. Consequently, the nucleus is reconstituted with twice the normal number of chromosomes, without any nuclear or cell division.

Ryegrass is one of the species that has been successfully improved by the induction of autopoloidy. Rye (*Secale cereale*) is perhaps the only grain-producing crop for which synthetic autopoloids have been developed. Meristematic tissue is most susceptible to colchicine treatment. Hence a germinating seed, a young seedling, or a developing bud, are the commonly used plant material for autopoloid induction. The chemical may be applied in aqueous solution or through various media (e.g., agar lanolin paste). Seeds may be soaked in

aqueous colchicine at a concentration of 0.05–0.4% for 30 minutes to 3 hours. Buds are treated differently, for example, by intermittently exposing the selected plant material for 2–6 days at concentrations of 0.2–0.5%. The breeder should determine the best treatment condition by experimentation. The material treated should be thoroughly washed after application to remove excess chemicals.

### Breeding autopoloids

In developing and using autopoloids in plant breeding, certain general guidelines may be observed.

- 1 Generally, species tend to have an optimum chromosome number (optimum ploidy number) at which they perform best. Because chromosome doubling instantly and drastically increases chromosome number, selecting parents with a low chromosome number for autopoloid breeding would reduce the risk of meiotic complications that are often associated with large chromosome numbers. This would increase the chance of obtaining fertile autopoloids.
- 2 Autopoloids tend to have gigas features and a high rate of infertility. Consequently, autopoloidy is more useful for breeding species in which the economic product is not seed or grain (e.g., forage crops, vegetables, ornamental flowers).
- 3 Producing autopoloids from cross-fertilizing species promotes gene recombination among the polyploids, with a better chance of obtaining a balanced genotype.

D. R. Dewey summarized the properties of a species suited for the induction of polyploidy as follows:

- 1 The species has low chromosome number.
- 2 The economic part of the plant is the vegetative material (e.g., forage grasses).
- 3 The plant is cross-pollinated (allogamous).
- 4 The plant is perennial in habit.
- 5 The plant has the ability to reproduce vegetatively.

### Autotetraploids and autotriploids

Tetraploid rye ( $2n = 4x = 28$ ) has about 2% more protein than diploid cultures and has superior baking qualities. However, it also has about 20% higher incidence of sterility per spike, resulting in lower grain yield than diploids. Autotriploids of commercial importance include sugar beet ( $2n = 2x = 18$ ;  $2n = 3x = 27$ ,  $2n = 4x = 36$ ). Triploidy is associated with the genetic consequences

of sterility because of the odd chromosome number that results in irregular meiosis. The sterility favors species that are grown for vegetative commercial parts (e.g., grasses) and ornamentals and fruits (seedless). In sugar beet, triploid cultivars of monogerm types have significantly impacted the sugar industry.

Triploid hybrids are produced by crossing diploids with tetraploids. Breeders use three kinds of genotypes. The diploid is male-sterile (female, cytoplasmic male-sterile), while the tetraploid is the pollinator. The third component is a male-sterility maintainer (a diploid,  $N$ ). The tetraploid is derived from a diploid by colchicine treatment of the seed (soak in 0.2% for 15 hours at 30°C). Seedless watermelon ( $3x = 33$ ) is also produced by crossing a diploid ( $2n = 2x = 22$ ) with a tetraploid ( $2n = 4x = 44$ ).

### Natural allopoloids

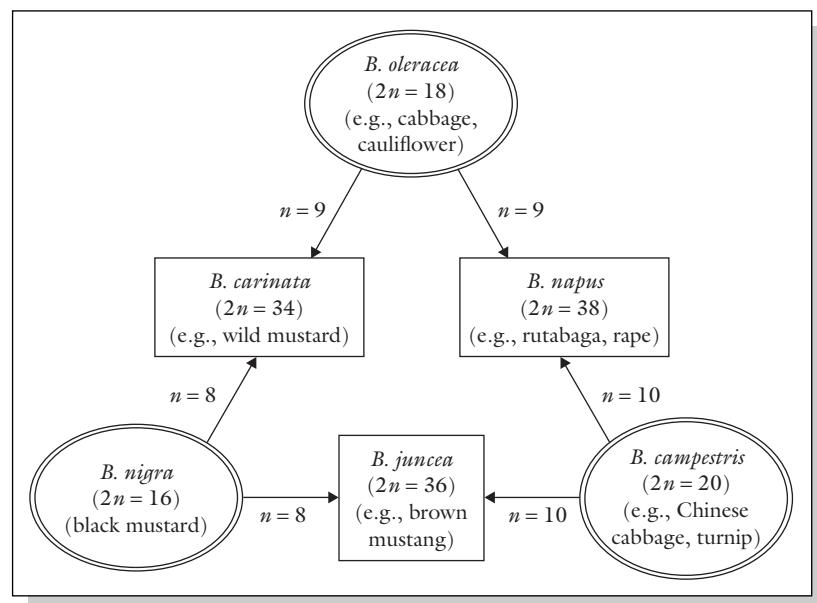
A number of economically important crops are allopoloids. These include food crops (e.g., wheat, oat), industrial crops (e.g., tobacco, cotton, sugarcane), and fruits crops (e.g., strawberry, blueberry). These crops, by definition, contain a combination of different genomes. Researchers over the years have attempted to elucidate the ancestral origin of some allopoloids. One of the most widely known successes was the work of

Nagaharu U, the Japanese scientist who described the genomic relationships among naturally occurring mustard (*Brassica*) species (Figure 13.5). Dubbed the **triangle of U**, it describes the origins of three *Brassica* species by allopoloidy. The diploid species involved are turnip or Chinese cabbage (*B. campestris*,  $n = 10$ ), cabbage or kale (*B. oleracea*,  $n = 9$ ), and black mustard (*B. nigra*,  $n = 8$ ). For example, rutabaga (*B. napus*) has  $2n = 38$ , being a natural amphiploid of *B. oleracea* and *B. campestris*.

In cereal crops, wheat is a widely studied allotetraploid that comprises genomes from three species. Cultivated common wheat (*Triticum aestivum*) is a hexaploid with 21 pairs of chromosomes and is designated *AABBDD*. The *AA* genome comes from einkorn (*T. monococcum*). Tetraploid wheats have the genomic formula *AABB*. Emmer wheat (*T. dicoccum*) crossed naturally with *Aegilops squarrosa* (*DD*) to form common wheat.

### Genetics of allopoloids

As previously indicated, allopoloids arise from the combination and subsequent doubling of different genomes, a cytological event called **allopoloidy**. The genomes that are combined differ in degrees of homology, some being close enough to pair with each other, whereas others are too divergent to pair. Sometimes, only segments of the chromosomes of the component genomes are different, a condition that is called **segmental allopoloidy**. Some of



**Figure 13.5** The triangle of U showing the origins of various allopoloids in *Brassica*.

the chromosomes of one genome may share a function in common with some chromosomes in a different genome. Chromosomes from two genomes are said to be **homeologous** when they are similar but not **homologous** (identical).

Most allopoloids have evolved certain genetic systems that ensure that pairing occurs between chromosomes of the same genome. A classic example occurs in wheat ( $2n = 6x = 42$ ) in which a gene on chromosome 5B, designated *Pb*, enforces this diploid-like paring within genomes of the allopoloid. When this gene is absent, pairing between homeologous chromosomes, as well as corresponding chromosomes of the three genomes,

occurs, resulting in the formation of multivalents at meiosis I.

Allopoloids exhibit a variety of meiotic features. Sometimes chromosomes pair as bivalents and thereby produce disomic ratios. Where the component genomes have genes in common, duplicate factor ratios will emerge from meiosis, an event that sometimes is an indication of allopoloid origin of the species. Whereas significant duplications of genetic material have been found in wheat, the genomes of upland cotton have little duplication. Tetrasomic ratios are expected for some loci where multivalent associations are found in allotetraploids.



## Industry highlights

### *Application of tissue culture for tall wheatgrass improvement*

Kanyand Matand and George Acquaah

Department of Agriculture and Applied Sciences, Langston University, Langston, OK 73050, USA

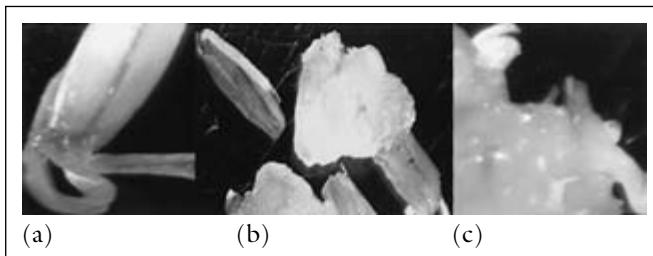
#### **Introduction**

Cultivar "Jose" of tall wheatgrass (*Thinopyrum ponticum* Posp.) is a perennial, cross-pollinating, bunchgrass, used as a cool-season livestock forage. It is adapted to the Oklahoma environment primarily because of its relative ease-to-establish and tolerance of a wide range of soil pH. This bunchgrass can produce high levels of dry matter per season (Coleman 1999) and exhibit superior characteristics such as high biomass production, persistence through extended periods of grazing, and saline, flood, and drought tolerance (Moser et al. 1996; Coleman 1999; Redmon 1999). Although used successfully to improve plants for many years, limitations associated with traditional breeding approaches coupled with the large genome of tall wheatgrass ( $2n = 10x = 70$ ) makes it improbable and difficult to improve this crop by relying on traditional breeding methods alone (Kindiger 2002). Because tissue culture is a prerequisite for the successful integration of classic plant breeding and genetic engineering for genetic crop improvement, researchers at Langston University, Oklahoma, and the US Department of Agriculture–Agricultural Research Service (USDA-ARS) Grazinglands Research Laboratory at El Reno, Oklahoma, collaborated to develop an efficient *in vitro* method for propagating this crop as an additional tool for its improvement. The discussion in this box focuses primarily on a one-step, reliable approach to microclone tall wheatgrass for germplasm development for its improvement; it also highlights cautions and pitfalls associated with the proposed method in particular, and tissue culture in general. This one-step method may be adapted for tissue culture research in other forage species.

#### **Basic approaches to plant regeneration**

There are a variety of ways to regenerate plants *in vitro*, but all of them can be grouped into two basic approaches, direct and indirect.

- 1 **Direct plant regeneration.** Direct plant regeneration occurs without an intervening callus phase. Potential growth focal structures consist of apical and axillary meristems, nodes, leaves, stem, roots, zygotic embryos, cotyledons, the whole seed, and other structures. These organs can be excised from the mother plant to induce somatic embryos or adventitious shoots, or both, without an intervening callus phase.
- 2 **Indirect plant regeneration.** This process is characteristic of the callus phase mediation for somatic plant formation from explants. Explants similar to those listed for direct plant regeneration can also be used to induce somatic embryos or adventitious shoots, or both, via callus formation.



**Figure 1** (a) Germinated seed ( $\times 50$ ), (b) callusing seeds ( $\times 50$ ), and (c) multiple shoot-forming callus ( $\times 50$ ).

soft, young seeds are easily bruised during hulling, thereby allowing the alcohol to be absorbed into the tissues, with lethal consequences.

Unlike cytokinin, auxin may be used alone to induce a callus, or in combination with cytokinin. Although treatment combinations of kinetin (KIN) or 6-benzylaminopurine (BA) and  $\alpha$ -naphthalene acetic acid (NAA) or indole-3-acetic acid (IAA) were used to induce healthy callus from the seeds, the greatest amount of friable callus was obtained by using treatment combinations of equal concentration (2–20  $\mu\text{M}$ ) of kinetin or BA and 2,4-dichlorophenoxyacetic acid (2,4-D) (Table 1). Further, the amount of callus formed increased with the growth regulator concentration. It was discovered that 2,4-D could also be successfully used alone to induce friable callus, although in smaller amounts. Callus formation may occur prior to, or after, seed germination, but often both occurred simultaneously. It was also observed that “organogenic” callus formation potential decreased with seed germination capacity. Seeds that had lost their germination capacity could also form callus; however, such callus growth was only temporary, and it could not form either friable callus or organogenic structures. This may suggest the existence of a correlation between “totipotency” of seed cells (cotyledonary or embryo) and zygotic embryo viability. When Murashige and Skoog (MS) (1962) and Gresshoff and Doy (GD) (1974) nutrient media were evaluated, the MS medium caused the greater rate of callusing of seeds (86% and 88%) over 4–6 weeks of culturing (Table 1). This rate could be improved to 100% by using freshly harvested seeds and/or including about 2% wood charcoal. The effect of darkness was later evaluated and found not to be a major factor for callus formation. Cultures grew well when cultured at 26–28°C under an 8-hour photoperiod. The morphogenetic growth responses were enhanced when the nutrient media were supplemented with 3% sucrose and 500 mg/l magnesium chloride. The pH of the media was adjusted to 5.7 with potassium hydroxide prior to autoclaving. All attempts to induce callus in leaf and root explants were unsuccessful. However, under unspecified conditions callus formed sporadically from the tips of non-severed shoots and roots. When such callus was collected and subcultured on fresh media, it formed normal shoots.

**Table 1** Organogenic responses of tall wheatgrass tissues to the combinations of 20  $\mu\text{M}$  6-benzylaminopurine (BA) or kinetin to 20  $\mu\text{M}$  auxins on Murashige and Skoog (MS) or Gresshoff and Doy (GD) media.

<b>Cytokinin</b>	% Seed callusing		% Shoot-forming callus		Shoot average per callus		
	<b>GD</b>	<b>MS</b>	<b>GD</b>	<b>MS</b>	<b>GD</b>	<b>MS</b>	<b>Auxin</b>
Kinetin	86	88	82	90	8.2	9.2	2,4-D
	14	15	10	13	4.4	6.1	NAA
	3	4	1	3	2.2	3.2	IAA
BA	79	86	78	84	6.2	8.3	2,4-D
	11	10	5	6	3.0	6.2	NAA
	4	5	2	3	2.1	2.4	IAA
Standard error least squares mean	1.57	1.57	1.64	1.64	0.42	0.42	

Coefficient of variation (CV): 30.70%, seed callusing; CV: 25.86%, shoot-forming callus; CV: 18.33%, average shoots per callus. 2,4-D, 2,4-dichlorophenoxyacetic acid; IAA, indole-3-acetic acid; NAA,  $\alpha$ -naphthalene acetic acid.

### Callus formation

Sterile, husked, dry, mature seeds were used as an explant organ for callus induction in tall wheatgrass (Figure 1a, b). Although grass hulled seeds could be sterilized using standard sterilization methods, tall wheatgrass hulled seeds were surface-sterilized with 10% sodium hypochlorite for 15 minutes and 70% ethanol for 5 minutes, and rinsed three to four times in distilled water. The sterilization method may vary depending upon the physiological state of the seed. Dry mature seeds are the most suitable for this sterilization method; alcohol exposure should be reduced to 1–2 minutes or eliminated from the protocol when using immature seeds. This is necessary because the alcohol to be absorbed into the tissues, with lethal consequences.

### ***Microcloning of tall wheatgrass for breeding***

Within 6 weeks of culturing sterile, hulled seeds of tall wheatgrass on MS or GD media to which kinetin or BA and 2,4-D, NAA, or IAA were added, adventitious shoot development could be observed from the callus (Figure 1c), irrespective of whether the callus originated from germinated or non-germinated seeds. Generally, shoot development was preceded by callus greening, which was a sign of meristem development. Such a physical manifestation has precedents in the first report on micropropagation of tall wheatgrass (Kindiger 2002) and reports on other crops such as cassava (Matand et al. 2004) and peanut (Matand et al. 1994). It has been previously reported that such plants, or other indirectly formed shoots, could result from a mono- or multicellular organogenetic process (George 1993). In this report, light microscopy observations suggested that most, if not all, tall wheatgrass shoots that were observed resulted from an aggregation of several adjacent callus cells. Single and multiple adventitious shoots were observed on both MS and GD semisolid media (Figure 1c).

As in the case of callus formation, treatments containing 2,4-D in combination with either kinetin or BA induced a significantly greater amount of shoots (9.2 or 8.3 shoots per callus on MS; and 8.2 or 6.2 shoots per callus on GD) and percentage of shoot formation (90% or 84% on MS; and 82% or 78% on GD) than those induced with either IAA or NAA combined with similar levels of cytokinin (Table 1). Overall, kinetin outperformed BA for morphogenetic responses; however, morphogenetic responses observed on the MS basal medium were generally comparable to those observed on the GD basal medium.

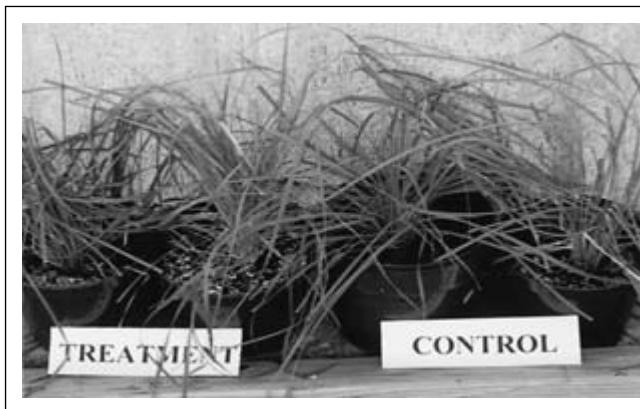
One of the main contributions of this report is that it delineates, for the first time, a one-step method of microcloning for the improvement of tall wheatgrass. In the first report on tissue culture of tall wheatgrass, Kindiger (2002) depicted a three-treatment microcloning approach corresponding to the three phases of callus, shoot, and root induction, which is a common practice in tissue culture of other crop species such as cassava as described by Matand et al. (2004). Callus was induced from seeds in the dark, on MS medium containing basal salts, 3% sucrose, 0.5 g/l casein, 8 g/l phytigel, and 0.005 g/l 2,4-D and benzylaminopurine (BAP). Then, callus was transferred to the shoot-inducing medium containing an increased concentration of sugar (5%), a reduced concentration of phytigel (5 g/l), a growth regulator combination treatment with a lower concentration of BAP (0.001 g/l) and a similar concentration of NAA (0.005 g/l) without 2,4-D and casein. Lastly, shoots were rooted on half-strength MS medium containing decreased concentrations of sucrose (1.5%) and 2,4-D (0.75 mg/l), an addition of IAA (3.75 mg/l) and kinetin (1.075 mg/l), and an increased concentration of phytigel (10 g/l) without casein. The present report expounds on a related but one-step microcloning approach in which all the three morphogenetic phases including callus, shoot, and root formation were controlled under a single treatment. Not only could it reduce a long culturing period, but it could also eliminate the cost of increased concentrations of, and/or additional, chemicals.

However, when producing *in vitro* plants through indirect adventitious shoot formation, it should be borne in mind that there is a potential for indefinite rejuvenation of shooting callus. Maintaining such a property requires that the callus be subcultured in relatively mid-sized amounts. Frequent break-offs of callus may enhance the wounding effect that normally stimulates callus formation in plant tissues in nature. This induces plant wound healing that is underlain by intense cell division. After the callus has formed shoots, it is necessary to frequently break them off about every 2–3 weeks to stimulate rapid shoot growth. Further, it was observed that a cluster of young shoots would grow uniformly, with new shoot formation at their base, as long as none of the shootlets significantly outgrew the others. Based upon our laboratory experience, it was surmised that any more rapidly growing shoot might exert an inhibitory effect on its immediate surrounding shootlets, similar to the apical dominance. This was further corroborated by the finding that the removal of the tallest shootlet of the cluster resulted in a rapid elongation of all the surrounding shootlets and an initiation of new shoot development on the same callus.

### **Shoot rooting**

Generally, *in vitro*-induced shoots of appropriate height, according to the desires of the researcher, may be rooted in a variety of ways: (i) shoots may be transferred to fresh basal media; (ii) shoots may be transferred to fresh media with different growth regulator treatments, primarily auxin alone, or with a trace amount of cytokinin; and (iii) shoots may be repeatedly transferred onto fresh shoot-inducing media. When root-inducing treatment media are different from shoot-inducing treatment media, it is recommended that the concentrations of growth regulators be judiciously decreased. As is the experience of other scientists working under other experimental conditions, it is strongly recommended that root-inducing media be devoid of cytokinins but have a low concentration of a single auxin. However, because each species may respond differently to similar treatments, it is recommended that researchers apply their own best judgment based upon their knowledge of the crop.

*In vitro* adventitious shoots that were formed in tall wheatgrass were initially rooted in media devoid of growth regulators. However, it was further observed that shoot-inducing media could also easily induce roots by extending the incubation of shoots eligible for rooting on shoot-inducing media by 1 or 2 weeks. Following the standard practice, *in vitro*-induced shoots are systematically acclimatized to the environmental conditions prior to their transfer into the greenhouse. During the course of our investigation with tall wheatgrass, it was observed that an *in vitro* rooting system could eventually supply nutrients to support plants in the greenhouse, when the acclimatization step was omitted. Thus, all *in vitro*-formed rooted plants were thereafter successfully transferred into the greenhouse directly from their test tubes without incurring any loss. Similar observations were also recently made on cassava (Matand et al. 2004). Plants regenerated *in vitro* developed normally in the greenhouse as well as in the field (Figure 2).



**Figure 2** *In vitro* regenerated plants and seed-germinated plants (control) growing in the greenhouse.

types can be observed in new plants while they are still growing in test tubes prior to being transferred to the greenhouse. This was the case in parallel studies on tall wheatgrass, which were conducted at the USDA-ARS Grazinglands Research Laboratory. White stripe mutations were observed as soon as sizable new plants developed from callus (Kindiger 2002). The cause of those mutations remains cryptic.

### Conclusion

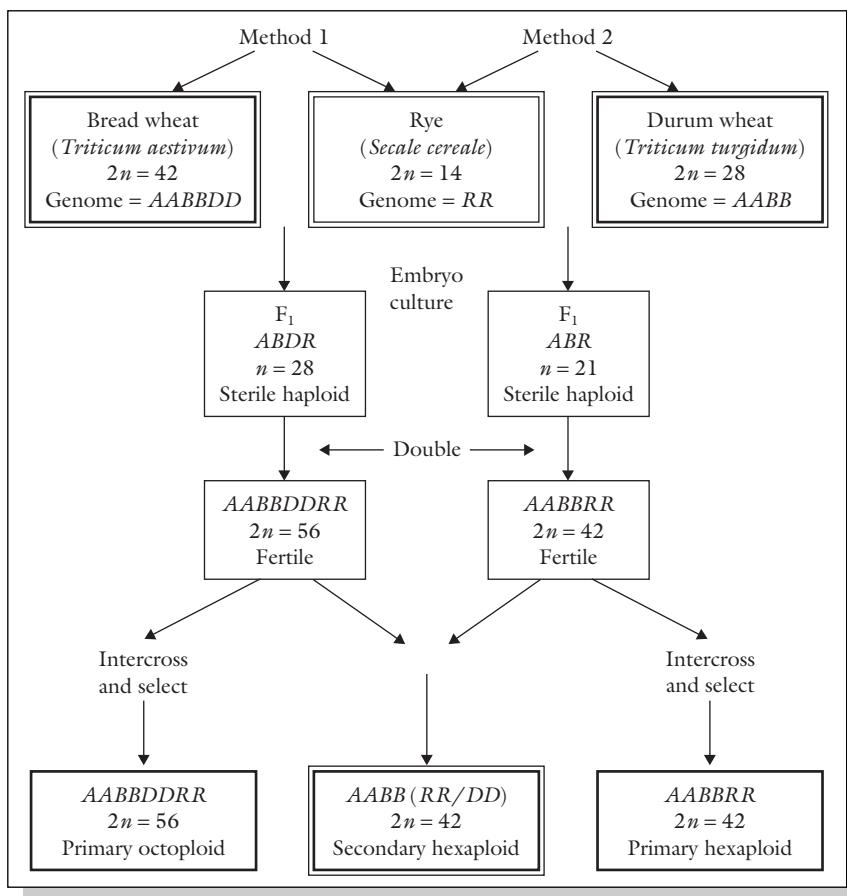
Classic breeding and clonal methods may be used to produce genetically uniform populations of plants. However, the purity of such populations is more difficult to maintain through seed than clonal multiplication. Although the two crop improvement methodologies may be used in complementary fashion in a single breeding program, the manner in which they are used may depend on the breeding objectives and the species, among other factors. In the improvement program of tall wheatgrass, outlined above, *in vitro* cloning was resorted to in the earlier stages of the improvement program. Seeds from agronomically selected individual plants were used to initiate cloning by indirect regeneration. The selection of seed as the culture-initiating explant was based upon our experience with the crop. Resulting clones were identical to their parents and were cataloged as germplasm for the breeding program. A single treatment was used to initiate callus, shoots, and roots. All attempts to initiate callus from leaf and root tissues were unsuccessful.

### References

- Coleman, S. 1999. Performance and management of cool-season perennial grasses in the Southern Plains Regions. USDA-ARS, El Reno, OK.
- George, E.F. 1993. Plant propagation by tissue culture, Part I – The technology. Exegetics Ltd, Edington, UK. 574 pp.
- Greshoff, P.M., and C.H. Doy. 1974. Development and differentiation of haploid *Lycopersicum esculentum* (tomato). Plant 107:161–170.
- Kindiger, B. 2002. Callus induction and plant regeneration in tall wheatgrass (*Thinopyrum ponticum* (Podp.) Barkw. & D.R. Dewey). Grassland Sci. 48(4):362–365.
- Matand, K., G. Acquaah, and M. Burns. 2004. A one-step *in vitro* procedure for induction of callus and shoots in cassava (*Manihot esculenta* Crantz). Plant Growth Reg. Soc. Am. 32(4):124–131.
- Matand, K., A.D. Porobo, and C.S. Prakash. 1994. Thidiazuron promotes high frequency regeneration system for peanut (*Arachis hypogaea*, L.) plants *in vitro*. Plant Cell Rep. 14:1–5.
- Moser, L.E., D.R. Buxton, and M.D. Casler (eds). 1996. Cool-season forage grasses. Monograph No. 34. American Society of Agronomy, Madison, WI.
- Murashige, T., and F. Skoog. 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. 15:473–479.
- Redmon, L. 1999. Cool-season perennial grasses in Oklahoma. Production Technology Bulletin No. 11(1). Agronomy Department, Oklahoma State University, Stillwater, OK.

### Nursery stock clones for a breeding program

After *in vitro* plants were adapted to greenhouse conditions for 2 weeks, they were used in a new, separate study for 6 months. This study was designed to investigate whether there could be late development of somaclonal variants (off-types) amongst the *in vitro*-produced plants. Some of the screenable variables that were considered included the plant phenotype, plant growth (as indicated by plant height, leaf size, and root system development), and the number of shoots formed by a single tuft arising from a single *in vitro* plant. Once the plants were cleared of any detectable abnormalities, they were then evaluated in studies at Langston University, or at the USDA-ARS Grazinglands Research Laboratory at El Reno, focusing primarily on the screening of *F*<sub>2</sub> progenies for potential segregation of genetically controlled variations. No phenotypic reversal has so far been observed in plants, which have been declared genetically invariant following *in vitro* culturing. Generally, off-



**Figure 13.6** Steps in the development of triticale.

### Breeding allopoloids

Allopoloids may be induced by crossing two species with different genomes, followed by chromosome doubling of the hybrid. Compared to autopoloids, inducing allopoloids is not commonly done by plant breeders. If successful, the newly induced amphiploid instantly becomes a new species (unable to cross to either parent). It also becomes reproductively isolated from its parents. Success of induced allopoloids is enhanced by the proper choice of parents. In particular, using parents with low ploidy levels increases the chance of high fertility and seed set in the amphiploid. Commercially successful induced allopoloids are few. The most noted success with induced allopoloidy is the commercially grown amphiploid, **triticale** ( $\times$  *Triticosecale*), derived from a cross between wheat (*Triticum*) and rye (*Secale*) (Figure 13.6). The objective of developing triticale is to obtain a product that com-

bines the qualities of wheat with the hardiness of rye. In lieu of doubling the  $F_1$  to produce the desired synthetic product, a wheat  $\times$  rye cross may be undertaken. The  $F_1$  plant possesses 28 chromosomes and exhibits intermediate traits that favor rye (hairy neck, spike length). All  $F_1$ s are sterile because of the formation of univalents and irregular gametogenesis.  $F_1$ s are backcrossed to wheat to produce progenies containing 42 chromosomes (seven from rye and the rest from wheat). The wheat chromosomes form bivalents at meiosis, while the rye chromosomes form univalents. The bivalent wheat chromosomes are irregularly arranged. Fertilization of an ovule with  $21 + 7$  chromosomes by pollen with the same genomic constitution will contain the full complement of chromosomes for wheat and rye (56 chromosomes). This product is the synthetic allotetraploid called triticale. Hexaploid triticale ( $AABBRR$ ,  $2n = 6x = 42$ ) is superior agronomically to octoploid triticale

(*AABBDDRR*,  $2n = 8x = 56$ ), but it requires embryo culturing to obtain  $F_1$ s between durum wheat and rye.

All amphiploid breeding is a long-term project because it takes several cycles of crossing and selection to obtain a genotype with acceptable yield and product quality. Common undesirable features encountered in triticale breeding include low fertility, shriveled seeds, and weak straw. Even though tetraploid ( $2n = 2x = 28$ ), hexaploid ( $2n = 6x = 42$ ), and octoploid ( $2n = 8x = 56$ ) forms of triticale have been developed, the hexaploid forms have more desirable agronomic traits and hence are preferred. Allopoloids have been used to study the genetic origins of species. Sometimes, amphiploidy is used by breeders as bridge crosses in wide crosses.

## Aneuploidy

Whereas polyploidy entails a change in ploidy number, **aneuploidy** involves a gain or a loss of one or a few chromosomes that make up the ploidy of the species (i.e., one or a few chromosomes less or more than the complete euploid complement of chromosomes). Just like polyploidy, aneuploidy has its own nomenclature (Table 13.7).

### Cytogenetics of autopoloids

The diploid complement of chromosomes is designated  $2n$ . A **nullisomic**, for example, is an individual with a

**Table 13.7** Aneuploidy nomenclature.

Chromosome number	Term	Nature of chromosomal change
$2n$	Diploid	Normal
		<b>Aneuploidy</b>
$2n - 1$	Monosomy	One of a pair of chromosomes missing
$2n - 2$	Nullisomy	Two chromosomes missing
$2n + 1$	Trisomy	Three copies of one chromosome (i.e., an extra copy)
$2n + 2$	Tetrasomy	Four copies of one chromosome (i.e., two extra copies)
$2n + 3$	Pentasomy	Five copies of one chromosome (i.e., three extra copies)

The individual with the condition, e.g., trisomy, is called a trisomic

missing pair of chromosomes ( $2n - 2$ ), while a **tetrasomic** has gained a pair of chromosomes ( $2n + 2$ ). Similarly, a **monosomic** has lost one chromosome from a homologous pair ( $2n - 1$ ), while a **trisomic** has gained an extra chromosome ( $2n + 1$ ).

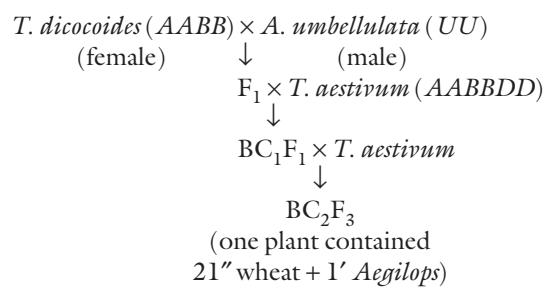
Aneuploidy commonly arises as a result of irregular meiotic mechanisms such as **non-disjunction** (failure of homologous chromosomes to separate) leading to an unequal distribution of chromosomes to opposite poles (Figure 13.7). Consequently, gametes resulting from such aberrant meiosis may have a loss or gain of chromosomes. Furthermore, chromosome additions often cause chromosome imbalance and reduced plant vigor.

### Applications of aneuploidy

Aneuploidy is used in various genetic analyses as described next.

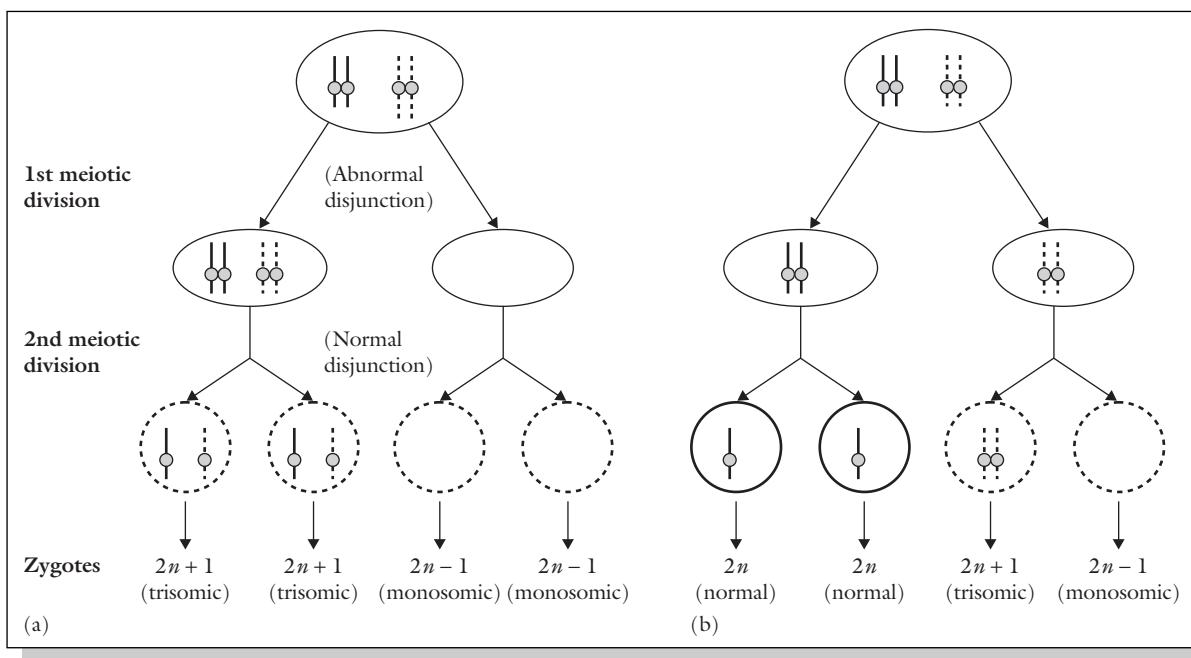
#### Chromosome additions

Chromosome addition lines are developed by back-crossing the synthetic allopolloid ( $F_1$ ) as seed parent to a cultivated species as pollen parent. This strategy is preferred because male gametogenesis is more readily perturbed by chromosomal or genic disharmonies than is the case in the female gametophyte. For example, E. R. Sears transferred the resistance to leaf rust of *Aegilops umbellulata* to *Triticum aestivum* (bread wheat) via bridge crossing with *T. dicoccoides* as follows:



However, it had drawbacks (sterile pollen, brittle spike axis, etc.). Subjecting this chromosome addition line to irradiation successfully translocated the segment of the *Aegilops* chromosome with the desired resistance genes to chromosome 6B of wheat, effectively removing the negative effects. The new genotype has been used in breeding as a source of resistance to leaf and stem rust.

Trisomics are important in genetic analysis. There are several types of trisomics. The term **primary trisomic** is used to refer to a case in which the euploid complement



**Figure 13.7** The origin of aneuploidy. Abnormal disjunction may occur at the first meiotic division (a) or at the second meiotic division (b) producing gametes with a gain or loss in chromosomes.

is increased by one complete chromosome. A **secondary trisomic** is one in which the extra chromosome has identical arms (i.e., one particular arm occurs as a quadruplicate). Such a chromosome is called an **isochromosome**. Sometimes, the extra chromosome added is derived from parts of different chromosomes. Such additives arise from chromosome breakage–fusion events.

Primary trisomics may be used to assign genes to chromosomes. Theoretically, there are as many possible trisomics as there are chromosome pairs. Scientists can generate trisomic stocks for a species. To assign a gene, the mutant (e.g., *a*) that is homozygous for the allele of interest is crossed to all the trisomic tester stocks. Assuming that all stocks are homozygous for the wild type and assuming normal meiotic segregation, two  $F_1$  types will be produced. Those produced from the union of normal gametes ( $n$ ) will segregate with the normal diploid ratio  $3A : 1aa$ . However, where a trisomic plant ( $n + 1$  gamete) is involved, an aberrant ratio would result. Because a trisomic stock is unique, the gene of interest would be located on the chromosome that the trisomic stock represents. These results assume random segregation of the three chromosomes of the trisomic

plant and equal viability of pollen regardless of genetic constitution. In reality there is a preponderance of  $n$  gametes and reduced function of  $n + 1$  gametes. The consequence of reduced functionality is that, sooner or later, a trisomic would revert to a diploid, unless the scientist makes special efforts to maintain it. Trisomics have been applied in creative ways in plant breeding, including their use in hybrid seed production in barley, using a balanced tertiary trisomic that carries a recessive male-sterility gene. The addition of chromosomes from other species (called **alien addition lines**) has been explored in interspecies crosses such as wheat  $\times$  rye. Chromosome addition lines may be unstable enough to be developed as cultivars.

#### *Chromosome deletions*

Unlike chromosome addition in which gene duplication occurs (hence an implied duplication in function), **chromosome deletion** leads to a loss of function. The consequence of a deletion depends on the functional roles of the genes of the chromosome that is lost. Invariably, surviving plants have less vigor and more sterility problems. However, in polyploids, the presence of homeologous

(in allopoloids) or homologous (in autopoloids) chromosomes may make up for the missing functions.

Monosomics ( $2n - 1$ ) may be used just like trisomics to assign genes to chromosomes in a polyploid species. This requires the development of monosomics for all the existing chromosome pairs in the species, as was done by E. R. Sears for the Chinese spring cultivar of wheat. Nullisomics ( $2n - 2$ ) may also be used in this fashion, but with less success because of severe reduction in vigor and fertility.

#### *Chromosome substitution*

Whereas alien chromosome addition entails adding an alien chromosome to the genome of an existing genotype, a **chromosome substitution** entails replacing or substituting a chromosome of the recipient species with an alien chromosome. Intervarietal (between varieties of the same species) and interspecific chromosome substitutions are more important in plant breeding than the addition of chromosomes. One of the well-known substitutions involves chromosome 1B of wheat and chromosome 1R of rye. The resultant wheat cultivar provided resistance to disease (leaf rust, stripe rust, powdery mildew). To use this technique, there have to be monosomic lines for the species (lines are available for wheat, cotton, tobacco, and oats).

The backcross breeding procedure may be used to substitute one chromosome for another in monosomics or nullisomics. Such chromosome substitution may be done within the species or involve other species (i.e., alien substitution). Researchers such as Sears have used the technique to assign numerous genes to chromosomes. However, the technique is challenging and requires a great amount of cytological analysis.

#### *Supernumerary chromosomes*

Also called **accessory** or **B-chromosomes, supernumerary chromosomes** are natural additions of varying numbers of small chromosomes to the normal genome. They have been found in all major taxonomic groups of organisms. These chromosomes are often predominantly heterochromatic and unstable in behavior. Although largely considered as genetically inert, studies in some species have indicated that the B-chromosomes increase the recombination frequency of **A-chromosomes** (the normal set of chromosomes) in species in which they occur. It is possible to use certain breeding techniques to increase their number. In some species such rye, fertility is reduced by the presence of one or two supernumerary chromosomes. However, corn plants can accumulate at least 10 such chromosomes before an adverse effect on fertility is noticeable.

---

#### **References and suggested reading**

- Bingham, E.T. 1980. Maximizing heterozygosity in auto-tetraploids. *In: Polyploidy: Biological relevance* (Lewis, W.H., ed.). Plenum Press, New York.
- Borojevic, S. 1990. Principles and methods of plant breeding. Elsevier, New York.
- Chase, S.S. 1964. Analytic breeding of amphiploid plant varieties. *Crop Sci.* 4:334–337.
- Dewey, D.R. 1980. Some applications and misapplication of induced polyploidy to plant breeding. *In: Polyploidy: Biological relevance* (Lewis, W.H., ed.). Plenum Press, New York.
- Harlan, J.R., and J.M.J. de Wet. 1975. On a wing and a prayer: The origin of polyploids. *Bot. Rev.* 41:361–390.
- Haynes, K.G. 1992. Some aspects of inbreeding in derived tetraploids of potatoes. *J. Hered.* 83:67–70.
- Hermsen, J.G.Th. 1984. Nature, evolution, and breeding of polyploids. *Iowa State J. Res.* 58:411–420.
- Hill, R.R. 1971. Selection in autotetraploids. *Theor. Appl. Genet.* 41:181–186.
- Poehlman, J.M., and D.A. Sleper. 1995. Breeding field crops, 5th edn. Iowa State University Press, Ames, IA.
- Sears, E.R. 1954. The aneuploids of common wheat. *Missouri Agric. Exp. Stn. Res. Bull.* 572:1–58.
- Singh, A.K., J.P. Moss, and J. Smartt. 1990. Ploidy manipulations for interspecific gene transfer. *Adv. Agron.* 43:199–240.
- Sterett, S.B., M.R. Henninger, G.C. Yencho, W. Lu, B.T. Vinyard, and K.G. Haynes. 2003. Stability of internal heat necrosis and specific gravity in tetraploid × diploid potatoes. *Crop Sci.* 43:790–796.
- Thomas, H. 1993. Chromosome manipulation and polyploidy. *In: Plant breeding: Principles and prospect* (M.D. Hayward, N.O. Bosemark, and I. Ramagosa, eds). Chapman & Hall, London.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 A genomic formula of  $2n - 1$  refers to a trisomic.
- 2 The regular set of chromosomes of a species is called A-chromosomes.
- 3 An individual in which the euploid complement of chromosomes is increased by one complete set of chromosomes is called a secondary trisomic.
- 4 The triangle of U describes genomic relationships among naturally occurring species of wheat.
- 5 Triticale is a euploid.
- 6 Colchicine is used for reducing the number of chromosomes in a cell.
- 7 Aneuploids have a duplicate of the entire chromosome set.
- 8 The genotype *AAAA* represents a duplex tetraploid.
- 9 The genotype *AADDEE* represents an allotetraploid.
- 10 A hexaploid consists of six genomes.

### Part B

Please answer the following questions:

- 1 Describe the triangle of U.
- 2 Distinguish between homologous and homeologous chromosomes.
- 3 Distinguish between a primary trisomic and a secondary trisomic.
- 4 Discuss a common mechanism of aneuploidy.
- 5 Distinguish between an aneuploid and euploid.
- 6 ..... is polyploidy with chromosomes from different genomes.
- 7 Write the genetic formula for a triploid genotype: .....
- 8 What is a segmental allotetraploid?

### Part C

Please write a short essay on each of the following topics:

- 1 Discuss the effect of polyploidy on plants.
- 2 Discuss, with an example, a polyploidy series.
- 3 Discuss the artificial induction of polyploids.
- 4 Discuss the importance of doubled haploids to plant breeding.



# 14

# Biotechnology in plant breeding

---

## Purpose and expected outcomes

*Genes can be transferred from one parent to another through crossing. Such a gene transfer is restrictive in the sense that it occurs only when individuals can be successfully crossed. The recombinant DNA (rDNA) technology, coupled with other molecular breeding tools, theoretically allows scientists to transfer genes from one organism to any other, circumventing the sexual process. For example, a gene from a bacterium can be transferred into a corn genome. Consequently, rDNA technology allows scientists to treat all living things as theoretically belonging to one giant breeding gene pool. The discussion in this chapter is meant to be an overview of rDNA technology and how it is applied in plant breeding. After studying this chapter, the student should be able to:*

- 1 Describe the basic steps in genetic engineering or rDNA technology.
  - 2 Discuss the enabling technologies of genetic engineering.
  - 3 Discuss the importance of microorganisms in genetic engineering.
  - 4 Outline the fundamental difference between conventional breeding and genetic engineering.
  - 5 Discuss the concept of molecular breeding.
- 

## What is biotechnology?

Etymologically, **biotechnology** is the study of tools from living things. In its current usage, the term is defined either broadly or narrowly. It may be defined broadly as the use of techniques based on living systems to make products or improve other species. This would include the use of microbes to make products via fermentation, an age-old practice. In a narrower definition, biotechnology refers to the genetic manipulation of organisms for specific purposes. The term **genetic engineering** is sometimes used to describe this practice. Some argue that classic plant breeding is genetic engineering, since the genetics (DNA) of plants are manipulated by breeders, albeit *indirectly*. Consequently, a much narrower definition of genetic engineering is used to describe the manipulation of organisms at the

molecular level, *directly* involving the DNA. However, it is the revolutionary technology of **recombinant DNA (rDNA)**, which enables researchers to transfer genes from any organism to another, that some accept as genetic engineering. The term **molecular breeding** is used to describe the use of a variety of tools for manipulating the DNA of plants (which may or may not involve rDNA) to improve them for specific purposes.

## General steps in rDNA technology

Even though crossing of two different parents produces new recombinants in the segregating population, the term recombinant DNA is restricted to the product of the union of DNA segments of different biological origins. A cultivar developed by the rDNA procedure is

called a **transgenic cultivar** or a **genetically modified (GM) cultivar**. Generally, an organism developed by the rDNA procedure is called a **genetically modified organism (GMO)**.

Certain basic steps are common to all rDNA projects:

- 1 The DNA of interest that is to be transferred (the **transgene**) is extracted from the source organism. The specific DNA sequence of interest is cut out using special enzymes.
- 2 The transgene is inserted into a special DNA molecule (a **cloning vector**) and joined to produce a new rDNA molecule.
- 3 The rDNA is transferred into and maintained in a host cell (bacterium) by the process of **transformation**. The vector replicates, producing identical copies (called **clones**) of the insert DNA.
- 4 The host cells with the cloned transgene are identified and isolated from untransformed cells.
- 5 The cloned transgene can be manipulated such that the protein product it encodes is expressed by a host cell.

### Landmark discoveries in the application of molecular biology

Numerous discoveries have been made over the years that plant breeders and other researchers may use for manipulating plants. An overview of the most influential tools is given here.

#### Restriction enzymes

The gene of interest must first be excised from the source genome before it can be transferred into another genome. The most common method of cutting DNA is by using special bacterial enzymes called **restriction endonucleases** (or simply **restriction enzymes**). These enzymes are base-specific and cut between specific DNA bases after identifying a short sequence of bases unique to the enzyme (called the **recognition sequence** or **site**). An example is 5' . . . TCGA . . . 3' for the bacterium *Thermus aquaticus* (only one chain is shown).

#### Southern hybridization

**Southern hybridization** (or **Southern blotting**) was discovered by E. Southern as a technique for isolating a specific DNA from a mixture of fragments following a restriction enzyme digest. The DNA digest is denatured into single strands. The fragments are transferred to a DNA-binding material (nylon or nitrocellulose), a pro-

cess called **blotting**. The blot is exposed to a solution containing a labeled probe (a nucleotide sequence that is complementary to the DNA sequence of interest) to bind (hybridize) to it, thereby positively identifying the gene or DNA sequence of interest.

#### Reverse transcriptase

As discussed in Chapter 3, the genetic information of the DNA is transcribed into a template (mRNA) by an enzyme called **DNA transcriptase**, for onward translation into protein. The discovery of the enzyme **reverse transcriptase** was a major breakthrough in molecular biology because it allowed researchers to synthesize a complementary copy of a given DNA (called **complementary DNA** or **cDNA**). Researchers, knowing the protein product, can now work backwards to synthesize its DNA.

#### Polymerase chain reaction

The **polymerase chain reaction (PCR)** is an enzymatic amplification of a DNA fragment. This discovery, among other applications, allows researchers to make millions of copies of DNA from just one copy. In its basic form, the PCR technique uses enzymes to amplify a DNA fragment by flanking it with two oligonucleotide primers that hybridize to opposite strands of the target sequence. A basic PCR cycle consists of three stages: (i) heat **denaturation** (at about 95°C) to separate the double strands of the target DNA; (ii) **renaturation** at a cooler temperature (about 55°C) to allow the primers to anneal to their complementary sequence in the single-stranded target DNA; and (iii) **extension** (at about 75°C) of the primers to create copies of the target DNA. The PCR mixture consists of DNA polymerase and the four DNA nucleotides, along with the source DNA. The PCR occurs in an automated unit called a **thermocycler**.

The technique can also be used where only one sequence suitable as a primer-binding site is known (called **anchored PCR**), or to amplify a region of DNA of unknown sequence that flanks a known sequence (called **inverse PCR**). There are modifications of the PCR (e.g., real time or **RT-PCR**).

#### Rapid sequencing

**DNA sequencing** allows researchers to determine the complete order (sequence) of bases in a DNA molecule. Sequencing provides information on gene organization

and expression (including codon usage, comparing known sequences, gene searches, etc.).

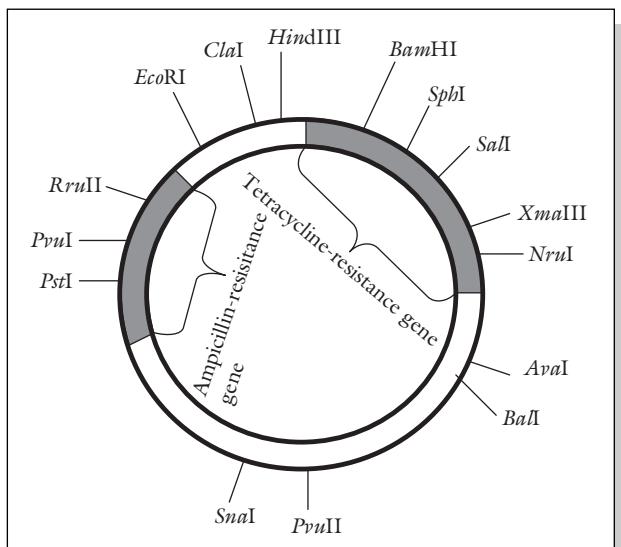
### Identifying and cloning genes of interest

Before a gene can be transferred from one organism to another, it must be identified and isolated from its source's genome. By itself, a piece of DNA cannot self-replicate. The isolated gene (or for that matter any DNA sequence) is maintained in a cell so that copies can be generated as the cell divides (called **cloning**).

#### Cloning vectors

**Cloning vectors** (or simply vectors) are replicating units into which isolated fragments of DNA may be integrated for maintenance. A vector has three basic features – a selectable marker, a replication origin, and a cloning site (restriction enzyme site) (Figure 14.1). Vectors differ in the size of fragments they can stably incorporate, the procedure for screening the insert DNA or target DNA, and the number of recombinant copies they can produce per cell. Examples are:

- 1 **Plasmid vectors.** Plasmids are double-stranded, self-replicating extrachromosomal molecules with antibiotic selectable markers. They are small molecules that can handle small inserts of about 10 kb in size.



**Figure 14.1** A cloning vector showing restriction enzyme sites (no shading) and selection markers sites (dark shading).

2 **Bacteriophage  $\lambda$ -derived vectors.** These are viruses that infect bacteria. They have a cloning capacity of about 23 kb.

3 **Cosmids.** Cosmids are plasmid vectors with *cos*-sequences (for packing) from bacteriophages. They are capable of carrying about 42 kb inserts.

4 **Yeast artificial chromosomes (YAC).** These are useful for cloning large fragments in the megabase range.

5 **Phosphoinositide (PI)-derived artificial chromosome (PAC) and bacterial artificial chromosome (BAC).** PAC vectors can handle about 350 kb, while BACs can handle about 150 kb inserts.

#### Gene isolation and cloning

Gene isolation starts with the construction of a library representing a set of recombinant molecules that contain DNA fragments. There are two basic kinds of DNA libraries (or gene banks) – **genomic** and **cDNA**.

##### Genomic library

DNA is extracted and purified (preferably nuclear DNA). Fragments of DNA are cloned into appropriate vectors. The number of recombinant clones needed to create a gene bank holding the complete genome is given by the equation:

$$N = [\ln(1 - P)] / [\ln(1 - t)n]$$

where  $N$  = number of independent recombinant clones,  $P$  = probability that a particular sequence is represented (e.g., 95%),  $t$  = average length of the fragment to be cloned (kb), and  $n$  = total amount of DNA per cell.

##### cDNA library

As discussed in Chapter 3, a eukaryotic gene comprises coding and non-coding segments (exons and introns). A genomic library contains both elements. A cDNA library consists of only coding segments of the genome. This is obtained from the mRNA by reverse transcription. A cDNA library consists of recombinant molecules containing all the mRNA in the organism. It is important to note that only genes that are functionally expressed in the cell from which the mRNA was isolated would be included in the library. Consequently, breeders interested in drought genes, for example, may extract DNA from a drought-stressed plant, since that condition is likely to trigger the expression of drought-resistance genes.

## Gene identification

With a library constructed, the breeder may search it to identify a specific gene of interest using one of several techniques based on two approaches:

- 1 Based on gene product.** The breeder may purify the protein expressed by the gene and from that obtain the amino acid sequence (e.g., by 2D gel electrophoresis). The information is used to design synthetic oligonucleotides corresponding to the sequence. This is labeled and used in a Southern hybridization to identify the corresponding gene. The method of antibodies is also effective.
- 2 Based on sequence characteristics.** The breeder may use nucleotide sequence information from studies in other species. The probes designed from such information are called heterologous probes and represent a complete or partial gene sequence. Genes may also be isolated by techniques such as chromosome walking (map-based cloning), molecular tagging (transposon tagging), or differential screening.

## Gene transfer

Once the desired gene has been identified from the library, it is ready to be transferred into a host cell, a process called **genetic transformation**. There are two categories of transgene transfer or delivery procedures – direct and mediated transfer.

### Direct gene transfer

#### *By particle acceleration or bombardment*

One of the commonly used direct gene transfer method is **microprojectile bombardment** (or **biostatic**). A biostatic device (called a **gene or particle gun**) is used to literally shoot the target DNA into intact cells (hence the nickname of **shotgun transformation**). The widely used particle delivery system is marketed by DuPont (called the DuPont Biostatic® PDS1000/He device) and uses helium as the propellant gas. Small amounts (about 50 µg) of micron-size (1–5 µm diameter) carrier particles (tungsten or gold) are coated with the target DNA and propelled in the barrel of the gene gun at energies high enough to penetrate plant cells. The rate of acceleration may be up to 430 m/s in a partial vacuum. The carrier particles pass through a mesh, hitting

the target tissue (e.g., callus) in a Petri dish below the biostatic device. A low penetration number of projectiles (1–5 per cell) is desirable. More than 80% of bombarded cells may die if particle penetration reaches 21 projectiles per cell.

#### *Electroporation*

Callus culture (or explants such as immature embryos of protoplasts) is placed in a cuvette and inserted into a piece of equipment called an electroporator, for electroporation. This procedure widens the pores of the protoplast membrane by means of electrical impulses. The widened pores allow DNA to enter through them to be integrated with nuclear DNA.

#### *Other methods*

Other direct methods are available, including micro-injection and silicon carbide procedures.

### Mediated gene transfer

In directed or mediated gene transfer, intermediary agents are used to act as couriers (vectors) for carrying the target DNA into recipient cells. In plants, the *Agrobacterium*-mediated gene transfer is a common practice. The bacterium causes crown gall tumors in dicots. Upon infection, the bacterium transfers a part of its DNA into the host. The tumor-inducing part of the specially designed bacterium is deleted. It has been determined that the oncogenic (tumor-inducing) properties of *A. tumefaciens* reside on a large **tumor-inducing (Ti) plasmid** (called **Ri plasmid** in *A. rhizogenes*). The Ti plasmid has two regions, the T-region (called T-DNA or transferred DNA) is what is transferred into the plant cell and integrated into the host chromosome (i.e., natural genetic engineering!). A second region on the Ti/Ri plasmid, called the virulence (*vir*) region, carries the genes for tumor induction and is also involved in the transfer of T-DNA. T-DNA also carries genes for synthesizing metabolic substrates for bacteria called **opines**. Further, the T-DNA is flanked or bordered by short (25 bp) direct repeats of DNA. The deletion of the tumor-inducing segment of the DNA does not prevent the transfer of T-DNA.

#### *Requirements for transformation*

The following are some key considerations for a successful *Agrobacterium*-mediated transformation project:

- 1 Efficient plant regeneration system.** Without an efficient regeneration system to grow full plants from cells, it is futile to undertake any transformation. In effect, it may be said that “thou shall not transform unless thou can regenerate”! A small number of cells in a plant are competent for both transformation and regeneration.
- 2 Determination that the cells are susceptible to *Agrobacterium* transformation.** The *Agrobacterium*-mediated transfer system should be efficient and be able to transform most of the cells at the target site.
- 3 An efficient and sensitive selection method.** Similarly, an efficient selection system should be available to readily identify and select the transformed cells from among untransformed cells.
- 4 Stable transformation.** Transformation is successful if the cells regenerate, and the transgene is expressed in subsequent generations. The fewer the number of cells producing the regenerated plant the better, otherwise a chimeric product will result.

#### Procedure of Agrobacterium transformation

*Agrobacterium* cells containing the DNA inserts are co-cultivated (co-cultured) (e.g., with leaf disks, cotyledons, or other materials) on a regeneration medium for 2–3 days. The bacteria bind to the wounded sites and subsequently transfer the insert DNA into the host genome. The explant is transferred to a regeneration medium containing carbenicillin (to kill *Agrobacterium*) and an antibiotic (e.g., kanamycin) to inhibit growth of untransformed cells. Successful transformants are cultured to full plants. The bacterium gains entrance into the plant material through the wounds. Thus, gene delivery is more successful with dicots than monocots because the former are more wound responsive than cereals.

#### Transient versus stable transformation

As previously stated, stable expression of a transgene is desired so that the transgene would be heritable and expressed from one generation to the next. However, the breeder may need to have an earlier indication of successful transformation and hence may evaluate the expression of the transgene at an intermediate stage in the transformation process. For example, protoplasts may be transformed with the target DNA and the protoplasts isolated after only 1–2 hours for the evaluation of expression of encoded genes. This preliminary expression of the transgene is described as **transient expression**. Transient expression occurs when the insert

gene is not integrated into the host genome. Such an expression dissipates with time.

#### Tissue culture and selection of transformation events

Tissue culture is a critical part of a genetic engineering project. This is because transgenes are usually delivered into cells before they differentiate. Common explants used are protoplasts, cell suspension, immature embryos, shoot meristem, and immature inflorescences. Because a transgene may not be successfully delivered into all cells in a mass of callus, it is important to have a selection system to discriminate among cells to identify and isolate only genuine transformants.

Cloning vectors, as previously discussed, are designed for specific purposes. Selectable marker systems built into vectors may be grouped as follows: antibiotic selection (e.g., kanamycin), herbicide selection (e.g., acetolactate synthase genes are also used in herbicide selection (for glyphosate or sulfonylurea herbicides)), scorable gene-mediated selection, and positive selection. The selection agent is included in the tissue culture medium. Selection systems differ in their ease of use and efficiency.

#### Antibiotic selection

This is the first generation selection marker system in biotechnology. Like antibiotic markers, herbicide markers are used for establishing preferential growth of transgenic cells as in other marker systems.

#### Scorable gene-mediated selection

Reporter genes are expressed in cells without integration into the genome and are assayed in a variety of ways. In a transformation project, it is always desirable, whenever researchers can, to have a preliminary evaluation of success of transformation. **Scorable marker genes (reporter genes)** are used for rapid visual confirmation for transient expression following transformation. Reporter gene assay is a transient assay and can be conducted within 24 hours after transformation of the cells. The commonly used scorable marker genes encode enzymes that have distinct substrate specificities, thereby enabling researchers to monitor changes with various visualization systems. Examples of scorable marker genes are **GUS** ( $\beta$ -glucuronidase), which is visualized by a histochemical or fluorometric protocol, and **LUC** (luciferase), which is visualized by a luminescence system and a green fluorescent protein (GFP) system that allows

the non-destructive, visual identification of transgenic cells by standard fluorescence microscopy.

### The search for new selection systems

Antibiotic selection systems have been severely criticized by activists for being environmentally risky. Consequently, scientists are searching for new marker systems with no perceived or real adverse environmental consequences (so-called **benign markers**). Some of these selection systems (called **positive selection**) are based on metabolic pathways. An example is the mannose selection system, where the phosphomannose isomerase gene (*pmi*) is used as a selectable marker, and mannose as the selecting agent. Transgenic cells have a metabolic advantage over untransformed cells and hence grow and develop better. Various approaches are also used to select transformation events without markers. These include cotransformation, site-specific recombination, and use of transposable elements.

### Confirmation of transformation

After selecting transformation events and culturing cells into transgenic plants, it is necessary to confirm the success of transformation in the putative transgenic plants. Several methods are used, including phenotypic assays, PCR analysis, enzyme assays, and Southern or western blots. When *A. rhizogenes* is used, the transgenic roots are very hairy.

### Transgene integration into the host genome

Once transferred into a host cell, transgenes integrate into the genome at sites located randomly throughout the genome, but predominantly in transcriptionally active regions. Further, as previously discussed, multiple copies of the transgene may be incorporated into one cell. The number of insertions at a locus may vary between one and five, a factor that is influenced by the methodology (e.g., the preparation and physical status of the plasmid DNA (coiled, linear)), and the concentration and amount of DNA delivered. However, *Agrobacterium*-mediated transformation generally produces fewer insertions per locus.

### Transgene expression in transgenic plants

Once integrated into the genome, the success of the project is determined not only by the stability of transformation, but also by the desirability of expression.

In genetic engineering research, **promoters** are the “engines” that drive DNA expression by determining the level of transcription of a selectable coding sequence. Sources of promoters include viruses, bacteria, and plants. Promoters may be grouped into three major classes.

**1 Constitutive promoters.** As previously indicated, some genes are “turned on” continuously, while others are turned on as needed. In other words, structural genes are transcribed either continuously (constitutive expression) or periodically as the gene product is required (regulated expression). Constitutive promoters have high affinity for RNA polymerase and consequently promote frequent transcription of the adjacent region. This category of promoters is often used to drive the expression of the selectable marker gene for identification of transgenic tissues *in vitro*. A widely used constitutive promoter is derived from the cauliflower mosaic virus, and is called CaMV 35S.

**2 Tissue-specific and developmentally regulated promoters.** Certain genes are expressed only in certain tissues (even though all cells are equally genetically endowed). In animals, milk is expressed only in the mammary glands. Similarly, some genes need to be expressed only at certain developmental stages, and hence need to be regulated for such specific roles. In plants, genes for grain quality, for example, should be targeted for expression in the endosperm. Promoters for targeted gene expression are available for plant tissues such as endosperm (e.g., for storage protein), anther tapetum (e.g., for engineering male sterility), embryo (e.g., for engineering grain quality traits), and phloem (e.g., for engineering pest resistance to sucking insects). Some promoters are responsive to environmental stimuli (e.g., light). Some genes need to be expressed only at certain developmental stages, and hence need to be regulated for such specific roles. The choice of a promoter is made according to the trait being improved.

**3 Inducible promoters.** Some promoters are designed to drive the expression of genes in response to injury. For example, a transgene may be developed to trigger the expression of a gene in response to pathogenic invasion or wounding. Wound-induced promoters have been isolated in species such as rice (e.g., the rice basic chitinase, RC24, which drives the expression of the *nudA* gene in wounded roots and stems).

### Stability of transgene expression

There is no guarantee that once stably transferred and expressed that the condition will be permanent.

Structural defect in the gene construct may cause transgenic breakdown. The phenomenon of **progressive transgenic silencing or failure** is not completely understood. It is suspected that ancillaries of an expression system (e.g., terminal sequence promoters) may have a role in the event. Especially, systems using heterologous sequences from non-related sources (e.g., plants having bacterial or viral sequences, instead of native sequences, or those from closely related species) are more prone to malfunction, as they become recognized as foreign DNA and consequently excised or methylated. It is also suspected that the structural integrity of the introduced foreign DNA sequence may be responsible. This relates to the presence or absence of introns and the nature of the 5' and 3' sequences. Transgenics produced from cDNA clones instead of genomic sequences lack introns and possibly other 3' gene sequences. This may reduce the accumulation of mRNA as some researchers have reported.

### Plant genomics

The basic set of chromosomes of an organism is called its **genome**. Traditional geneticists generally investigate single genes, one at a time, as snapshots. Studying genetics whereby the totality of the genes in an individual are considered together would be advantageous. Genes seldom, if ever, work independently in higher organisms.

#### What is genomics?

**Genomics** is the approach of investigating the totality of genes in an individual as a dynamic system, not as snapshots, but rather over time, and to determine how these genes interact and impact biological pathways and the general physiology of an organism, from a global perspective (i.e., the “big picture”).

Genomics may be broadly categorized into two – **structural** and **functional** – each with its set of tools and functions. However, as the field advances, new terms, categories, and subcategories continue to emerge in the literature.

#### Structural (classic) genomics

Structural genomics, at the most basic level, is concerned with activities at the initial phase of genome analysis – **mapping** (the construction of high resolution genetic, physical, and transcript maps of an organism). The

initial phase of genome analysis is the establishment of a physical map of the genome. The ultimate physical map of an organism that can be achieved is the complete sequence of its total DNA (**genomic sequence**). Genomic sequencing projects yield linear amino acid sequences. The most visible genome sequencing project in recent times is the Human Genome Project. Genome sequencing projects have been initiated or completed for some major food crops such as rice, maize, and wheat.

Genome mapping entails determining the order of genes (or other genetic markers) and the spacing between them on each chromosome. Two categories of mapping strategies are employed in genome mapping: (i) genetic linkage mapping; and (ii) physical mapping.

#### Genome sequencing

The purpose of **DNA sequencing** is to determine the order (sequence) of bases in a DNA molecule. Classic mapping approaches involve generating and mapping mutants, thus limiting characterization to those genes for which mutants have been isolated. rDNA techniques provide a direct approach to genetic analysis, whereby a genomic library is created, from which overlapping clones are then assembled to construct genetic and physical maps for the entire genome. Ultimately, the entire genome is sequenced such that all genes are identified by both their location in the chromosome as well as their nucleotide sequence.

Allan Maxam and Walter Gilbert invented the **chemical degradation method** of DNA sequencing. Another method, the Sanger **enzymatic (dideoxy) method (chain terminating)**, named after its inventor, is more commonly used. However, in these days of more advanced technology, there are high-throughput methods of DNA sequencing, making it more convenient for most researchers to contract sequencing jobs to outside companies. **Capillary array electrophoresis (CAE)** systems coupled with high-sensitivity detection provided by energy-transfer labeling reagents are now the standard for high-throughput DNA sequencing facilities.

Genome sequencing is involved and hence often undertaken on a collaborative basis (i.e., consortium) over several years. Modern supercomputers and powerful algorithms are making genome sequencing more efficient.

#### Comparative genomics

Biotechnology makes use of certain organisms as **models** for comparative research. A fundamental premise of

rDNA is that DNA is universal and hence the transfer of DNA is possible, in theory, across biological barriers. Eukaryotes are genetically complex organisms, and not directly amenable to certain biotechnological procedures. Consequently, scientists pick plants and animals that are easy to manipulate, and invest in a thorough genetic analysis of their genomes. The information derived from such studies is extrapolated to more complex organisms. *Arabidopsis thaliana* was the first plant to be completely sequenced. This flowering plant has a short life cycle of about 6 weeks from germination to seed maturity. It produces seed profusely. Because of its small size, *Arabidopsis* can be cultured in trays or pots in a limited space. Two of the most important higher plants used for food that are being developed as model plants are rice (*Oriza sativa*) and maize (*Zea mays*). There are several subspecies of rice, the most widely eaten being *O. indica*, followed by *O. japonica*. Through both public and private sector efforts, the genome sequences of the two subspecies have been developed. Genomes of all cereals (wheat, barley, rice, corn, etc.) are structurally similar (see next Section).

### Genomic colinearity and its application in plant breeding

Genomic colinearity is a term used to refer to the conservation (through evolutionary history) of gene content, order, and orientation between chromosomes of different species, or between non-homologous chromosomes within a single species. The term is sometimes used as synonymous with synteny, the phenomenon of conservation of gene order within related genomes. Synteny has been observed in the Brassicaceae using molecular techniques. An important discovery is the synteny found between dicots and monocots. The sequence information for *Arabidopsis thaliana* (a *Brassica* species) and rice have revealed a conserved gene order across this vast evolutionary distance. It is estimated that about 75% of genes in dicot species and about 40% in monocots occur in regions that have colinearity with *Arabidopsis* that could be exploited by breeders.

Synteny has also been discovered in Fabaceae involving soybean, peanut, mungbean, lentil, common bean, common pea, and alfalfa. Some conservation of the genomic region in this family with *Arabidopsis* has been established. Synteny in the Poaceae is one of the clearest to be documented. Using the rice genome as the inner circle, researchers have graphically aligned the genomes of several cereals (sorghum, maize, wheat) by colinearity

into concentric circles such that colinear regions in different species lie along any radius. Rice has become a model species for comparative genomic studies of species with large genomes (e.g., wheat, sugarcane) that are difficult to study by traditional genetics. Another advantage is that the so-called “orphan crops” (e.g., millet), which would otherwise not receive significant attention as the major cereals do, can now benefit from the information obtained from rice.

The applications of colinearity to plant breeding include the following:

- 1 The prediction of the location of genes controlling a particular function in crop species.
- 2 The facilitation of genome mapping by transferring markers from a well-mapped genome to a less studied one.
- 3 Mapping will facilitate marker-assisted selection or map-based cloning.
- 4 After cloning a locus of agronomic importance, colinearity may provide opportunities for accumulating alleles of that locus from other distantly related species. This collection of genes may be a source for genetic engineering applications for crop improvement.

### Functional genomics

Once DNA sequences have been obtained, the next important task is to understand their function. Structural genomics focuses on sequencing of the genome; **functional genomics** focuses on gene function. The genome is essentially a set of instructions for making proteins of various kinds. Because most genes are expressed as proteins, one of the common ways of understanding gene function is by tracking protein expression by cells (called **proteomics**). Genes may provide instructions for making specific proteins. However, in the process of carrying out these instructions, additional proteins can be produced, as has already been noted. Linking gene to function is a complex undertaking. Understanding the genome structure alone is insufficient; it is critical to identify the proteins they encode. Many techniques have been developed for deciphering gene function.

### Bioinformatics in plant breeding

**Bioinformatics** may be defined as a knowledge-based theoretical discipline that attempts to make predictions about biological function using data from DNA

sequence analysis. It is an application of information science to biology. It uses supercomputers and sophisticated software to search and analyze databases accumulated from genome sequencing projects and other similar efforts. Bioinformatics allows scientists to make predictions based on previous experiences with biological reality. In one application, the biological information bank is searched to find sequences with known function that resemble the unknown sequence and thereby predict the function of the unknown sequence. Databases are critical to bioinformatics and hence repositories exist in various parts of the world for genetic sequence data.

### Types of bioinformatics databases

Information used in bioinformatics research may be grouped into two categories:

- 1 **Primary databases.** These databases consist of original biological data such as raw DNA sequences and protein structure information from crystallography.
- 2 **Secondary databases.** These databases contain original data that have been processed to suit certain specific applications.

To be useful, a good database should have two critical parts: (i) the **original sequence**; and (ii) and an **annotation description** of the biological context of the data. It is critical that each entry be accompanied by a detailed and complete annotation, without which a bioinformatics search becomes an exercise in futility since it would be difficult to assign valid meaning to any relationships discovered. Some databases include taxonomic information such as the structural and biochemical characteristics of organisms.

The bulk of data in repositories consist of primary data. Three major entities are collaboratively responsible for maintaining gene sequence databases. These entities are the European Molecular Biology Lab (EMBL) of Cambridge, UK, the GeneBank of the National Center for Biotechnology Information (NCBI) that is affiliated with the National Institutes of Health, USA, and the DNA Databank of Japan.

Databases for both protein sequences and structure are being maintained. The Department of Medical Biochemistry (University of Geneva) and European Bioinformatics Institute collaboratively maintain properly annotated translations of sequences in the EMBL databases. This is called the SWISSPROT. TREMBL is another protein database consisting solely of protein-coding regions of the EMBL database (called the translated EMBL or TREMBL). The NCBI also maintains a

database of the translations of the GeneBank. Another kind of protein database consisting of experimentally derived 3D structures of proteins is kept at the protein databank where these structures are determined by X-ray diffraction and nuclear magnetic resonance.

### General steps in a bioinformatics project

One purpose of searching a bioinformatics database is to determine if the researcher's unknown sequence, DNA or protein, matches any sequence in the database in terms of structure or function. This requires the proper choice and skillful use of software to align the unknown sequence with the known. Gene-seeking software varies in capability and ease of use. They have certain properties in common: (i) algorithms for pattern recognition use statistical probability analysis to determine the similarity between two sequences; (ii) data tables contain information on consensus sequences for various genetic elements; (iii) taxonomic differences are included because consensus sequences vary between different taxonomic classes to facilitate analysis and minimize errors; and (iv) specific instructions describe how the algorithms should be applied in an analysis and how the results should be interpreted.

A search involves two key activities:

- 1 **Sequence alignment scoring matrices.** The unknown sequence is aligned with those in the bank. Scores are assigned on the basis of the sequence homology detected. It is most useful to align sequences such that the largest scores are assigned to the most biologically significant matches.
- 2 **Comparing sequences against a database.** One of the most common searches of bioinformatics databases is to compare an unknown sequence against those in the database to discover similarities. Typical homology search algorithms are used in this activity. Some of the most widely used software for this search is the Basic Local Alignment Search Tool (BLAST) and FASTA. BLAST uses a strategy based on short sequence fragments between the unknown sequence and those in the database. It is designed to match only continuous sequences (no gaps from deletion or insertion mutations are taken into account).

### DNA microarrays in plant breeding

Genes are not only variably expressed, but the level of expression varies during a physiological change, some



## Industry highlights

### *Bioinformatics for sequence and genomic data*

Hugh B. Nicholas, Jr., David W. Deerfield II, and Alexander J. Ropelowski

Pittsburgh Supercomputing Center, Pittsburgh, PA 15213, USA

#### ***Introduction***

Bioinformatics is the application of informatic techniques to biological data. These techniques include acquiring, annotating, analyzing, and archiving the biological data, using concepts from biology, computer science, and mathematics. Bioinformatics has its roots in the work of people like Margaret Dayhoff (collecting known sequences (Dayhoff et al. 1965) and a mathematical model of protein evolution (Dayhoff et al. 1978)) and David Sankoff (sequence alignment and statistical tests for homology) in molecular phylogenetics.

Bioinformatic analyses aim to discover precise, testable hypotheses to supplement or redirect biology experiments. The results of these bioinformatics-derived experiments then influence the next round of computer-based studies. This interaction between experiment and computation speeds scientific progress.

The first large collections of biological data were protein sequences, followed by nucleic acid sequences, and were collected by individual laboratories and stored on computer punch cards, along with annotations, such as species, biochemical function, and physiological role, and functional and structural domains. As the amount of sequence data grew and the need to share this data between multiple laboratories grew, the collection and curation were taken over by specialized groups such as National Biomedical Research Foundation/Protein Information Resource (NBRF/PIR), GeneBank, EMBL, and Swiss-Prot, where curation included standardizing the format and ancillary data. With increased size also came the need for tools to search, analyze, and annotate these databases.

#### ***Pairwise alignment and database searching***

Molecular biologists commonly isolate and sequence molecules based on their association with particular biological phenomena, such as disease resistance in plants. Typically, the biochemical function of the newly determined sequence is not known and one compares the newly determined sequence to all known sequences whose biochemical functions are known to generate a testable hypothesis about its function. Thus, an early bioinformatics tool was to search a database of annotated sequences with a newly determined sequence to find all similar sequences. Sequences that were similar enough were inferred to be homologous, that is to have descended from a common evolutionary ancestor. The inference of homology generates the hypothesis that the two molecules carry out the same biochemical function and perhaps the same physiological role. A complete discussion of database searching is given in Nicholas et al. (2000).

The power of a successful database search is demonstrated by comparing the histories of cystic fibrosis (CF) and type I neurofibromatosis (NF-1) research. Both disease genes were isolated in 1988. The CF gene was identified as a chloride ion transport protein, which led to the development of a number of therapies, with many now in final clinical trials. In 1988 the database search with the NF-1 gene failed and no homologues were found. It was not until 1998 that NF-1 was identified as a growth suppressor, which has rapidly lead to improved diagnosis and many potential therapies for which clinical trials are just beginning. Thus, the successful identification of the CF gene as a chloride ion transporter accelerated this research area by a decade compared to the time required to discover the biochemical function of the NF-1 gene through biological experiments.

#### ***Multiple sequence alignment***

A database search results in discovering many similar sequences from which one would like to create a multiple sequence alignment that simultaneously shows the relationship among its homologous residues in the other sequences. This alignment is a map of the evolution of the protein family. The multiple sequence alignment is a rich source of hypotheses to guide experimental work since the alignment contains patterns of conservation and variation of residues among the sequences, which provides insights into functional and structural positions for either the family of proteins or the genes encoding them. Such inferences are strongest if the alignment contains sequences from widely diverse species.

Multiple sequence alignment implies that the residues in each column of the alignment are all evolutionarily related to each other. Thus accuracy is most commonly considered to be improved by maximizing the observed degree of conservation in the alignment as a whole, as discussed in Nicholas et al. (2002).

### Analysis

Two evolutionary concepts underlie the inferences from analyses of multiple sequence alignments. The first is that the mutation of one sequence residue to another is a random event in nature and that all residues are more or less equally subject to mutation. The second is that the conservation of specific residues is maintained through evolutionary selection; that is, mutations that adversely affect the ability of a molecule to carry out its biochemical activity or physiological role will be eliminated by reducing the ability of the organism to live. Details of the evolutionary history between different pairs of sequences can lead to different inferences about the properties of the sequences.

In the evolutionary history of some pairs of sequences – orthologues or orthologous sequences – the sequences have only speciation events in their evolutionary history. Other pairs of sequences – paralogues or paralogous sequences – have one or more gene duplication events in their common evolutionary history as well as having speciation events. In general, paralogues will carry out the same biochemistry on different substrates or with different cofactors; while orthologues will carry out the same biochemistry on the same substrates and will often serve an identical physiological role in the same pathways in different organisms. Homologues include both orthologues and paralogues. Since paralogues carry out the same basic biochemistry, for example reduce an aldehyde, the residues responsible for this activity are conserved. But, the residues responsible for the physiological role (e.g., substrate recognition, which specific aldehyde to reduce or lipid to bind) will be under different evolutionary pressures and will often diverge. A complete analysis of the multiple sequence alignment includes identifying residues responsible for the common, shared properties of the entire family and the residues that discriminate between paralogous groups (Figure 1).

The results from the analysis of multiple sequence alignments is illustrated in the results obtained in the decades' long quest to understand how each of the 20 aminoacyl tRNA synthetase (aaRS) enzymes map each of the 60 tRNAs to its correct cognate amino acid. The problem, as outlined by Sir Francis Crick in 1957, is where is the information that allows an aaRS to recognize only the tRNAs encoding the correct amino acid? Experimental groups have employed a number of creative and innovative approaches towards answering this question (Söll & Schimmel 1974). McClain, Nicholas, and coworkers applied computational techniques, beginning with creating an accurate multiple sequence alignment of the tRNAs.

The tRNA multiple sequence alignment was divided into 20 subsets, each based on their amino acid acceptor activity (McClain & Nicholas 1987). First these authors identified the sequence positions in the 66 tRNA sequences from *Escherichia coli*, its bacteriophages, and near relatives that were conserved. They then examined the remaining positions and asked the question: what

Bpi_Bovin :	VTCSSCSSHINSVHVHISKSK -	VGWL IQLFSKKIESALRNKMN SQVCEKV	TNSVSSKLQPYFQTL P
Bpi_Human :	ITCSSCSSSDIADVEVDM SGD --	SGWLLNLFHQNQIESKFQKVLES RICEMIQKS VSSDLQPYLQTL P	
Lbp_Human :	GYCLSCSSSDIQNVELDIEGD --	LEELLNLLQS QIDARLRE REVLES KICRQI EEA VTAHLQPYLQTL P	
Lbp_Rabbit :	VTTSSCSSL RIRDLELHVSGN --	VGWL LNLFHQNQIESKLQKVLES KICEMIQKS VTS D LQPYLQTL P	
Lbp_Rat :	VTASGCNSFHKLLLHLQGEREP G WIKQLFTNF FISFTLKLVLKGQICKEI	-NVISNIMADFVQTRA	
Cetp_Human:	TDAPDCYLSFHKL LHLQGEREP G WIKQLFTNF FISFTLKLVLKGQICKEI	-NIISNIMADFVQTRA	
Cetp_Macfa:	TDAPDCYLA FHKL LHLQGEREP G WIKQLFTNF FISFTLKLILKRQVCNEI	-NTISNIMADFVQTRA	
Cetp_Rabbit:	TNAPDCYLA FHKL LHLQGEREP G WIKQLFTNF FISFTLKLILKRQVCNEI	-NTISNIMADFVQTRA	
Pltp_Human:	VSNVSCQASVSRMHA AFGGT --	FKKV YDFLSTFITS GMRF LLNQQICP VL YHAGTVLLNSLLD YVP	
Pltp_Mouse:	VSNVSCEAVSKMMA FGGT --	FRRM YNFFSTFITS GMRF LLNQQICP VL YHAGTVLLNSLLD TVP	
Lplc3_Rat :	LI LKRCNT ---	LLGHISLT -- SGLLPTPI FGLVEQTLCKVLPGLLC PVV - DSVLSVV NELLGATL	
Lplc4_Rat :	LVIERCDT ---	LLGGIKVKLLRG LPLPNLVDNLVN RVL ANVLPD LLLCP IV - DVVLGLVNDQLGLVD	
Consensus	C	i 1 C	t

**Figure 1** A section of a multiple sequence alignment of 12 sequences from the Bpi/Lbp/Lp1 superfamily of lipid-binding protein sequences. The 12 sequences belong to six different paralogous families, each of which binds a different lipid and performs a different physiological role, generally a transport function. Each sequence is identified by a label that identifies the paralogous family (they have a gene duplication event separating the families) before the underscore character in the protein name and the species after the underscore. The protein names are taken from the Swiss-Prot database. The annotation associated with each sequence in the Swiss-Prot database describes the lipid bound and the physiological role of the protein. The sequences within each paralogous family are orthologous (related to each other only by speciation events). The dash sequence character indicates an alignment column where either the amino acids replaced by dashes have been lost through a deletion event or amino acids shown by the single letter codes have been inserted into the sequence during the course of evolution. The section marked with a black background and white letters is a moderately well-conserved motif. Such conservation often marks regions important to the structure or function of the protein. Amino acids highlighted with black letters on a light gray background are identical within the highlighted family of orthologous sequences and have quite different chemical or physical properties to those in the same column in other families. They may mark positions that are responsible for the differences among the paralogous families.

position or combination of positions discriminate one subset from the others? Using a number of statistical techniques (e.g., multivariate analysis, group theory), they were able to develop a model for how tRNA sequences allowed the aaRS enzymes to identify the correct tRNA molecules and reject the incorrect ones (McClain 1995). Experimental verification (started prior to the computational work in one case) of these results were published subsequently (Hou & Schimmel 1988; McClain & Foss 1988).

### Additional use of the multiple sequence alignment

In addition to the above analysis of a well-crafted multiple sequence alignment, there are two additional areas that can use the information contained within a multiple sequence alignment. The first is the creation of a phylogenetic tree for evolutionary studies. The second is to allow more sensitive database searches using a representation that incorporates the pattern of substitution seen in the multiple sequence alignment to allow researchers to find more highly diverged homologous sequences.

### Pattern identification

Pattern identification has been developed to identify small, unique sections of the several unaligned sequences. Often, these contiguous regions of conserved residues, called motifs, are important for molecular interactions, such as regulatory regions or binding sites. Thus, motifs are often essential for the correct functioning of the molecule.

The classic example of pattern identification is to collect DNA sequences from the region just upstream (on the 5' side) of the coding region of a gene and examine these for a conserved pattern of nucleotides (Sadler et al. 1983) involved in regulating the transcription of the genes. What distinguishes this problem from global multiple sequence alignment is that outside the conserved patterns there is no expectation that the sequence is conserved and thus alignable.

Modern pattern identification programs (Bailey & Elkan 1994) make use of a modern statistical process designed to deal with the fact that we do not know where the patterns are located (expectation maximization) and a sophisticated sampling routine (stochastic sampling) that reduces the number of combinations that must be tried.

### Other techniques

As the biomedical sciences have expanded their repertoire of research methods and the kinds of data that can be collected, the field of bioinformatics has created techniques for dealing with these new kinds of data. The advent of the complete genome sequences for many organisms has been accompanied by software to allow the manipulation, annotation, analysis, and comparison of these large sequences. Complex mathematical models of genes try to find and identify all of the genes in each genome (Rogic et al. 2001).

Techniques have been developed to measure the change in expression for cDNA (microarrays) or the amount of proteins in cells over time or between mutants and wild types. It is not unusual for a research group to monitor thousands of molecules simultaneously; looking for either increases or decreases in the relative levels between the standard and the state under investigation. These large-scale experiments are being analyzed with a number of statistical techniques (Wetzel et al. 2000) such as analysis of variance, which produces a statistical model of the changes observed (Kerr & Churchill 2001). Other researchers are using multivariate statistical techniques to identify which molecules vary their presence in a coordinated manner in response to changing conditions. Interestingly, a number of these techniques were first developed many years ago to study the factors influencing crop growth.

### Conclusion

Ultimately, though, the field of bioinformatics does have some general themes that should continue to run throughout it in the future. First, the bioinformaticist tool chest is not complete – the tool chest of tomorrow will have only minimal relationship to today's set of tools, with better and more sensitive tools continuing to be developed. Second, the numbers and types of databases of experimental data will continue to expand at an alarming rate. The majority of the databases will be developed to describe one type of experimental data, like sequence data or microarray data, with only minimal references or consistencies (vocabulary) to the other databases. Third is that diverse data must be integrated across ranges of scale, both temporal and spatial. For example, a single point mutation in a mouse might cause kidney deformations that result in blood chemistry being incorrect. Thus, you have a single point mutation causing effects at the cellular and organ levels. Biological scientists must learn the techniques necessary to manage and make use of the new data resources that their research is creating.

### References

- Bailey, T.L., and C. Elkan. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *In: ISMB-94: Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology* (Altman, R., D. Brutlag, P. Karp, R. Lathrop, and D. Searls, eds), pp. 28–36. AAAI Press, Menlo Park, CA.
- Crick, F.H.C. 1957. The structure of nucleic acids and their role in protein synthesis. *In: Biochemical Society Symposium*, No. 14 (Crook, E.M., ed.), pp. 26–36. Cambridge University Press, Cambridge, UK.

- Dayhoff, M.O., R.V. Eck, M.A. Chang, and M.R. Sochard. 1965. Atlas of protein sequence and structure, Vol. 1. National Biomedical Research Foundation, Silver Spring, MD.
- Dayhoff, M.O., R.M. Schwartz, and B.C. Orcutt. 1978. A model of evolutionary change in proteins. In: *Atlas of protein sequences and structure*, Vol. 5 (Dayhoff M.O., ed.), pp. 345–352. National Biomedical Research Foundation, Washington, DC.
- Hou, Y.M., and P. Schimmel. 1988. A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature* 333:140–145.
- Kerr, K.M., and Churchill, G.A. 2001. Statistical design and the analysis of gene expression microarrays. *Genet. Res.* 77:123–128.
- McClain, W.H. 1995. The tRNA identity problem: Past, present, and future. In: *tRNA: Structure, biosynthesis and function* (Söll, D., and U.L. RajBhandary, eds), pp. 335–347. ASM Press, Washington, DC.
- McClain, W.H., and K. Foss. 1988. Changing the identity of a tRNA by introducing a G-U wobble pair near the 3' acceptor end. *Science* 240:793–796.
- McClain, W.H., and H.B. Nicholas Jr. 1987. Differences between transfer RNA molecules. *J. Mol. Biol.* 194:635–642.
- Nicholas, H.B. Jr., D.W. Deerfield II, and A.J. Ropelewski. 2000. Strategies for searching sequence databases. *Biotechniques* 28:1174–1191.
- Nicholas Jr., H.B., A.J. Ropelewski, and D.W. Deerfield II. 2002. Strategies for multiple sequence alignment. *Biotechniques* 32:592–603.
- Rogic, A., A.K. Mackworth, and F.B. Oullette. 2001. Evaluation of gene-finding programs on mammalian sequences. *Genome Res.* 11:817–832.
- Sadler, J.R., M.S. Waterman, and T.F. Smith. 1983. Regulatory pattern identification in nucleic acid sequences. *Nucleic Acids Res.* 11:2221–2231.
- Söll, D., and P.R. Schimmel. 1974. Aminoacyl tRNA synthetases. *Enzymes X*:489–538.
- Wetzel, A.W., J. Gilbertson, L. Zheng, et al. 2000. Three-dimensional reconstruction for genomic analysis of prostate cancer. *Proc. SPIE* 3905:204–212.

genes becoming more active while others become less active. Because more than one gene is usually involved in most biological processes, it would be best to examine many of the numerous life processes of an organism simultaneously, to see how they respond to changes over time. The microarray technology allows such an approach to be used to understand how an organism functions.

The rationale of **DNA microarray** (also called **DNA chips**, **genome chips**, **gene array**, and **biochips**) technology is that a large number of genes and their products (RNA, proteins) work together in a complex fashion to make an organism function as an integral whole. The technology allows researchers to adopt the “whole picture” approach in biological experimentation. That is, an ordered array allows the sum of all interactions across the full set of gene sequences to be measured simultaneously and calculated instantly.

The microarray technology is a tool for exploring the genome in a systematic and comprehensive fashion, to survey DNA and RNA variation. The underlying principle of the technology is specificity and affinity of complementary base pairs. The experimental methods for exploring attributes vary in complexity. Studying the differential expression at the mRNA level is relatively straightforward. Measuring the differential hybridization to a DNA microarray of fluorescently labeled

cDNAs prepared from the two mRNA samples can be used to compare the relative abundance of mRNA from each gene.

### Microarray fabrication

A DNA microarray consists of an ordered set of DNA molecules of known sequences, usually arranged in a rectangular fashion on a microscope slide or CMT-GAPS amino-salanized slide. Fabrication of DNA microarrays is usually automated and involves the use of high speed robotics. A spot of the DNA may be less than 200 µm in diameter and is placed at a precisely desired location. Each spot contains a specific sequence. The test material usually consists of RNA that has been amplified by using PCR methodology. A microarray set consists of four primary components: media or material, spotter (robotics analyzer), labeling and detection, and analytical software.

Labeling and detection are usually fluorescence-based systems. Each probe is labeled with a fluor of different color that is different enough to be distinguishable by “reading” devices equipped with optical filters. Readers (scanners) commonly use high intensity white light or laser-induced fluorescence that may be focused confocally or not.

### Steps in a DNA microarray experiment

To conduct a DNA microarray experiment, six general steps are followed:

- 1 Probe: the experimenter first selects the genetic material of known identity (e.g., cDNAs or small oligos) to be used as probes.
- 2 Fabrication of array: a fabrication format is selected.
- 3 Sample preparation: the sample (cDNA, RNA) to be used to interrogate the spotted and immobilized probes is prepared and fluorescently labeled.
- 4 Assay: the assay is then conducted.
- 5 Readout: the results are read by, for example, electronic devices.
- 6 Informatics: the data may be submitted to a variety of data management systems to obtain useful and desired information according to the objectives of the researcher.

### Applications of DNA microarrays in plant breeding

DNA microarrays may be used in two ways: (i) for sequence identification (normal genes and detection of mutations); and (ii) to examine gene expression (level or abundance of gene expression). One of the attributes of genes of great interest to microarray research is their expression. The expression pattern of a gene provides indirect information about its function.

The general steps followed in conducting a gene expression study are summarized as follows. Consider a study in which gene expression between two samples, A and B, are being compared.

- 1 Prepare fluorescently labeled cDNA of the total pool of mRNA from each cell population by reverse transcription in the presence of fluorescently labeled cDNA precursors. Use different fluors to allow distinction between their effects.
- 2 Mix two fluorescently labeled cDNAs.
- 3 Hybridize with a DNA microarray in which a distinct spot of DNA represents each gene. The cDNA sequences representing each individual transcript will hybridize with only the corresponding gene sequence in the array, regardless of the fluorescent labels.

The result is a pattern in which the relative abundance of the transcripts from each gene corresponds to the ratio of the two fluors used. It should be mentioned that gene expression data have certain limitations. For example, mRNA levels do not always reflect protein levels, and also, the expression of a protein may not always have a physiological consequence.

### Genetic use restriction systems

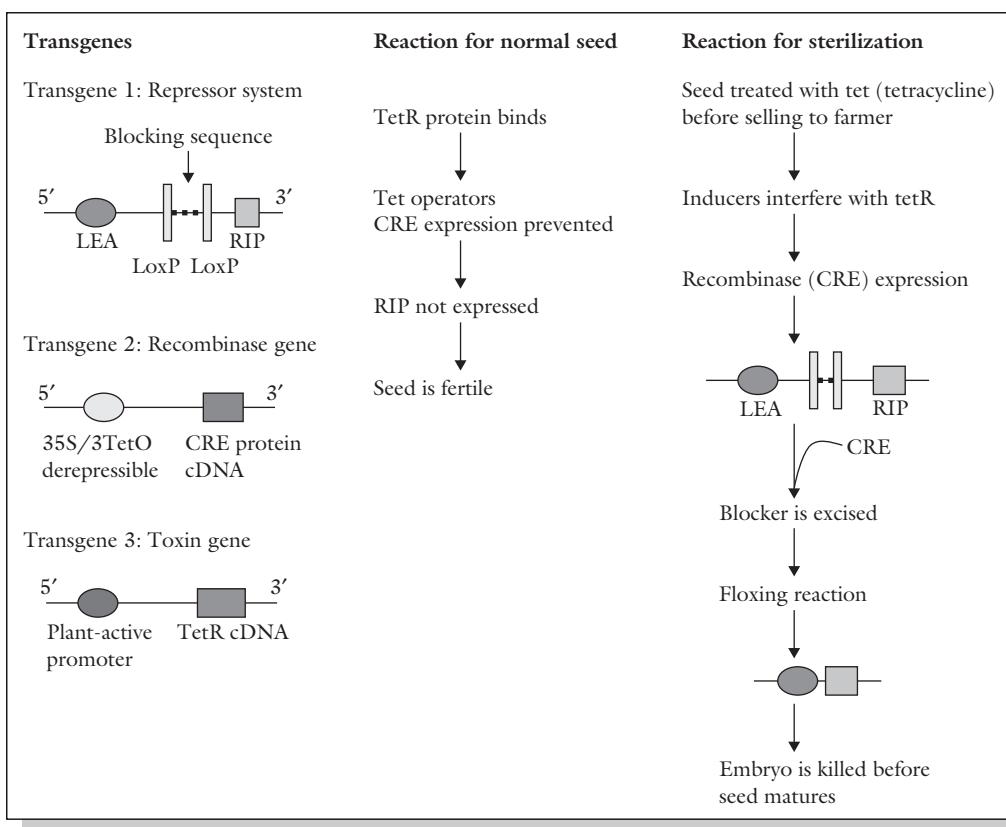
Plant breeders may protect their inventions (cultivars) by seeking patents or plant variety protection. However, the legal provisions are effective in protecting proprietary material from abuse only if they are enforced or the farmers are trusted to abide by the legal restrictions. Researchers have been working on protection systems that are self-regulatory, needing no policing for enforcement. The first of such systems was unveiled in 1998. Developed jointly by the US Department of Agriculture (USDA) and the Delta and Pine Land Company, the **technology protection system (TPS)** was awarded a patent in 1998. The nature of the patent allows each party to act independently from the other. The original genetic molecular switch was inserted into tobacco, and then cotton. Delta and Pine Land is the world leader in cotton seed production. Soon after the announcement, the technology was greeted by negative attacks from activist and other sources. The Rural Advancement Foundation International (RAFI) (now the ETC Group) described it in near derogatory terms as “**terminator technology**”, a term that appears to have stuck. To avoid this unscientific term, a new term was proposed and introduced in 1999, the **genetic use restriction system (GURT)**. The term is broadly used to describe the use of exogenous substances as inducers to control the expression of a plant’s genetic traits (e.g., trait for sterility, color, ripening, and cold tolerance). The restriction of a specific trait in a plant is called the T-GURT (also derided by activist as “traitor technology”); the V-GURT refers to the use of genetic engineering of plants to produce sterile seeds (i.e., the terminator technology).

### How the technology protection system works

TPS may be deployed in three basic steps:

- 1 The terminator genes are spliced into the genome of the target crop.
- 2 The seed company initiates the terminator process prior to selling the seed to farmers, by treating the seed with a substance (an inducer).
- 3 Farmers plant and harvest the seed in the usual way; however, the seed is sterile and will not germinate upon replanting.

Seed sterilization by TPS may be accomplished by one of three scientific approaches. Generally, all approaches use known gene mechanisms to control the expression



**Figure 14.2** A diagrammatic presentation of how the technology protection system (TPS) works as described by Marvin Oliver and colleagues. LEA (late embryogenesis abundant) promoter is derived from cotton. It is active only during the late stages of embryogenesis. RIP (recombinase inhibitory protein) cDNA is derived from soapwort. It inhibits cellular translation, resulting in cell death. Its expression is inhibited by the blocking sequence. LoxP sites are derived from bacteriophage P1. It comprises direct repeats recognized by CRE (recombinase gene) protein and mediates the removal of the blocker sequence. Column 1 shows three alternative genetic systems, while columns 2 and 3 show the application of only one genetic system – the repressor system – in normal seed and in inducing seed sterility.

of genes of known functions. Three genes with on/off switches are strategically engineered into the plant to interact in a predetermined sequence, the last gene becoming activated very late in seed development. At this stage, the gene is turned on by an inducer, causing a toxin to be produced that kills the embryo. The three proposed approaches to accomplishing seed sterility are described in Figure 14.2:

- 1 Transfection of target plant cells with three different transgenes.** Three different but functionally related transcriptional units are used – a repressor gene, a recombinase gene, and a toxin gene (Figure 14.2). The repressor gene codes for a protein called recombinase inhibitory protein (RIP). A DNA fragment that is a binding site to the repressor gene is

located between the promoter and the recombinase gene. In the absence of an exogenous inducer, the repressor binds to the binding site to prevent the plant from producing the recombinase protein. The toxin gene is controlled by a late promoter. A blocker sequence is located between this promoter and the toxin gene to interfere with the ability of the promoter to turn the gene on. An inducer is needed to release the recombinase enzyme that has the capacity to snip out the blocker gene to allow the late promoter to turn on the toxin gene.

- 2 Creation of a sterile hybrid.** In this approach, two fertile transgenic plants ( $A \times B$ ) containing two different sets of transgenes are developed. Plant A has a transcriptional unit made of the LEA (late embryogenesis abundant) promoter, LoxP sequences, blocking

sequence, and RIP coding sequence. Plant B consists of a promoter that is active during germination and a CRE coding sequence. When the seed from a cross of A × B is planted, the floxing reaction occurs. Because the LEA promoter is inactive after late embryogenesis, the expression of the RIP is restricted to the seeds of the resulting mature plants from the hybridization.

- 3 **The use of an inducible promoter.** This strategy is similar to using the hybridization process. However, in this case, the germination-specific promoter is replaced by a promoter that is controlled directly by an exogenous substance.

### Current status of GURT

Since the first GURT patent was awarded jointly to USDA and Delta and Pine Land, various entities, including universities and private corporations, have pursued the development of a variety of technologies for seed sterilization. These include Syngenta (with at least eight GURT patents), Dupont, Monsanto, BASF, Iowa State University, and Cornell University. The TPS technology has so far not been commercially exploited. However, it appears various companies are working towards this objective. The Convention on Biological Diversity continues to discuss the issue. Like every technology, there are those who see the promise of TPS and those who describe it in the most unflattering terms. Some of the stated potential advantages are:

- 1 TPS would be an incentive for further research and development of value-added cultivars.
- 2 It could possibly reduce the unintended gene flow from transgenic cultivars to conventional cultivars.
- 3 It could reduce the incidence of volunteer weeds.

Distracters counter that:

- 1 The only reason for developing and deploying the technology is to maximize the profits of seed companies.
- 2 Poor farmers cannot afford the seed; further, they cannot save seed to plant if they wanted to.
- 3 The protection provided lasts longer than any other similar protection system already in place.

### Molecular plant breeding

**Molecular breeding** may be defined as the use of molecular markers, in conjunction with linkage maps

and genomics, to select plants with desirable traits on the basis of genetic assays. The potential of indirect selection in plant breeding was recognized in the 1920s, but indirect selection using markers was first proposed in 1961 by Thoday. The lack of suitable markers slowed the adoption of this concept. Molecular breeding gained new momentum in the 1980s and has since made rapid progress, with the evolution of DNA marker technologies.

Molecular markers are used for several purposes in plant breeding.

- 1 **Gaining a better understanding of breeding materials and breeding system.** The success of a breeding program depends to a large extent on the materials used to initiate it. Molecular markers can be used to characterize germplasm, develop linkage maps, and identify heterotic patterns. An understanding of the breeding material will allow breeders to select the appropriate parents to use in crosses. Usually, breeders select genetically divergent parents for crossing. Molecular characterization will help to select parents that are complementary at the genetic level. Molecular markers can be especially useful in identifying markers that co-segregate with QTLs (quantitative trait loci) to facilitate the breeding of polygenic traits.
- 2 **Rapid introgression of simply inherited traits.** Introgression of genes into another genetic background involves several rounds of tedious backcrosses. When the source of desirable genes is a wild species, issues of linkage drag becomes more important because the dragged genes are often undesirable, requiring additional backcrosses to accomplish breeding objectives. Using markers and QTL analysis, the genome regions of the wild genotype containing the genes encoding the desirable trait can be identified more precisely, thereby reducing the fragment that needs to be introgressed, and consequently reducing linkage drag.
- 3 **Early generation testing.** Unlike phenotypic markers that often manifest in the adult stage, molecular markers can be assayed at an early stage in the development of the plant. Breeding for compositional traits such as high lysine and high tryptophan genes in maize can be advanced with early detection and selection of desirable segregants.
- 4 **Unconventional problem-solving.** The use of molecular markers can bring about novel ways of solving traditional problems, or solving problems traditional breeding could not handle. When linkage drag is recessive and tightly linked, numerous rounds of backcrosses may never detect and remove it. Disease resistance is often a recessive trait. When the

source of the gene is a wild germplasm, linkage drag could be difficult to remove by traditional backcross procedures. Marker analysis can help to solve the problem, as was done by J. P. A. Jansen when he introgressed resistance to the aphid *Nasonovia ribis-nigri* from a wild lettuce *Lactuca virosa* by repeated backcrosses. The result of the breeding was a lettuce plant of highly undesirable quality. The recessive linkage drag was removed by using DNA markers flanking the introgression to preselect for individuals that were recombinant in the vicinity of the gene.

The lifespan of new cultivars can be extended through the technique of **gene pyramiding** (i.e., transferring multiple disease-resistance genes into one genotype) for breeding disease-resistant cultivars. Marker-assisted backcross can be used to achieve this rapidly, especially for genes with indistinguishable phenotypes.

**5 Plant cultivar identification.** Molecular markers are effective in cultivar identification for protecting proprietary rights as well as authenticating plant cultivars. The types of molecular markers are discussed next.

### Molecular markers

Plant breeders use **genetic markers** (or simply markers) to study genomic organization, locate genes of interest, and facilitate the plant breeding process.

#### Concept of markers

Genetic markers are simply landmarks on chromosomes that serve as reference points to the location of other genes of interest when a genetic map is constructed. Breeders are interested in knowing the association (linkage) of markers to genes controlling the traits they are trying to manipulate. The rationale of markers is that an easy-to-observe trait (marker) is tightly linked to a more difficult-to-observe and desirable trait. Hence, breeders select for the trait of interest by indirectly selecting for the marker (that is readily assayed or detected or observed). When a marker is observed or detected, it signals that the trait of interest is present (by association).

Genetic markers can be detected at both the morphological level and the molecular or cellular level – the basis for classification of markers into two general categories as **morphological markers** and **molecular markers**. Morphological markers are manifested on the outside of the organism as a product of the interaction

of genes and the environment (i.e., an adult phenotype). On the other hand, molecular markers are detected at the subcellular level and can be assayed before the adult stage in the life cycle of the organism. Molecular markers of necessity are assayed by chemical procedures and are of two basic types – **protein** and **DNA markers**. Markers are indispensable in genetic engineering, being used in selection stages to identify successful transformation events.

#### Brief history of molecular markers

Chemical assays for **isozymes** (multiple forms of enzymes) marked the beginning of the practical application of molecular markers. The assays detect variations in protein products (products of translation) not variation in DNA *per se*. Their use is limited by the insufficient number of assays available (only about several dozen protocols exist), and their uneven distribution on the genetic map. **DNA markers** arrived on the scene with the discovery of **restriction fragment length polymorphisms (RFLPs)** in the 1980s, and are superior to both morphological and isozyme markers. Since then, several other molecular markers have been developed. DNA markers that are currently widely used include RFLP, AFLP (amplified fragment length polymorphism), SNPs (single nucleotide polymorphisms), and microsatellites or simple sequence repeats (SSRs). These markers are discussed next.

#### Classification of molecular markers

Molecular markers are classified in various ways, including a genetic basis and an operational basis. Various systems are used to assay molecular markers. There are **enzyme-based markers** (isozymes) and **DNA-based markers** (require hybridization between a probe and homologous DNA segment(s) within the genome). **PCR-based molecular markers** have the advantage of requiring small amounts of DNA and being relatively quick to assay. On the basis of genetic characteristics, molecular markers may be grouped into two general categories:

- 1 **Single-locus, multiallelic, codominant markers.** Examples are RFLPs and microsatellites (SSRs). Microsatellites are capable of detecting higher levels of polymorphisms than RFLPs.
- 2 **Multilocus, single-allelic, dominant markers.** Examples are AFLPs and RAPD (random amplified polymorphic DNA).

### Isozymes

Enzymes are macromolecular compounds that catalyze specific biochemical reactions. Most enzymes are proteins. Scientists have developed methods that allow the coupling of certain chemical reactions to the biochemical processes for colorimetric detection and location of specific enzymes. **Isozymes** are multiple forms of an enzyme that differ from each other by the substrate they act on, their maximum activity, or their regulatory properties. The term refers to enzyme polymorphisms that result from different loci. Another term, **allozyme**, is reserved for allelic enzymes.

As previously discussed, proteins can exist at one of four levels of structural complexity, of which the primary structure is the simplest. The most complex form, the quaternary structure, is attained through the folding and aggregation of polypeptide units. When an enzyme comprises one polypeptide chain, it is called a **monomer**. An enzyme comprising aggregates of polypeptide chains is called a **multimer** (or polymer). If an allozyme is multimeric, both homomers and heteromers will be produced in heterozygous individuals.

Isozyme technology has certain limitations, the major ones being the paucity of isozymes in plants and their tendency to be limited to certain chromosomes (not evenly distributed in the genome). Also, isozymes are sensitive to tissue type and age. However, the technology is inexpensive and relatively easy to apply. Some of the earlier successful applications were made in tomato (e.g., tagging of the *Aps-1* locus for acid phosphatase, and the exploitation of its linkage to nematode resistance). In spite of advances in molecular marker technology, isozymes are still used for certain purposes, such as the authentication of hybridity in hybrid development.

### Restriction fragment length polymorphisms (RFLPs)

RFLP markers are the first generation of DNA markers and one of the best for plant genome mapping. The RFLP variations are codominantly inherited. Mutation events (e.g., insertion, deletion) cause natural variations to occur in the genome. These variations may cause alterations (abolish) in the recognition sites for restriction enzymes. As a consequence, when homologous chromosomes are subjected to restriction enzyme digestion, different restriction products are produced upon electrophoresis (hence the term restriction fragment length polymorphisms). RFLPs are randomly distributed throughout the genome of an organism and may occur in both exons and introns. The DNA profiles or

fingerprints produced are specific to the combination of the restriction enzyme and probe (used to detect the polymorphism, using Southern blotting). Probes may be derived from random genomic DNA libraries, cDNA libraries, or minisatellites from other organisms.

One of the advantages of RFLPs is that the sequence used as a probe need not be known. All that a researcher needs is a genomic clone that can be used to detect the polymorphism. Very few RFLPs have been sequenced to know what sequence variation is responsible for the polymorphism. In the absence of sequence information, interpreting complex RFLP allelic systems may be problematic.

There are different types of RFLP polymorphisms, the simplest being the two-allele system involving the presence or absence of a recognition site for a single restriction enzyme. Screening reveals three different types of banding patterns: a large band (homozygous), two smaller bands (restriction site occurs on both homologues), and all three bands (heterozygous). It is assumed that a single base pair change within the recognition site will result in a chromosome that would either have the restriction site or not. In another allele system, one band corresponds to one allele. This system is also easy to score. One variable band corresponds to a homozygote. An individual inherits only two of the variant types of fragment sizes. Tomato was one of the first plant species to be characterized by RFLPs. The disadvantage of this marker system is that it is expensive and has low throughput.

### Random amplified polymorphic DNA (RAPD)

PCR is a technology discovered in 1986 for directly amplifying a specific short segment of DNA without the use of a cloning method. This eliminates the tedious process of repeated cloning to obtain ample quantities of DNA for a study. An attractive feature of a PCR-based marker system is that only a minute amount of DNA is needed for a project. Also, it has a higher throughput than RFLP. Because of the sensitivity of PCR technology to contamination, it is common to observe a variety of bands that are not associated with the target genome but are artifacts of the PCR condition. Consequently, certain bands may not be reproducible.

RAPD is a PCR-based marker system. In RAPD, the total genomic DNA is amplified using a single, short (about 10 bases), random primer. The PCR product is electrophoresed. This method yields high levels of polymorphism and is simple and quick to conduct. When using RAPD markers, using only the reproducible major

bands for identification may minimize its shortcomings. Further, one may include parental genomes where available to help determine bands of genetic origin.

### DNA amplification fingerprinting (DAF)

DAF is a variation of the RAPD methodology. It produces more variation than RAPD because it uses very short (5–8 bases) random primers. Because of the great capacity for producing polymorphisms, DAF is best used where plants are genetically closely related (e.g., used to distinguish among GM cultivars that differ only in transgenes). It is less effective for distinguishing among species of plants at a higher taxonomic level where genetic variation is already pronounced. The procedure can be made more effective and efficient by digesting the template DNA with restriction enzymes prior to conducting the PCR technique and optimizing the PCR environment for reproducible results.

### Simple sequence repeats (SSRs)

Repetitive DNA sequences are common in the eukaryotic genome. These short repetitive sequences are called **microsatellites** (or **variable nucleotide tandem repeats – VNTRs** – with tandem repeats of about 9–100). SSRs are random tandem repeats of 2–5 nucleotides (e.g., GT, GACA) that occur in microsatellites. The copy number of these repeats varies among individuals and is a source of polymorphism in plants. The SSR technique is also PCR based. Because the DNA sequences that flank microsatellite regions are usually conserved, primers specific for these regions are designed for use in the PCR reaction.

The SSR and RFLP techniques are more tedious to conduct, but they are more reliable than the RAPD and DAF techniques. These procedures require nucleotide information for primer design for the polymerase chain reaction, and sophisticated electrophoresis systems and computer software for accurate separation and scoring of bands.

### Amplified fragment length polymorphisms (AFLPs)

AFLPs are simply RFLPs visualized by selective PCR amplification of DNA restriction fragments. The technique uses primers that are 17–21 nucleotides in length and are capable of annealing perfectly to their target sequences (the adapter and restriction sites) as well as a small number of nucleotides adjacent to the restriction sites. This property of the AFLP technology makes it

very reliable, robust, and immune to small variations in PCR amplification parameters (e.g., thermal cyclers, template concentration). The technique also produces a high marker density. A typical AFLP fingerprint (the restriction fragment patterns generated by the technique) contains 50–100 amplified fragments, of which up to 80% may serve as genetic markers. Another advantage of the technology is that it does not require sequence information or probe collections prior to generating the fingerprints. This is particularly useful when DNA markers are scarce. The markers generated are unique DNA fragments (usually exhibit Mendelian inheritance) and are mostly monoallelic (the corresponding allele is not detected).

Some of the applications of AFLP markers include biodiversity studies, analysis of germplasm collections, genotyping of individuals, identification of closely linked DNA markers, construction of genetic DNA marker maps, construction of physical maps, gene mapping, and transcript profiling

### Sequence characterized amplified regions (SCARs) and sequence tagged sites (STSs)

SCAR and STS markers are derived from PCR-based markers by sequencing the ends of fragments to develop primers. SCAR markers are obtained by sequencing the ends of RAPD fragments, whereas STS markers are obtained by sequencing the ends of RFLP markers.

### Single nucleotide polymorphisms (SNPs)

An SNP is a single base pair site in the genome that is different from one individual to another. The more common the marker, the more likely it would be for scientists to discover difference among individuals in the population. SNPs are often linked to genes, making them very attractive subjects to study by scientists interested in locating, for example, disease genes. Sometimes, SNPs have no detectable phenotypic effect. However, in other cases, SNPs are responsible for dramatic changes.

### Introduction to genetic mapping: RFLP mapping

RFLP makers are inherited in a Mendelian fashion. Gene mapping simply entails obtaining a set of crossover frequencies between genes (see Chapter 3). Recombination between homologous chromosomes is the basis of

genetic mapping. As previously discussed, the recombination frequency between genes is a function of the distance between them. Consequently, a map distance is a function of the recombination that occurs during meiosis. Genetic maps are more useful in plant breeding if they are dense (i.e., the entire genome is covered with closely spaced markers). There are certain basic steps in mapping.

#### *Select parents and screen for polymorphism*

The first step in mapping is to select parents that are genetically divergent at many loci, and where desirable agronomic traits are segregating. It is important that useful polymorphisms be present in each pair of parents used for mapping traits of interest. The basic steps in the screening process are DNA extraction → restriction enzyme digestion → electrophoresis → Southern blotting → clone hybridization → RFLP detection. The probe and enzyme combination used for RFLP detection between the parents is used again on the F<sub>2</sub> population. The frequency of polymorphisms depends on the nature of the population. Usually, polymorphism frequency between parents may be about 80% per restriction enzyme for species that are cross-compatible, and about 20% per enzyme between varieties that are genetically distinct. Polymorphism between backcross varieties may be 0–1%. The researcher's goal is to find as many useful probes as possible (probes that can be used to distinguish among genotypes). Some probes detect RFLPs more frequently than others. About 5–10 restriction enzymes are used to screen about 500 parents to find about 100 useful probes.

#### *Generate a mapping population*

From the results of the parental screening, desirable parents are crossed to produce F<sub>1</sub> and F<sub>2</sub> plants. BC<sub>1</sub> populations may be generated, but F<sub>2</sub> populations give more information per sample. More advanced generations (F<sub>3</sub>–F<sub>x</sub>) may be used if agronomic traits are scored over different environments and over years. About 50–100 plants are mapped, more plants giving closer map distances.

#### *Score RFLPs*

DNA extracted from P<sub>1</sub>, P<sub>2</sub>, F<sub>1</sub>, and a random sample of F<sub>2</sub> plants are digested with a set of restriction enzymes. Several filters are obtained for each enzyme (following electrophoresis and blotting). A filter may be probed several to many times (10 times or more) with different

probes (washing between probes). About 100 or more probes may be made.

#### *Linkage analysis*

Segregation of RFLP markers in F<sub>2</sub> plants is used in constructing the linkage map. The markers are arranged in linkage groups (the linkage of markers to chromosomes is established). These analyses are made possible by using statistical packages and computer mapping software (MapMaker, Map Manager XP). Linkage relationship among markers with a recombination frequency of less than 50% LOD (logarithm of the odds, or LOD score) of 3 or greater indicates that a gene and an RFLP marker are linked. LOD is the ratio of two probabilities: no linkage/a certain degree of linkage. The gene and RFLP marker are first assumed to be unlinked. Then, the probability that the observed pattern of inheritance of the gene and the marker have a certain degree of linkage is calculated.

### **QTLs in plant breeding**

Quantitative traits are of a major concern to plant breeders because most of the traits of economic importance to breeders and consumers are quantitative traits.

#### **What are quantitative traits?**

These traits, as previously indicated, are controlled by polygenes or QTLs. QTLs are genetic factors that are responsible for a part of the observed phenotypic variation for a quantitative trait (polygenic trait). A QTL is similar to a gene, but in actuality, it merely indicates a region on the genome that may comprise one or more functional genes. QTL mapping entails an analysis of the association between observed trait values and the presence or absence of alleles of markers that have been mapped onto a linkage map. A QTL is declared to exist when it is determined that the association did not occur by some random process.

QTLs have come a long way from being statistical abstractions that breeders used to explain the performance of traits, to their current status where they are being isolated and cloned as DNA fragments, sequenced, and attempts are being made to use their allelic variation to elucidate them. A lot remains to be known about QTLs. It is not generally known if they are structural or regulatory genes or how many there are. Researchers are working to better understand the precise nature of the allelic variation, their genomic

location, genetic effects, and the interactions of these QTLs for better use in plant breeding. Knowledge of the location of a QTL of agronomic importance will allow breeders to target specific chromosome intervals in other species that have been less studied (the concept of synteny).

### Methods of QTL mapping

The basic approach to QTL analysis is to attempt to correlate their genetic variation in a specific quantitative trait with polymorphic genomic regions that are identified by molecular markers. Just like Mendelian genes, the degree of co-segregation of genes at different loci indicates the genetic distance between the loci of interest. In Mendelian analysis, there is a one to one correspondence between phenotype and genotype, making it relatively easy to construct a gene map based on the frequency of recombination. In the case of quantitative traits, the phenotype is the result of several to many genes, plus the environment, making the one to one relationship between phenotype and genotype not valid. Consequently, phenotypic variation only provides partial information about the segregation of the genes that control them, making it necessary to use statistical methods to acquire additional genotypic information about each QTL–marker relationship.

As summarized by M. J. Kearsey and Z. W. Luo, the three key requirements for mapping QTLs are: (i) trait phenotype; (ii) polymorphic markers; and (iii) genetic structure of populations. The phenotypic record of an individual for traits of interest reflects the genetic effects of QTL alleles present in the individual and the environmental contribution to the development of the trait. Molecular markers can be tracked and mapped like major genes, because they have uniquely identifiable effects. The third resource for mapping, the genetic structure of the mapping population, defines the domain in which the genes at a specific QTL segregate, as well as the pattern of recombination between the genes at linked loci. Appropriate statistical tools are used to bridge the relationship between the trait phenotype and the genotype at the genomic region that is specified by marker loci.

The choice of the mapping population is critical in QTL mapping. The breeder generates a segregating population by crossing lines with extreme phenotypic performance for the quantitative trait of interest. The most frequently used populations are derived from crossing two inbred lines that are assumed to be homozygous with different alleles at both QTLs and genetic markers. These materials include  $F_2$ s, backcrosses,

recombinant inbred lines, and doubled haploids. Strong linkage disequilibrium at marker loci and alleles of linked loci controlling the trait is needed, making the  $F_2$ , the generation with the strongest expression of linkage disequilibrium, the one most desirable to use.

Two basic types of analysis are used in QTL mapping to determine the association between different marker genotypes and their trait mean values.

#### *Single marker analysis*

The simplest way of identifying QTLs is to compare homozygous marker classes to the marker loci (i.e., compare the trait means of different classes for each marker locus individually in the form of a single factor analysis of variance (ANOVA); the error term in this procedure is the individual value, while the marker classes are the factor). Statistical methods commonly used for single marker analysis are the simple *t*-test, ANOVA, linear regression, and the maximum likelihood estimation. Single marker analysis has certain drawbacks (produces ambiguities regarding both location of QTLs and the estimates of their effects). Specifically, a significant association of trait mean with a specific marker indicates the presence of a QTL at or near the marker, but does not indicate on which side of the marker it is located.

#### *Interval mapping*

The method of flanking markers is based on the hypothesis that a QTL lies between linked marker loci. The statistical package, MapMaker/QTL developed by Eric Lander and colleagues, is used for this analysis. Several modifications have been proposed to interval mapping that consider a single segregating QTL on a chromosome. This has drawbacks. By assessing the likelihood for a single putative QTL at each map location on the genome separately, the method ignores the effects of other mapped or yet to be mapped QTLs. The hypothesis of one putative QTL is unrealistic because a large number of genes distributed throughout the genome are suspected to be involved in quantitative trait expression. Subsequently, methods for mapping multiple QTLs simultaneously have been proposed.

### **Marker-assisted breeding**

Molecular markers may be used in several ways to make the plant breeding process more efficient. The adoption of a **marker-assisted selection** (MAS) or **marker-aided**

**selection** in a breeding program hinges on the availability of useful molecular markers. Fortunately, this resource is becoming increasingly available to many species, thanks to the advances in biotechnology. This breeding approach is applicable to improving both simple and complex traits, as a means of evaluation of a trait that is difficult or expensive to evaluate by conventional methods. The basic requirement is to identify a marker that co-segregates with a major gene of the target trait. MAS is more beneficial to breeding quantitative traits with low heritability.

The key steps in the implementation of MAS in breeding QTLs are summarized by D. W. Podlich and his colleagues as follows:

- 1 Creation of a dense genetic map of molecular markers.
- 2 Detection of QTLs based on statistical association between markers and phenotypic variability.
- 3 Definition of a set of desirable marker alleles based on the results of the QTL analysis.
- 4 The use and/or extrapolation of this information to the current set of breeding germplasm to enable marker-based selection decisions to be made.

However, as these researchers noted, MAS for more complex traits is still challenging, partly because of the difficulty of effective detection, estimation, and utility of QTLs and their effects. The problem is more significant with more complex traits (such as grain yield) that are controlled by many genes under the influence of epistasis (gene-by-gene interaction) and gene-by-environment ( $G \times E$ ) interaction effects. On the contrary, most researchers engaged in evaluations of mapping and MAS tend to assume that QTLs act independently (i.e., no interaction with other genes and/or environment).

To overcome this problem, Podlich and colleagues proposed a new approach to MAS from the conventional one, which assumes that desirable QTL alleles, once identified, will remain relevant throughout many cycles of selection during plant breeding. In the conventional procedure, researchers tend to estimate QTL effects at the beginning of the project and continue to apply the estimates to new germplasm created during the breeding process (i.e., mapping start only). The assumption of fixed QTL values is appropriate if the traits are controlled by additive genes. This will allow MAS to be conducted by independently assembling or stacking desirable alleles. This assumption is not applicable to situations in which context dependencies (changes in genetic background) occur. On such occasions, the value of QTL alleles can change depending on

the genetic structure of the current set of germplasm in the breeding program. In other words, QTL values will change over the cycles of selection, as the background effects change. These progressive changes in genetic structure may make the initial combinations of alleles no longer the best target or no longer significant in increasing the trait performance in future breeding cycles.

Podlich and colleagues proposed the “mapping as you go” approach to MAS of complex traits. QTL effects are cyclically re-estimated each time a new set of germplasm is created. This ensures that the basis of MAS remains relevant to the current set of germplasm.

S. D. Tanksley and T. C. Nelson developed a procedure, **advanced backcross breeding**, for the simultaneous discovery and transfer of desirable QTLs from unadapted germplasm into elite lines. Basically, this procedure postpones QTL mapping until the  $BC_2$  or  $BC_3$ , applying negative selection during these generations to reduce the occurrence of undesirable alleles from the donor (unadapted genotype). The advantage of this strategy is that  $BC_2/BC_3$  provides adequate statistical power for QTL identification, while at the same time being sufficiently similar to the recurrent parent to allow selection for QTL-NIL (near-isogenic line) in a short time (1–2 years). The QTL discovered can be verified and the NILs used directly as improved cultivars or as parents for hybrid breeding.

### Important applications of molecular markers in plant breeding

Molecular markers may be used by breeders to increase the rate of genetic gain, verification or identification of parentage, characterization of germplasm, and quantification of genetic variability.

- 1 **Screening single traits.** Molecular markers associated with many agronomic traits of importance (e.g., nematode resistance in sugar beet, blast resistance in rice) have been discovered and used for screening for these traits in breeding. MAS is useful for simple traits, or those for which genetic gain per unit time (rate of genetic gain) is of high economic return.
- 2 **Speeding up breeding programs.** Markers add speed and precision to backcross breeding programs. They help to identify the gene of interest to be transferred and facilitate the elimination of the undesirable genome in the donor parent (reduces linkage drag). A fewer number of backcrosses are needed to recover the genotype of the adapted parent.

**3 Germplasm evaluation.** Markers can distinguish homology and analogy. They can provide a common measure of assessing divergence in germplasm. Molecular markers have been used to assist in the discovery of heterotic patterns in order to guide the selection of parents for use in a hybrid breeding program (e.g., in maize). Marker analysis (e.g., by Roger's distance) may be used to ascertain genetic dissimilarity among lines of different heterotic groups to enable the breeder to predict the performance of hybrids to be developed from different intergroup crosses.

**4 Cultivar identification and protection.** Molecular markers may be used to fingerprint genotypes for the protection of proprietary materials.

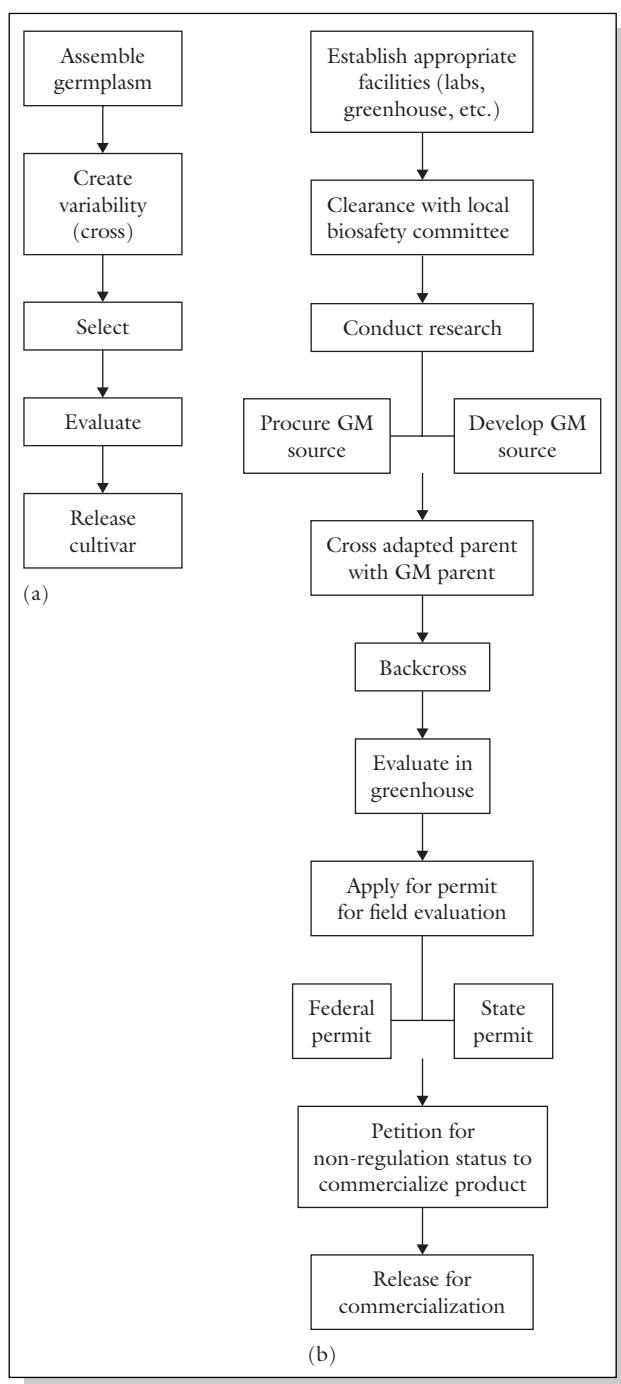
### Breeding genetically modified cultivars

Steps in the breeding of conventional cultivars are summarized in Figure 14.3a. The breeder starts by assembling germplasm to create the base population. Selection is practiced in the segregating population to identify and advance the more desirable genotypes, which are evaluated to identify and release the most promising one as a commercial cultivar. A breeding program can be implemented without oversight by any entity, except at the stage of certification, when certifying agencies inspect the product for conformity with set standards.

On the other hand, breeding genetically modified (GM) cultivars is a highly regulated activity, from inception to completion of the project (Figure 14.3b). Before biotechnology development and application activities can proceed in a country, there has to be a national biosafety policy in place (see Chapter 16). Furthermore, the institution at which the breeding program is to be conducted must also have its own institutional research policy guidelines, prescribing procedures for conducting research. Some of the guidelines pertain to the use of hazardous materials. There are specific guidelines for rDNA research (some variation may occur from one institution to another).

#### Clearance

The first step in a genetic engineering project in the USA is to submit the proposal to the local Institutional Review Board and other specialized committees. (e.g., Institutional Biosafety Committee). These bodies review the project materials and methods for scientific quality and compliance with established safety and ethical



**Figure 14.3** A comparison of the general steps involved in breeding cultivars by (a) the conventional method and (b) the use of genetic engineering technology. The specific steps vary among breeders and situations.

guidelines. There has to be some provision for containment and/or disposal of hazardous materials. Depending on the nature of the project, a biosafety level is assigned by the committee according to federal guidelines.

The biosafety levels and their restrictions are as follows:

- BL1-P** Basic containment level: restricted access to greenhouse; insect, weed, and rodent control mechanisms required; screens recommended.
- BL2-P** For agents of moderate potential hazard: BL1-P requirements; plus concrete floor; screens restricting movement of small insects but not pollen; autoclave to sterilize transgenic materials before removal.
- BL3-P** For agents of serious potential hazard: BL2-P requirements; plus collection and sterilization of liquid runoff; sealed windows; ventilation filters; security fence; protective clothing.
- BL4-P** For work with extremely hazardous agents, including certain exotic plant pathogens: similar to BL3-P but more stringent.

#### *Conduct research*

Large seed companies such as Monsanto and Pioneer develop their own GM breeding stock. They transfer the gene of interest into the appropriate genetic background. Other breeding programs may procure the cloned gene or genetic stock at a fee, for use in their projects. Developers of transgenic breeding lines must evaluate them initially for:

- 1 Activity of the introduced gene (transgene).
- 2 Stable inheritance of the gene.
- 3 Unintended effects of the growth and development of the plant.

#### *Hybridize (cross)*

Once an appropriate GM breeding line has been obtained, it is crossed with a conventional adapted cultivar to transfer the transgene into the desirable cultivar.

#### *Backcross*

The first cross is followed by cycles of repeated backcrosses to the adapted cultivar to recover the traits of the parent as much as possible. All these activities are conducted under a restricted environment (greenhouse).

#### *Evaluation*

The breeder evaluates the success of the project in the greenhouse. This may include evaluating the proper expression of the transgene, yield and quality, and overall performance of the new product.

#### *Field testing*

When the breeder is satisfied that the breeding project is successful, the next step is to apply for permission to field test the cultivar. An application to transport or field test a transgenic plant is submitted to the Animal and Plant Health Inspection Service (APHIS). Under the Federal Plant Pest Act, APHIS must determine if a transgenic plant variety has the potential to become a pest (see Chapter 15).

Before permission is granted for field testing, some of the basic criteria to be satisfied by the breeder include presenting the following pieces of evidence to show:

- 1 Stable integration in the plant chromosome of the transgene.
- 2 Non-pathogenicity to animals or humans.
- 3 Unlikely to be toxic to other non-target organisms.
- 4 Low risk of creating new plant viruses.

Field testing, if permission is granted, is conducted at a number of locations and over several years. The breeder should conduct the test in a manner that will not permit contamination of the environment or food supply system. At the end of the evaluation, a comprehensive report is submitted to APHIS, including data on the gene construct, effects on plant biology, effects on the ecosystem, and spread of the gene to other species. Depending on the product and its intended use, other federal agencies (Food and Drug Administration, Environmental Protection Agency) may be involved in the process of field evaluation (see Chapter 15). In addition to satisfying federal agency requirements, some states may have their own regulations that must be satisfied for field testing of GM varieties.

#### *Commercialization*

Once the field evaluation is successfully completed, the breeder may apply for non-regulatory status to commercialize the cultivar. If successful, the cultivar may be released as a commercial cultivar to producers. Of course, the breeder may seek appropriate protection of the invention through patent application or other appropriate measures.

## Engineering pest resistance

To date, the most widespread use of genetic modification in practical breeding is in the development of

pest-resistant cultivars, specifically, the *Bt* cultivars (e.g., in corn, cotton), and the development of herbicide tolerance (specifically, Roundup Ready® products). These applications are discussed in Chapter 20.

## References and suggested reading

- Acquaah, G. 1997. Practical protein electrophoresis for genetic research. Dioscoredes Press, Portland, OR.
- Barcelo, P., S. Rasco-Gaunt, C. Thorpe, and P.A. Lazzeri. 2001. Transformation and gene expression. In: Advances in botanical research incorporating advances in plant pathology, Vol. 34, Biotechnology of cereals (Shewry, P.R., P.A. Lazzeri, K.J. Edwards, and J.A. Callow, eds). Academic Press, New York.
- Dawson, E. 1999. SNP maps: More markers needed? Mol. Med. Today 10:419–420.
- Devos, K.M., and M.D. Gale. 2000. Genome relationships: The grass model in current research. Plant Cell 12:637–646.
- Dunwell, J.M. 2000. Transgenic approaches to crop improvement. J. Exp. Biol. 51:487–496.
- Glick, B.R., and J.J. Pasternak. 2002. Molecular biotechnology, principles and applications of recombinant DNA. ASM Press, Washington, DC.
- Goff, S.A., D. Ricke, T. Lan, et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. spp. *japonica*). Science 296:92–100.
- Goodman, M.M. 2005. Plant breeding requirements for applied molecular biology. Crop Sci. 44:1913–1914.
- Harvey, M.J. 2005. Application of genomic technology to crop plant: Opportunities and challenges. Crop Sci. 44:1893–1909.
- Jannick, J.L. 2005. Selective phenotyping to accurately map quantitative trait loci. Crop Sci. 45:901–908.
- Knapp, S.J. 1998. Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Sci. 38:1164–1174.
- Lander, E.S. 1999. Array of hope. Nature Genet. (Suppl.) 21:3–4.
- Lemieux, B., A. Aharoni, and M. Schena, 1998. Overview of DNA chip technology. Mol. Breed. 4:277–289.
- Lockhart, D.J., and E.A. Winzeler. 2000. Genomics, gene expression and DNA arrays. Nature 405(6788):827–836.
- Newbury, H.J., and A.H. Paterson. 2000. Genomic colinearity and its implications in crop improvement. In: Plant molecular breeding (H.J. Newbury, ed.). Blackwell Publishing, Oxford and CRC Press, Boca Raton, FL.
- Podlich, D.W., C.R. Winkler, and M. Cooper. 2004. Mapping as you go: An effective approach for marker-assisted selection of complex traits. Crop Sci. 44:1560–1571.
- Powell, W.M., M. Morgante, C. Hudre, M. Hanafey, J. Vogel, S. Tingey, and A. Rafalski. 1996. The comparison of RFLP, RAPD, ALFP, and SSR (microsatellites) markers for germplasm analysis. Mol. Breed. 2:225–238.
- Xiao, J., J. Li, L. Yuan, S.R. McCouch, and S.D. Tanksley. 1996. Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. Theor. Appl. Genet. 92:637–643.
- Zhong, G.-Y. 2001. Genetic issues and pitfalls in transgenic breeding. Euphytica 118:137–144.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Isozymes markers are DNA markers.
- 2 AFLP is a PCR-based marker.
- 3 SCAR markers are derived from RFLP markers.
- 4 RFLP is a multilocus marker.
- 5 CaMV 35S is a constitutive promoter.
- 6 A transgene cannot function without a promoter.
- 7 Restriction enzymes have a common recognition site.

**Part B**

Please answer the following questions:

- 1 Give the full name for each of the following acronyms:  
RFLP: .....  
AFLP: .....  
SCAR: .....
- 2 Give two specific disadvantages of isozyme markers.
- 3 Give two specific applications of molecular markers in plant breeding.
- 4 What is marker-assisted selection?
- 5 What is a transgenic organism?
- 6 What is a transgene?
- 7 Describe how microprojectile bombardment is used in biotechnology.
- 8 Give two examples of scorable markers used in biotechnology.
- 9 What is a promoter?
- 10 Discuss the types of promoters used in genetic engineering.
- 11 What is genomics?
- 12 What is bioinformatics, and what is its role in biotechnology?

**Part C**

Please write a brief essay on each of the following topics:

- 1 Give the rationale of markers.
- 2 Discuss the use of markers in plant cultivar identification.
- 3 Discuss the advantages of AFLP technology.
- 4 Discuss how breeding GM cultivars differs from breeding conventional cultivars.
- 5 Describe gene transfer by *Agrobacterium* mediation.
- 6 Give the general characteristics of restriction endonucleases.
- 7 Discuss plasmid vectors used in rDNA research.
- 8 Describe the method of gene transfer by biolistics.



# Issues in the application of biotechnology in plant breeding

---

## Purpose and expected outcomes

*Intellectual property rights are necessary to preserve the competitive edge over the competition by a company that owns an invention. They also protect a valuable resource that the property owner can later license to a third party for profit. The development and application of biotechnology raises ethical questions, some of which are serious enough to generate significant opposition from the consuming public. Breeders need to be aware of the local and international issues pertaining to public acceptance and rights and regulations affecting the biotechnology industry. The purpose of this chapter is to discuss intellectual property and ethical issues associated with the breeding of plants, especially as they pertain to the use of biotechnology. After studying this chapter, the student should be able to:*

- 1 Discuss the concept and importance of intellectual property in plant breeding.
  - 2 Explain what a patent is, discuss the types of patents, and what can be patented.
  - 3 Discuss patent infringement.
  - 4 Discuss how ethics impacts the development and application of plant biotechnology.
  - 5 Discuss the agencies and their specific roles in the regulation of the biotechnology industry.
  - 6 Discuss international biotechnology regulation issues.
  - 7 Discuss public perception of biotechnology and its implication to breeding.
- 

## Concept of intellectual property

**Intellectual property** consists of principles that a society observes to ensure that an inventor is protected from unfair use of his or her invention by others. To achieve this, a variety of legal provisions are made to protect against improper use of another's original ideas and creations. The most common of such provisions are **copyrights**, **confidential information**, **breeders' rights**, **trademarks**, and **patents**.

### Copyrights ©

Copyright protection is commonly sought by people or entities for the protection of such things as aesthetic

creations, music, painting, works of literature, and computer software. It is of little use to plant breeders except for the protection of published research results, or inventions such as computer-based tools.

### Confidential information

This protection is limited in its scope and pertains to an organization. Its success depends on the extent people can be trusted to keep a secret. The term "confidential information" is used to apply to the variety of strategies used by companies to protect their unpatented inventions, in the hope that it will not be leaked into the public domain. The common strategies include **trade secrets** and **proprietary information**. An invention

may not be patented because it is unpatentable, or perhaps the inventor may choose not to, or for financial reasons. Keeping trade secrets is cheaper than applying for a patent. To ensure that employees who know the privileged information do not divulge it for any purpose, some companies enforce a policy that all employees sign a confidentiality agreement.

#### *Breeders' rights*

This is one of the common protections sought by plant breeders. It was specifically designed for the benefit of plant and animal breeders to protect them from the unlawful use of their creation by other scientists, competitors, or producers. Exclusive rights to the sale and multiplication of the reproductive material may be obtained, provided the cultivar is new and not previously marketed, is different from all other varieties, and is uniform (all plants in the cultivar are the same) and stable (remains the same from one generation to the next).

#### *Trademarks™*

Pharmaceutical companies use trademark protection extensively. Trademarks are important to all businesses.

#### *Patents*

Plant breeders protect their inventions by seeking patents. The right to patent an invention is one of the most widely applicable distinct rights provided by intellectual property.

### **Patents**

#### **Definition**

A **patent** may be defined as an exclusive right granted for an invention of a product or process that provides a novel way of doing something, or offers a new technical solution to a problem. The owner of a patent is the **patentee**. A patent provides protection for the invention to the patentee for a predetermined period. A patent may be described as a "negative right" because it confers upon the patentee the right to exclude others from commercially exploiting the invention without the owner's authorization. The protection stipulates that without permission, no one should make, use, sell, offer for sale, or import the invention. The duration of such exclusive rights is 20 years, after which the invention is released

into the public domain. It should be pointed out that while a patent prohibits unauthorized use by others, it does not authorize limitless use by the patentee. A patent is not a "positive right" because it does not empower or obligate the patentee to do something that he or she would otherwise be prohibited from doing. In other words, in making, using, or selling the invention, the patentee must operate within the limits of existing laws of society. Further, it goes to emphasize the fact that an inventor is under no obligation to patent an invention in order to commercially exploit it. However, without a patent for protection of the invention, a third party can commercially exploit the invention without authorization and with no legal consequence.

#### **Importance of patents to society**

Patents are so pervasive in society that it is nearly impossible to find anything created by humans that is not directly or indirectly associated with a patent. As the US constitution indicates (US Constitution, Article 1, Section 8, cl. 8) the **patent law** serves to "promote the progress of science and the useful arts". In exchange for exclusive rights to the invention, the patentee is obligated to disclose the invention and provide information that expands the existing technical knowledge base. In other words, whereas others cannot imitate the invention during the life of the patent, they can utilize the divulged information toward further innovations, advancing science and arts to enhance the quality of life of humans. Patents are not only prestigious to own, but they have potential financial value. They provide effective incentives for creativity and innovation, recognizing the achievement of the inventor and in cases, where the invention can be commercialized, financial reward. Because companies invest large sums of money in the R&D (research and development) of new products, patents ensure that they can enjoy a monopoly for a period in which they hopefully recoup their investments, at the least. Patents are hence strong incentives for continued R&D. It goes without saying that for-profit breeding companies invest resources in improving crop species with high market potential. Hence, crops of low market potential but of high social value (e.g., important to poor developing countries) are often ignored (the so-called "orphaned crops").

#### **What can be patented?**

Patent law specifies what can be patented and the conditions under which a patent can be granted. One of

the key functions of a patent is to define the scope of the protection. Ideas and suggestions, no matter how brilliant and creative, cannot be patented. Similarly, mixtures of ingredients (e.g., medicines) are also excluded, unless such mixtures produce synergistic effects or some unique and unexpected advantage. Defining the scope of protection is straightforward in certain cases (e.g., a simple device) and very complex in others (e.g., biotech inventions). There are five basic classes of patentable inventions:

- 1 Compositions of matter.** A new chemical entity produced from the combination of two or more compounds (common in pharmaceutical and agrochemical research).
- 2 Processes or procedures.** A series of steps that are followed to synthesize a new compound or a new method of making a product.
- 3 Articles of manufacture.** Nearly every manmade object.
- 4 Machines.** Any mechanical or electrical apparatus or device.
- 5 Improvements.** Improvements on any of the previous four categories.

### Types of patents

Patents may be classified into three basic types – **utility**, **design**, and **plant**. The duration of each of these patents is 20 years.

- 1 Utility patent.** A utility patent is one of the most common and also most difficult patents to obtain. The applicant is required to submit a comprehensive description of how to make and use the invention, including detailed drawings, where appropriate, among other requirements. The scope of protection includes the functional characteristics of machines, electronic devices, manufacturing processes, chemical compounds, composition of medical treatment, and manufactured articles.
- 2 Design patent.** A design patent may be sought for the protection of the shape of ornamental or artistic features of an article (e.g., unique shape of a bottle, the grill of an automobile).
- 3 Plant patent.** The plant patent protects the invention or discovery of distinct and new plant varieties via asexual reproductive methods.

### National and international patents

Patent laws vary among the nations of the world. National patents pertain to the applicant's country of

origin or operation. However, applicants may also seek worldwide protection for an invention (e.g., from the European Patent Convention (EPC)). The EPC allows patent rights to be obtained in one or more European countries that are party to the EPC by making a single European patent application. If successful, the applicant can then use the patents in the countries that the applicant designated. **International patents** may be obtained by filing a single patent and designating the countries in which protection is desired. Within a specified time limit, this single patent will be copied to the designated countries where they will be processed as national applications. Countries vary in the diligence with which they honor and enforce patent laws. Patent violations cost companies huge losses each year. Sometimes, it takes political pressure to encourage the enforcement of international patents.

### Scope of patent protection

The applicant is responsible for defining the **scope of protection** desired. However, there must be ample evidence to justify the scope of protection being sought. The scope of protection may be defined narrowly or broadly, each scenario with its consequences. Sometimes, in an attempt to prevent the competition from copying the invention, an applicant may make claims to encompass the embodiments of the invention that were not existent at the time of the invention. Such a practice favors the inventor to the disadvantage of the competition. An example of a broad scope of protection is one sought by Monsanto to protect its highly successful herbicide, Roundup®.

### Criteria for patentability

- 1 Conception.** The applicant should be able to paint a mental picture of the invention that is detailed enough to allow a person knowledgeable in the subject to which the invention relates to make and use the invention.
- 2 Reduction to practice.** The inventor should make or construct the invention and test it to demonstrate its usefulness.
- 3 Utility.** An invention must be useful (not merely aesthetic) to the user. That is, the invention must take some practical form (applicability).
- 4 Novelty.** The invention must not be a copy or repetition of an existing one. Among other things, it should not have been known, published, or used publicly anywhere previously. In the UK, an invention must not have been in the public domain anywhere in the world prior to application of a patent. Some countries

(e.g., USA) allow a grace period during which an invention that has already been introduced to the public, under certain conditions, could still be patented.

**5 Obviousness.** This is a very difficult criterion to satisfy. The invention or product should neither be expected nor obvious. A person knowledgeable in the subject matter (skilled in the art) should not be able to readily figure out how to piece together the component parts to make the product (Section 103 of the US Patent Law).

### Applying for a patent

In the USA, patent applications may be submitted to the Patent and Trademarks Office, US Department of Commerce, Washington, 20231 DC. Copyright applications may be submitted to the US Copyright Office, Library of Congress, Washington, 20559 DC. In Europe, applications may be submitted to the European Patent Office. The services of patent attorneys may be engaged in preparing and filing an application. An applicant must define the problem or objective addressed by the invention very clearly and in detail. There are certain general steps involved in applying for a patent:

- 1 Filing fee.** An applicant for a patent is required to pay a fee for the processing of the application.
- 2 Search and examination.** The patent examiner will conduct a “prior art” search to ascertain the novelty and non-obviousness of the invention. The claims defining the scope and monopoly being sought are rigorously examined. The examiner may accept or reject the application based upon the search and examination results. The applicant has the right to argue against an adverse judgment, or to amend the claims to the satisfaction of the examiner.
- 3 Publication.** Successful applications will be published along with the claims of the applicant.
- 4 Maintenance fees.** Most countries require a successful applicant to pay a periodic maintenance fee to guarantee or prevent the patent from lapsing.

### Exploiting intellectual property

Apart from exploiting an intellectual property personally, the owner of the patent has the right to sell, license, or give it away.

#### Assignment

A patent or patent application of an invention may be sold or assigned to another party just like a piece of property. The method of payment for use or acquisition

of the property is entirely up to the owner of the property (e.g., one-time payment or periodic royalties).

#### License

Rather than outright sale, the intellectual property may be licensed to another party according to specified terms and conditions. The license agreement usually guarantees the licensor royalties from the licensee for use of the invention. License agreements lay out limitations to the extent of exploitation of the invention permitted, and may limit the application of the invention to certain uses. In this era of genetic engineering, a wide variety of genes have been cloned. Also, various biological tools (e.g., vectors, promoters) have been developed. Some of those undertakings were expensive and tedious. Rather than reinvent the wheel, it might be easier to pay a fee to use the technology.

#### Freedom of use

**Freedom of use** consideration may hinder a patent from being exploited without infringing upon existing patents. Such restrictions on free exploitation of an inventor’s patent may derive from the scope of the patent. Scopes that are too narrow are susceptible to such infringements. An example of such a situation involving freedom of use may arise when an inventor patents a process that requires another patented compound in order to exploit the invention.

### Patents in plant breeding and biotechnology: unique issues and challenges

#### Patenting organisms

Patenting in the biological sciences or non-empirical sciences is very challenging, especially when life is involved (e.g., organisms). Whereas it is not easy to duplicate and proliferate a mechanical invention, it is easy to reproduce an organism (just like computer software or copyrighted music). Consequently, the early application of patent laws to biology tended to be strongly and broadly interpreted in favor of inventors.

Generally, it is easier to satisfy the patent requirements when inventions are the results of empirical discovery as obtains in the pharmaceutical and agro-chemical industries. It is easy to lay claim to the discovery of a compound with potential for use as an active ingredient in pesticides, antibiotics, and other therapeutics.

The development of truly new and unexpected phenomenon is not common. Progress is made incrementally. It is often a challenge to satisfy the traditional criteria stipulated by the patent office – obvious phenomenon, specific utility, and teaching others how to make and use the invention. Another limitation is that ideas and properties of nature are not patentable.

### Patenting hereditary material

The turning point in patenting genes and other biological resources occurred in 1980, with the US Supreme Court decision in Diamond vs Chakrabarty to grant a patent for an oil-dissolving microbe. The technologies of genetic engineering and genomics have resulted in the discovery of millions of genes and fragments of genes (expressed sequence tags or ESTs) that have been submitted for patenting. However, not all players are satisfied with the scope of protection provided by the patent laws. A microscopic view will allow nearly anything novel to be patentable, while opening up the doors for competitors to easily circumvent the narrow claims. Some scientists are opposed to the granting of broad patents to what they describe as the early stages of the biotechnology game. Some of the genes submitted for patents have not been characterized, neither have the applicants determined their functions and specific uses. The concern is that large-scale and wholesale patenting of genes by biotechnology researchers or companies, who have no clue about the functions of these genes, is tantamount to staking a claim to all future discoveries associated with those genes (the so-called “**reach-through patents**”).

This concern is a genuine one. Until recently, genomic companies had a field day staking claims to the genome landmark (the “genome run”). But with the focus now on understanding gene function, the proteomic companies now have their chance to do likewise. This is stirring up new controversies in the patenting of biotechnology inventions.

Another issue with biotechnology patents is “**patent stacking**”, a situation in which a single gene is patented by different scientists. This situation is not favorable to product development because users are deterred by the possibility that they would have to pay multiple royalties to all owners of the patent. Further, because patent applications are secret, it is possible for an R&D team in a different company to be working on the development of a product only to be surprised at a later date by the fact that a patent (called a **submarine patent**) has already been granted.

### Patenting proteins

A patent on a specific DNA sequence and the protein it produces may not cover some biologically important variant. It is estimated that the top genomic companies have collectively filed over 25,000 DNA-based patents. The business rationale to their strategy includes the potential to receive royalties from third parties that use any of them. But this may not be as simple as it sounds, unless one gene makes one mRNA, which in turn makes one protein – something that is not true anymore (see Chapter 3, p. 50).

It is most likely and perhaps inevitable that some protein discovery projects will turn out proteins that correlate better with a disease than those for which patent claims are already in existence. In such cases, litigation seems the likely recourse. However, it is also likely that potential litigants may opt for the less costly route of **cross-licensing**, whereby each party can cross-license another’s patents.

### Patenting products of nature

Patent laws protect the public by enforcing the “product of nature” requirement in patent applications. The public is free to use things found in nature. That is, if for example a compound occurs naturally but it is also produced commercially by a company via a biotechnology method, the genetically engineered product is technically identical to the natural product. However, in the case of *Scripp vs Genentech*, a US court ruled that a genetically engineered factor VIIIic infringed a claim to VIIIic obtained by purification of a natural product. This indicates that a previously isolated natural product had first claim to patent rights over a later invention by genetic engineering. If a company seeks to apply for a patent for an invention to produce a rare, naturally occurring compound in pure form, the argument will have to be made for the technique used for extraction, purification, or synthesis, not for the material *per se*.

### Moral issues in patenting

Biotechnology also faces a moral dilemma in patent issues. Specifically, is it moral to patent any form of life? Further, if the discovery has medical value, should it be patented? Then there is the issue of the poor. Is it moral to demand that the poor pay royalties they can ill afford for using patented products for survival purposes? A debated issue is the plant breeder’s rights. Should breeders be permitted to incorporate seed-sterilizing

technology (e.g., the so-called “terminator technology”) in their products to prevent farmers from using seed from harvested proprietary material for planting in the field the following season? For balance, is it fair to expect a company to invest huge amounts of resources in an invention and not recoup its investments? There are no easy answers to these questions.

### International issues in patenting

Courts in Europe and the USA, as well as other parts of the world, differ in their positions on patent issues. Patent laws and how they are enforced may also differ among nations. For example, the European Directive on the Legal Protection of Biotechnological Inventions passed in 1998 declares that a mere discovery of the sequence or partial sequence of a gene does not constitute a patentable invention. Genes are not patentable while they are in the body (*in situ*). However, genes isolated from the organism or artificial copies of the genes produced by some technical process may be patentable, provided the novelty, inventive step, and utility are clearly demonstrated. The US laws have been tightened to include a clause to the effect that the utility of the invention must be “specific, substantial, and credible” (i.e., readily apparent, well-established utility).

In addressing the issue of morality, the European Directive also specifically excludes certain inventions from patentability. These include processes for reproductive cloning of human beings, processes for modifying the germline genetic identity of human beings, and uses of human embryos for industrial or commercial purposes. Essentially, if the publication or exploitation of an invention would generally be considered immoral or contrary to public order, it cannot be patented.

### Protecting plant varieties: a brief history of US efforts

Intellectual property rights issues impact plant breeders, researchers, producers, seed companies, as well as consumers. A formal specific government protection for plant varieties in the USA was first implemented in 1930. Called the **Plant Patent Act**, it was (and still is) limited to clonally (vegetatively) propagated plants. The **International Union for the Protection of New Varieties of Plants** (UPOV) was established in 1961. Protection of all new plant varieties was first adopted in the USA in 1970 under the **Plant Variety Protection**

(PVP) Act of 1970. This act was amended in 1994, when the US implemented the UPOV Act of 1991. As previously indicated, the use of utility patents for plant variety protection was made possible following the US Supreme Court ruling in favor of Charkrabarty in 1980, declaring that “anything under the sun that is made by man” may be patented. This ruling originally pertained to microorganisms. However, the US Patent and Trademark Office extended patent protection to plant varieties in 1985. Plant variety protection in Canada is presented in the second industry highlights box in Chapter 24 (p. 442).

A plant breeder wishing to patent a plant variety in the USA currently has three options:

- 1 **Plant patents.** This is limited to vegetatively propagated varieties.
- 2 **Plant variety protection.** This is applicable to all sexually or asexually propagated varieties, pure lines, and hybrids produced from pure lines.
- 3 **Utility patent.** This is applicable to all plant varieties (including pure lines and hybrids produced from pure lines).

These plant patents are honored in the US only, but members who are signatories to UPOV operate similar laws. Utility patents are honored mainly in the USA, Japan, and Australia. Most plant breeders protect their inventions under either plant variety protection or utility patents. A major difference between the two protection systems is that PVP-protected varieties may be used as breeding material by a researcher without requiring the permission of the rights holder. Should the breeder develop a product that is distinct from the parental germplasm, it may qualify for protection in its own right. On the other hand, the claims in a utility patent may be so designed to exclude breeding use without express permission from the owner of the property. Breeding companies use utility patents to protect their germplasm base. This is because breeding progress usually proceeds by serial improvement on the current optimal materials. Whereas this is advantageous to commercial companies, the research exemption it denies to the wider scientific community is viewed by some as undesirable. However, patents would allow the company to recoup its investments in developing a breeding material that can be utilized by its competitor. For example, developing a breeding material by prebreeding (germplasm enhancement, see Chapter 6) can be risky, expensive, and of long duration. The developer needs to protect his or her invention.

## Ethics in plant breeding

Manipulating plants through conventional plant breeding efforts have generally proceeded without fanfare and objections from the consumers. Significant protestations from the consuming public started when plant breeders added biotechnology to their tool bag.

### The biotechnology debate

Public perceptions about biotechnology products are rooted in the perceived risks that these products pose to social and personal values. Public reaction to biotechnology often is influenced by activism and special interest propaganda. The biotechnology debate is rooted in three fundamental disagreements:

- 1 **Scientific disagreements.** Society is concerned about the potential risk that the development and application of biotechnology poses to humans, animals, and environmental health. These issues involve empirical questions and are usually resolved by scientific methods. However, they are not exclusively resolvable by the scientific method of enquiry. Sometimes, value judgment is critical in their resolution. For example, the way to handle uncertainties in scientific data and definition of the levels of risk deemed acceptable are both value judgments.
- 2 **Political disagreements.** Political disagreements are generally about the social and economic impacts of biotechnology based on the various political viewpoints. Political disagreements usually intensify during an election season and may swing in favor of one side, depending on the dominant political ideology of the day.
- 3 **Religious, ethical, and philosophical disagreements.** These disagreements are often faith-based and include issues about morality and whether scientists are playing God, or whether the biotechnology products are natural. Faith-based disagreements are difficult to resolve in a pluralistic society. Sometimes, such issues incite extremist acts or vigilantism (e.g., destroying field tests or laboratory projects).

### Concepts of ethics, morals, and values

The following brief introduction to ethics, morals, and values, is designed to help the student become better equipped to participate in debates that center around these issues in relation to genetic manipulation of plants. **Ethics** is the science of morals in human conduct (i.e., study of moral principles). **Morals** are concerned with

the accepted rules and standards of human behavior in a society. They involve the concept of right or wrong, and the goodness or badness of human character or behavior. **Value** is basically the worth attached to something. In other words, ethics is evaluative of the decisions people make and the actions they take as they are presented with dilemmas. Morality depends on values in order to determine the goodness or badness of an action. In a pluralistic society, there are differences in the sense of values (i.e., relativism). Consequently, there are a variety of moral theories that do not necessarily constitute truth. Further, law, religion, and custom should be distinguished from morality. In law, lawmakers define what is right or wrong. Those who break the law are subject to punishment prescribed by the legislature. In religion, right or wrong is based on revelation or scriptural authority. Whatever choice that is made has eternal consequences. In the case of custom, tradition determines what is acceptable or not, and society expresses approval or disapproval of an action.

### Plant biotechnology: ethical and value issues

The issues raised in this section are only a sample of public opinions intended to generate discussion, not to support the author's point of view. The issue of ethics, morals, and values in society is more important in a society in which religion is important to a significant portion of the population. Where there is a belief in God the Creator, a major source of discord in society regarding biotechnology is the notion that scientists are playing God when they fail to respect human limitations. God, humanity, and nature are linked, God being the creator of both the latter. Some people see nature as God's creation for the benefit of humans who therefore can use plants, animals, and the ecosystem for their purposes, as they deem necessary. Others see nature as a sacred creation that must be respected and not tempered with. Does this respect mean that humans cannot manipulate nature? What cannot be denied is that the Creator has endowed humans with considerable creative genius. The obvious question then is whether exercising creativity through plant breeding and biotechnology is within the scope of this endowment or whether it is tantamount to an infringement on divine prerogative? For those who see nature as a gift to humans for their use, recombining genetic materials may be justified as just another way of using natural resources.

In order for us to be correctly evaluative of our choices, decisions, and acts as they pertain to genetic manipulation, there is the need for certain basic sets of

information to be available. One set pertains to the values we attribute to things and acts we perform, the other set being value-free. Scientists, traditionally, generate value-free information. However, both kinds of information (tested empirically and experientially) and their impact need to be accumulated for use in making choices and decisions about biotechnology.

The ethical issues and the passion with which they are debated in the public arena vary among applications. Manipulation of the food chain seems to attract more attention than clinical applications (e.g., xenografts). For example, heart valves from pigs have been used in humans without fanfare. However, genetically modified (GM) grains have encountered considerable public opposition from certain quarters. In general, the ethical issues of concern to the public are the impacts of biotechnology on human health and safety, environmental impacts, intrusions into the natural order, invasion of privacy, issues of rights and justice, economics, and others. It is important that both the benefits and risks of biotechnology be considered in making ethical decisions about the discipline.

The problem is that, at the moment, we are limited in our knowledge about the full benefits and risks of biotechnology. Consequently, we are in danger of either underestimating or overestimating the potential of biotechnology for good or evil. Further, public reaction may be rooted in undue fear or hope stemming from misunderstanding, misinformation, or lack of information about various aspects of the genetic manipulation of plants.

New technologies often tend to tip the scales in favor of those with resources to acquire them. They are most likely to be adopted if they increase profitability to producers while lowering the cost to consumers. There is also the issue of the developing countries. Many of the germplasm resources used in plant and animal improvement are derived from these regions of the world. The debate over patenting biological material is often linked to this fact.

## Risk analysis of biotechnology

Risk analysis of biotechnology is complicated by the fact the activity is unique for the crop species, the genetic modification, and the production environment. A more useful and fair analysis of the impact of biotechnology would be obtained if risk analysis of a biotech product were done in comparison with competing products or technologies. Examples of fair analysis

would be to compare chemical pesticides with *Bacillus thuringiensis* (*Bt*) products; the use of glyphosate herbicide with glyphosate-resistant crops, compared with the use of the herbicide atrazine or other weed management methods; or planting GM crops with high productivity compared to clearing new land to plant conventional lower productivity cultivars. In conducting risk assessment, it is important that the process enhances consumer confidence and trust, without which marketing GM products is bound to be problematic. In part, public perceptions and attitudes about biotechnology are shaped by concerns about the risks and safety (acceptability of risk) of genetically engineered foods and other products. These biotechnology products are perceived as posing risks to a variety of social and personal values.

An expert panel on the future of food biotechnology commissioned by the Canadian Food Inspection Agency and Environment Canada categorized the values that are perceived by the public as being placed at risk by biotechnology into three categories:

- 1 Potential risks to the health of human beings, animals, and natural environment.** The risks to human health and the environment are at the top of the list of public concerns about the impact of biotechnology on society.
- 2 Potential risks to social, political, and economic relationships and values.** Commonly, the public is concerned about the monopoly of certain industries (e.g., seed) by multinational corporations to the detriment of small producers and the risk or increased dependency of developing economies on these monopolies. It is the opinion of many experts that the level of risk acceptable by the public depends on the overriding benefits to be achieved (risk-cost benefit).
- 3 Potential risks to fundamental philosophical, religious, or metaphysical values held by different individuals and groups.** This category addresses the issue the public takes with the *process* of biotechnology rather than the product or *impacts*. The concern is the risk of playing God by implementing processes that are unnatural to alter nature.

The extent to which the public is willing to be exposed to unknown or uncertain risks, and how much risk is acceptable, is influenced by social, economic, and philosophical factors. People will be more willing, for example, to accept a higher risk level if they are strongly convinced about the benefits of adoption of biotechnology products, or, on the other hand, the adverse consequences of not adopting biotechnology products.



## Industry highlights

### *The intersection of science and policy in risk analysis of genetically engineered plants*

David A. Lee<sup>1</sup>, and Laura E. Bartley<sup>2\*</sup>

<sup>1</sup> EPA Office of Research and Development, 8623N, Washington, DC 20460, USA;

<sup>2</sup> USDA-APHIS Biotechnology Regulatory Services, Riverdale, MD 20737, USA

#### **Introduction**

In our view as scientists recently entrenched in the policy world, the regulation of genetically engineered (GE) plants illustrates the challenges of applying science to policy for a number of reasons. The laws and implementing regulations regarding the oversight of GE products require that regulation be science-based; however, risk analysis always requires some judgment on the part of decision-makers. In addition, the United States does not have an overarching biosafety law and instead uses existing laws to regulate GE organisms. Further, despite the objective of US policy that regulation of GE organisms be product-, not process-, based, in some instances different regulations and standards are used to oversee organisms that are GE as opposed to those that are not. Below we elaborate on these ideas and discuss the data used in risk assessment. Despite the tensions that exist within the US biotechnology regulatory system, our educated opinion is that the system is generally working to allow responsible use of this technology.

#### **Risk analysis: a science-based endeavor that relies on values**

Risk assessment fits into a larger process known as risk analysis, which has the following three components: risk assessment, risk management, and risk communication. Risk assessment involves the highly scientific and analytical process of identifying hazards that might result from an action and evaluating their probability and consequences, as simplified in the following equation: risk = (probability of a hazard) × (consequences of a hazard). However, human judgments are often required to complete the assessment of risk because typically gaps exist in the data regarding the probability and consequences of potential hazards. Risk management is the process in which decisions regarding the action are made and any identified risks are managed, mitigated, or monitored for. Value-based judgments are particularly required in decision-making to balance the risks of action with those of inaction. Risk communication is the part of the iterative risk analysis process in which information regarding the assessment and management steps is made available to, and discussed with, interested parties.

To briefly illustrate risk analysis for plant genetic engineering, an example of a hazard that is often discussed is that of a gene introduced through genetic engineering being passed to a wild relative of a crop, termed “gene flow”, and causing a change in the population of the wild relative. In this case, the probability of gene flow occurring has been shown in many cases to be significantly greater than zero (Ellstrand 2003). However, while geneticists can describe scenarios that would lead to changes in plant populations due to the introduction of a single gene, the probability of those changes and the long-term consequences of such changes on ecosystems and for the preservation of crop biodiversity are unknown in most cases. In addition, some would argue that these risks need to be weighed with the difficult-to-assess risk of not using a particular GE variety, which could lead to changes in wild relative populations and ecosystems through other means, such as habitat destruction due to inefficient agriculture. Thus, despite the legal requirement that decisions regarding GE products be science-based, there are many decisions regarding data gaps and value judgments that must be made in the course of risk analysis.

#### **Law meets science: product- versus process-based regulation and some consequences of the use of particular laws for the regulation of biotechnology in the USA**

The 1986 Coordinated Framework for Regulation of Biotechnology lays out the general US policy on regulation of GE products and highlights that regulation should be science-based and product-based, rather than treating products created through a particular process (i.e., genetic engineering) as inherently different. Consistent with this approach, the Coordinated Framework outlines how the USDA, FDA, and EPA are expected to regulate GE products using existing laws. However, to implement the Coordinated Framework and apply existing laws to GE products, the agencies established separate regulatory staffs, and USDA and EPA also found it necessary to write regulations that are in essence, if not in letter, specific to GE products. Thus a process-based system has been established, even though there is no *scientific* justification for regulating GE plants differently than conventional plants (National Research Council 2002). For example, plants that are genetically engineered to be herbicide tolerant fall under the scrutiny of the USDA; whereas, conventionally bred herbicide-tolerant varieties do not, even if the conventionally bred lines are resistant to the same herbicide as the GE ones.

Under the different laws and implementing regulations for GE products, the agencies have varying abilities to assert authority over GE products, require particular data, and use particular information in decision-making. The Animal Plant Health Inspection

\* The authors are solely responsible for the content of this piece and any views expressed do not necessarily represent those of the EPA, the USDA, or of AAAS.

Service (USDA-APHIS) uses the authority of the Plant Protection Act to regulate GE organisms. This law gives USDA the authority to restrict introduction into the environment of plant pests, which are defined as living organisms that cause disease in or damage to plants not including humans and non-parasitic plants (US Congress 2000). The current USDA-APHIS regulations use this “plant pest authority” to regulate GE organisms based on the potential plant pest risk caused by the use of plant pest (e.g., viral) sequences or vectors (e.g., *Agrobacterium*) in the creation of many GE plants (USDA-APHIS 1997). For plants created through biotic transformation that do not have plant pest sequences, the regulations can be imposed on articles that USDA has “reason to believe” pose a plant pest risk (USDA-APHIS 1997). Technically, it may be possible that the use of plant pest components in the creation of a GE plant could create a new plant pest or increase the GE plant’s susceptibility to a disease. However, the rarity of these effects in GE plants and the dependence on the reason to believe clause, causes some to be concerned that the current regulations are tenuous (National Research Council 2002). In order for a GE plant to be released into the environment in an unconfined manner and thus sold commercially, USDA-APHIS will evaluate a petition for non-regulated status to determine if the GE product does not pose a plant pest risk or cause other environmental harm, as evaluated under the National Environmental Policy Act. Because granting non-regulated status takes the product out from all USDA-APHIS oversight, this feature has drawn criticism as it may limit USDA-APHIS actions in terms of monitoring and other risk management activities. Also, while much information in petitions is available to the public, applicants may claim portions are confidential business information under the Freedom of Information Act, reducing the transparency of the USDA-APHIS system.

In part to address some of the issues discussed above, USDA-APHIS initiated a process in 2004 to revise its regulations based on the so-called “noxious weed authority” in the Plant Protection Act (USDA-APHIS 2004). This authority gives USDA-APHIS the ability to restrict introduction into the environment of noxious weeds, which are defined quite broadly as “any plant or plant product that can directly or indirectly injure or cause damage to crops, . . . other interests of agriculture, . . . natural resources of the United States, public health, or the environment” (US Congress 2000). The revised regulations would regulate based on potential noxious weed risk of GE plants, greatly expanding the reasons for USDA-APHIS to assert its authority.

In contrast to the somewhat limited abilities of USDA-APHIS under its current regulations, the EPA broadly uses its authority under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) to regulate GE plants that contain pesticidal elements. To use this law to regulate GE plants, the EPA defined a new pesticide type, the plant-incorporated protectant (PIP), as “a pesticidal substance produced by the plants and the genetic material necessary for them to produce the substance” (EPA 2001). This allows the EPA to retain authority over approved products and it often requests additional data from applicants as a condition of continued registration, thereby reducing the data gaps present during risk assessment. FIFRA provides special protections for information regarding the health effects of products that might be claimed as confidential under other statutes, and requires that the EPA submit major decisions to a panel of external experts, called a scientific advisory panel. Thus, relative to other agencies, the EPA has a more transparent and understandable regulatory process. However, some have criticized the EPA for too restrictively regulating PIPs by establishing requirements that are not commensurate with the risks posed by GE plants and are unnecessarily burdensome for applicants, especially since conventionally bred PIPs are exempted from regulation. Nonetheless, the EPA’s efforts may improve acceptability of the technology, and, in the case of *Bt* crops, extend the lifetime of their benefits to agriculture by requiring insect-resistance management.

The FDA regulates foods derived from GE products under the Federal Food Drug and Cosmetic Act (FFDCA). As published in 1992, FDA policy is that it will regulate foods derived from GE products in the same way as those derived from conventionally developed products (FDA 1992). The FDA’s regulation is based on whether the product has altered nutritional properties or contains a food additive, which is defined as a substance introduced into food that is not a pesticide and is not “generally recognized as safe” (GRAS) (FDA 1992). As for conventionally developed foods, FFDCA makes it the responsibility of the developer to determine that GE-derived foods are safe and any substances new to the variety are GRAS, but the FDA provides a voluntary consultation process to help developers determine this. The consultation process, through which developers submit data to FDA scientists until the FDA has no more questions regarding safety, is available for both conventional and GE products. The voluntary nature of the consultations makes some observers very uncomfortable with the FDA’s regulation of GE foods. On the other hand, FDA records show that all GE products currently approved in the United States have completed such a consultation, but, because it is voluntary, none of the information submitted to the FDA is available for public scrutiny. Therefore, while the FDA’s regulation is most genuinely product-based, the transparency of the system is the lowest among the three agencies.

### ***GE risk assessment concerns and typical data and information evaluation***

Given the peculiarities of the US regulatory system for GE crops, how does it actually work? Here, we discuss the risk assessment concerns and information that the regulatory agencies regularly examine before approving a GE variety for widespread use.

To evaluate the safety of GE crops before approval for food, feed, or planting in the United States, the USDA, EPA, and FDA consider the potential impact of a large number of environmental- and health-related effects of the GE crops. Notably, the list of concerns that the agencies evaluate includes all of the potential environmental hazards a recent National Research Council (2002) panel identified that GE plants could cause:

- 1 The GE trait could be passed to a wild or weedy relative and increase its weediness or invasiveness, or the GE plant itself could become weedy or invasive.

- 2 The GE trait could negatively impact non-target organisms in the environment.
- 3 Organisms that the GE trait is intended to harm could develop resistance to the trait.

Currently, agency scientists and risk managers evaluate each submission on a case-by-case basis and determine the specific data the applicant should submit depending on the product. However, as shown in Table 1, there are a number of concerns

**Table 1** Risk assessment concerns and information used by the Food and Drug Administration (FDA), Environmental Protection Agency (EPA), and US Department of Agriculture (USDA).

Agency	Risk assessment concerns	Typical data <sup>1</sup> and information <sup>2</sup> used in assessment
FDA, EPA, and/or USDA-APHIS <sup>3</sup>	Characterization of inserted DNA	Data from Southern blots and/or sequencing
	Characterization of expressed protein(s)	Data from western blots or ELISA assays often with multiple plant tissues; data from phenotypic analyses; data on peptide modification <i>in vivo</i>
	Stable inheritance of the transgene	Data from Southern blots and/or phenotypic analysis over multiple generations
	Plant composition	Data on measurements of amino acids, minerals, fatty acids, carbohydrates, water, etc., including amounts of any toxins and antinutrients that are typically found in the plant
	Allergenicity of expressed protein	Data on <i>in vitro</i> digestibility, heat stability; data on sequence similarity to known allergens
EPA and USDA-APHIS	Effects on non-target organisms	Data from toxicity assays on non-target insects, soil organisms, birds, mammals, and fish; in some instances, data from surveys of non-target invertebrates in the field
	Gene flow to wild and weedy relatives	Information on hybridization potential and distributions of wild and weedy relatives; in some instances, data on hybridization potential
	Potential weediness of engineered plant	Data describing plant characteristics relative to a comparator plant line
	Threatened and endangered species impact	Information on distributions of threatened and endangered species that are related to the target organism or could otherwise be affected
	Plant pest risk	Data on incidence of plant pests in field trials
USDA-APHIS	Agronomic management	Agency assessments <sup>4</sup> ; information on alternative agricultural management options
	Acute toxicity	Mammalian toxicity assay
	Transgenic protein fate in environment	Half-life of protein under soil conditions
	Cost benefit analysis	Agency assessments; information on alternative agricultural management options with respect to pesticide use
	Insect Resistance Management	Agency assessments; data on effective dosage compared with protein expression in the plant; information on target insect biology and behavior

<sup>1</sup> "Data" refers to original data, in the form of formal or informal observations, submitted by the applicant.

<sup>2</sup> "Information" refers to information that is not typically generated by the applicant (i.e., not original data), typically public information from the scientific or agronomic literature.

<sup>3</sup> All three agencies evaluate molecular characterization data. FDA and USDA-APHIS examine plant composition. FDA and EPA evaluate allergenicity potential of the expressed protein(s).

<sup>4</sup> "Agency assessments" signify a particular reliance on agency risk assessment expertise and information that is not generated by the applicant.

APHIS, Animal and Plant Health Inspection Service; ELISA, enzyme-linked immunosorbent assay.

that the agencies typically evaluate for every product that falls under their purview. All of the agencies evaluate a detailed molecular and genetic characterization of the product to obtain information about the identity of the GE plant and confirm that the inserted gene is functioning as intended. In addition, the FDA and USDA-APHIS examine plant composition to gauge unintended, pleiotropic changes due to transgenesis, and the FDA and EPA evaluate the allergenicity potential of the expressed protein(s).

USDA-APHIS determines the potential of a GE plant to become an agricultural weed, or to cause damage to agriculture through the introduction of a novel plant pathogen produced by the transgenic plant or a change in plant susceptibility to pests. Under FIFRA the benefits of products as well as their potential risks are evaluated in the registration process, so the EPA assesses the potential economic impact of the introduction of the PIP product, along with the environmental and human health benefits of the altered pesticide-use regime. When appropriate, the EPA evaluates insect resistance management (IRM) plans proposed by applicants to confirm that the plan will be sufficient to delay resistance development to *Bt*. Both EPA and USDA-APHIS are concerned with the potential for gene flow to occur from the transgenic plant to wild relatives (Table 1). In the case of plants engineered to produce a PIP, USDA-APHIS and EPA evaluate whether there will be toxicity to non-target organisms that might come into contact with the crop or its residues.

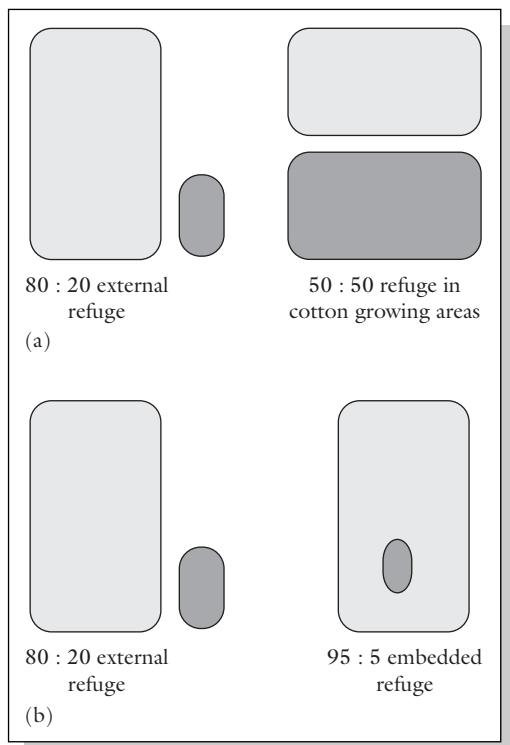
#### *Using science for risk management: the insect-resistance management example*

In addition to the data provided to the agencies during their assessment of specific products during the approval process, both the EPA and USDA fund active research programs to continue studying the environmental and human health impacts of GE crop plants. An example of the use of science in determining regulatory policy is that of the EPA's IRM plan for PIPs utilizing proteins produced by the bacterium *Bacillus thuringiensis* (*Bt*), which are the most common PIPs engineered into plants.

Insect populations exposed to pesticides over a long enough time-frame will develop resistance (Feyereisen 1995), so because preparations of the bacteria that express *Bt* toxins are an important pest management tool for the organic farming industry, concerns were raised that the development of resistance to *Bt* would deprive the agricultural community of a safe, environmentally friendly pesticide. Due to the adverse health and environmental effects of having to use conventional pesticides instead of *Bt*, the EPA has required a very stringent IRM plan for the use of *Bt*-PIP-containing crops to delay resistance development, in contrast to almost all other pesticides.

After studying insect-resistance models and experimental data, the EPA developed a program to delay resistance development based on a "high dose/structured refuge" approach. This strategy relies on resistance to *Bt* being a genetically recessive trait and the initial frequency of the resistance allele being very low. When this is the case, refuges for susceptible insects can be designed so that in principle any resistant insects that arise in the population will almost certainly mate with a susceptible individual so that the heterozygous offspring will be susceptible to the PIP. The high dose requirement for PIP products necessitates that the plant expresses a level of *Bt* protein at least 25-fold greater than that needed to kill 99% of susceptible insects in laboratory assays. The basic structured refuge requirements for *Bt* crops are satisfied in general by planting 20% of the field as a contiguous non-*Bt* refuge that should be located within 0.8 km of the *Bt* crop fields (Figure 1). However, if *Bt* corn is planted in cotton-producing areas then the non-*Bt* refuge should be 50% of the corn acreage because cotton pests could feed on both cotton and corn and develop resistance more rapidly.

Monitoring for insect-resistance development has always been a requirement by the EPA for registrants, who inform the EPA of the results of their monitoring program on an annual basis, along with any grower observations of increased crop damage by insects normally susceptible to *Bt* toxins. In addition, academic researchers have performed recent studies in Arizona, North Carolina, and Iowa to measure resistance development in *Bt* corn and cotton fields over multiple growing seasons (Tabashnik et al. 2003). In most cases, the initial frequency of resistance alleles to particular GE *Bt*-containing plants in the target populations was very low (<0.001), and the frequency dropped over time, in some cases becoming undetectable in the populations analyzed.



**Figure 1** Examples of refuge strategies for *Bt* crops that are acceptable to the EPA: (a) corn refuge requirements and (b) cotton refuge requirements. Light areas represent fields of *Bt* crops, which the dark areas represent non-*Bt* refuges. In all cases, the refuge can be sprayed with non-*Bt* pesticides if economically necessary.

These results are consistent with the notion that the EPA's high dose/structured refuge plans are effective, but scientists caution that it may be too soon to tell whether or not resistance would have developed by this time without the implementation of IRM.

### **Conclusion**

Through the oversight provided by USDA-APHIS, EPA, and FDA, GE plants have been more heavily scrutinized than any other comparable agricultural product. While it is not possible to conclusively demonstrate that an agricultural product (GE or not) will have no environmental or human health impact, there has been no documented case of a negative impact on human health or an observed effect on the environment after a decade of growing GE plants.

However, new GE products currently under development, such as plants engineered to tolerate abiotic stress or to produce pharmaceuticals or industrial compounds, may have significant impacts on (and benefits for) humans or the environment. Regulators may need to continue to creatively use their legal authorities to accomplish the goal of improving agriculture, the environment, and health as they are faced with increasingly complex risk assessments. It will also be the responsibility of regulatory agencies to continue supporting research to assess the potential for these risks to occur, the means to mitigate them, and to continue monitoring GE crops to ensure that no significant unexpected effects arise from their use.

### **References**

- Ellstrand, N.C. 2003. Dangerous liaisons? When cultivated crops mate with their wild relatives. Johns Hopkins University Press, Baltimore, MD.
- EPA. 2001. Regulations under the Federal Insecticide, Fungicide, and Rodenticide Act for Plant-Incorporated Protectants. Code of Federal Regulations, Section 40, Parts 152 and 174. (Available at: [http://www.epa.gov/pesticides/biopesticides/pips/pip\\_rule.pdf](http://www.epa.gov/pesticides/biopesticides/pips/pip_rule.pdf).)
- FDA. 1992. Statement of policy: Foods derived from new plant varieties. Fed. Reg. 57:22984–23005.
- Feyereisen, R. 1995. Molecular biology of insecticide resistance. *Toxicol. Lett.* 82/83:83–90.
- National Research Council. 2002. Environmental effects of transgenic plants: The scope and adequacy of regulation. National Academy Press, Washington, DC.
- Tabashnik, B., Y. Carrière, T. Dennehy, S. Morin, A. Sisterson, and J.-Z. Zhao. 2003. Insect resistance to transgenic Bt crops: Lessons from the laboratory and field. *J. Econ. Entomol.* 96:1031–1038.
- US Congress. 2000. The Plant Protection Act (Public Law 106–224). (Available at: <http://www.aphis.usda.gov/brs/pdf/PlantProtAct2000.pdf>.)
- USDA-APHIS. 1997. Introduction of organisms and products altered or produced through genetic engineering which are plant pests or which there is reason to believe are plant pests. Code of Federal Regulations, Section 7, Part 340. (Available at: <http://www.aphis.usda.gov/brs/7cfr340.html>.)
- USDA-APHIS. 2004. Environmental impact statement: Introduction of genetically engineered organisms. Fed. Reg. 69:3271–3272.

### **Overview of the regulation of the US biotechnology industry**

The rationale for regulation of biotechnology is to protect consumers from product risk, and to promote and retain their confidence in biotechnology products, as well as to promote trade. The agencies with regulatory oversight in biotechnology in the USA are the US Department of Agriculture (USDA), Food and Drug Administration (FDA), and Environmental Protection Agency (EPA). The products they regulate are summarized in Table 15.1. Manufacturers and developers of biotech products are required to meet certain minimum product standards stipulated in state and federal marketing statutes. These include state seed certification laws – the Federal Food and Drug and Cosmetic Act (FFDCA), Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Toxic Substances Control Act (TSCA), Plant

**Table 15.1** US agencies and the biotechnology they regulate.

Agency	Products regulated
US Department of Agriculture	Plant pests, plants, veterinary biologics
Environmental Protection Agency	Microbial/plant pesticides, new uses of existing pesticides, novel microorganisms
Food and Drug Administration	Food, feed, food additives, veterinary drugs, human drugs, medical devices

Source: USDA.

Protection Act, and National Environmental Policy Act (NEPA). Depending upon the product, an agency may review it for its safety to grow, safety to eat, or safety to the environment. The Animal and Plant Health

Inspection Service (APHIS) conducts the USDA biotech evaluation. The EPA ensures the safety of pesticides and enforces FIFRA (to regulate the distribution, sale, use, and testing of plants and microbes producing pesticidal substances) and FFDCA (to set tolerance limits or exemptions from tolerances for pesticides used in food, and feed production). The FDA is part of the Department of Health and Human Services, and enforces the FDA's regulations (to regulate foods and feeds derived from new plant varieties).

### USDA-APHIS

APHIS is authorized to regulate the confined release, in particular the interstate movement importation and field testing, of organisms and products altered or produced through biotechnology processes that are plant pests or are suspected of being so, or have incorporated genetic components from organisms that are plant pests. An individual or an entity seeking to conduct any of the above-mentioned activities must apply for and receive one of the permits from APHIS before proceeding.

- 1 Permit for movement and importation.** This requires the applicant to disclose the nature of the organism, its origin, and its intended use.
- 2 Permit for release into the environment.** APHIS oversees field testing of biotech products. The applicant is required to provide information on the plant (including new genes and new gene products), their origin, the purpose of the test, experimental design, and precautions to be taken to prevent escape of pollen, plants, or plant parts from the experimental site.

For low risk or familiar organisms or classes of modification, the developer may apply for an expedited permit through the notification process. Furthermore, in order to be allowed to release a previously regulated organism or product in an unconfined manner and commercially sell the organism, the developer must petition APHIS for non-regulated status. When evaluating petitions for non-regulated status, APHIS considers the potential risks to agriculture due to release of the organism, under the Plant Protection Act, and to the environment, under NEPA and the Threatened and Endangered Species Act.

### FDA

The decision to subject all biotech products to the same standards of regulation of traditional products was made

by the FDA in 1997. In the Federal Register, Vol. 57, the FDA directs that companies or researchers whose products meet one of the following criteria should submit them for testing:

- 1 Unexpected effects:** that is, the product produces unexpected genetic effects.
- 2 Known toxicants:** the product has higher than normal levels of toxicants in other edible varieties of the same species.
- 3 Nutrient level:** the product has altered levels of essential nutrients.
- 4 New substances:** the chemical composition of the product is significantly different from the existing normal products.
- 5 Allergenicity:** the product contains proteins that have allergenic properties.
- 6 Antibiotic resistance selectable marker:** the product is produced by a biotech process that utilizes genetic markers that could adversely impact current clinically useful antibiotics.
- 7 Plants developed to make specialty non-food substance:** plants are engineered to produce pharmaceuticals or industrial compounds.
- 8 Issues specific to animal feed:** the product's chemical composition regarding nutrient and toxins is significantly different from levels in similar products used for feed.

In addition to these federal regulatory activities, individual states are at liberty to develop and implement additional regulations. Exempted from premarket approval are products that are classified as **GRAS** ("generally recognized as safe"). Such food products may have been engineered to express proteins. However, a GRAS substance is excluded from the definition of a food additive. FDA encourages developers of engineered foods and feeds to consult with them throughout the development process to ensure that the product does not need to be regulated by the FDA.

### EPA

The EPA regulates pesticides through a registration process. Its definition of a pesticide is any substance or mixture of substances intended for preventing, destroying, repelling, or mitigating pests. A new category of pesticides is the **plant incorporated protectants (PIPs)**. These are substances that have been genetically engineered into plants so that the plant becomes protected from pests by producing the substance

**Table 15.2** Some viral-coated proteins that have exemption from the Environmental Protection Agency.

Papaya ringspot virus coat protein
Potato leaf roll virus (PLRV) replicase protein as produced in plants
Potato virus Y coat protein
Watermelon mosaic virus (WMV2) coat protein in squash
Zucchini yellow mosaic (ZYMV) coat protein
WMV2 and ZYMV in ASGROW ZWO

(pesticides) in its tissues (e.g., *Bt* crops). Although plants engineered to be herbicide resistant are not classified as PIPs and thus are not regulated by the EPA, the use of herbicide-tolerant plants is nonetheless subject to EPA regulation. The authority for such regulation is provided under FIFRA and FFDCA. If a plant producing a plant pesticide is intended to be used for food, the EPA must establish a “safe level” of the pesticide residue allowed. The EPA defines a safe level as one where there is a reasonable certainty that no harm will result from aggregate exposure to the pesticide chemical residue, including all anticipated dietary exposures and all other exposures for which there is reliable information. If the EPA determines that no harm will result from the aggregate exposure to the pesticide residue, the EPA can exempt the pesticide from needing a tolerance. The EPA has categorically exempted the nucleic acids that encode PIPs, and has also exempted genetically engineered products that utilize viral-coated proteins to protect plants against viral infections from the requirement of a tolerance prior to being used. EPA has also exempted all *Bt* products currently registered for use as pesticides with the EPA. Some of the products are listed in Table 15.2.

### GM foods and the issue of food allergy

Food allergy or food sensitivity is a common medical condition in society. It is described as an adverse immunological reaction resulting from the ingestion, inhalation, or contact with a food or food additive. The specific factor that triggers the allergic reaction is the **allergen**. Food allergy should be distinguished from food intolerance (e.g., lactose intolerance). The most widely studied mechanism of food allergy is that mediated by **immunoglobulin E (IgE)**. Allergic reactions are produced immediately when IgE, an antibody, is

exposed to an allergen (usually a protein substance). Such an exposure causes immune cells in the body (mast cells and basophils) to release different kinds of toxic mediators (e.g., histamine and leukotrienes), which then trigger an allergic reaction. The reaction may manifest in a variety of ways, ranging from minor itches to anaphylactic shock and death. There are some allergic reactions to foods that are not mediated by IgE (e.g., celiac disease or gluten-sensitive enteropathy).

The only current treatment for food allergy is avoidance. An estimated 34% of emergency room visits in the USA are for the treatment of anaphylaxis related to food allergy. Because of the rise in allergic disorders, the public is concerned about the potential risk posed by bioengineered foods and food products. The most common foods associated with food allergies (and accounting for over 90% of reported food allergies worldwide) are peanuts, tree nuts (e.g., almond, brazil nut, cashew, macadamia, hazelnut, pecan, pistachio, walnut), cow’s milk, fish, shellfish (crustaceans and mollusks), soy, wheat, and sesame seed.

It is known that allergenic proteins can be transferred by genetic engineering from one organism to another. Such a transfer was confirmed in the case of the brazil nut 2S albumin storage protein that was transferred by rDNA technology to soybean to increase its methionine contents (legumes like soybean are low or deficient in the essential amino acid methionine). Tests for allergenic potential to humans by radioallergosorbent test (RAST) and skin prick test showed that the brazil nut allergen had been transferred. It should be mentioned that commercial development of this food product was summarily discontinued. Further, there are no validated reports of allergic reactions to any of the currently marketed GM foods as a result of transgenic protein.

Genetic engineering of foods may have potential collateral changes (or pleiotropic effects) resulting from the transgene having an effect simultaneously on more than one characteristic of the host. In principle this may include an alteration of the intrinsic allergenicity of the protein by, for example, glycosylation, or an alteration of the amount of allergenic protein produced. It is hence important to consider the potential changes in endogenous host allergens subsequent to gene transfer. Therefore, many of these physicochemical characteristics are assessed before registration of PIP products by the EPA and during the consultation process with the FDA. In the past, the EPA has denied approval for human consumption of a PIP that demonstrated some allergenic characteristics.

## Concept of substantial equivalence in the regulation of biotechnology

A major challenge facing regulatory agencies the world over is in deciding what constitutes a meaningful difference between a conventional crop cultivar and its genetically modified derivative. The concept of **substantial equivalence** originates from the general position taken by regulators to the effect that conventional cultivars and their GM derivatives are so similar that they can be considered “substantially equivalent”. This concept, apparently, has its origins in conventional plant breeding in which the mixing of the genomes of plants through hybridization may create new recombinants that are equivalent. However, critics are quick to point out that this is not exactly what obtains with genetic engineering. In conventional breeding, the genes being reshuffled between domesticated cultivars have had the benefit of years of evolution during which undesirable traits have been selected out of most of the major crops, though the use of a wild relative for breeding purposes could introduce genes potentially detrimental to human health. In contrast, the genes involved in GM crop production have been introduced into unfamiliar genetic backgrounds.

The traditional method of evaluating new cultivars from breeding programs is to compare them with the existing cultivars they would replace, if successful. To this end, evaluations include gross phenotype, general performance, quality of product, and chemical analysis (especially if the goal of the breeding program is to improve plant chemical content). It is expected (or at least hoped) that the new cultivar will not be identical to the existing cultivar, because of the investment of time and effort in breeding. Nonetheless it is not expected that the new cultivar would produce any adverse effects in users of the product. On the other hand, there are still genes detrimental to human health in wild relatives used for breeding, and, in principle, mutagenic methods used to generate diversity in breeding programs could also cause changes that impact food safety or quality.

In 1995, the World Health Organization (WHO) released a report in which the concept of substantial equivalence was endorsed and was promoted as the basis for safety assessment decisions involving genetically modified organisms and products. Since then, the concept has attracted both supporters and opponents. Some feel that, as a decision threshold, the concept is vague, ambiguous, and lacks specificity, setting the standard of evaluation of GM products as low as possible. On the other hand, supporters believe that what is intended is

for regulators to have a conceptual framework, not a scientific formulation, that does not limit what kinds and amounts of tests regulators may impose on new foods. The concept of substantial equivalence has since been revisited by the Food and Agricultural Organization (FAO) and WHO and amended. It appears that opposition to the use of this concept as a regulatory tool would be minimized if a product is declared substantially equivalent after rigorous scientific analysis has been conducted to establish the GM product poses no more health or environmental risk than its conventional counterpart. Of course the cost of such an evaluation may prohibit the use of biotechnology varieties unless there are great benefits from their use.

## Issue of “novel traits”

Sometimes, conventional plant breeding may introduce a “novel trait” into the breeding program through wide crosses or mutagenesis. This notwithstanding, the new cultivar produced is still considered substantially equivalent to other cultivars of the same crop. On the other hand, even though the presence of a transgene in a GM cultivar is considered to be an incorporation of a “novel trait”, the GM product and the conventional product differ. The novel traits incorporated in the major commercially produced GM crops are derived from non-plant origins (mainly from microorganisms). Secondly, only a single gene separates a GM product from its derivative. In conventional breeding, the desired genes are transferred along with numerous other unintended genes.

The question then is whether the more precise gene transfer of genetic engineering means that a GM crop and its traditional counterpart differ only in the transgene and its products. If this were so, a simple linear model would be adequate to predict the phenotype of a GM organism. Unfortunately, because of the role of the environment in gene expression, and the complex interactions that occur in a biological system, linear models are seldom adequate in predicting complex biological systems. Further, it is known that single mutations often produce pleiotropic effects (collateral changes) in the organism. Similarly, collateral effects of a transgene have been demonstrated in transgenic salmon carrying the transgene coding for human growth hormone, in which researchers found a range of phenotypes. Also, it is important to mention that the altered phenotypes may appear at particular growth times in the growth cycle of the organism, or in response to specific environmental

conditions. Further, these phenotypic changes induced by the transgene may only be minor alterations.

In view of the foregoing, it appears that the best way to assess any adverse effect of a transgene is to directly test for harmful outcomes. In food biotechnology, some believe that this should include testing for both short- and long-term human toxicity and allergenicity, among others. There should also be an assessment of environmental impact over time and across relevant sites. Then, a final assessment should be made regarding the extent to which the transgenic cultivar deviates from the parental genotype. It is important that the analysis indicates whether such deviations, if any, are biologically significant. Otherwise, the GM cultivar would be substantially equivalent to the existing cultivar and would not need prior approval for introduction into the food chain.

However, as the transgene is derived from a source that would not necessarily be consumed with the crop or be present in the environment where the crop is grown, there is a novel exposure to the transgene product. This could provide justification for an increased assessment of human health and environmental effects of the transgenic crop compared with a conventionally bred crop.

### **Concept of the precautionary principle**

The **precautionary principle** is an approach to handling uncertainty in the assessment and management of risk. This principle recommends that uncertainty, when it exists, be handled in favor of certain values (health and environment) over others. In other words, when our best predictions turn out to be in error, it is better to err on the side of safety. Another way of putting it is that, all things being equal, it is better to have foregone the important benefits of a technology by wrongly predicting its risks to health or the environment, than to have experienced harmful consequences by wrongly failing to predict them. In statistical terms, if an error in scientific prediction should occur, it is better to commit a Type I error of declaring a false positive (that is, erroneously predict an adverse effect where there is none), than a Type II error (erroneously predict no such effect when there actually is one). However, it is the custom of science that it is more serious a flaw in analysis to commit a Type I error (make a premature claim, e.g., reject the null hypothesis that a GM crop poses no significantly greater risk than its conventional counterpart) without adequate scientific evidence.

In view of the foregoing, it is not difficult to see why the precautionary principle has both proponents and

opponents. Proponents see it as a proactive and anticipatory strategy for protecting the public, environment, and animals from potential harm that is hard to predict by even the best science available. On the other hand, opponents view the precautionary principle as unscientific, a tool that promotes unfounded fear in the public and mitigates against research and development of new technologies.

This principle emerged in the 1970s and is currently invoked in numerous international laws, treaties, and protocols (e.g., the Cartegena Protocol on Biosafety of 2000). It is more cautiously interpreted in Europe than the USA. There are certain common criticisms of the precautionary principle. Some feel it is ambiguous and lacks uniform interpretation. Also, it marginalizes the role of scientists in that, whenever it is invoked, it usually tends to relax the standards of proof normally required by the scientific community. Others see the precautionary principle as a veiled form of trade protectionism. Specifically, nations may invoke this principle to circumvent the science-based decisions established in trade agreements and enforced by the World Trade Organization. Such rules generally require that a nation provide reliable scientific evidence to support its decision (e.g., to ban importation of a product). For example, the decision by the European markets to ban American and Canadian beef treated with rBST (growth hormone) is considered to be colored by protectionism.

### **Regulation and the issue of public trust**

The public needs to trust those who develop and implement the regulations that govern the development and application of technologies. It is widely accepted that even the most minimal risks may be unacceptable if levels of public trust in those who manage these risks are low or eroding. In Europe, the general public apprehension about the risks of GM foods is blamed to a large extent on the loss of public trust in scientists and regulatory bodies resulting from the bovine spongiform encephalopathy (BSE) crisis in Britain.

It is claimed that the assessment of biotechnology risks is a science-based activity. Consequently, it is important that the process be above reproach. The science should be of high quality, and the conduct of the assessment be independent and objective. There should be no conflict of interest in the regulatory process. Any association between producers and regulators is bound to cast doubt on the integrity of the process.

One factor in boosting public confidence is transparency of the regulatory process. During the application process, the applicant is required to submit certain data to the regulatory agency, though it can be claimed that parts of the submitted data are confidential business information (CBI) and should not be viewed by the public. Different regulatory agencies (in the US government and other governments) have somewhat different restraints on what sorts of information can be claimed as CBI. The question then is how much of the information should be divulged to the public and how much should remain proprietary information? Because the regulatory process is claimed to be science based, and because the custom of scientific enquiry is to be open and completely transparent, and further, because the decision of the regulatory authority is based on the scientific evidence, any attempt to withhold information involved in the decision-making process may cast doubt on the integrity of the process.

### Biosafety regulation at the international level

Because biotech products are accepted to varying extents in various countries, and because international trade involves crops that are targets for biotechnology, it is imperative that trading nations develop a consensus for biosafety regulation. An international delegation was convened to draft global regulatory guidelines, called the **Cartagena Protocol on Biosafety**. An outgrowth of the Convention on Biological Diversity (which was adopted at the 1992 Earth Summit in Rio de Janeiro), the Biosafety Protocol, which entered into force in September 2004, is designed to provide guidelines for signatories of the Protocol and their trading partners on the transfer, handling, and use of what are described as living modified organisms (LMOs) that have the potential to impact the conservation and sustainable use of biodiversity. The Cartagena Protocol on Biosafety has been interpreted by some to mean that LMOs intended for food, feed, or processing must be identified as LMOs.

Basically, an exporter of a product will be under obligation to provide the importer with information about the LMO regarding risk assessment and obtain consent prior to shipment. Critics of the Biosafety Protocol say that its implementation will adversely impact international trade by imposing severe trade barriers on a wide variety of biotech products (bulk

grain, processed food, drugs, etc.). The cost of goods will increase as shippers will have to segregate products, thereby increasing handling costs. Scientific development progress will also be impacted as scientists are compelled to pay more attention to special interest groups.

Biosafety regulation stringency is variable from one nation to another. In Japan, the Ministry of Agriculture, Forestry, and Fisheries is responsible for assessing environmental and feed safety, while the Ministry of Health and Welfare is responsible for food safety assessment. Basically, a product is subject to scrutiny if it was developed by rDNA technology. In Canada, the basis of assessment is the safety of the novel traits that have been incorporated, regardless of the technology used to produce the product. Gaining access to the European Union (EU) market is a complicated task. However, once approved, the product becomes legal in all the member countries of the EU. The product manufacturer or importer must submit a notification to the competent authority of the member state of the EU where the product is intended to be marketed. In China, the State Science and Technology Commission has the responsibility of developing a regulatory system for GM organisms. Regulations in developing countries are generally lacking. The Biosafety Protocol might be beneficial in this regard to assist the less industrialized economies in gaining market access to developed economies. There is no denying that a unified regulatory system of GM organisms would facilitate international trade involving these products. Unfortunately, a consensus that will be fully acceptable to all nations will be difficult to achieve in the near future.

### Labeling of biotechnology products

Some propose that consumers should have the right of “informed choice” about exposure to the risks of GM products. This push for labeling is partly because of the perception of lack of transparency from regulatory agencies, and the absence of balanced risk/benefit analyses. Because the first generation products of biotechnology benefited the food-producing industry directly, as previously indicated, consumers tend to view GM crops as geared towards enriching large corporations. Some consumer advocates would like to see all biotech foods labeled as such. The argument against labeling advanced by the biotechnology industry is viewed as an attempt to conceal information from the public. Opponents do not see a need for labeling since the FDA has ruled that

there is no inherent health risk in the use of biotechnology to develop new food products. The food industry opposes mandatory labeling because of the concern that such labeling could be interpreted as being “warning labels” implying that biotech foods are less safe or nutritious than their conventional counterparts.

The FDA requires a food product (including biotechnology foods) to be labeled if the following apply:

- 1 It contains a protein known to pose allergenic risk (e.g., milk, eggs, peanuts, tree nuts). Consequently, any genetic engineering involving gene transfer from any of these organisms must be labeled.
- 2 Its nutrient content as a result of the genetic manipulation is significantly different from what occurs in a normal product. For example, if a high level of protein is engineered into a cereal or root crop, the product must be labeled.

Opponents argue that labeling all biotechnologically produced foods would increase the cost of products as a result of the added cost of product segregation for the purpose of the so-called **identity preservation** of GM and non-GM products. To avoid contamination, biotech and conventional products must be kept apart at all phases of production, storage, processing, and distribution at additional cost. This would impact bulk or commodity products like grains (corn, wheat, soybean). However, specialty and high value fruits and vegetables are already identity preserved for premium prices.

Labeling of all products might be helpful to those who practice certain lifestyles or religious beliefs that impose strict dietary observances. A plant with an animal gene may not be acceptable to a strict vegetarian. However, studies have shown that both the kosher (Jewish) and halal (Muslim) communities have mechanisms in place to determine which products are acceptable to their adherents. Leaders of both religious groups have ruled that simple gene additions that lead to one or a few new components in a species are acceptable for their religious practices. However, the Muslim community has not resolved the issue regarding acceptability of gene transfer from swine into species, should that happen. Both Jewish and Muslim communities accept the use of bioengineered chymosin (rennin) in the cheese production.

Many countries have some form of labeling regulations or guidelines, which can be mandatory or voluntary. The primary forum for the discussion of food labeling at the international level is the Codex Alimentarius Commission. Mandatory labeling has been

implemented in the EU and is being implemented in Japan. In Europe, all products containing GM organisms must be labeled as such. Even where mixtures of conventional and GM organisms are concerned, a label must be provided to indicate that GM organisms may be present. The USA and Canada require GM food products that could have health or safety effects (possible allergens or changes in nutritional content from acceptable levels) to be labeled.

In North America, labeling is generally thought to be necessary only when there is some feature of the product itself that needs to be brought to consumers’ attention (e.g., health risk or nutritional issue). The process by which the product is produced (e.g., by genetic modification) is considered inconsequential. This is described as product-based (as opposed to process-based) regulation. An exception to this approach in the USA and Canada is the requirement that food subjected to the processes of irradiation be labeled. In the USA, the FDA and the courts generally consider reference should be made to a “material fact” about the product that is pertinent to nutritional value or safety. This affirms the concept of substantial equivalence in which a new food product that is substantially equivalent to existing products is exempt from labeling.

### Economic impact of labeling and regulations

The economic impact of food regulations and labeling on trade depends on the products involved, the cost of labeling, and sometimes how consumers use such information. The cost of labeling will depend on the stringency imposed. That is, whether “zero tolerance” or “minimum tolerance” of GM product is the goal. Implementing the former standard would require expensive safeguards to be implemented to avoid cross-contamination. Harvesting, processing, shipping, and other product handling would require modification.

Government approval can have severe adverse consequences on trade. For example, sale of US corn in EU countries was dealt a devastating blow in 1999 because certain GM corn varieties were not approved for sale in the EU. This action caused US corn export to the EU to drop from US\$190 million in 1997 to \$35 million in 1998, and then to a low of \$6 million in 1999. Consumer response to labeling has an impact on product demand. Sometimes, products intended for use as feed may not require labeling.

## Public perceptions and fears about biotechnology

Stakeholders generally acknowledge that no technology is without potential for adverse impacts. Another truth is that the perception of technologies varies between the scientists who develop them and the public who use them. It is tempting to contrast public perception, which is largely subjective, with scientific knowledge, which is largely objective. Further, public perception is viewed by some as basically illogical, irrational, and based on emotions rather than scientific facts. There is the temptation to condescend and disregard public opinion as uninformed. However, scientists also have perceptions of the technologies they develop that are not necessarily rooted in facts, because the knowledge available may be scanty or inconclusive. There is the tendency, therefore, to either overestimate or underestimate the benefits and risks of technology.

Because of the imperfections that plague both camps, there are some misconceptions about biotechnology in the general community that hinder its development and application. A few of these will be introduced for the sole purpose of initiating discussion on the subject.

### Techniques of biotechnology are alien, unnatural, and too radical

It was made clear in Chapter 1 that the definition of biotechnology can be broad or narrow. In the broad sense, organisms have been used to make products for thousands of years (e.g., yeast in bakery products and bacteria in fermented products). In the narrow sense, biotechnology allows genes to be transferred unrestricted among living things, in effect disregarding natural genetic barriers. Whereas such interspecies gene exchange is not the norm in nature, there are examples of various degrees of such horizontal gene transfers, especially on evolutionary time scales. Cross-pollinated species propagate through gene mixing, normally within the species. A more dramatic natural mixing occurred in wheat – common wheat (*Triticum aestivum*) is an allopolyploid (hexaploid) consisting of three genomes of three different species. Certain microbes have the capacity to transfer some of their genetic material into the hosts they infect, even though the outcome may be undesirable for the hosts.

In biotechnology, the gene transfer system of choice in plants is *Agrobacterium*-mediated transfer. This bacterium naturally transfers a portion of its genome into the plant it infects, a bacterium to plant gene transfer

that forms galls in the plants. In biotechnology, scientists capitalize on this natural process to transfer genes of choice, only the result is not disease because the bacterium is disarmed prior to use.

The normal direction of genetic information transfer is from DNA to RNA to protein (the so-called central dogma of molecular biology). However, there are certain viruses that have RNA as genetic material. In retroviral infections, the single-stranded RNA is reverse transcribed to single-stranded DNA and then doubled to double-stranded DNA. This is incorporated into the host genome. Scientists can go a step further, and replace the disease-causing components of the retroviral genome with the desired genes for incorporation.

Mutations or heritable genetic changes occur naturally as spontaneous events. Such natural gene alterations produce variability for evolutionary processes to occur. Instead of haphazard and random events, scientists are able to induce the genetic changes that are desired.

It is obvious from these selected examples that science merely imitates nature after studying to understand it. Rather than random events, scientists attempt to nudge nature purposefully to the advantage of humans. One may argue that just because nature does it, does not mean humans should do the same. But then another may argue why not?

### Genetic engineering is an exact science

It is true that specific genes can be identified, isolated, and characterized. However, the current gene transfer systems leave much to be desired. Once the DNA is delivered into the cell, scientists are not able to direct or predict where it will be inserted in the genome. Consequently, scientists cannot predict precisely the outcome of a transformation event. Where the gene inserts itself in the genome has a bearing on its expression. Even though this appears to be an uncontrollable event that could lead to unexpected phenotypes, scientists screen the products of transformation to identify the individual(s) in which the transgene apparently has been properly inserted and is functioning as desired. It should be pointed out, that, compared with traditional breeding in which transfer of a desirable gene is usually accompanied by the transfer of numerous others, genetic engineering is relatively very precise.

### Pesticide resistance in the agroecosystem as a result of the use of biotech crops is unavoidable

There are major crops in production with engineered, targeted resistance to pests and herbicides. Pesticide-

resistance concerns (or “collateral resistance”) center around four main aspects: (i) creating weeds out of cultivated cultivars; (ii) creating “super weeds” from existing weeds; (iii) creating resistant pests; and (iv) creating antibiotic resistance in harmful microbes.

The problems vary in the probability of occurring. The fact is that pests always manage to eventually adapt to any pest management strategy that is implemented repeatedly over a long period. This is especially true when the organism has a short life cycle (e.g., bacteria and many insects). The fear of creating a weed out of cultivated crops engineered to be herbicide tolerant stems from the fact that most modern herbicides have a broad-spectrum action (i.e., kill many plant species). Bioengineered herbicide-resistant crops are consequently resistant to broad-spectrum herbicides. If, for example Roundup Ready® soybean follows Roundup Ready® corn, volunteer corn plants will resist Roundup® and be a weed problem. Whereas crop rotations are desirable, it should not involve crops engineered with identical herbicide resistance. Further, using the same herbicide repeatedly for a long time is not a recommended agronomic practice, anyhow. Should a farmer decide not to heed this advice, there are herbicides besides glyphosate that can control Roundup Ready® corn or soybean “weeds”.

On the issue of biotechnology contributing to the development of “super weeds”, the potential exists for gene escape from cultivated species engineered for herbicide tolerance to interbreed with wild relatives, thereby creating more competitive and difficult to control weeds (so-called **super weeds**). While the movement of transgenes into wild relatives of transgenic crops is possible, this would occur only if the cultivated crop species are grown where their weedy relatives with which they can interbreed also occur. This is not the case for the major crops that are transgenic for herbicide resistance in the USA (e.g., corn, soybean, potato). However, it is the case for squash and canola. It should be pointed out that the development of resistance to herbicides by weeds and other plants does occur, following prolonged exposure to certain herbicides. Further, irrespective of the herbicide or weed management tactic used, resistance to the chemical over the long haul is inevitable. This is why new herbicides will continue to be needed. There is no evidence to suggest that development of resistance is more problematic with the use of transgenic crops than direct use of herbicides. The issue of insects developing resistance to transgenic crops is similar to the situation in weeds. Pests routinely overcome management tactics used against them.

To reduce the rate of insect-resistance development to PIPs, producers of transgenic crops require customers to grow a refuge of non-biotech and untreated crops around the *Bt* crops. The idea is that the refuge maintains an insect population susceptible to *Bt* that can interbreed with any pests on the biotech product that may develop resistance, and prevent the resistance from becoming established in the population.

Certain biotechnology techniques utilize antibiotic-resistance markers in developing transgenic crops. Consequently, the products contain these genes that can possibly be transferred to microbes in the environment. The fact is that the antibiotic-resistance markers used in crop development do not provide resistance to most of the antibiotics used in the clinical setting.

### Biotechnology products are unnatural and unsafe

Nature, and therefore anything natural, is perceived by some as superior to anything artificial. Modern foods that have been processed and modified in all and sundry ways are making us sick, they perceive; hence, the booming organic food and health food markets. Herbal medicine is being actively promoted in Western societies. The public is concerned about biotechnology inadvertently introducing undesirable and unnatural chemicals into the food chain. The fact is that the public has embraced artificial components in food and medicine for a long period of time. Western therapy is almost exclusively dependent upon synthetic pharmaceuticals. Food additives and coloring are used routinely in both home and industrial food preparations. Instead of adding these materials to food during preparation, biotechnology seeks to make plants and animals produce nutrition-augmenting materials via natural processes to be included in the plant and animal tissues. For example, instead of postharvest vitamin-enriching of the product (e.g., rice) for value added, the plant is engineered with the capacity to produce pro-vitamin A (“Golden Rice”).

### Some concerns of plant breeders

Apart from these social concerns, some of the key concerns of interest to plant breeders include the following:

- 1 Genetic engineering technology produces transgenic plants that are often useful as breeding materials. The transgenic trait must be transferred into adapted cultivars (via backcrossing) to develop useful cultivars. Breeding by genetic engineering, hence, can be a lengthy process and expensive.

- 2 The transfer of proprietary genes into public germplasm has the potential to restrict the free access and distribution of germplasm among researchers.

### **Legal risks that accompany adoption of GM crops**

As biotech products enter the food and agriculture production chain, their adoption is accompanied by a variety of legal risks. Some of these legal issues are briefly discussed next.

#### **Tort liability versus regulatory approval**

Biotechnology research and application are highly regulated by the federal government as previously discussed. These regulations include the conditions under which a certain GM crop or GM organism or products will be approved for the safe use by humans and concerning the safety of the environment. Regulations also affect how the products are produced, marketed, and used.

However, before or after regulatory approval, a GM plant or organism could cause damage to property, persons, markets, environment, or social structure. **Tort** is a civil legal action whereby the claimant alleges injury or wrong, arising independent of contact, to the person or property of the claimant. Both producers and users of biotech products are subject to these kinds of liability, which may exist for these kinds of possible damage, as discussed by D. L. Kershen.

#### **Damage to property**

##### *Trespass*

The most likely grounds on which property damage may be alleged are as follows:

- 1 Pollen flow from a GM crop to a non-GM crop.
- 2 Mechanical mixing of seeds or plant parts during harvesting, storage, or transportation.

Producers of specialty crops (e.g., organic crop producers) may allege this damage and bring a common law cause of action based on **trespass** (the physical invasion of the possessory interests of the property (land) of the person claiming damages). Physical spread of GM pollen to a neighbor's field is enough for this infringement to occur. Because pollen flow is a natural process that occurs freely, it is important that in such a law suit,

the plaintiff establish that the physical invasion caused damage.

##### *Strict liability*

Damage to property may also be alleged and a tort claim in **strict liability** brought by a plaintiff. This liability is a case of liability without fault despite the exercise of utmost care, and can be considered if it is shown that the activity of growing a GM crop is abnormally dangerous. The plaintiff would have to show: (i) a high degree of risk of some harm to the person, land, or chattels of others exists; (ii) there is a likelihood that the harm that results will be great; (iii) there is an inability to eliminate the risk by the exercise of reasonable care; (iv) the extent to which the activity is not a matter of common usage; (v) there is inappropriateness of the activity to the place where it is conducted; and (vi) the extent to which its value to the community is outweighed by its dangerous attributes.

##### *Negligence*

**Negligence** is a fault-based claim that alleges that property damage resulted because a neighbor growing a GM crop failed to take adequate precautions. In this case, both the farmer and the agrobiotech company that developed the GM cultivar are liable. It is for this reason that refuge rows are required by growers of GM crops.

##### *Private nuisance*

The owner of a property has a right to the private use and enjoyment of land free from non-trespassory invasion by another. The plaintiff must prove that nearby fields of GM crops have unreasonably interfered with the use and enjoyment of one's own land.

#### **Damage to person**

A **damage to person claim** may be brought by a plaintiff alleging damage arising from the toxicity of a GM crop or its product, or an allergic response to these crops or their food products. The claim may also be based on an alleged damage from long-term exposure to GM crops or that their food products cause ill effects to health. A case in point is the Starlink® corn in human food, involving Aventis Corporation. Pharmaceutical GM crops have the highest likelihood to present such a risk.

### Damage to economic interest (market)

Some producers have brought claims to the effect that the introduction of GM crops increased their production costs and equipment costs, while decreasing the prices of their agricultural products. In other words, the presence *per se* of GM crops in the agricultural sector has affected market access and market prices for their non-

GM crops in a general way. Of relevance in such a case is the level or type of approval given by the regulating authority for the GM crop (i.e., limited use, all use, limited to USA, etc.). A case in point, again, is the Starlink® GM cultivar that was approved for limited use (animal feed), but found its way into the food chain. Such a case is relatively easy to prosecute with zero tolerance as the regulatory standard.

### References and suggested reading

- Acquaah, G. 2004. Understanding biotechnology: An integrated and cyber-based approach. Prentice Hall, Upper Saddle River, NJ.
- Baumgardt, B.R., and M.A. Martin (eds). 1991. Agricultural biotechnology: Issues and choices. Purdue University Agricultural Experiment Station, West Lafayette, IN.
- Borlaug, N.E. 2000. Ending world hunger. The promise of biotechnology and the threat of antiscience zealotry. *Plant Physiol.* 124:487–490.
- Brill, W. 1985. Safety concerns and genetic engineering in agriculture. *Science* 227:381–384.
- Crespi, R.S. 1988. A basic guide to patenting in biotechnology. Cambridge University Press, Cambridge, UK.
- Gaythwaite, D.M. 1999. Intellectual property and technical know-how. *In:* Biotechnology. The science and business (Moses, V., and R.E. Cape, eds). Harwood Academic Publishers, London.
- Kershen, D.L. 2004. Legal liability issues in agricultural biotechnology. *Crop Sci.* 44:456–463.
- Lesser, W., and M.A. Mutschler. 2004. Balancing investment incentives and social benefits when protecting plant varieties: Implementing initial variety systems. *Crop Sci.* 44:1113–1120.
- National Research Council. 1997. Intellectual property rights and plant biotechnology. National Academy Press, Washington, DC.
- Royal Society of Canada Expert Panel on the Future of Food Biotechnology. 2001. Report. Royal Society of Canada, Ottawa.
- Sanders, P.R., C.L. Thomas, M.E. Groth, J.M. Astwood, and R.L. Fuchs. 1998. Safety assessment of insect-protected corn. *In:* Biotechnology and safety assessment, 2nd edn (Thomas, J.A., ed.). Taylor & Francis, Philadelphia.
- Sechley, K.A., and H. Schroeder. 2002. Intellectual property protection of plant biotechnology innovations. *Trends Biotechnol.* 20:456–461.
- UPOV. 1996. Model law on the protection of new varieties of plants. Union for the Protection of New Varieties of Plants, Geneva.

### Outcomes assessment

#### Part A

Please answer the following questions true or false:

- 1 When a single gene is patented by more than one company, the patent is called a reach-through patent.
- 2 Mixtures of ingredients can be patented.
- 3 A plant patent is awarded for 30 years.
- 4 A patent is described as a “negative right”.
- 5 The FDA is responsible for monitoring foods for allergenicity.
- 6 The Biosafety Protocol was drafted in 2002.
- 7 *Bt* crop products are monitored by the EPA.

**Part B**

Please answer the following questions:

- 1 What is a patent?
- 2 What is the importance of intellectual property rights in biotechnology?
- 3 Not all inventions are patentable. Explain.
- 4 What is a utility patent?
- 5 Patents are not always the best protection of an invention against illegal use. Explain.
- 6 A patent is not a “positive right”. Explain.
- 7 Give two specific public concerns about the development and application of biotechnology.
- 8 What agencies oversee the regulation of biotechnology in the USA?
- 9 Describe how the FDA regulates products.
- 10 How is regulation of biotechnology in the USA different from what obtains in the EU?
- 11 What are the pros and cons of labeling for the biotechnology industry?
- 12 Give four common sources of allergens. Discuss the importance of allergenicity in risk assessment.

**Part C**

Please write a brief essay on each of the following topics:

- 1 Discuss the concept of “freedom of use” in patent law.
- 2 List and discuss the basic steps in applying for a patent.
- 3 Distinguish between ethics, morals, and values.
- 4 In terms of ethics, what is your opinion about the development and application of biotechnology?
- 5 What role should the public play in the way the biotechnology industry operates?
- 6 Discuss the concept of substantial equivalence in risk assessment.
- 7 Discuss the precautionary rule and its application in risk regulation.
- 8 The public must view the regulating agencies of biotechnology as trustworthy. Explain why this is necessary.
- 9 List and discuss two common misconceptions about biotechnology.
- 10 No technology is without potential for adverse impacts. Explain.

# **Section 6**

## **Classic methods of plant breeding**

---

**Chapter 16 Breeding self-pollinated species**  
**Chapter 17 Breeding cross-pollinated species**  
**Chapter 18 Breeding hybrid cultivars**

---

Methods of breeding (or precisely, methods of selection) crops vary according to the natural method of reproduction of the species. Generally, there are two categories of breeding methods: those for self-pollinated species and those for cross-pollinated species. In practice, there is no hard distinction between the two; breeders crossover and use methods as they find useful. Furthermore, plant breeders may use a combination of several methods in one breeding program, using one procedure at the beginning and switching to another along the way. It should be mentioned also that the steps described in the various chapters for each selection method are only suggested guidelines. Breeders may modify the steps, regarding the number of plants to select, the number of generations to use, and other aspects of breeding, to suit factors such as budget and the nature of the trait being improved.

---



# 16

# Breeding self-pollinated species

---

## Purpose and expected outcomes

*As previously discussed, self-pollinated species have a genetic structure that has implication in the choice of methods for their improvement. They are naturally inbred and hence inbreeding to fix genes is one of the goals of a breeding program for self-pollinated species in which variability is generated by crossing. However, crossing does not precede some breeding methods for self-pollinated species. The purpose of this chapter is to discuss specific methods of selection for improving self-pollinated species. After studying this chapter, the student should be able to discuss the characteristics, application, genetics, advantages, and disadvantages of the following methods of selection:*

- 1 Mass selection.
- 2 Pure-line selection.
- 3 Pedigree selection.
- 4 Bulk population.
- 5 Single-seed descent.

*And to:*

- 6 Describe the technique/method of backcrossing.
  - 7 Discuss the method of multiline breeding.
  - 8 Discuss the method of breeding composites.
  - 9 Discuss the method of recurrent selection.
- 

## Types of cultivars

At the beginning of each project, the breeder should decide on the type of cultivar to breed for release to producers. The breeding method used depends on the type of cultivar to be produced. There are six basic types of cultivars that plant breeders develop. These cultivars derive from four basic populations used in plant breeding – **inbred pure lines, open-pollinated populations, hybrids, and clones**. Plant breeders use a variety of methods and techniques to develop these cultivars.

### *Pure-line cultivars*

**Pure-line cultivars** are developed for species that are highly self-pollinated. These cultivars are homogeneous and homozygous in genetic structure, a condition attained through a series of self-pollinations. These cultivars are often used as parents in the production of other kinds of cultivars. Pure-line cultivars have a narrow genetic base. They are desired in regions where uniformity of a product has a high premium.

### *Open-pollinated cultivars*

Contrary to pure lines, **open-pollinated cultivars** are developed for species that are naturally cross-pollinated. The cultivars are genetically heterogeneous and heterozygous. Two basic types of open-pollinated cultivars are developed. One type is developed by improving the general population by **recurrent** (or repeated) **selection** or bulking and increasing material from selected superior inbred lines. The other type, called a **synthetic cultivar**, is derived from planned matings involving selected genotypes. Open-pollinated cultivars have a broad genetic base.

### *Hybrid cultivars*

**Hybrid cultivars** are produced by crossing inbred lines that have been evaluated for their ability to produce hybrids with superior vigor over and above those of the parents used in the cross. Hybrid production exploits the phenomenon of hybrid vigor (or heterosis) to produce superior yields. Heterosis is usually less important in crosses involving self-pollinated species than in those involving cross-pollinated species. Hybrid cultivars are homogeneous but highly heterozygous. Pollination is highly controlled and restricted in hybrid breeding to only the designated pollen source. In the past, physical human intervention was required to enforce this strict pollination requirement, making hybrid seed expensive. However, with time, various techniques have been developed to capitalize on natural reproductive control systems (e.g., male sterility) to facilitate hybrid production. Hybrid production is more widespread in cross-pollinated species (e.g., corn, sorghum), because the natural reproductive mechanisms (e.g., cross-fertilization, cytoplasmic male sterility) are more readily economically exploitable than in self-pollinated species.

### *Clonal cultivars*

Seeds are used to produce most commercial crop plants. However, a significant number of species are propagated by using plant parts other than seed (vegetative parts such as stems and roots). By using vegetative parts, the cultivar produced consists of plants with identical genotypes and is homogeneous. However, the cultivar is genetically highly heterozygous. Some plant species sexually reproduce but are propagated clonally (vegetatively) by choice. Such species are improved through hybridization, so that when hybrid vigor exists it can be fixed (i.e., the vigor is retained from one generation to

another), and then the improved cultivar propagated asexually. In seed-propagated hybrids, hybrid vigor is highest in the  $F_1$ , but is reduced by 50% in each subsequent generation. In other words, whereas clonally propagated hybrid cultivars may be harvested and used for planting the next season's crop without adverse effects, producers of sexually reproducing species using hybrid seed must obtain a new supply of seed, as previously indicated.

### *Apomictic cultivars*

**Apomixis** is the phenomenon of the production of seed without the benefit of the union of sperm and egg cells (i.e., without fertilization). The seed harvested is hence genetically identical to the mother plant (in much the same way as clonal cultivars). Hence, apomictic cultivars have the same benefits of clonally propagated ones, as previously discussed. In addition, they have the convenience of vegetative propagation through seed (versus propagation through cuttings or vegetative plant parts). Apomixis is common in perennial forage grasses.

### *Multilines*

**Multilines** are developed for self-pollinating species. These cultivars consist of a mixture of specially developed genotypes called **isolines** (or **near isogenic lines**) because they differ only in a single gene (or a defined set of genes). Isolines are developed primarily for disease control, even though these cultivars could, potentially, be developed to address other environmental stresses. Isolines are developed by using the techniques of back-crossing in which the  $F_1$  is repeatedly crossed to one of the parents (recurrent parent) that lacked the gene of interest (e.g., disease resistance).

## **Genetic structure of cultivars and its implications**

The products of plant breeding that are released to farmers for use in production vary in genetic structure and consequently the phenotypic uniformity of the product. Furthermore, the nature of the product has implications in how it is maintained by the producers, regarding the next season's planting.

### **Homozygous and homogeneous cultivars**

A cultivar may be genetically homozygous and hence produce a homogeneous phenotype or product.

Self-pollinated species are naturally inbred and tend to be homozygous. Breeding strategies in these species are geared toward producing cultivars that are homozygous. The products of economic importance are uniform. Furthermore, the farmer may save seed from the current season's crop (where legal and applicable) for planting the next season's crop, without loss of cultivar performance, regarding yield and product quality. This attribute is especially desirable to producers in many developing countries where the general tradition is to save seed from the current season for planting the next season. However, in developed economies with well-established commercial seed production systems, intellectual property rights prohibit the reuse of commercial seed for planting the next season's crop, thus requiring seasonal purchase of seed by the farmer from seed companies.

### Heterozygous and homogeneous cultivars

The method of breeding of certain crops leaves the cultivar genetically heterozygous yet phenotypically homogeneous. One such method is hybrid cultivar production, a method widely used for the production of especially outcrossing species such as corn. The heterozygous genetic structure stems from the fact that a hybrid cultivar is the  $F_1$  product of a cross of highly inbred (repeatedly selfed, homozygous) parents. Crossing such pure lines produces highly heterozygous  $F_1$  plants. Because the  $F_1$  is the final product released as a cultivar, all plants are uniformly heterozygous and hence homogeneous in appearance. However, the seed harvested from the  $F_1$  cultivar is  $F_2$  seed, consequently producing maximum heterozygosity and heterogeneity upon planting. The implication for the farmer is that the current season's seed cannot be saved for planting the next season's crop for obvious reasons. The farmer who grows hybrid cultivars must purchase fresh seed from the seed company for planting each season. Whereas this works well in developed economies, hybrids generally do not fit well into the farming systems of developing countries where farmers save seed from the current season for planting the next season's crop. Nonetheless, the use of hybrid seed is gradually infiltrating crop production in developing countries.

### Heterozygous and heterogeneous cultivars

Other approaches of breeding produce heterozygous and homogeneous (relatively) cultivars, for example, synthetic and composite breeding. These approaches

will allow the farmer to save seed for planting. Composite cultivars are suited to production in developing countries, while synthetic cultivars are common in forage production all over the world.

### Homozygous and heterogeneous cultivars

Examples of such a breeding product are the mixed landrace types that are developed by producers. The component genotypes are homozygous, but there is such a large amount of diverse genotypes included that the overall cultivar is not uniform.

### Clonal cultivar

Clones, by definition, produce offspring that are not only identical to each other but also to the parent. Clones may be very heterozygous but whatever advantage heterozygosity confers is locked in for as long as propagation is clonally conducted. The offspring of a clonal population are homogeneous. Once the genotype has been manipulated and altered in a desirable way, for example through sexual means (since some species are flowering, but are vegetatively propagated and not through seed), the changes are fixed for as long as clones are used for propagation. Flowering species such as cassava and sugarcane may be genetically improved through sex-based methods, and thereafter commercially clonally propagated.

### Types of self-pollinated cultivars

In terms of genetic structure, there are two types of self-pollinated cultivars:

- 1 Those derived from a single plant.
- 2 Those derived from a mixture of plants.

Single-plant selection may or may not be preceded by a planned cross but often it is the case. Cultivars derived from single plants are homozygous and homogeneous. However, cultivars derived from plant mixtures may appear homogeneous but, because the individual plants have different genotypes, and because some outcrossing (albeit small) occurs in most selfing species, heterozygosity would arise later in the population. The methods of breeding self-pollinated species may be divided into two broad groups – those preceded by hybridization and those not preceded by hybridization.

## Common plant breeding notations

Plant breeders use shorthand to facilitate the documentation of their breeding programs. Some symbols are standard genetic notations, while others were developed by breeders. Unfortunately there is no one comprehensive and universal system in use, making it necessary, especially with the breeding symbols, for the breeder to always provide some definitions to describe the specific steps in a breeding method employed in the breeding program.

### Symbols for basic crosses

- 1 **F**. The symbol F (for **filial**) denotes the progeny of a cross between two parents. The subscript ( $x$ ) represents the specific generation ( $F_x$ ). If the parents are homozygous, the  $F_1$  generation will be homogeneous. Crossing of two  $F_1$  plants (or selfing an  $F_1$ ) yields an  $F_2$  plant ( $F_1 \times F_1 = F_2$ ). Planting seed from the  $F_2$  plants will yield an  $F_2$  population, the most diverse generation following a cross, in which plant breeders often begin selection. Selfing  $F_2$  plants produces  $F_3$  plants, and so on. It should be noted that the seed is one generation ahead of the plant, that is, an  $F_2$  plant bears  $F_3$  seed.
- 2  $\otimes$ . The symbol  $\otimes$  is the notation for selfing.
- 3 **S**. The S notation is also used with numeric subscripts. In one usage  $S_0 = F_1$ ; another system indicates  $S_0 = F_2$ .

### Symbols for inbred lines

Inbred lines are described by two systems. System I describes an inbred line based on the generation of plants that are being currently grown. System II describes both the generation of the plant from which the line originated as well as the generation of plants being currently grown. Cases will be used to distinguish between the two systems.

- Case 1.** The base population is  $F_2$ . The breeder selects an  $F_2$  plant from the population and plants the  $F_3$  seeds in the next season.

System I: the planted seed produces an  $F_3$  line.  
 System II: the planted seed produces an  $F_2$  derived line in  $F_3$  or an  $F_{2:3}$  line.

If seed from the  $F_3$  plants is harvested and bulked, and the breeder samples the  $F_4$  seed in

the next season, the symbolism will be as follows:

System I: the planted seed produces an  $F_4$  line.  
 System II: the planted seed produces an  $F_2$  derived line in  $F_4$  or an  $F_{2:4}$  line.

- Case 2.** The breeder harvests a single  $F_4$  and plants  $F_5$  seed in a row.

System I: the planted row produces an  $F_5$  line.  
 System II: the planted row constitutes an  $F_4$  derived line in  $F_5$  or an  $F_{4:5}$  line.

Similarly the S notation may be treated likewise. Taking case 1 for example:

System I:  $S_1$  line.  
 System II:  $S_0$  derived line in  $S_1$  or an  $S_{0:1}$  line.

### Notation for pedigrees

Knowing the **pedigree** or ancestry of a cultivar enables the plant breeder to retrace the steps in a breeding program to reconstitute a cultivar. Plant breeders follow a short-hand system of notations to write plant pedigrees. Some pedigrees are simple, others are complex. Some of the common notations are as follows:

- 1 A slash, /, indicates a cross.
- 2 A figure between slashes, //, indicates the sequence or order of crossing. A //2/ is equivalent to // and indicates the second cross. Similarly, / is the first cross, and /// the third cross.
- 3 A backcross is indicated by \*; \*3 indicates the genotype was backcrossed three times to another genotype.

The following examples will be used to illustrate the concept.

#### Pedigree 1: MSU48-10/3/Pontiac/Laker/2/MS-64.

Interpretation:

- (a) The first cross was Pontiac (as female)  $\times$  Laker (as male).
- (b) The second cross was [Pontiac/Laker (as female)]  $\times$  MS-64 (as male).
- (c) The third cross was MSU48-10 (as female)  $\times$  [Pontiac/Laker//MS-64 (as male)].

**Pedigree 2:** MK2-57\*3/SV-2.

Equivalent formula: MK2-57/3/MK2-57/2/MK2-57/SV-2.

Interpretation: the genotype MK2-57 was backcrossed three times to genotype SV-2.

## Mass selection

Mass selection is an example of selection from a biologically variable population in which differences are genetic in origin. The Danish biologist, W. Johansen, is credited with developing the basis for mass selection in 1903. Mass selection is often described as the oldest method of breeding self-pollinated plant species. However, this by no means makes the procedure outdated. As an ancient art, farmers saved seed from desirable plants for planting the next season's crop, a practice that is still common in the agriculture of many developing countries. This method of selection is applicable to both self- and cross-pollinated species.

### Key features

The purpose of mass selection is population improvement through increasing the gene frequencies of desirable genes. Selection is based on plant phenotype and one generation per cycle is needed. Mass selection is imposed once or multiple times (recurrent mass selection). The improvement is limited to the genetic variability that existed in the original populations (i.e., new variability is not generated during the breeding process). The goal in cultivar development by mass selection is to improve the average performance of the base population.

### Applications

As a modern method of plant breeding, mass selection has several applications:

- 1 It may be used to maintain the purity of an existing cultivar that has become contaminated, or is segregating. The off-types are simply rogued out of the population, and the rest of the material bulked. Existing cultivars become contaminated over the years by natural processes (e.g., outcrossing, mutation) or by human error (e.g., inadvertent seed mixture during harvesting or processing stages of crop production).
- 2 It can also be used to develop a cultivar from a base population created by hybridization, using the procedure described next.

- 3 It may be used to preserve the identity of an established cultivar or soon-to-be-released new cultivar. The breeder selects several hundreds (200–300) of plants (or heads) and plants them in individual rows for comparison. Rows showing significant phenotypic differences from the other rows are discarded, while the remainder is bulked as breeder seed. Prior to bulking, sample plants or heads are taken from each row and kept for future use in reproducing the original cultivar.
- 4 When a new crop is introduced into a new production region, the breeder may adapt it to the new region by selecting for key factors needed for successful production (e.g., maturity). This, hence, becomes a way of improving the new cultivar for the new production region.
- 5 Mass selection can be used to breed horizontal (durable) disease resistance into a cultivar. The breeder applies low densities of disease inoculum (to stimulate moderate disease development) so that quantitative (minor gene effects) genetic effects (instead of major gene effects) can be assessed. This way, the cultivar is race-non-specific and moderately tolerant of disease. Further, crop yield is stable and the disease resistance is durable.
- 6 Some breeders use mass selection as part of their breeding program to rogue out undesirable plants, thereby reducing the materials advanced and saving time and reducing costs of breeding.

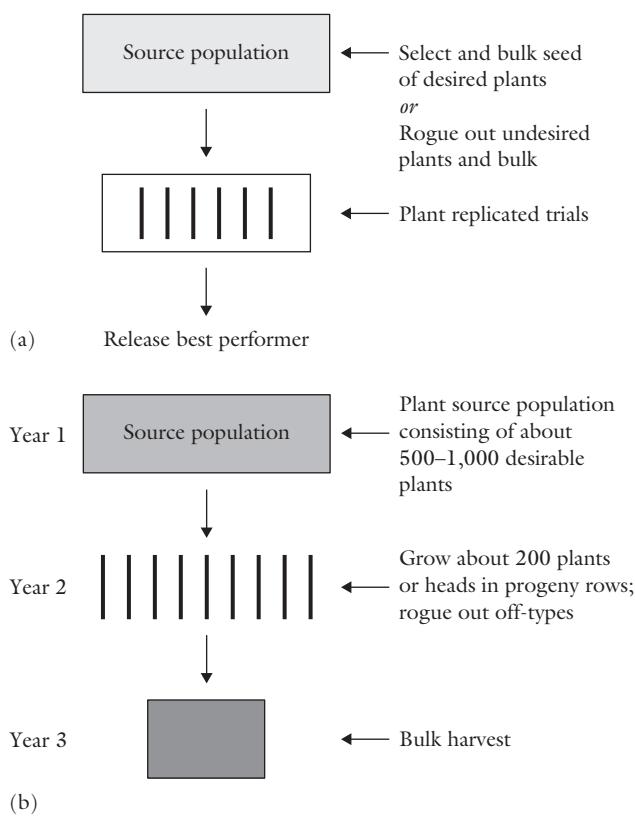
### Procedure

#### Overview

The general procedure in mass selection is to rogue out off-types or plants with undesirable traits. This is called by some researchers, **negative mass selection**. The specific strategies for retaining representative individuals for the population vary according to species, traits of interest, or creativity of the breeder to find ways to facilitate the breeding program. Whereas roguing out and bulking appears to be the basic strategy of mass selection, some breeders may rather select and advance a large number of plants that are desirable and uniform for the trait(s) of interest (**positive mass selection**). Where applicable, single pods from each plant may be picked and bulked for planting. For cereal species, the heads may be picked and bulked.

#### Steps

The breeder plants the heterogeneous population in the field, and looks for off-types to remove and discard



**Figure 16.1** Generalized steps in breeding by mass selection for (a) cultivar development, and (b) purification of an existing cultivar.

(Figure 16.1). This way, the original genetic structure is retained as much as possible. A mechanical device (e.g., using a sieve to determine which size of grain would be advanced) may be used, or selection may be purely on visual basis according to the breeder's visual evaluation. Further, selection may be based on targeted traits (direct selection) or indirectly by selecting a trait correlated with the trait to be improved.

**Year 1** If the objective is to purify an established cultivar, seed of selected plants may be progeny-rowed to confirm the purity of the selected plants prior to bulking. This would make a cycle of mass selection have a 2-year duration instead of 1 year. The original cultivar needs to be planted alongside for comparison.

**Year 2** Evaluate composite seed in a replicated trial, using the original cultivar as a check. This test may be conducted at different locations and over several years. The seed is bulk harvested.

### Genetic issues

Contamination from outcrossing may produce heterozygotes in the population. Unfortunately, where a dominance effect is involved in the expression of the trait, heterozygotes are indistinguishable from homozygous dominant individuals. Including heterozygotes in a naturally selfing population will provide material for future segregations to produce new off-types. Mass selection is most effective if the expression of the trait of interest is conditioned by additive gene action.

Mass selection may be conducted in self-pollinated populations as well as cross-pollinated populations, but with different genetic consequences. In self-pollinated populations, the persistence of inbreeding will alter population gene frequencies by reducing heterozygosity from one generation to the next. However, in cross-pollinated populations, gene frequencies are expected to remain unchanged unless the selection of plants was biased enough to change the frequency of alleles that control the trait of interest.

Mass selection is based on plant phenotype. Consequently, it is most effective if the trait of interest has high heritability. Also, cultivars developed by mass selection tend to be phenotypically uniform for qualitative (simply inherited) traits that are readily selectable in a breeding program. This uniformity notwithstanding, the cultivar could retain significant variability for quantitative traits. It is helpful if the selection environment is uniform. This will ensure that genetically superior plants are distinguishable from mediocre plants. When selecting for disease resistance, the method is more effective if the pathogen is uniformly present throughout the field without "hot spots".

Some studies have shown correlated response to selection in secondary traits as a result of mass selection. Such a response may be attributed to linkage or pleiotropy.

### Advantages and disadvantages

Some of the major advantages and disadvantages of mass selection for improving self-pollinated species are given here.

#### Advantages

- 1 It is rapid, simple, and straightforward. Large populations can be handled and one generation per cycle can be used.
- 2 It is inexpensive to conduct.

- 3 The cultivar is phenotypically fairly uniform even though it is a mixture of pure lines, hence making it genetically broad-based, adaptable, and stable.

### Disadvantages

- 1 To be most effective, the traits of interest should have high heritability.
- 2 Because selection is based on phenotypic values, optimal selection is achieved if it is conducted in a uniform environment.
- 3 Phenotypic uniformity is less than in cultivars produced by pure-line selection.
- 4 With dominance, heterozygotes are indistinguishable from homozygous dominant genotypes. Without progeny testing, the selected heterozygotes will segregate in the next generation.

### Modifications

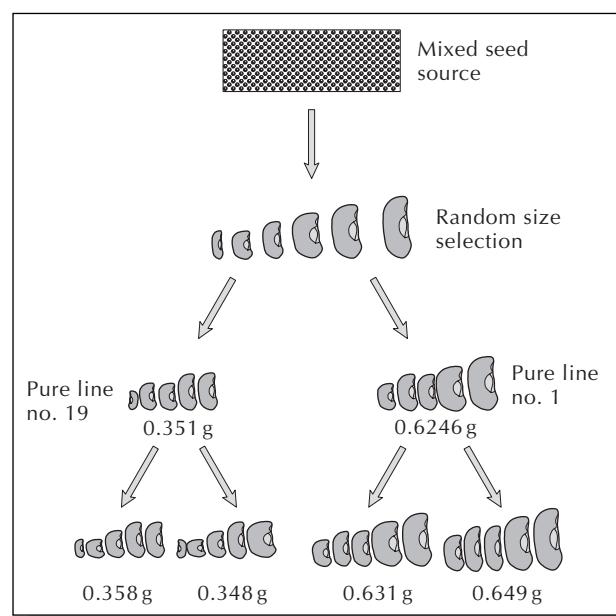
Mass selection may be direct or indirect. Indirect selection will have high success if two traits result from pleiotropy or if the selected trait is a component of the trait targeted for improvement. For example, researchers improve seed protein or oil by selecting on the basis of density separation of the seed.

### Pure-line selection

The theory of the pure line was developed in 1903 by the Danish botanist Johannsen. Studying seed weight of beans, he demonstrated that a mixed population of self-pollinated species could be sorted out into genetically pure lines. However, these lines were subsequently non-responsive to selection within each of them (Figure 16.2). Selection is a passive process since it eliminates variation but does not create it. The pure-line theory may be summarized as follows:

- 1 Lines that are genetically different may be successfully isolated from within a population of mixed genetic types.
- 2 Any variation that occurs within a pure line is not heritable but due to environmental factors only. Consequently, as Johannsen's bean study showed, further selection within the line is not effective.

Lines are important to many breeding efforts. They are used as cultivars or as parents in hybrid production (inbred lines). Also, lines are used in the development of genetic stock (containing specific genes such as disease



**Figure 16.2** The development of the pure-line theory by Johannsen.

resistance or nutritional quality) and synthetic and multi-line cultivars.

### Key features

A **line cultivar**, by definition, is one that has a coefficient of parentage of at least 0.87. A **pure line** suggests that a cultivar has identical alleles at all loci. Even though plant breeders may make this assumption, it is one that is not practical to achieve in a breeding program. What plant breeders call pure-line cultivars are most aptly called "near" pure-line cultivars, because as researchers such as K. J. Frey observed, high mutation rates occur in such genotypes. Line cultivars have a very narrow genetic base and tend to be uniform in traits of interest (e.g., height, maturity). In cases of proprietary dispute, lines are easy to unequivocally identify.

### Applications

Pure-line breeding is desirable for developing cultivars for certain uses:

- 1 Cultivars for mechanized production that must meet a certain specification for uniform operation by farm machines (e.g., uniform maturity, uniform height for location of economic part).

- 2 Cultivars developed for a discriminating market that puts a premium on visual appeal (e.g., uniform shape, size).
- 3 Cultivars for the processing market (e.g., demand for certain canning qualities, texture).
- 4 Advancing “sports” that appear in a population (e.g., a mutant flower for ornamental use).
- 5 Improving newly domesticated crops that have some variability.
- 6 The pure-line selection method is also an integral part of other breeding methods such as pedigree selection and bulk population selection.

## Procedure

### Overview

The pure-line selection in breeding entails repeated cycles of selfing, following the initial selection from a mixture of homozygous lines. Natural populations of self-pollinated species consist of mixtures of homozygous lines with transient heterozygosity originating from mutations and outcrossing.

### Steps

#### Year 1

The first step is to obtain a variable base population (e.g., introductions, segregating populations from crosses, landrace) and space plant it in the first year, select, and harvest desirable individuals (Figure 16.3).

#### Year 2

Grow progeny rows of selected plants. Rogue out any variants. Harvest selected progenies individually. These are experimental strains.

#### Years 3–6

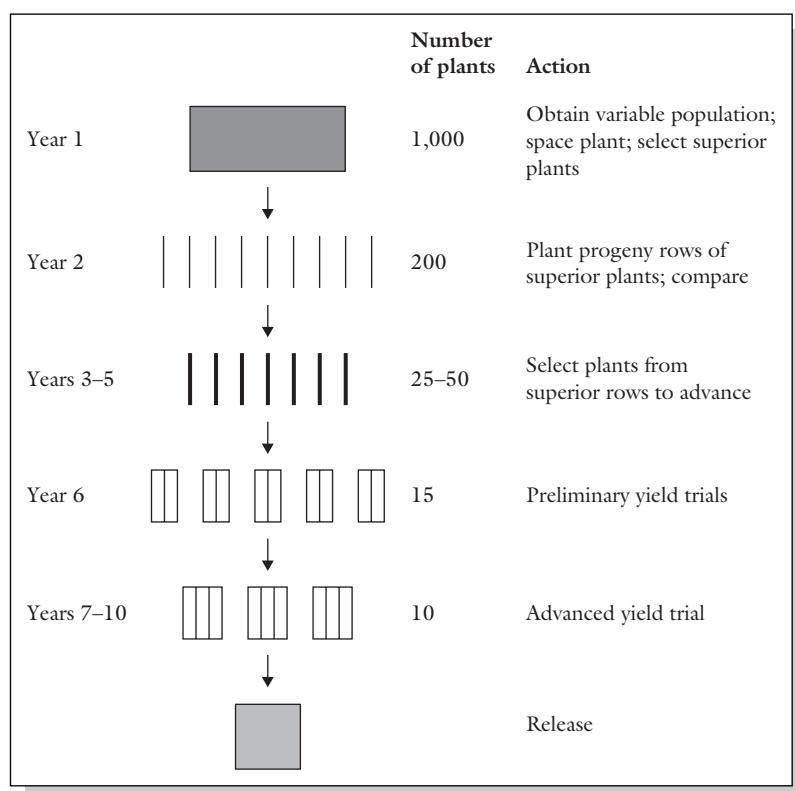
Conduct preliminary yield trials of the experimental strains including appropriate check cultivars.

#### Years 7–10

Conduct advanced yield trials at multiple locations. Release highest yielding line as new cultivar.

### Genetic issues

Pure-line breeding produces cultivars with a narrow genetic base and hence less likely to produce stable



**Figure 16.3** Generalized steps in breeding by pure-line selection.

yields over a wider range of environments. Such cultivars are more prone to being wiped out by pathogenic outbreaks. Because outcrossing occurs to some extent within most self-pollinated cultivars, coupled with the possibility of spontaneous mutation, variants may arise in commercial cultivars over time. It is tempting to select from established cultivars to develop new lines, an action that some view as unacceptable and unprofessional practice. As previously discussed, pure-line cultivars depend primarily on phenotypic plasticity for production response and stability across environments.

### **Advantages and disadvantages**

Some of the major advantages and disadvantages of the application of the pure-line method for improving self-pollinated species are given here.

#### *Advantages*

- 1 It is a rapid breeding method.
- 2 The method is inexpensive to conduct. The base population can be a landrace. The population size selected is variable and can be small or large, depending on the objective.
- 3 The cultivar developed by this method has great “eye appeal” because of the high uniformity.
- 4 It is applicable to improving traits of low heritability, because selection is based on progeny performance.
- 5 Mass selection may include some inferior pure lines. In pure-line selection, only the best pure line is selected for maximum genetic advance.

#### *Disadvantages*

- 1 The purity of the cultivar may be altered through admixture, natural crossing with other cultivars, and mutations. Such off-type plants should be rogued out to maintain cultivar purity.
- 2 The cultivar has a narrow genetic base and hence is susceptible to devastation from adverse environmental factors, because of uniform response.
- 3 A new genotype is not created. Rather, improvement is limited to the isolation of the most desirable or best genotype from a mixed population.
- 4 The method promotes genetic erosion because most superior pure lines are identified and multiplied to the exclusion of other genetic variants.
- 5 Progeny rows take up more resources (time, space, funds).

### **Pedigree selection**

Pedigree selection is a widely used method of breeding self-pollinated species (and even cross-pollinated species such as corn and other crops produced as hybrids). A key difference between pedigree selection and mass selection or pure-line selection is that hybridization is used to generate variability (for the base population), unlike the other methods in which production of genetic variation is not a feature. The method was first described by H. H. Lowe in 1927.

#### **Key features**

Pedigree selection is a breeding method in which the breeder keeps records of the ancestry of the cultivar. The base population, of necessity, is established by crossing selected parents, followed by handling an actively segregating population. Documentation of the pedigree enables breeders to trace parent–progeny back to an individual  $F_2$  plant from any subsequent generation. To be successful, the breeder should be able to distinguish between desirable and undesirable plants on the basis of a single plant phenotype in a segregating population. It is a method of continuous individual selection after hybridization. Once selected, plants are reselected in each subsequent generation. This process is continued until a desirable level of homozygosity is attained. At that stage, plants appear phenotypically homogeneous.

The breeder should develop an effective, easy to maintain system of record keeping. The most basic form is based on numbering of plants as they are selected, and developing an extension to indicate subsequent selections. For example, if five crosses are made and 750 plants are selected in the  $F_2$  (or list the first selection generation), a family could be designated 5-175 (meaning, it was derived from plant 175 selected from cross number 5). If selection is subsequently made from this family, it can be named, for example, 5-175-10. Some breeders include letters to indicate the parental sources or the kind of crop (e.g., NP-5-175-10), or some other useful information. The key is to keep it simple, manageable, and informative.

#### **Applications**

Pedigree selection is applicable to breeding species that allow individual plants to be observed, described, and harvested separately. It has been used to breed species including peanut, tobacco, tomato, and some cereals,

especially where readily identifiable qualitative traits are targeted for improvement.

### General guides to selection following a cross

The success of breeding methods preceded by hybridization rest primarily on the parents used to initiate the breeding program. Each generation has genetic characteristics and is handled differently in a breeding program.

***F<sub>1</sub>* generation** Unless in hybrid seed programs in which the *F<sub>1</sub>* is the commercial product, the purpose of the *F<sub>1</sub>* is to grow a sufficient *F<sub>2</sub>* population for selection. To achieve this, *F<sub>1</sub>* seed is usually space planted for maximum seed production. It is critical also to be able to authenticate hybridity and identity and remove seeds from self-pollination. Whenever possible, plant breeders use genetic markers in crossing programs.

***F<sub>2</sub>* generation** Selection in the plant breeding program often starts in the *F<sub>2</sub>*, the generation with the maximum genetic variation. The rate of segregation is higher if the parents differ by a larger number of genes. Generally, a large *F<sub>2</sub>* population is planted (2,000–5,000). Fifty percent of the genotypes in the *F<sub>2</sub>* are heterozygous and hence selection intensity should be moderate (about 10%) in order to select plants that would likely include those with the desired gene combinations. The actual number of plants selected depends on the trait (its heritability) and resources. Traits with high heritability are more effectively selected, requiring lower numbers than for traits with low heritability. The *F<sub>2</sub>* is also usually space planted to allow individual plants to be evaluated for selection. In pedigree selection, each selected *F<sub>2</sub>* plant is documented.

***F<sub>3</sub>* generation** Seed from individual plants are progeny-rowed. This allows homozygous and heterozygous genotypes to be distinguished. The homozygosity in the *F<sub>3</sub>* is 50% less than in the *F<sub>2</sub>*. The heterozygotes will segregate in the rows. The *F<sub>3</sub>* generation is the beginning of line formation. It is helpful to include check cultivars in the planting to help in selecting superior plants.

***F<sub>4</sub>* generation** *F<sub>3</sub>* plants are grown in plant-to-row fashion as in the *F<sub>3</sub>* generation. The progenies become more homogeneous (homozygosity is 87.5%). Lines are formed in the *F<sub>4</sub>*. Consequently, selection in the *F<sub>4</sub>* should focus more on progeny rather than on individuals plants.

***F<sub>5</sub>* generation** Lines selected in the *F<sub>4</sub>* are grown in preliminary yield trials (PYTs). *F<sub>5</sub>* plants are 93.8% homozygous. These PYTs are replicated trials with at least two replications (depending on the amount of seed available). The seeding rate is the commercial rate (or as close as possible), receiving all the customary cultural inputs. Evaluation of quality traits and disease resistance can be included. The PYT should include check cultivars. The best performing lines are selected for advancing to the next stage in the breeding program.

***F<sub>6</sub>* generation** The superior lines from *F<sub>5</sub>* are further evaluated in competitive yield trials or advanced yield trials (AYTs), including a check.

***F<sub>7</sub>* and subsequent generations** Superior lines from *F<sub>6</sub>* are evaluated in AYTs for several years, at different locations, and in different seasons as desirable. Eventually, after *F<sub>8</sub>*, the most outstanding entry is released as a commercial cultivar.

### Procedure

#### Overview

The key steps in the pedigree selection procedure are:

- 1 Establish a base population by making a cross of selected parents.
- 2 Space plant progenies of selected plants.
- 3 Keep accurate records of selection from one generation to the next.

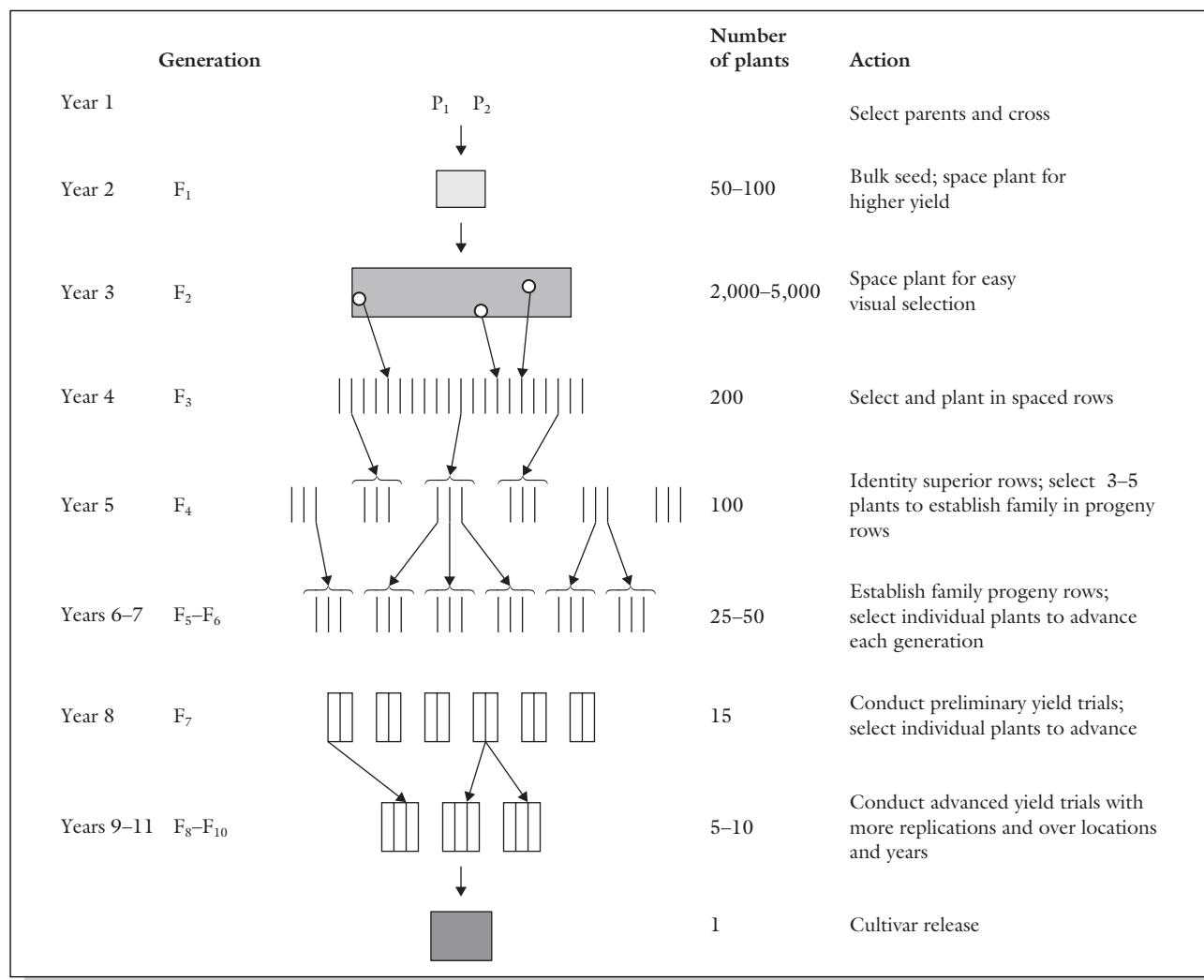
#### Steps

**Year 1** Identify desirable homozygous parents and make about 20–200 crosses (Figure 16.4).

**Year 2** Grow 50–100 *F<sub>1</sub>* plants including parents for comparison to authenticate its hybridity.

**Year 3** Grow about 2,000–5,000 *F<sub>2</sub>* plants. Space plant to allow individual plants to be examined and documented. Include check cultivars for comparison. Desirable plants are selected and harvested separately keeping records of their identities. In some cases, it may be advantageous not to space plant *F<sub>2</sub>*s to encourage competition among plants.

**Year 4** Seed from superior plants are progeny-rowed in the *F<sub>3</sub>*–*F<sub>5</sub>* generations, making sure to space plant the rows for easy record keeping. Selection at this stage is both within and



**Figure 16.4** Generalized steps in breeding by pedigree selection.

between rows by first identifying superior rows and selecting 3–5 plants from each progeny to plant the next generation.

**Year 5** By the end of the F<sub>4</sub> generation, there should be between 25–50 rows with records of the plant and row. Grow progeny of each selected F<sub>3</sub>.

**Year 6** Family rows are planted in the F<sub>6</sub> to produce experimental lines for preliminary yield trials in the F<sub>7</sub>. The benchmark or check variety is a locally adapted cultivar. Several checks may be included in the trial.

**Year 7** Advanced yield trials over locations, regions, and years are conducted in the F<sub>8</sub>–F<sub>10</sub> genera-

tions, advancing only superior experimental material to the next generation. Ultimately, the goal is to identify one or two lines that are superior to the check cultivars for release as a new cultivar. Consequently, evaluations at the advanced stages of the trial should include superior expression of traits that are deemed to be of agronomic importance for successful production of the particular crop (e.g., lodging resistance, shattering resistance, disease resistance). If a superior line is identified for release, it is put through the customary cultivar release process (i.e., seed increase and certification).

### Comments

- 1 Growing parents, making a cross, and growing  $F_1$  plants may take 1–2 years, depending on the facilities available for growing multiple experiments in a year (e.g., greenhouse) and the growing period of the crop.
- 2 The number of plants selected in the  $F_2$  depends on resources available (labor, space, time), and can even be 10,000 plants.
- 3  $F_3$  family rows should contain a large enough number of plants (25–30) to permit the true family features to be evident so the most desirable plant(s) can be selected. Families that are distinctly inferior should be discarded, while more than one plant may be selected from exceptional families. However, generally, the number of plants advanced does not exceed the number of  $F_3$  families.
- 4 From  $F_3$  to  $F_5$ , selection is conducted between and within rows, identifying superior rows and selecting 3–5 of the best plants in each family. By  $F_5$ , only about 25–50 families are retained.
- 5 By  $F_5$ , plant density may reflect the commercial seedling rate. Further, the plants from this generation and future ones would be sufficiently homozygous to warrant conducting preliminary and, later, advanced yield trials.

### Genetic issues

Detailed records are kept from one generation to the next regarding parentage and other characteristics of plants. The method allows the breeder to create genetic variability during the process. Consequently, the breeder can influence the genetic variation available by the choice of parents. The method is more conducive for breeding qualitative disease resistance, than for quantitative resistance. The product (cultivar) is genetically relatively narrow based but not as extremely so as in pure-line selection. The records help the breeder to advance only progeny lines with plants that exhibit genes for the desired traits.

### Advantages and disadvantages

The pedigree method of breeding has advantages and disadvantages, the major ones include the following.

#### Advantages

- 1 Record keeping provides a catalog of genetic information of the cultivar unavailable from other methods.

- 2 Selection is based not only on phenotype but also on genotype (progeny row) making it an effective method for selecting superior lines from among segregating plants.
- 3 Using the records, the breeder is able to advance only the progeny lines in which plants that carry the genes for the target traits occur.
- 4 A high degree of genetic purity is produced in the cultivar, an advantage where such a property is desirable (e.g., certification of products for certain markets).

#### Disadvantages

- 1 Record keeping is slow, tedious, time-consuming, and expensive. It places pressure on resources (e.g., land for space planting for easy observation). Seeding and harvesting are tedious operations. However, modern research plot equipment for planting and harvesting are versatile and sophisticated to allow complex operations and record taking to be conducted, making pedigree selection easier to implement and hence be widely used. Large plant populations can now be handled without much difficulty.
- 2 The method is not suitable for species in which individual plants are difficult to isolate and characterize.
- 3 Pedigree selection is a long procedure, requiring about 10–12 years or more to complete, if only one growing season is possible.
- 4 The method is more suited for qualitative than for quantitative disease-resistance breeding. It is not effective for accumulating the number of minor genes needed to provide horizontal resistance.
- 5 Selecting in the  $F_2$  (early generation testing) on the basis of quantitative traits such as yield may not be effective. It is more efficient to select among  $F_3$  lines planted in rows than to select individual plants in the  $F_2$ .

#### Modifications

As previously indicated, the pedigree selection method is a continuous selection of individuals after hybridization. A discontinuous method (called the  $F_2$  progenies test) has been proposed but is not considered practical enough for wide adoption. The breeder may modify the pedigree method to suit specific objectives and resources. Some specific ways are as follows:

- 1 The numbers of plants to select at each step may be modified according to the species, the breeding objective, and the genetics of the traits of interest, as well as the experience of the breeder with the crop, and resources available for the project.

- 2 The details of records kept are at the discretion of the breeder.
- 3 Off-season planting (e.g., winter nurseries), use of the greenhouse, and multiple plantings a year (where possible), are ways of speeding up the breeding process.

Early generation selection for yield in pedigree selection is not effective. This is a major objection to the procedure. Consequently, several modifications have been introduced by breeders to delay selection until later generations (e.g.,  $F_5$ ). Mass selection or bulk selection is practiced in the early generations.

### Bulk population breeding

Bulk population breeding is a strategy of crop improvement in which the natural selection effect is solicited more directly in the early generations of the procedure by delaying stringent artificial selection until later generations. The Swede, H. Nilsson-Ehle, developed the procedure. H. V. Harlan and colleagues provided an additional theoretical foundation for this method through their work in barley breeding in the 1940s. As proposed by Harlan and colleagues, the bulk method entails yield testing of the  $F_2$  bulk progenies from crosses and discarding whole crosses based on yield performance. In other words, the primary objective is to stratify crosses for selection of parents based on yield values. The current application of the bulk method has a different objective.

#### Key features

The rationale for delaying artificial selection is to allow natural selection pressure (e.g., abiotic factors such as drought, cold) to eliminate or reduce the productivity of less fit genotypes in the population. Just like the pedigree method, the bulk method also applies pure-line theory to segregating populations to develop pure-line cultivars. Genetic recombination in the heterozygous state cannot be used in self-pollinated species because self-pollination progressively increases homozygosity. By  $F_6$  the homozygosity is about 98.9%. The strategy in plant breeding is to delay selection until there is a high level of homozygosity.

#### Applications

It is a procedure used primarily for breeding self-pollinated species, but can be adapted to produce inbred

populations for cross-pollinated species. It is most suitable for breeding species that are normally closely spaced in production (e.g., small grains – wheat, barley). It is used for field bean and soybean. However, it is not suitable for improving fruit crops and many vegetables in which competitive ability is not desirable.

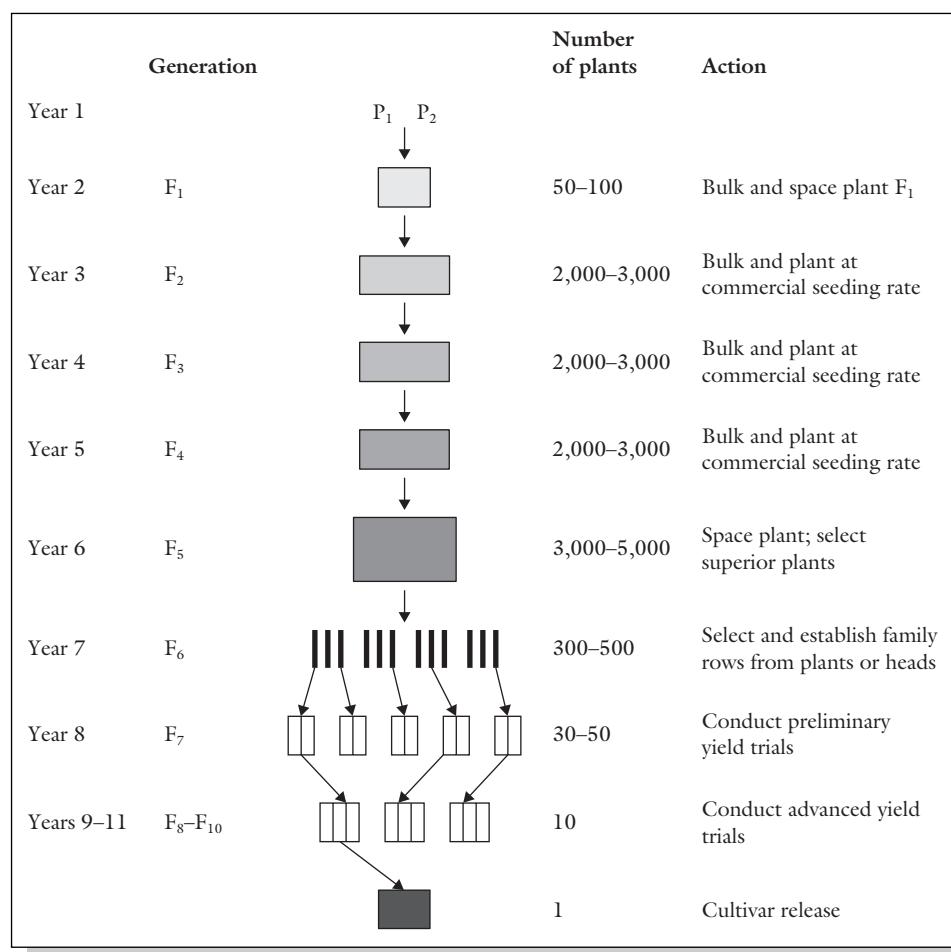
#### Procedure

##### Overview

After making a cross, several hundreds to several thousands of  $F_2$  selections are planted at a predetermined (usually conventional rate), close spacing. The whole plot is bulk harvested. A sample of seed is used to plant another field block for the next selection, subjecting it to natural selection pressure through the next 2–3 generations. In the  $F_5$ , the plants are space planted to allow individual plant evaluation for effective selection. Preliminary yield trials may start in the  $F_7$  followed by advanced yield trials, leading to cultivar release.

##### Steps

- |                         |   |
|-------------------------|---|
| <b>Year 1</b>           | Identify desirable parents (cultivars, single crosses, etc.) and make a sufficient number of crosses between them (Figure 16.5).  |
| <b>Year 2</b>           | Following a cross between appropriate parents, about 50–100 $F_1$ plants are planted and harvested as a bulk, after roguing out selfs.  |
| <b>Year 3</b>           | The seeds from the second year are used to plant a bulk plot of about 2,000–3,000 $F_2$ plants. The $F_2$ is bulk harvested.  |
| <b>Years 4–6</b>        | A sample of the $F_2$ seed is planted in bulk plots, repeating the steps for year 2 and year 3 until the $F_4$ is reached or when a desired level of homozygosity has been attained in the population. Space plant about 3,000–5,000 $F_5$ plants and select about 10% (300–500) superior plants for planting $F_6$ progeny rows. |
| <b>Year 7</b>           | Select and harvest about 10% (30–50) progeny rows that exhibit genes for the desired traits for planting preliminary yield trials in the $F_7$ .  |
| <b>Year 8 and later</b> | Conduct advanced yield trials from $F_8$ through $F_{10}$ at multiple locations and regions, including adapted cultivars as checks. After identifying a superior line, it is put through the customary cultivar release process.  |



**Figure 16.5** Generalized steps in breeding by bulk selection.

#### Comments

- 1 Space planting of the F<sub>1</sub> will increase the yield of F<sub>2</sub> seed.
- 2 The breeder may screen the bulk population under different natural environments in a rotation (e.g., soil condition – salinity, acidity; disease resistance; temperature – winter kill, etc.). There may be an increase in broad adaptation of the cultivar. However, care should be exercised to avoid the evaluation of plants under a condition that could eliminate genotypes that are of value at different sets of environmental conditions.
- 3 Screening for photoperiodic response is desirable and advantageous in the early stages to eliminate genotypes that are incapable of reproducing under the environmental conditions.

**4** Natural selection may be aided by artificial selection. Aggressive and highly competitive but undesirable genotypes may be physically rogued out of the population to avoid increasing the frequency of undesirable genes, or to help select benign traits such as seed color or fiber length of cotton. Aiding natural selection also accelerates the breeding program.

**5** The degree of selection pressure applied, its consistency, duration, and the heritability of traits, are all factors that impact the rate at which unadapted segregates are eliminated from the bulk population.

#### Genetic issues

Applying the theories of population genetics (see Chapter 7), repeated self-pollination, and fertilization will result in three key outcomes:

- 1 At advanced generations, the plants will be homozygous at nearly all loci.
- 2 The mean population performance will be improved as a result of natural selection.
- 3 Genotypes with good agricultural fitness will be retained in the population.

Bulk selection promotes intergenotypic competition. By allowing natural selection to operate on early generations, the gene frequencies in the population at each generation will depend upon:

- 1 The genetic potential of a genotype for productivity.
- 2 The competitive ability of the genotype.
- 3 The effect of the environment on the expression of a genotype.
- 4 The proportions and kinds of genotypes advanced to the next generation (i.e., sampling).

The effects of these factors may change from one generation to the next. More importantly, it is possible that desirable genotypes may be outcompeted by more aggressive undesirable genotypes. For example, tall plants may smother short desirable plants. It is not possible to predict which  $F_2$  plant's progeny will be represented in the next generation, nor predict the genetic variability for each character in any generation.

The role of natural selection in bulk breeding is not incontrovertible. It is presumed to play a role in genetic shifts in favor of good competitive types, largely due to the high fecundity of competitive types. Such an impact is not hard to accept when traits that confer advantage through resistance to biotic and abiotic stresses are considered. For example, if the bulk populations were subjected to various environments (e.g., salinity, cold temperature, water logging, drought, photoperiod), fecundity may be drastically low for ill-adapted genotypes. These are factors that affect adaptation of plants. Some traits are more neutral in competition (e.g., disease resistance). If two genotypes are in competition, their survival depends on the number of seed produced by each genotype as well as the number of seeds produced by their progeny.

Using the natural relationship developed by W. Allard for illustration, the survival of an inferior genotype may be calculated as:

$$A_n = \alpha \times S^{n-1}$$

where  $A_n$  = proportion of inferior genotypes,  $n$  = generation,  $\alpha$  = initial proportion of the inferior genotype, and  $S$  = selection index. Given two genotypes, A (superior)

and B (inferior), in equal proportions in a mixture (50% A : 50% B), and of survival capacities  $A=1$ ,  $B=0.9$ , the proportion of the inferior genotype in  $F_5$  would be:

$$\begin{aligned} A_5 &= (0.5) \times (0.9)^{5-1} \\ &= 0.3645 \text{ (or } 36.45\%) \end{aligned}$$

This means the inferior genotype would decrease from 50% to 36.45% by  $F_5$ . Conversely, the proportion of the superior genotype would increase to 63.55%.

As previously indicated, the bulk selection method promotes intergenotypic competition; it is important to point out that the outcome is not always desirable because a more aggressive inferior genotype may out-compete a superior (desirable) but poor competitor. In a classic study by C. A. Suneson, equal mixtures (25%) of four barley cultivars were followed. After more than five generations, the cultivar "Atlas" was represented by 88.1%, "Club Mariot" by 11%, "Hero" by 1%, while "Vaughn" was completely eliminated. However, in pure stands, "Vaughn" outyielded "Atlas". It may also be said that if the genotypes whose frequency in the population increased over generations are the ones of agronomic value (i.e., desired by the breeder), then the competition in bulking is advantageous to plant breeding. The effect of natural selection in the bulk population can be positive or negative, and varies according to the traits of interest, the environment under which the population is growing, and the degree of intergenotypic competition (spacing among plants). If there is no competition between plants, genotype frequencies would not be changed significantly. Also, the role of natural selection in genetic shifts would be less important when the duration of the period is smaller (6–10 generations), as is the case in bulk breeding. This is so because natural selection acts on the heterozygotes in the early generations. However, the goal of bulk breeding is to develop pure lines. By the time the cultivar is released, the breeding program would have ended, giving natural selection no time to act on the pure lines.

### Advantages and disadvantages

Some of the key advantages and disadvantages of bulk breeding method are as follows.

#### Advantages

- 1 It is simple and convenient to conduct.
- 2 It is less labor intensive and less expensive in early generations.

- 3 Natural selection may increase frequency of desirable genotypes by the end of the bulking period.
- 4 It is compatible with mass selection in self-pollinated species.
- 5 Bulk breeding allows large amounts of segregating materials to be handled. Consequently, the breeder can make and evaluate more crosses.
- 6 The cultivar developed would be adapted to the environment, having been derived from material that had gone through years of natural selection.
- 7 Single-plant selections are made when plants are more homozygous, making it more effective to evaluate and compare plant performance.

### *Disadvantages*

- 1 Superior genotypes may be lost to natural selection, while undesirable ones are promoted during the early generations.
- 2 It is not suited to species that are widely spaced in normal production.
- 3 Genetic characteristics of the populations are difficult to ascertain from one generation to the next.
- 4 Genotypes are not equally represented in each generation because all the plants in one generation are not advanced to the next generation. Improper sampling may lead to genetic drift.
- 5 Selecting in off-season nurseries and the greenhouse may favor genotypes that are undesirable in the production region where the breeding is conducted, and hence is not a recommended practice.
- 6 The procedure is lengthy, but cannot take advantage of off-season planting.

### **Modifications**

Modifications of the classic bulk breeding method include the following:

- 1 The breeder may impose artificial selection sooner ( $F_3$  or  $F_4$ ) to shift the population toward an agriculturally more desirable type.
- 2 Rouging may be conducted to remove undesirable genotypes prior to bulking.
- 3 The breeder may select the appropriate environment to favor desired genotypes in the population. For example, selecting under disease pressure would eliminate susceptible individuals from the population.
- 4 Preliminary yield trials may be started even while the lines are segregating in the  $F_3$  or  $F_4$ .
- 5 The **single-seed descent** method may be used at each generation to reduce the chance of genetic drift. Each

generation, a single seed is harvested from each plant to grow the next bulk population. The dense planting makes this approach problematic in locating individual plants.

- 6 **Composite cross bulk population** breeding, also called the **evolutionary method of breeding**, was developed by C. A. Suneson and entails systematically crossing a large number of cultivars. First, pairs of parents are crossed, then pairs of  $F_1$ s are crossed. This continues until a single hybrid stock containing all parents is produced. The method has potential for crop improvement, but it takes a very long time to complete.

### **Single-seed descent**

The method of single-seed descent was born out of a need to speed up the breeding program by rapidly inbreeding a population prior to beginning individual plant selection and evaluation, while reducing a loss of genotypes during the segregating generations. The concept was first proposed by C. H. Goulden in 1941 when he attained the  $F_6$  generation in 2 years by reducing the number of generations grown from a plant to one or two, while conducting multiple plantings per year, using the greenhouse and off-season planting. H. W. Johnson and R. L. Bernard described the procedure of harvesting a single seed per plant for soybean in 1962. However, it was C. A. Brim who in 1966 provided a formal description of the procedure of single-seed descent, calling it a **modified pedigree method**.

### **Key features**

The method allows the breeder to advance the maximum number of  $F_2$  plants through the  $F_5$  generation. This is achieved by advancing one randomly selected seed per plant through the early segregating stages. The focus on the early stages of the procedure is on attaining homozygosity as rapidly as possible, without selection. Discriminating among plants starts after attainment of homozygosity.

### **Applications**

Growing plants in the greenhouse under artificial conditions tends to reduce flower size and increase cleistogamy. Consequently, single-seed descent is best for self-pollinated species. It is effective for breeding small

grains as well as legumes, especially those that can tolerate close planting and still produce at least one seed per plant. Species that can be forced to mature rapidly are suitable for breeding by this method. It is widely used in soybean breeding to advance the early generation. One other major application of single-seed descent is in conjunction with other methods.

## Procedure

### Overview

A large  $F_1$  population is generated to ensure adequate recombination among parental chromosomes. A single seed per plant is advanced in each subsequent generation until the desired level of inbreeding is attained. Selection is usually not practiced until  $F_5$  or  $F_6$ . Then, each plant is used to establish a family to help breeders in selection and to increase seed for subsequent yield trials.

### Steps

- |                         |   |
|-------------------------|---|
| <b>Year 1</b>           | Crossing is used to create the base population. Cross selected parents to generate an adequate number of $F_1$ for the production of a large $F_2$ population.  |
| <b>Year 2</b>           | About 50–100 $F_1$ plants are grown in a greenhouse in the ground, on a bench, or in pots. They may also be grown in the field. Harvest identical $F_1$ crosses and bulk.   |
| <b>Year 3</b>           | About 2,000–3,000 $F_2$ plants are grown. At maturity, a single seed per plant is harvested and bulked for planting $F_3$ . Subsequently, the $F_2$ plants are spaced enough to allow each plant to produce only a few seeds. |
| <b>Years 4–6</b>        | Single pods per plant are harvested to plant the $F_4$ . The $F_5$ is space planted in the field, harvesting seed from only superior plants to grow progeny rows in the $F_6$ generation.                                     |
| <b>Year 7</b>           | Superior rows are harvested to grow preliminary yield trials in the $F_7$ .   |
| <b>Year 8 and later</b> | Yield trials are conducted in the $F_8$ – $F_{10}$ generations. The most superior line is increased in the $F_{11}$ and $F_{12}$ as a new cultivar.   |

### Comments

- 1 If the sample is too small, superior genetic combinations may be lost because only one seed from each plant is used.
- 2 It may be advantageous to use progeny rows prior to yield testing to produce sufficient seed as well as to help in selecting superior families.
- 3 The breeder may choose to impose some artificial selection pressure by excluding undesirable plants from contributing to the subsequent generations (in the early generations). This is effective for qualitative traits.
- 4 Record keeping is minimal and so are other activities such as harvesting, especially in the early generations.

### Genetic issues

Each individual in the final population is a descendent from a different  $F_2$  plant. Each of these plants undergoes a decrease in heterozygosity at a rapid rate, each generation. Barring the inability of a seed to germinate or a plant to set seed, the effect of natural selection is practically non-existent in the single-seed descent procedure. Only one seed per plant is advanced, regardless of the number produced. That is, a plant producing one seed is as equally represented in the next generation as one producing 1,000 seeds. Selection is conducted on homozygous plants rather than segregating material. An efficient early generation testing is needed to avoid genetic drift of desirable alleles. Single-seed descent is similar to bulk selection in that the  $F_6$ / $F_7$  comprises a large number of homozygous lines, prior to selection among progenies. A wide genetic diversity is carried on to relatively advanced generations ( $F_6$ / $F_7$ ).

### Advantages and disadvantages

Single-seed descent has certain advantages and disadvantages, the major ones including the following.

#### Advantages

- 1 It is an easy and rapid way to attain homozygosity (2–3 generations per year).
- 2 Small spaces are required in early generations (e.g., can be conducted in a greenhouse) to grow the selections.
- 3 Natural selection has no effect (hence it can not impose an adverse impact).
- 4 The duration of the breeding program can be reduced by several years by using single-seed descent.

- 5 Every plant originates from a different  $F_2$  plant, resulting in greater genetic diversity in each generation.
- 6 It is suited to environments that do not represent those in which the ultimate cultivar will be commercially produced (no natural selection imposed).

#### Disadvantages

- 1 Natural selection has no effect (hence no benefit from its possible positive impact).
- 2 Plants are selected based on individual phenotype not progeny performance.
- 3 An inability of seed to germinate or a plant to set seed may prohibit every  $F_2$  plant from being represented in the subsequent population.
- 4 The number of plants in the  $F_2$  is equal to the number of plants in the  $F_4$ . Selecting a single seed per plant runs the risks of losing desirable genes. The assumption is that the single seed represents the genetic base of each  $F_2$ . This may not be true.

#### Modifications

The procedure described so far is the classic single-seed descent breeding method. There are two main modifications of this basic procedure. The multiple seed procedure (or **modified single-seed descent**) entails selecting 2–4 seeds per plant, bulking and splitting the bulk into two, one for planting the next generation, and the other half held as a reserve. Because some soybean breeders simply harvest one multiseeded pod per plant, the procedure is also referred to by some as the **bulk pod method**.

Another modification is the **single hill method** in which progeny from individual plants are maintained as separate lines during the early generations by planting a few seeds in a hill. Seeds are harvested from the hill and planted in another hill the next generation. A plant is harvested from each line when homozygosity is attained.



W. T. B. Thomas

Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK

## Industry highlights

### *Barley breeding in the United Kingdom*

**Table 1** Characters listed in the current UK recommended lists of barley ([www.hgca.com](http://www.hgca.com)).

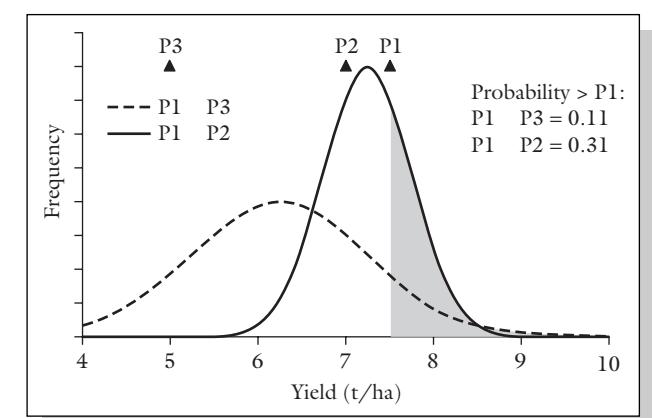
Character	Spring barley	Winter barley
Yield (overall and regional with fungicide)	Yes	Yes
Yield without fungicide	Yes	Yes
Height	Yes	Yes
Lodging resistance	Yes	Yes
Brackling resistance	Yes	
Maturity	Yes	Yes
Winter hardiness		Yes
Powdery mildew resistance	Yes	Yes
<i>Rhynchosporium</i> resistance	Yes	Yes
Yellow rust resistance	Yes	Yes
Brown rust resistance	Yes	Yes
Net blotch resistance		Yes
BaYMV complex resistance		Yes
BYDV resistance	Yes	
Grain nitrogen content	Yes	Yes
Hot water extract	Yes	Yes
Screenings (2.25 and 2.5 mm)	Yes	Yes
Specific weight	Yes	Yes

#### Targets

Barley breeding in the UK aims to produce new cultivars that offer an improvement in one or more of the key characters for the region (Table 1). New cultivars must have a good yield, preferably in excess of the current established cultivars, if targeted solely at the feed market. To be accepted for malting use, a new cultivar must offer improvement in one or more key facets of malting quality, primarily hot water extract, with no major defects in, for instance, processability characters. Additionally, new cultivars must have minimum levels of disease resistance, which equates to being no worse than moderately susceptible, to the key diseases listed in Table 1.

#### Crossing to commercialization

Barley breeders therefore design crosses in which the parents complement each other for these target characters and attempt to select out recombinants that offer a better balanced overall phenotype. Whilst a wide cross may offer a better chance of producing superior



**Figure 1** Frequency distribution of two crosses with a common parent (P1) and alternative second parents (P2 and P3). P2 is a slightly lower yielding parent, thus progeny from the cross will have a high mid-parent value and small variation. P3 is comparatively high-yielding unadapted parent and the cross has a lower mid-parent value but much greater variance. Areas under the shaded portion of both curves represent the fraction selected for high-yield potential ( $> P1$ ). Thus, while the extreme recombinant of  $P1 \times P3$  has a greater yield potential than that of  $P1 \times P2$ , the probability of identifying superior lines for just this one character is far greater for the latter.

eries for shuttle breeding for the spring crop or doubled haploidy (DH) or single-seed descent (SSD) for the winter crop. The length of the breeding cycle is thus fairly well defined with occasional reduction by a year when a cultivar from a highly promising cross is speculatively advanced by a breeder. A breeder may also delay submitting a line for official trials for an extra season's data but breeders now aim to submit the majority of their lines to official trials within 4–5 years of making a cross. Given that many breeders would have begun recrossing such selections by this stage of their development, the approximate time for the breeding cycle in the UK is 4 years.

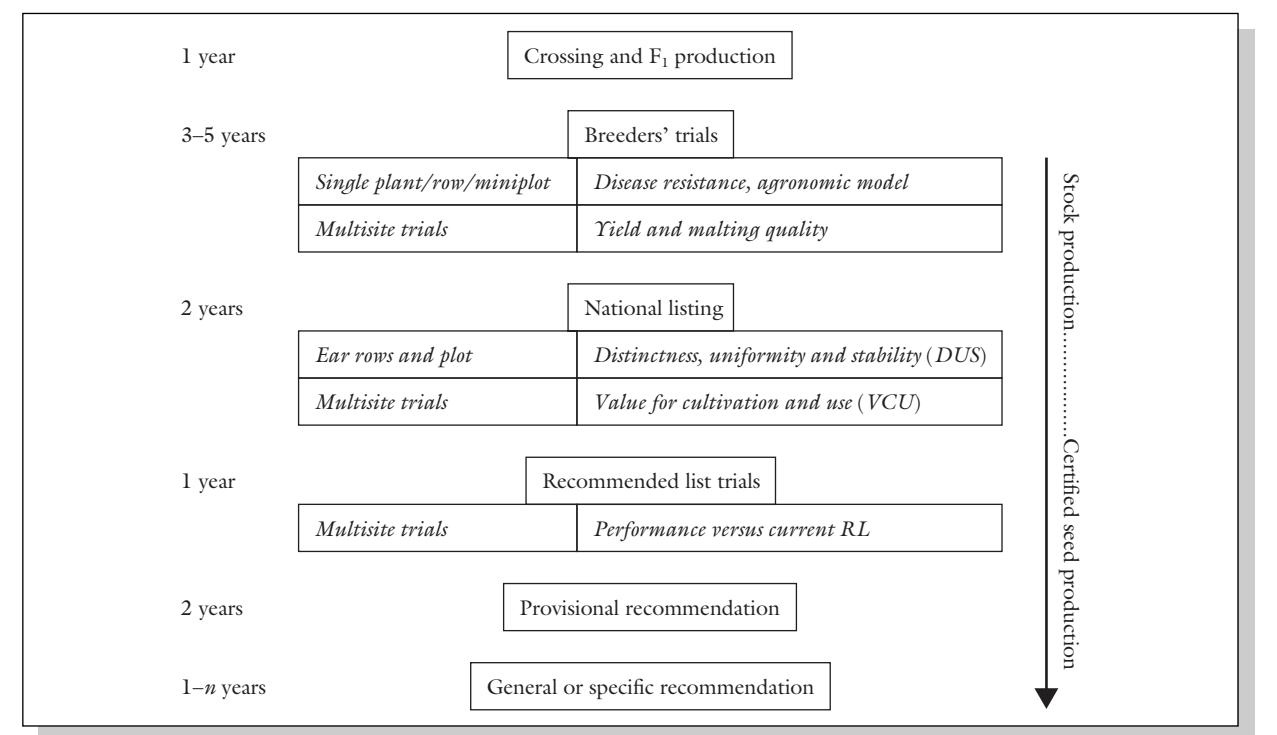
During the 2 years of national list trials (NLTs), potential cultivars are tested for distinctness, uniformity, and stability (DUS) using established botanical descriptors. A submission therefore has to be distinct from any other line on the National List and not have more than a permitted level of off-types, currently equivalent to a maximum of three in 100 ear rows. Lines are tested over more than 1 year to ensure that they are genetically stable and do not segregate in a subsequent generation. DUS tests are carried out by detailed examination of 100 ear rows and three bulk plots (approximately 400 plants in total) submitted by the breeder. Thirty-three characters are examined routinely and there are three special and 59 approved additional characters. At the same time as plot trials are carried out to establish whether the submission has value for cultivation and use (VCU), the VCU and DUS submissions are checked to verify that they are the same. Occasionally, a submission may fail the DUS test in NLT<sub>1</sub> in which case the breeder has the option of submitting a new stock for a further 2 years of testing. Generally, the VCU results are allowed to stand and a cultivar can be entered into the recommended list trials (RLTs) before it has passed the DUS test in the anticipation that it will have succeeded by the time a recommendation decision has to be made. Full details can be obtained from [www.defra.gov.uk/planth/pvs/VCU\\_DUS.htm](http://www.defra.gov.uk/planth/pvs/VCU_DUS.htm).

### The UK barley breeding community

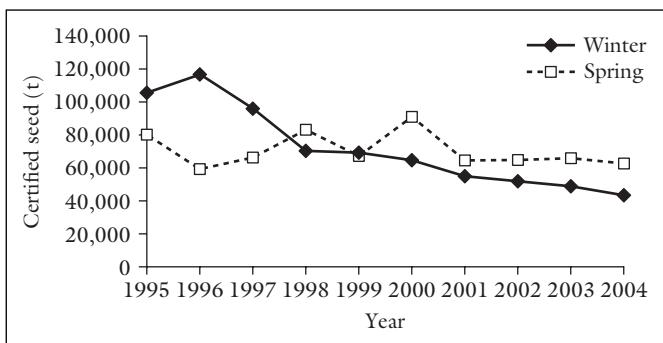
The Plant Varieties and Seeds Act of 1964, which enabled plant breeders to earn royalties on certified seed produced for their cultivars, led to a dramatic increase in breeding activity in the UK. Formerly, it was largely the province of state-funded improvement programs such as that of the Plant Breeding Institute (PBI), Cambridge, which had produced the highly successful spring cultivar "Proctor". The increase in breeding activity in the 1970s and early 1980s was largely as a result of a dramatic expansion in the commercial sector, initially led by Miln Marsters of Chester, UK, who produced "Golden Promise", which dominated Scottish spring barley production for almost two decades. The two sectors coexisted until the privatization of the breeding activity at PBI

recombinants, most barley breeders in the UK concentrate on narrow crosses between elite cultivars. The main reason for doing so is that a narrow cross between elite lines is more likely to produce a high midparental value for any one character and so the proportion of desirable recombinants is thus far greater in the narrow cross than in the wide cross (Figure 1). Thus, the chances of finding a desirable recombinant for a complex character such as yield in the wide cross is low and the chances of combining it with optimum expression for all the other characters is remote. As breeders are still making progress using such a narrow crossing strategy, it is possible that there is still an adequate level of genetic diversity within the elite barley gene pool in the UK. A similar phenomenon has been observed in barley breeding in the USA where progress has been maintained despite a narrow crossing strategy (Rasmusson & Phillips 1997). Rae et al. (2005) genotyped three spring barley cultivars ("Cocktail", "Doyen", and "Troon") on the 2005 UK recommended list with 35 simple sequence repeat (SSR) markers and found sufficient allelic diversity to produce over 21 million different genotypes. It would therefore appear that the breeding challenge is not so much to generate variation as to identify the best recombinants.

The progress of a potential new barley cultivar in the UK, in common with that of other cereals, proceeds through a series of filtration tests (Figure 2) and the time taken to pass through all but the first is strictly defined. The opportunity to reduce the time taken for breeders' selections is fairly limited given that the multiplication of material for, and the conducting of single- and multisite trials, takes at least 3 years, irrespective of whether one uses out-of-season nurseries



**Figure 2** Breakdown of the phases in the development of a successful new cultivar from crossing to commercialization, with the timescale for each step. The exact nature of the scheme adopted in breeders' trials varies according to the breeder and crop type, but is either based upon a version of the pedigree or a doubled haploid system. A cultivar may persist on the recommended list (RL) for  $n$  years, where  $n$  is the number of years where there is a significant demand for it.



**Figure 3** Tonnes of certified barley seed produced in the UK from 1995 to 2004.

declined by 43% since 1995 with most due to a reduction in winter barley seed (Figure 3). There are a number of potential reasons for this, such as an increase in farm-saved seed, but the principal feature has been a marked decrease in winter barley cropping over the period whereas spring barley has remained fairly static and winter wheat has increased. Over this period, certified seed production has exceeded 100,000 tonnes for two spring ("Opti" and "Chariot") and two winter ("Regina" and "Pearl") barley cultivars and these can be considered notable market successes. There has been substantial production of a number of others but

and the state marketing arm, the National Seed Development Organization, together with a change in government policy led to the withdrawal of the public sector from barley breeding in the UK. Barley breeding in the commercial sector in the UK is highly competitive with currently five UK-based crossing and selection programs. A number of other companies have their own selection programs based in the UK and many continental breeders have agency agreements for the testing and potential marketing of their products. For example, 41 spring and 34 winter barley lines were submitted for NLT<sub>1</sub> testing for harvest in 2004 and these were derived from 16 different breeders.

The amount of certified seed produced for each cereal variety in the UK is published by the National Institute of Agricultural Botany. The total annual production of certified barley seed has been in decline since its peak of over 250,000 tonnes in 1987, and has

total production exceeded 25,000 tonnes for only six spring and seven winter barley cultivars. When one considers that over 830 lines were submitted for NLTs over this period, the overall success rate is 1.6%. Nevertheless, real breeding progress is being made. Using yield data from the recommended list trials from 1993 to 2004 to estimate the mean yield of each recommended cultivar and then regressing that data against the year that it was first recommended, revealed that genetic progress was in the order of 1% per annum (Rae et al. 2005).

### ***Impact of molecular markers***

The first whole genome molecular maps of barley were published in 1991 (Graner et al. 1991; Heun et al. 1991) and were closely followed by QTL maps in 1992 (Heun 1992) and 1993 (Hayes et al. 1993) with well over 40 barley mapping studies now in the public domain. Despite this apparent wealth of information, barley breeders in the UK are largely relying on conventional phenotypic selection to maintain this progress. This is in marked contrast to the highly successful use of marker-assisted selection (MAS) in the Australian barley program (Langridge & Barr 2003), which is probably a reflection of the different breeding strategies in the two countries. In the UK, improvement is being achieved in the elite gene pool, as noted above, whereas MAS has been deployed in an introgression breeding strategy in Australia. Given that most barley mapping studies have concentrated on diverse crosses to maximize polymorphism and facilitate map construction, there are very few published QTL studies that are relevant to current UK barley breeding strategies. Surveying results from eight different barley mapping populations (Thomas 2003), found that there were very few instances where QTLs were co-located for three or more crosses for important characters such as yield and hot water extract.

### **Major gene targets**

Markers have been developed for a number of known major genes and could potentially be deployed in MAS by UK breeders. Many of these major gene targets are, however, disease resistances, many of which have been defeated by matching virulence in the corresponding pathogen population. UK barley breeders have been required to select for at least some resistance to the key foliar pathogens listed in Table 1 since the introduction of minimum standards, and have accordingly developed efficient phenotypic screens. There are exceptions, most notably the barley yellow mosaic virus (BaYMV) complex, which is transmitted by infection of the roots with the soil-borne fungus vector *Polymixa graminis*. A phenotypic screen therefore requires an infected site and the appropriate environment for infection and expression. Phenotypic screening can be expensive if a breeder is distant from an infected site and is subject to potential misclassification.

Resistance due to the *rym4* allele was initially found in "Ragusa" and was effective against BaYMV strain 1 and a number of cultivars carrying this allele have been developed, initially by phenotypic screening. Markers to select for this resistance have also been developed, beginning with the RFLP (restricted fragment length polymorphism) probe MWG838 (Graner & Bauer 1993), later converted to an STS (sequence-tagged site) (Bauer & Graner 1995), and were used in some breeding programs in the UK and Europe. BaYMV strain 2, which became more frequent in the 1990s, could overcome the *rym4* resistance, but another resistance, *rym5*, was identified in "Mokusekko 3" as being effective against both strains. This resistance was co-located with *rym4* and the SSR marker Bmac29 was found to be linked to it (Graner et al. 1999). Bmac29 could not only distinguish between resistant and susceptible alleles but also between the *rym4* and *rym5* alleles derived from "Ragusa" and "Mokusekko 3", respectively. However, as it is 1.3 cM from the gene locus, it is not effective in a wide germplasm pool as *Hordeum spontaneum* lines predicted to be resistant by the marker were found to be susceptible (R. P. Ellis, unpublished data). Bmac29 has, however, proved to be particularly effective for UK, and European, barley breeders as they are working with a narrow genetic base and just the two sources of resistance. Other resistance loci have been identified together with suitable markers to deploy in a pyramiding strategy in an attempt to provide durable resistance (Ordon et al. 2003). They provide a clear example of how the use of markers in MAS has evolved together with the pathogen.

Another example relates to a particular requirement of the Scotch whisky distilling industry. In grain and certain malt whisky distilleries, a breakdown product of the gynogenic glycoside epiheterodendrin can react with copper in the still to form the carcinogen ethyl carbamate, which can be carried over into the final spirit in distilling. This has lead to a demand for barley cultivars that do not produce epiheterodendrin. The character is controlled by a single gene with the non-producing allele originating in the mildew resistance donor "Arabische" used in the derivation of the cultivar "Emir". The phenotypic assay for the character involves the use of hazardous chemicals, and the finding of a linked SSR marker (Bmac213) offered a simpler and safer alternative (Swanston et al. 1999). The distance between the gene locus and the marker (6 cM) meant that, in contrast to Bmac29, Bmac213 was not reliable in the cultivated gene pool. For instance, the cultivar "Cooper" and its derivatives possess the non-producing allele yet are producers. However, the marker could still be used when the parents of a cross were polymorphic for both the phenotype and the marker. Recently, a candidate gene has been identified and markers used for reliable identification of non-producers have been developed (P. Hedley, personal communication).

### **QTL targets**

Currently, UK barley breeders do not use MAS for any other malting quality targets. A QTL for fermentability was detected in a cross between elite UK genotypes (Swanston et al. 1999) but the increasing allele was derived from the parent with relatively poor

malting quality. When this QTL was transferred into a good malting quality cultivar, the results were inconclusive (Meyer et al. 2004), probably because the effect of the gene was more marked in a poor quality background and any extra activity due to it was superfluous in a good quality background. This highlights one of the problems in developing MAS for complex characters such as yield and malting quality. Results from an inappropriate gene pool may well not translate to a target gene pool and it is therefore essential that QTL studies are carried out in the appropriate genetic background.

### **Future prospects**

The genotyping of entries from Danish registration trials coupled with associations of markers with yield and yield stability phenotypes demonstrated that QTLs can be detected in the elite gene pool (Kraakman et al. 2004) but the findings need validation before the markers can be used in MAS. At the Scottish Crop Research Institute, we will be undertaking extensive genotyping of UK RLT entries over the past 12 years in collaboration with the University of Birmingham, the National Institute of Agricultural Botany, the Home Grown Cereals Authority, UK, barley breeders, and representatives of the malting, brewing, and distilling industries in a project funded by the Defra Sustainable Arable LINK scheme. The RLT phenotypic data set represents an extensive resource that can discriminate between the fine differences in elite cultivars and will facilitate the identification of meaningful associations within the project for validation and potential use in MAS. How MAS is then utilized by commercial breeders in the UK might well vary but could range from early generation selection from an enriched germplasm pool upon which phenotypic selection can be concentrated, to identification of candidate submission lines carrying target traits.

### **Acknowledgments**

The author is funded by the Scottish Executive Environmental and Rural Affairs Department.

### **References**

- Bauer, E., and A. Graner. 1995. Basic and applied aspects of the genetic analysis of the *ym4* virus resistance locus in barley. *Agronomie* 15:469–473.
- Graner, A., and E. Bauer. 1993. RFLP mapping of the *ym4* virus-resistance gene in barley. *Theor. Appl. Genet.* 86:689–693.
- Graner, A., A. Jahoor, J. Schondelmaier, H. Siedler, K. Pillen, G. Fischbeck, and G. Wenzel. 1991. Construction of an RFLP map of barley. *Theor. Appl. Genet.* 83:250–256.
- Graner, A., S. Streng, A. Kellermann, et al. 1999. Molecular mapping and genetic fine-structure of the *rym5* locus encoding resistance to different strains of the barley yellow mosaic virus complex. *Theor. Appl. Genet.* 98:285–290.
- Hayes, P.M., B.H. Liu, S.J. Knapp, et al. 1993. Quantitative trait locus effects and environmental interaction in a sample of North-American barley germ plasm. *Theor. Appl. Genet.* 87:392–401.
- Heun, M. 1992. Mapping quantitative powdery mildew resistance of barley using a restriction-fragment-length-polymorphism map. *Genome* 35:1019–1025.
- Heun, M., A.E. Kennedy, J.A. Anderson, N.L.V. Lapitan, M.E. Sorrells, and S.D. Tanksley. 1991. Construction of a restriction-fragment-length-polymorphism map for barley (*Hordeum vulgare*). *Genome* 34:437–447.
- Kraakman, A.T.W., R.E. Niks, P.M.M.M. Van den Berg, P. Stam, and F.A. van Eeuwijk. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics* 168:435–446.
- Langridge, P., and A.R. Barr. 2003. Better barley faster: the role of marker assisted selection – Preface. *Aust. J. Agric. Res.* 54:i–iv.
- Meyer, R.C., J.S. Swanston, J. Brosnan, M. Field, R. Waugh, W. Powell, and W.T.B. Thomas. 2004. Can anonymous QTLs be introgressed successfully into another genetic background? Results from a barley malting quality parameter. In: Barley genetics IX, Proceedings of the 9th International Barley Genetics Symposium, June 20–26, Vol. 2 (Spunar, J., and J. Janikova, eds), pp. 461–467. Agricultural Research Institute, Kromeriz, Czech Republic.
- Ordon, F., K. Werner, B. Pellio, A. Schiemann, W. Friedt, and A. Graner. 2003. Molecular breeding for resistance to soil-borne viruses (BaMMV, BaYMV, BaYMV-2) of barley (*Hordeum vulgare* L.). *J. Plant Dis. Protect.* 110:287–295.
- Rae, S.J., M. Macaulay, L. Ramsay, et al. 2005. Molecular barley breeding. *Euphytica*, in press.
- Rasmusson, D.C., and R.L. Phillips. 1997. Plant breeding progress and genetic diversity from *de novo* variation and elevated epistasis. *Crop Sci.* 37:303–310.
- Swanston, J.S., W.T.B. Thomas, W. Powell, G.R. Young, P.E. Lawrence, L. Ramsay, and R. Waugh. 1999. Using molecular markers to determine barleys most suitable for malt whisky distilling. *Mol. Breed.* 5:103–109.
- Thomas, W.T.B. 2003. Prospects for molecular breeding of barley. *Ann. Appl. Biol.* 142:1–12.

## Backcross breeding

The application of this method in plants was first proposed by H. V. Harlan and M. N. Pope in 1922. In principle, **backcross breeding** does not improve the genotype of the product, except for the substituted gene(s).

### Key features

The rationale of backcross breeding is to replace a specific undesirable gene with a desirable alternative, while preserving all other qualities (adaptation, productivity, etc.) of an adapted cultivar (or breeding line). Instead of inbreeding the  $F_1$  as is normally done, it is repeatedly crossed with the desirable parent to retrieve (by “modified inbreeding”) the desirable genotype. The adapted and highly desirable parent is called the **recurrent parent** in the crossing program, while the source of the desirable gene missing in the adapted parent is called the **donor parent**. Even though the chief role of the donor parent is to supply the missing gene, it should not be significantly deficient in other desirable traits. An inferior recurrent parent will still be inferior after the gene transfer.

### Applications

The backcross method of breeding is best suited to improving established cultivars that are later found to be deficient in one or two specific traits. It is most effective and easy to conduct when the missing trait is qualitatively (simply) inherited, dominant, and produces a phenotype that is readily observed in a hybrid plant. Quantitative traits are more difficult to breed by this method. The procedure for transferring a recessive trait is similar to that for dominant traits, but entails an additional step.

Backcrossing is used to transfer entire sets of chromosomes in the foreign cytoplasm to create a cytoplasmic male-sterile (CMS) genotype that is used to facilitate hybrid production in species including corn, onion, and wheat. This is accomplished by crossing the donor (of the chromosomes) as male until all donor chromosomes are recovered in the cytoplasm of the recurrent parent.

Backcrossing is also used for the introgression of genes via wide crosses. However, such programs are often lengthy because wild plant species possess significant amounts of undesirable traits. Backcross breeding can also be used to develop **isogenic lines** (genotypes that differ only in alleles at a specific locus)

for traits (e.g., disease resistance, plant height) in which phenotypes contrast. The method is effective for breeding when the expression of a trait depends mainly on one pair of genes, the heterozygote is readily identified, and the species is self-fertilizing. Backcrossing is applicable in the development of multilines (discussed next).

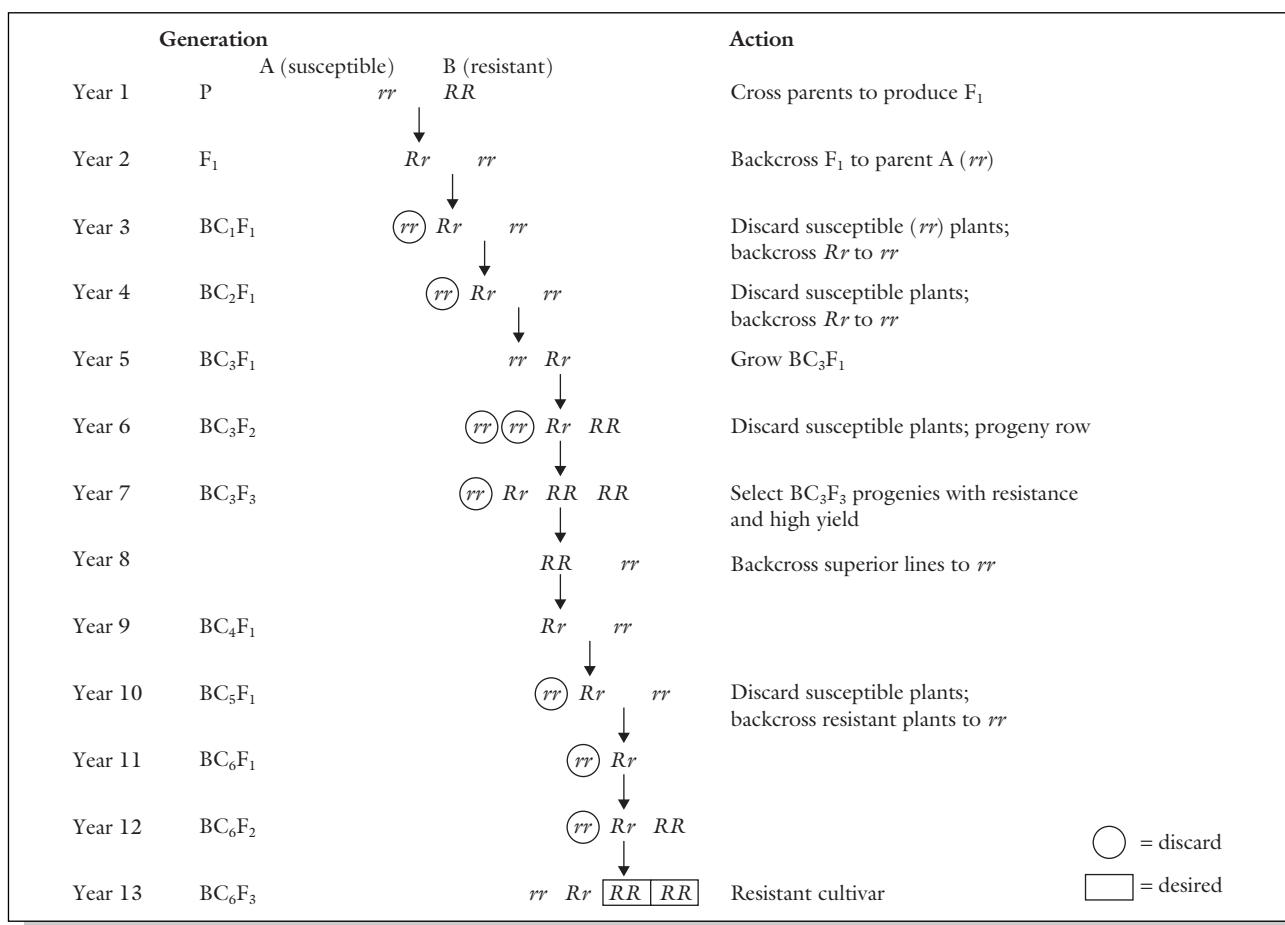
### Procedure

#### Overview

To initiate a backcross breeding program, the breeder crosses the recurrent parent with the donor parent. The  $F_1$  is grown and crossed with the recurrent parent again. The second step is repeated for as long as it takes to recover the characteristics of the recurrent parent. This may vary from two to five cycles (or more in some cases) depending on how easy the expression of the transferred gene is to observe, how much of the recurrent parental genotype the breeder wants to recover, and the overall acceptability of the donor parent. A selection pressure is imposed after each backcross to identify and discard the homozygous recessive individuals. Where the desired trait is recessive, it will be necessary to conduct a progeny test to determine the genotype of a backcross progeny before continuing with the next cross.

#### Steps: dominant gene transfer

- Year 1** Select the donor ( $RR$ ) and recurrent parent ( $rr$ ) and make 10–20 crosses. Harvest the  $F_1$  seed (Figure 16.6).
- Year 2** Grow  $F_1$  plants and cross (backcross) with the recurrent parent to obtain the first backcross ( $BC_1$ ).
- Years 3–7** Grow the appropriate backcross ( $BC_1$ – $BC_5$ ) and backcross to the recurrent parent as female. Each time, select about 30–50 heterozygous parents (backcrosses) that most resemble the recurrent parent to be used in the next backcross. The recessive genotypes are discarded after each backcross. The breeder should use any appropriate screening techniques to identify the heterozygotes (and discard the homozygous recessives). For disease-resistance breeding, artificial epiphytic conditions are created. After six backcrosses, the  $BC_5$  should very closely resemble the recurrent parent and express the donor trait. As generations advance, most plants would be increasingly more like the adapted cultivar.



**Figure 16.6** Generalized steps in breeding a dominant trait by the backcross method. The exact steps vary among breeding programs.

- Year 8** Grow  $BC_5F_1$  plants to be selfed. Select several hundreds (300–400) desirable plants and harvest them individually.
- Year 9** Grow  $BC_5F_2$  progeny rows. Identify and select about 100 desirable non-segregating progenies and bulk.
- Year 10** Conduct yield tests of the backcross with the recurrent cultivar to determine equivalence before releasing.

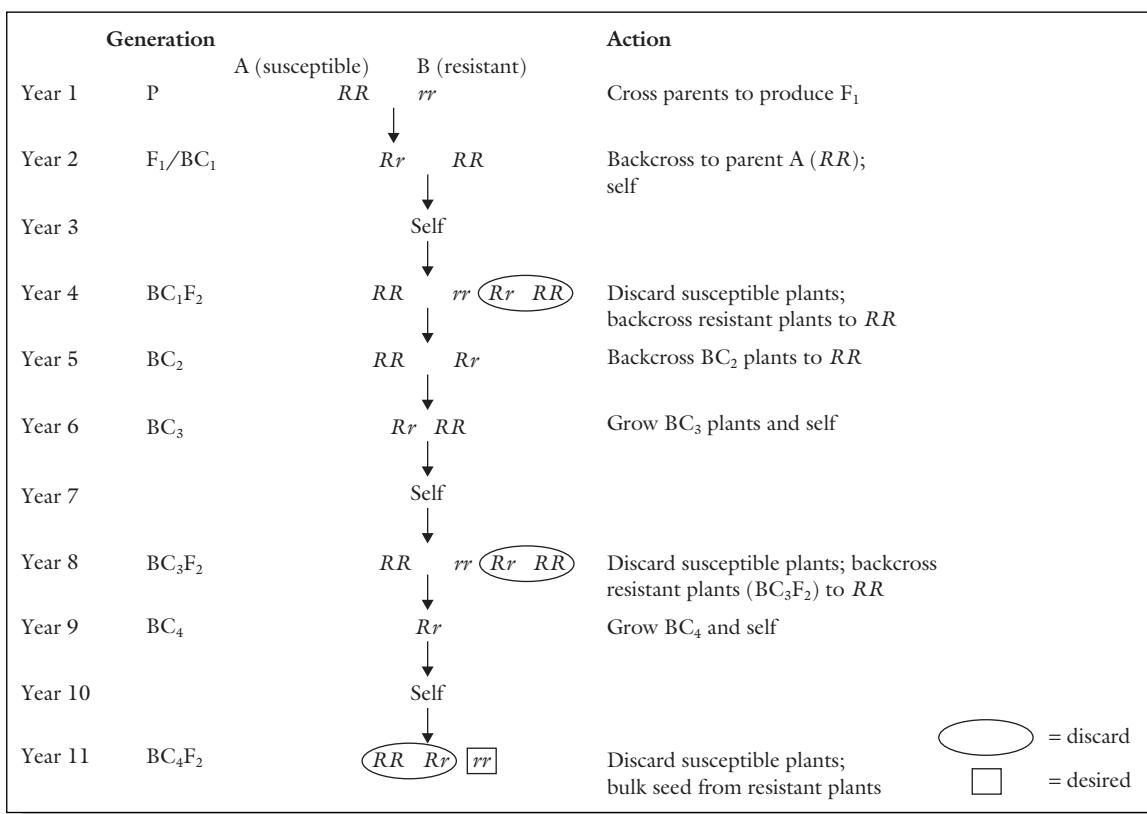
#### Comments: dominant gene transfer

The steps for transferring a dominant gene are straightforward. Following the first cross between the parents, phenotypic selection is adequate for selecting plants that exhibit the target trait. Recessive genotypes are discarded. The recurrent parent traits are not selected at

this stage. The next cross is between the selected  $F_1$  and the recurrent parent. This step is repeated for several cycles ( $BC_n$ ). After satisfactory recovery of the recurrent parent, the selected plant ( $BC_nF_1$ ) will be homozygous for other alleles but heterozygous for the desired traits. The last backcross is followed by selfing to stabilize the desired gene in the homozygous state. All homozygous ( $BC_nF_2$ ) recessive segregates are discarded.

#### Steps: recessive gene transfer

- Years 1–2** These are the same as for dominant gene transfer. The donor parent has the recessive desirable gene (Figure 16.7).
- Year 3** Grow  $BC_1F_1$  plants and self, harvest, and bulk the  $BC_1F_2$  seed. In disease-resistance breeding, all  $BC_1$ s will be susceptible.



**Figure 16.7** Generalized steps in breeding a recessive trait by the backcross method. The exact steps vary among breeding programs.

- Year 4** Grow BC<sub>1</sub>F<sub>2</sub> plants and screen for desirable plants. Backcross 10–20 plants to the recurrent parent to obtain BC<sub>2</sub>F<sub>2</sub> seed.
- Year 5** Grow BC<sub>2</sub> plants. Select 10–20 plants that resemble the recurrent parent and cross with the recurrent parent.
- Year 6** Grow BC<sub>3</sub> plants, harvest and bulk the BC<sub>3</sub>F<sub>2</sub> seed.
- Year 7** Grow BC<sub>3</sub>F<sub>2</sub> plants, screen, and select the desirable plants. Backcross 10–20 plants with the recurrent parent.
- Year 8** Grow BC<sub>4</sub> plants, harvest, and bulk the BC<sub>4</sub>F<sub>2</sub> seed.
- Year 9** Grow BC<sub>4</sub>F<sub>2</sub> plants, screen, and select the desirable plants. Backcross 10–20 plants with the recurrent parent.
- Year 10** Grow BC<sub>5</sub> plants, harvest, and bulk the BC<sub>5</sub>F<sub>2</sub> seed.
- Year 11** Grow BC<sub>5</sub>F<sub>2</sub> plants, screen, and backcross.
- Year 12** Grow BC<sub>6</sub> plants, harvest, and bulk the BC<sub>6</sub>F<sub>2</sub> seed.

- Year 13** Grow BC<sub>6</sub>F<sub>2</sub> plants and screen; select 400–500 plants and harvest separately for growing progeny rows.
- Year 14** Grow progenies of selected plants, screen, and select about 100–200 uniform progenies; harvest and bulk the seed.
- Years 15–16** Follow the procedure as in breeding for a dominant gene.

The key difference between the transfer of dominant and recessive alleles is that in the latter case, phenotypic identification is not possible after a cross. Each cross needs to be followed by selfing so that the progeny with the homozygous recessive genotype can be identified and backcrossed to the recurrent parent.

#### Comments: recessive gene transfer

- 1 Backcrossing does not have to be conducted in the environment in which the recurrent parent is adapted because all that is needed is to be able to identify and select the target trait.

- 2 Extensive advanced testing is not necessary in a backcross because the new cultivar already resembles the adapted cultivar, except for the newly incorporated trait.
- 3 It is possible to transfer two or more genes by simultaneous selection among the progeny. This undertaking requires a larger population than would be necessary if two genes are transferred independently.
- 4 Introgression of genes from weedy, adapted, exotic, or wild germplasm is possible by backcrossing. However, such transfers often take longer than typical transfers, because of the time needed to remove the undesirable agronomic traits brought in by these distantly related sources.

### Genetic issues

With each backcross, the progeny becomes more like the recurrent parent. In theory, the BC<sub>4</sub> genotype will be 93.75% identical to the recurrent parent. The mathematical relationship for the recovery of the recurrent parent is presented by W. Allard as  $1 - (\frac{1}{2})^{m-1}$ , where  $m$  is the number of generations of selfing or backcrosses. In another way, the proportion of the donor genes is reduced by 50% following each generation of backcrossing. This is obtained by the relationship  $\frac{1}{2}^{m+1}$ , where  $m$  is the number of crosses and backcrosses to the parent. For example, in the BC<sub>4</sub>, the value is  $\frac{1}{2}^5 = 3.125\%$ . To obtain the percentage of homozygotes for alleles of recurrent parents in any generation, the mathematical relationship is:

$$[(2^{m-1})/2^m]^n$$

where  $n$  = number of genes.

Because of cytoplasmic inheritance, it is sometimes critical which of the two parents is used as the female. For example, to use CMS in breeding, the male-fertile inbred lines with normal cytoplasm and non-restorer genes (B-lines) are converted to sterile cytoplasm (A-lines) to be used as male-sterile female lines in a cross.

The resulting cultivar from a backcross breeding program could differ from the starting cultivar beyond the transferred gene(s) because of **linkage drag** from the association of undesirable traits with the genes from the donor. Backcrossing is more effective in breaking linkages over selfing, especially where heritability is low for the undesirable trait.

A certain number of individuals are needed for a chance to recover the desired genes in a backcross program. This number increases as the number of genes controlling the donor trait increases. Furthermore, for

multiple gene traits, it will be necessary to grow backcross progeny through to F<sub>2</sub> or later generations to obtain the desired genotypes for advancing the program. When the trait is governed by a dominant gene, it is easy to identify plants carrying the desired gene. However, when the desired trait is conditioned by a recessive gene, an additional step is needed after each backcross to produce an F<sub>2</sub> generation in order to identify the recessive trait. The genetic advance in backcross breeding depends on several factors:

- 1 **Heritability of the trait.** As previously indicated, traits that are conditioned by major genes and have high heritability are easier to transfer by backcrossing.
- 2 **Sustainable intensity of trait expression.** Progress with selection will be steadier where the expression of the trait of interest remains at a high intensity throughout the program (i.e., no modifier gene action).
- 3 **Availability of selection aids.** The ability to identify and select desirable genotypes after the backcross is critical to the success of the procedure. Depending on the trait, special selection techniques may be needed. For disease-resistance breeding, artificial disease epiphytotics may be necessary. Molecular markers may be helpful in selection to reduce the number of backcrosses needed for the program.
- 4 **Number of backcrosses of the marker.** The genetic distance between the parents is important to the progress made in backcrossing. If both are closely related cultivars, fewer backcrosses will be needed than if the gene transfer is from a wild genotype to an adapted one.

### Advantages and disadvantages

The major advantages and limitations of backcross breeding include the following.

#### Advantages

- 1 The method reduces the number of field testings needed since the new cultivar will be adapted to the same area as the original cultivar (especially true when both parents are adapted).
- 2 Backcross breeding is repeatable. If the same parents are used, the same backcrossed cultivar can be recovered.
- 3 It is a conservative method that does not permit new recombination to occur.
- 4 It is useful for introgressing specific genes from wide crosses.
- 5 It is applicable to breeding both self-pollinated and cross-pollinated species.

### Disadvantages

- 1 Backcrossing is not effective for transferring quantitative traits. The trait should be highly heritable and readily identifiable in each generation.
- 2 The presence of undesirable linkages may prevent the cultivar being improved from attaining the performance of the original recurrent parent.
- 3 Recessive traits are more time-consuming to transfer.

### Modifications

When transferring a recessive gene (*rr*) the backcross will segregate for both homozygous dominant and heterozygous genotypes (e.g., *RR* and *Rr*). To identify the appropriate genotype to advance, it will be necessary to self the backcross to distinguish the two segregants for the *Rr*. Alternatively, both segregants may be used in the next cross, followed by selfing. The backcross progenies from the plants that produce homozygous (*rr*) segregates are heterozygous and are kept while the others are discarded. This is actually not a modification *per se*, since it is the way to transfer a recessive allele.

If a breeding program is designed to transfer genes for multiple traits, it will be more efficient to conduct separate backcross programs for each trait. The backcross-derived lines are then used as parents in a cross to develop one line that contains the multiple traits.

### Special backcross procedures

#### Congruency backcross

The **congruency backcross** technique is a modification of the standard backcross procedure whereby multiple backcrosses, alternating between the two parents in the cross (instead of restricted to the recurrent parent), are used. The technique has been used to overcome the interspecific hybridization barrier of hybrid sterility, genotypic incompatibility, and embryo abortion that occurs in simple interspecific crosses. The crosses and their genetic contribution are demonstrated in Table 16.1.

#### Advanced backcross QTLs

The **advanced backcross quantitative trait loci** (QTLs) method developed by S. D. Tanksley and J. C. Nelson allows breeders to combine backcrossing with mapping to transfer genes for QTLs from unadapted germplasm

**Table 16.1** The concept of congruency backcrossing.

Cross	Hybrid type	Genetic constitution A : B
A × B	F <sub>1</sub>	50 : 50
F <sub>1</sub> × A	BC <sub>1</sub>	75 : 25
BC <sub>1</sub> × B	CBC <sub>2</sub>	37.5 : 62.5
CBC <sub>2</sub> × A	CBC <sub>3</sub>	68.8 : 31.3
CBC <sub>3</sub> × B	CBC <sub>4</sub>	34.4 : 65.6
CBC <sub>4</sub> × A	CBC <sub>5</sub>	67.2 : 32.8

into an adapted cultivar. This method was developed for the simultaneous discovery and transfer of desirable QTLs from unadapted germplasm into elite lines. It was briefly discussed in Chapter 14.

### Multiline breeding and cultivar blends

N. F. Jensen is credited with first using this breeding method in oat breeding in 1952 to achieve a more lasting form of disease resistance. Multilines are generally more expensive to produce than developing a synthetic cultivar, because each component line must be developed by a separate backcross.

#### Key features

The key feature of a multiline cultivar is disease protection. Technically, a **multiline** or **blend** is a planned seed mixture of cultivars or lines (multiple pure lines) such that each component constitutes at least 5% of the whole mixture. The pure lines are phenotypically uniform for morphological and other traits of agronomic importance (e.g., height, maturity, photoperiod), in addition to genetic resistance for a specific disease. The component lines are grown separately, followed by composting in a predetermined ratio. Even though the term multiline is often used interchangeably with blend, sometimes the former is limited to mixtures involving **isolines** or near **isogenic lines** (lines that are genetically identical except for the alleles at one locus). The purpose of mixing different genotypes is to increase heterogeneity in the cultivars of self-pollinated species. This strategy would decrease the risk of total crop loss from the infection of one race of the pathogen, or some other biotic or abiotic factor. The component genotypes are designed to respond to different versions or degrees of an environmental stress factor (e.g., different races of a pathogen).

## Applications

One of the earliest applications of multilines was for breeding “variable cultivars” to reduce the risk of loss to pests that have multiple races, and whose incidence is erratic from season to season. Planting a heterogeneous mixture can physically impede the spread of disease in the field as resistant and susceptible genotypes intermingle.

Mixtures may be composed to provide stable performance in the face of variable environments. Mixtures and blends are common in the turfgrass industry. Prescribing plants for conditions that are not clear-cut is challenging. Using mixtures or blends will increase the chance that at least one of the component genotypes will match the environment.

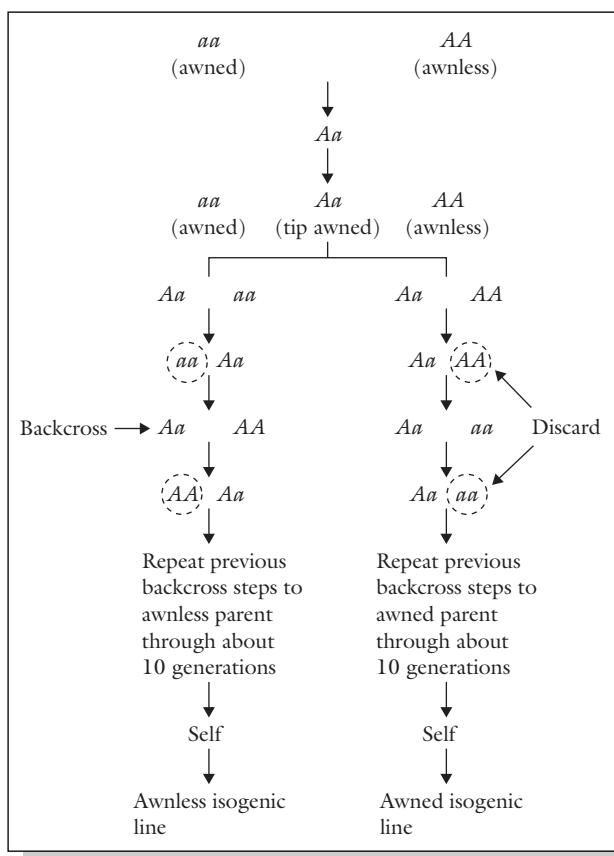
In backcross breeding, the deficiency in a high-yielding and most desirable cultivar is remedied by gene substitution from a donor. Similarly, the deficiency of an adapted and desirable cultivar may be overcome by mixing it with another cultivar that may not be as productive but has the trait that is missing in the desirable cultivar. Even though this strategy will result in lower yield per unit area in favorable conditions, the yield will be higher than it would be under adverse conditions if only a pure adapted cultivar was planted.

Multilines composed for disease resistance are most effective against airborne pathogens with physiological races that are explosive in reproduction. An advantage of blends and mixtures that is not directly related to plant breeding, is marketing. Provided a label “variety not stated” is attached to the seed bag, blends of two or more cultivars can be sold under various brand names, even if they have identical composition.

## Procedure

The backcross is the breeding method for developing multilines. The agronomically superior line is the recurrent parent, while the source of disease resistance constitutes the donor parent. To develop multilines by isolines, the first step is to derive a series of backcross-derived isolines or near-isogenic lines (since true isolines are illusive because of linkage between genes of interest and other genes influencing other traits). A method for developing multilines is illustrated in Figure 16.8. The results of the procedure are two cultivars that contrast only in a specific feature. For disease resistance, each isolate should contribute resistance to a different physiological race (or group of races) of the disease.

The component lines of multilines are screened for disease resistance at multilocations. The breeder then



**Figure 16.8** Generalized steps in breeding multiline cultivars.

selects resistant lines that are phenotypically uniform for selected traits of importance to the crop cultivar. The selected components are also evaluated for performance (yield ability), quality, and competing ability. Mixtures are composed annually based on disease patterns. It is suggested that at least 60% of the mixture comprises isolines resistant to the prevalent disease races at the time. The proportion of the component lines are determined by taking into account the seed analysis (germination percentage, viability).

## Genetic issues

Multiline cultivars consist of one genetic background but different genes for the trait of interest. A multiline is hence spatially differentiated, plant to plant. When planted, the cultivar creates a mosaic of genotypes in the field to provide a buffering against the rapid development of disease.

Two basic mechanisms are used by multiline cultivars to control disease – stabilization of the patterns of virulence genes and population resistance (see Chapter 20). By stabilizing the patterns of virulence genes in the pathogen, it is supposed that genes for resistance would retain their value in protecting the cultivar for an extended period. The concept of population resistance is the delay in the buildup of the pathogen in the multiline cultivar.

Spore trapping has also been proposed to explain disease buildup in the population of a multiline by reducing the effective inoculation load in each generation. Following the primary inoculation (the initial spores to infect the field), spores that land on resistant genotypes will not germinate. Similarly, progeny spores from susceptible genotypes landing on resistant genotypes will not germinate. The sum effect of these events is a reduction in the inoculum load in each generation.

### Advantages and disadvantages

Multilines have certain key advantages and disadvantages.

#### Advantages

- 1 A multiline provides protection to a broad spectrum of races of a disease-producing pathogen.
- 2 The cultivar is phenotypically uniform.
- 3 Multilines provide greater yield stability.
- 4 A multiline can be readily modified (reconstituted) by replacing a component line that becomes susceptible to the pathogen, with a new disease-resistant line.

#### Disadvantages

- 1 It takes a long time to develop all the isolines to be used in a multiline, making it laborious and expensive to produce.
- 2 Multilines are most effective in areas where there is a specialized disease pathogen that causes frequent severe damage to plants.
- 3 Maintaining the isolate is labor intensive.

#### Modifications

Cultivars can be created with different genetic backgrounds (instead of one genetic background). When different genetic lines (e.g., two or more cultivars) are combined, the mixture is a **composite** called a **variety blend**. Blends are less uniform in appearance than a pure-line cultivar. They provide a buffering effect against genotype × environment interactions.

### Composites

As previously stated, a **composite cultivar**, like a multiline, is a mixture of different genotypes. The difference between the two lies primarily in the genetic distance between the components of the mixture. Whereas a multiline is constituted of closely related lines (isolines), a composite may consist of inbred lines, all types of hybrids, populations, and other less similar genotypes. However, the components are selected to have common characters, such as a similar growth period, or degrees of resistance to lodging or to a pathogenic agent. This consideration is critical to having uniformity in the cultivar.

A composite cultivar should be distinguished from a composite cross that is used to generate multiple-parent crosses by successively crossing parents (i.e., single, double, cross, etc.) until the final parent contains all parents. Composites may serve as a continuous source of new entries for a breeding nursery. Any number of entries may be included in a composite, provided selection is judiciously made after evaluation. New entries may be added at any time. Technically, a composite may derive from a single diverse variety, a progeny from a single cross, or even several hundreds of entries. However, a good number of entries lies between 10 and 20. The breeder's objectives determine the kind of entries used for breeding a composite. Using elite and similar genotypes would make the composite more uniform, robust (at least initially), but less genetically diverse. The reverse would be true if diverse entries are included. As a population improvement product, the yield of a composite can be improved by advancing it through several cycles of selection.

In species such as sorghum, which are predominantly self-pollinated, a recessive male-sterility gene that is stable across environments may be incorporated into the composite (e.g., the *ms*<sub>3</sub> in sorghum) by crossing each entry to the source of the sterility gene prior to mixing. The F<sub>1</sub> (fertile) is first selfed and then backcrossed to the male-sterile segregates. The recurrent parents are then mixed to create the composite.

### Recurrent selection

**Recurrent selection** is a cyclical improvement technique aimed at gradually concentrating desirable alleles in a population. It is one of the oldest techniques of plant breeding. The name was coined by F. H. Hull in 1945. It was first developed for improving cross-pollinated

species (maize) and has been a major breeding method for this group of plants. Hence detailed discussion of this method of breeding is deferred to Chapter 17. It is increasingly becoming a method of improving self-pollinated species. It has the advantage of providing additional opportunities for genetic recombination through repeated intermating after the first cross, something not available with pedigree selection. It is effective for improving quantitative traits.

### Comments

- 1 Recurrent selection requires extensive crossing, which is a challenge in autogamous species. To overcome this problem, a male-sterility system may be incorporated into the breeding program. With male sterility, natural crossing by wind and/or insects will eliminate the need for hand pollination.
- 2 Adequate seed may be obtained by crossing under a controlled environment (greenhouse) where the crossing period can be extended.

### Advantages and disadvantages

There are several advantages and disadvantages of the application of recurrent selection to breeding autogamous species.

#### Advantages

- 1 Opportunities to break linkage blocks exist because of repeated intercrossing.
- 2 It is applicable to both autogamous grasses (monocots) and legumes (dicots).

#### Disadvantages

- 1 Extensive crossing is required, something that is a challenge in autogamous species. A male-sterility system may be used to facilitate this process.
- 2 Sufficient seed may not be available after intercrossing. This also may be resolved by including male sterility in the breeding program.
- 3 More intermatings may prolong the duration of the breeding program.
- 4 There is also the possibility of breaking desirable linkages.

### References and suggested reading

- Agrawal, R.L. 1998. Fundamentals of plant breeding and hybrid seed production. Science Publishers, Inc., Enfield, NH.
- Chahal, G.S., and S.S. Gosal. 2000. Principles and procedures of plant breeding: Biotechnological and conventional approaches. CRC Press, New York.
- Degago, Y., and C.E. Caviness. 1987. Seed yield of soybean bulk populations grown for 10–18 years in two environments. *Crop Sci.* 27:207–210.
- Eaton, D.L., R.H. Busch, and V.L. Youngs. 1986. Introgression of unadapted germplasm into adapted spring wheat. *Crop Sci.* 26:473–478.
- Ferh, W.R. 1987. Principles of cultivar development. McMillan, New York.
- Frey, K.J. 1983. Plant population management and breeding. In: *Cultivar breeding* (Wood, D.R., ed.). American Society of Agronomy, Madison, WI.
- Hamblin, J. 1977. Plant breeding interpretations on the effects of bulk breeding on four populations of beans (*Phaseolus vulgaris* L.). *Euphytica* 25:157–168.
- Jensen, N.F. 1978. Composite breeding methods and the DSM system in cereals. *Crop Sci.* 18:622–626.
- Poehlman, J.M., and D.H. Slepper. 1995. Breeding field crops. Iowa State University Press, Ames, IA.
- Tigchelaar, E.C., and V.W.D. Casali. 1976. Single seed descent: Applications and merits in breeding self-pollinated crops. *Acta Horticulturae* 63:85–90.
- Wilcox, J.R., and I.F. Cavins. 1995. Backcrossing high seed protein to a soybean cultivar. *Crop Sci.* 35:1036–1041.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 The adapted and highly desirable parent in a backcross is the donor parent.
- 2 With each backcross, the progeny becomes more like the donor parent.
- 3 Isogenic lines differ in alleles at a specified locus.
- 4 A composite may consist of hybrid cultivars.
- 5 The donor parent is used only once in a cross in a backcross program.
- 6 Single-seed descent is the oldest plant breeding method.
- 7 Record keeping is a critical part of the pedigree selection method of breeding.
- 8 Nilsson-Ehle developed the mass selection procedure.
- 9 Bulk population breeding is suited to breeding plants that are closely spaced in commercial planting.
- 10 Mass selection is most effective if the trait of interest has high heritability.

### Part B

Please answer the following questions:

- 1 ..... is the adapted parent in a backcross.
- 2 Give a specific advantage of multiline cultivars.
- 3 Give a specific disadvantage of the backcross breeding method.
- 4 ..... developed the pure-line method of plant breeding.
- 5 Discuss a specific genetic issue involved with mass selection.
- 6 Give a specific disadvantage of the pedigree method of breeding.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the key features of backcross breeding.
- 2 Discuss the application of multiline breeding.
- 3 Distinguish between composite breeding and multiline breeding.
- 4 Discuss the advantages of bulk breeding.
- 5 Discuss the advantages of the single-seed descent method of selection.
- 6 Compare and contrast the mass selection and pure-line selection methods of breeding.
- 7 Describe the steps involved in the mass selection method of breeding.



# Breeding cross-pollinated species

---

## Purpose and expected outcomes

*As previously noted, breeding cross-pollinated species tends to focus on population improvement rather than the improvement of individual plants as is the focus in breeding self-pollinated species. In addition to methods such as mass selection that are applicable to both self- and cross-pollinated species, there are specific methods that are suited to population improvement. Some methods are used less frequently in breeding. Further, certain methods are more effective and readily applied for breeding certain species than others. After studying this chapter, the student should be able to:*

- 1 Present the method of mass selection in cross-pollinated species.
  - 2 Discuss the concept of recurrent selection.
  - 3 Describe the methods of half- and full-sib selection.
  - 4 Discuss the method of S<sub>1</sub> and S<sub>2</sub> selection.
  - 5 Discuss the development of synthetic cultivars.
  - 6 Discuss the application of the backcross technique in cross-pollinated species.
- 

## Concept of population improvement

As stated in the introduction, the methods of selection (discussed in Chapter 16) for improving self-pollinated species tend to focus on improving individual plants. The methods of improving cross-fertilized species, on the other hand, tend to focus on improving a population of plants. As defined in Chapter 7, a population is a large group of interbreeding individuals. The application of the principles and concepts of population genetics are made to effect changes in the genetic structure of a population of plants. Overall, breeders seek to change the gene frequency such that desirable genotypes predominate in the population. Also, in the process of changing gene frequencies, new genotypes (that did not exist in the initial population) will arise. It is important for breeders to maintain genetic variability in these populations so that further improvements of the population may be achieved in the future.

To improve the population, breeders generally assemble germplasm, evaluate selected selfed plants, cross the progenies of the selected selfed plants in all possible combinations, and bulk and develop inbred lines from the populations. In cross-pollinated species, a cyclical selection approach, called recurrent selection, is often used for intermating. The cyclical selection was developed for increasing the frequency of favorable genes for quantitative traits. Various methods of recurrent selection are used for producing progenies for evaluation as will be discussed next.

The procedures for population improvement may be classified in several ways, such as according to the unit of selection – either individual plants or family of plants. Also, the method may be grouped according to the populations undergoing selection as either intrapopulation or interpopulation.

- 1 **Intrapopulation improvement.** Selection is practiced within a specific population for its improvement for

specific purposes. Intrapopulation improvement is suitable for:

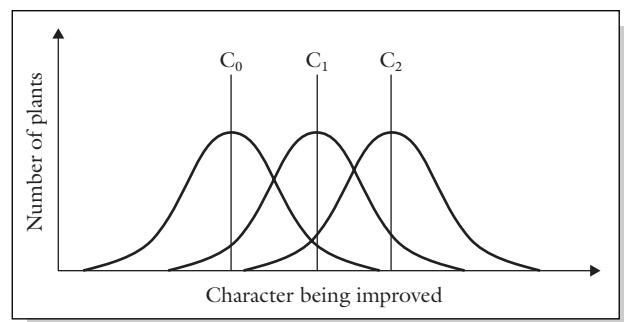
- Improving populations where the end product will be a population or synthetic cultivar.
- Developing elite pure lines for hybrid production.
- Developing mixed genotype cultivars (in self-fertilized species).

**2 Interpopulation improvement.** Methods of interpopulation improvement entail selection on the basis of the performance of a cross between two populations. This approach is suitable for use when the final product will be a hybrid cultivar. Interpopulation heterosis is exploited.

### Concept of recurrent selection

The concept of **recurrent selection** was introduced in Chapter 16 as a cyclical and systematic technique in which desirable individuals are selected from a population and mated to form a new population; the cycle is then repeated. The purpose of a recurrent selection in a plant breeding program is to improve the performance of a population with respect to one or more traits of interest, such that the new population is superior to the original population in mean performance and in the performance of the best individuals within it (Figure 17.1).

The source material may be random mating populations, synthetic cultivars, single cross, or double cross plants. The improved population may be released as a new cultivar or used as a breeding material (parent) in other breeding programs. The most desirable outcome of recurrent selection is that the improved population is produced without reduction in genetic variability. This way, the population can respond to future improvement.



**Figure 17.1** The concept of recurrent selection.

The success of a recurrent selection program rests on the genetic nature of the base population. Several key factors should be considered in the development of the base population. First, the parents should have high performance regarding the traits of interest and should not be closely related. This would increase the chance of maximizing genetic diversity in the population. It is also recommended to include as many parents as possible in the initial crossing to increase genetic diversity. Crossing provides opportunity for recombination of genes to increase genetic diversity of the population. More rounds of mating will increase the opportunity for recombination, but it increases the duration of the breeding program. The breeder should decide on the number of generations of intermating that is appropriate for a breeding program.

### Key features

A recurrent selection cycle consists of three main phases:

- Individual families are created for evaluation. Parents are crossed in all possible combinations.
- The plants or families are evaluated and a new set of parents selected.
- The selected parents are intermated to produce the population for the next cycle of selection.

This pattern or cycle is repeated several times (3–5 times). The first (original) cycle is labeled  $C_0$ , and is called the base population. The subsequent cycles are named consecutively as  $C_1$ ,  $C_2$  . . .  $C_n$  (Figure 17.1). It is possible, in theory, to assemble all the favorable genes in a population in a single generation if plant breeders could handle a population of infinite size. However, in practice, as J. K. Frey pointed out, the technique of recurrent selection is applied to breeding with the hope that desirable genes will be gradually accumulated until there is a reasonable probability of obtaining the ultimate genotype in a finite sample.

### Applications

Recurrent selection may be used to establish a broad genetic base in a breeding program. Because of multiple opportunities for intermating, the breeder may add new germplasm during the procedure when the genetic base of the population rapidly narrows after selection cycles. Research has indicated that recurrent selection is superior to classic breeding when linkage disequilibrium exists. In fact, the procedure is even more effective when

epistatic interactions enhance the selective advantage of new recombinants. Recurrent selection is applied to legumes (e.g., peanut, soybean) as well as grasses (e.g., barley, oats).

### **Genetic basis of recurrent selection**

Various recurrent selection schemes are available. They exploit additive partial dominance to dominance and overdominance types of gene action. However, without the use of testers (as in simple recurrent selection) the scheme is effective only for traits of high heritability. Hence, only additive gene action is exploited in the selection for the trait. Where testers are used, selection for general combining ability (GCA) and specific combining ability (SCA) are applicable, permitting the exploitation of other gene effects. Recurrent selection for GCA is more effective than other schemes when additive gene effects are more important; recurrent selection for SCA is more effective than other selection schemes when overdominance gene effects are more important. Reciprocal recurrent selection is more effective than others when both additive and overdominance gene effects are more important. All three schemes are equally effective when additive with partial to complete dominance effects prevail. The expected genetic advance may be obtained by the following general formula:

$$\Delta G = (C_i V_A) / y \sigma_p$$

where  $\Delta G$  = expected genetic advance,  $C$  = measure of parental control ( $C = 0.5$  if selection is based on one parent, and equals 1 when both parents are involved),  $I$  = selection intensity,  $V_A$  = additive genetic variance among the units of selection,  $y$  = number of years per cycle, and  $\sigma_p$  = phenotypic standard deviation among the units of selection. Increasing the selection pressure (intensity) will increase gain in selection provided the population advanced is not reduced to a size where genetic drift and loss of genetic variance can occur. Other ways of enhancing genetic advance per cycle include selection for both male and female parents, maximizing available additive genetic variance, and management of environmental variance among selection units. The formulae for various schemes are presented at the appropriate times in the textbook.

The role of parental control in genetic gain can be manipulated by the breeder through the exercise of control over parents in a breeding program. Genetic control over both parents will double the genetic gain that

can be achieved when the breeder has control over the females only. Both parents may be controlled in one of several ways: (i) selfing of selected individuals (instead of being pollinated by selected and unselected males); (ii) selection before pollination and recombination among selected plants only; and (iii) recombination occurs among selected clones.

### **Types of recurrent selection**

There are four basic recurrent selection schemes, based on how plants with the desired traits are identified.

- 1 **Simple recurrent selection.** This is similar to mass selection with 1 or 2 years per cycle. The procedure does not involve the use of a tester. Selection is based on phenotypic scores. This procedure is also called phenotypic recurrent selection.
- 2 **Recurrent selection for general combining ability.** This is a half-sib progeny test procedure in which a wide genetic-based cultivar is used as a tester. The testcross performance is evaluated in replicated trials prior to selection.
- 3 **Recurrent selection for specific combining ability.** This scheme uses an inbred line (narrow genetic base) for a tester. The testcross performance is evaluated in replicated trials before selection.
- 4 **Reciprocal recurrent selection.** This scheme is capable of exploiting both general and specific combining ability. It entails two heterozygous populations, each serving as a tester for the other.

### **Intrapopulation improvement methods**

Common intrapopulation improvement methods in use include mass selection, ear-to-row selection, and recurrent selection. Intrapopulation methods may be based on single plants as the unit of selection (e.g., as in mass selection), or family selection (e.g., as in various recurrent selection methods).

### **Individual plant selection methods**

Mass selection for line development (see Chapter 16) is different from mass selection for population improvement. Mass selection for population improvement aims at improving the general population performance by selecting and bulking superior genotypes that already exist in the population.

### Key features

The selection units are individual plants. Selection is solely on phenotypic performance. Seed from selected plants (pollinated by the population at large) are bulked to start the next generation. No crosses are made, but a progeny test is conducted. The process is repeated until a desirable level of improvement is observed.

### Procedure

- Year 1** Plant the source population (local variety, synthetic variety, bulk population, etc.). Rogue out undesirable plants before flowering, and then select several hundreds of plants based on the phenotype. Harvest and bulk.
- Year 2** Repeat year 1. Grow selected bulk in a preliminary yield trial, including a check. The check is the unselected population (original), if the goal of the mass selection is to improve the population.
- Year 3** Repeat year 2 for as long as progress is made.
- Year 4** Conduct advanced yield trials.

The mass selection may be longer, depending on the progress being made.

### Genetic issues

The effectiveness of the method depends on the heritability of the trait since selection is solely on the phenotype. It is also most effective where additive gene action operates. Effectiveness of mass selection also depends on the number of genes involved in the control of the trait of interest. The more additive genes are involved, the greater the efficiency of mass selection. The expected genetic advance through mass selection is given by the following (for one sex – female):

$$\begin{aligned}\Delta G_m &= [(1/2)i\sigma_A^2]\sigma_p \\ &= [(1/2)i\sigma_A^2]/[\sigma_A^2 + \sigma_D^2 + \sigma_{AE}^2 + \sigma_{DE}^2 \\ &\quad + \sigma_e^2 + \sigma_{me}^2]\end{aligned}$$

where  $\sigma_p$  = phenotypic standard deviation in the population,  $\sigma_A^2$  = additive variance,  $\sigma_D^2$  = dominance variance, and the other factors are interaction variances.  $\Delta G_m$  doubles with both sexes. This large denominator makes mass selection inefficient for low heritability traits. Selection is limited to only the female parents since there is no control over pollination.

### Advantages and disadvantages

The major advantages and disadvantages of individual plant selection methods include the following.

#### Advantages

See Chapter 16.

#### Disadvantages

- 1 Using phenotypic selection makes the selection of superior plants often difficult.
- 2 Lack of pollen control means both desirable and undesirable pollen will be involved in pollination of the selected plants.
- 3 If selection intensity is high (small population size advanced) the possibility of inbreeding depression is increased.

### Modifications

- 1 **Stratified or grid system.** Proposed by C. O. Gardener, the field is divided into small grids (or subplots) with little environmental variance. An equal number of superior plants is selected from each grid for harvesting and bulking.
- 2 **Honeycomb design.** Proposed by A. Fasoulas, the planting pattern is triangular rather than the conventional rectangular pattern. Each single plant is at the center of a regular hexagon, with six equidistant plants, and is compared to the other six equidistant plants.

There are other modifications that are sometimes complex to apply and have variable effects on selection response.

### Family selection methods

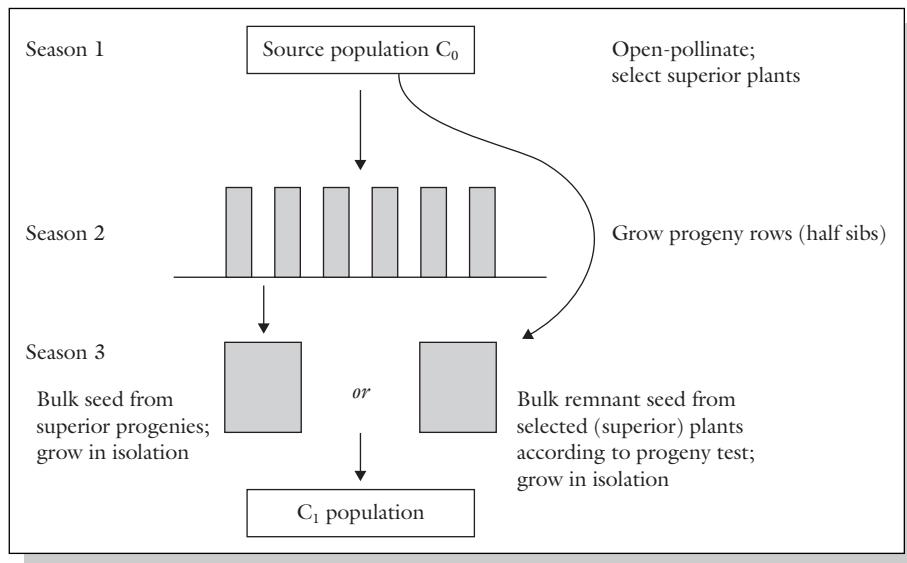
Family selection methods are characterized by three general steps:

- 1 Creation of a family structure.
- 2 Evaluation of families and selection of superior ones by progeny testing.
- 3 Recombination of selected families or plants within families to create a new base population for the next cycle of selection.

Generally, the duration of each step is one generation, but variations exist.

#### Half-sib family selection methods

The basic feature of this group of methods is that half-sib families are created for evaluation and recombination,



**Figure 17.2** Generalized steps in breeding by ear-to-row selection.

both steps occurring in one generation. The populations are created by random pollination of selected female plants in generation 1. The seed from generation 1 families are evaluated in replicated trials and in different environments for selection. There are different kinds of half-sib family selection methods including the following.

**Ear-to-row selection** This is the simplest scheme of half-sib selection applicable to cross-pollinated species (Figure 17.2).

**Applications** Half-sib selection is widely used for breeding perennial forage grasses and legumes. A poly-cross mating system is used to generate the half-sib families from selected vegetatively maintained clones. The families are evaluated in replicated rows for 2–3 years. Selecting traits of high heritability (e.g., oil and protein content of maize) is effective.

#### Procedure

**Season 1** Grow the source population (heterozygous population) and select desirable plants ( $S_0$ ) based on the phenotype according to the traits of interest. Harvest plants individually. Keep remnant seed of each plant.

**Season 2** Grow replicated half-sib progenies ( $S_0 \times$  tester) from selected individuals in one environment (yield trial). Select best progenies and bulk to create progenies for the next

cycle. The bulk is grown in isolation (crossing block) and random mated.

**Season 3** The seed is harvested and used to grow the next cycle.

Alternatively, the breeder may bulk the remnant seed of  $S_0$  plants whose progeny have been selected, and use that to initiate the next cycle.

**Genetic issues** The expected genetic gain from half-sib selection is given by:

$$\Delta G_{HS} = [(1/4)i\sigma_A]/\sigma_{PHS}$$

where  $\sigma_{PHS}$  = standard deviation of the phenotypic variance among half sibs. Other components are as before. The tester is the parental population and hence selection or control is over only sex. The genetic gain is hence reduced by half (the available additive genetic variance is also reduced by half because of the control over the female parent). Genetic gain can be doubled by selfing each parent to obtain  $S_1$ , then crossing to obtain half sibs.

**Modifications** The basic or traditional ear-to-row selection method did not show much gain over mass selection. An improvement was proposed by J. H. Lohnquist in which the creation of family structure, evaluation, and recombination are conducted in one generation. The half-sib families are evaluated in

replicated trials in many environments. The approach was to better manage the environmental and  $G \times E$  interactions.

**Modified half-sib selection** This is a modified version of the half-sib family selection method.

**Applications** This method of breeding has been applied to the improvement of perennial species as previously indicated for the traditional ear-to-row selection. It has been used in maize for yield gains of between 1.8% and 6.3% per cycle.

#### Procedure

**Season 1** Select desirable plants from source population. Harvest these open-pollinated (half sibs) individually.

**Season 2** Grow progeny rows of selected plants at multiple locations and evaluate for yield performance. Plant female rows with seed from individual half-sib families, alternating with male rows (pollinators) planted with bulked seed from the entire population. Select desirable plants (based on average performance over locations) from each progeny separately. Bulk the seed to start the next cycle.

**Genetic issues** The genetic gain has two components – among ear rows across environments (interfamily selec-

tion) and within families (intrafamily selection). The total genetic gain is given by:

$$\Delta G_{\text{mHS}} = [(1/8)i\sigma_A]/\sigma_{\text{PHS}} + [(3/8)i\sigma_A]/\sigma_{\text{we}}$$

where  $\sigma_{\text{we}}$  = square root of the plant-to-plant within-plot variance. Other components are as before.

#### Full-sib family selection

Full sibs are generated from biparental crosses using parents from the base population. The families are evaluated in a replicated trial to identify and select superior full-sib families, which are then recombined to initiate the next cycle.

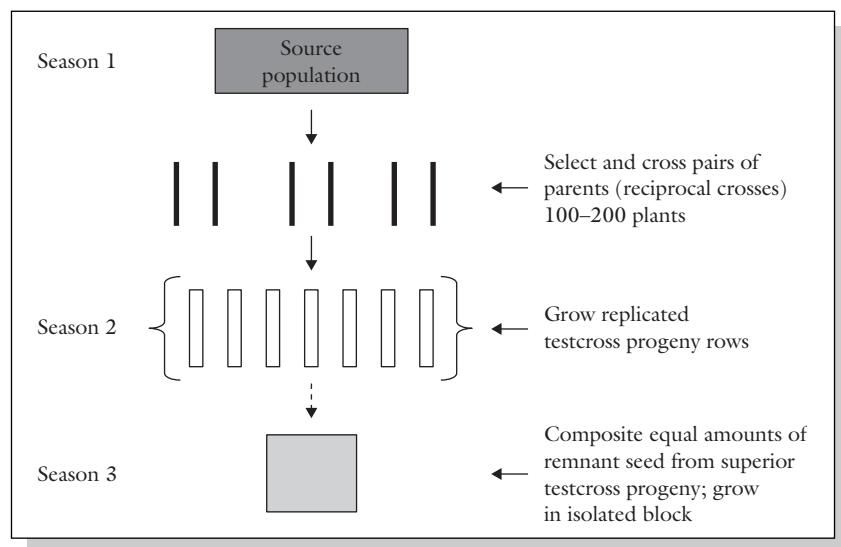
**Applications** Full-sib family selection has been used for maize improvement. A selection response per cycle of about 3.3% has been recorded in maize.

#### Procedure: cycle 0

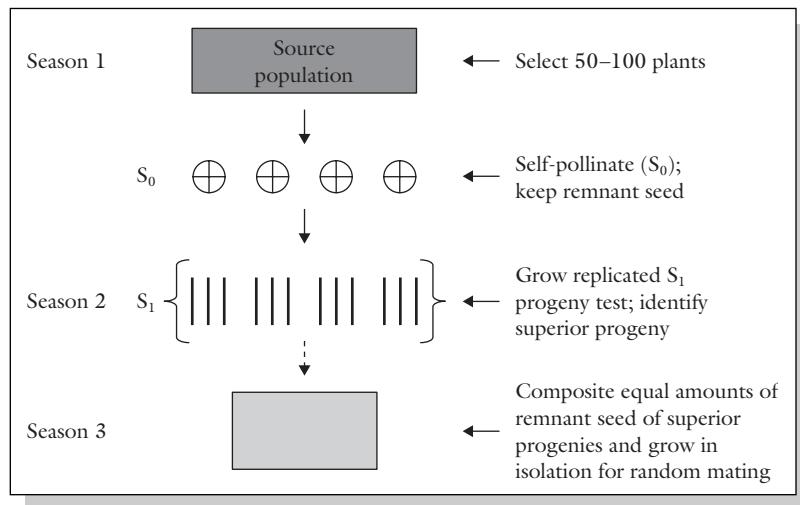
**Season 1** Select random pairs of plants from the base population and intermate, pollinating one with the other (reciprocal pollination). Make between 100 and 200 biparental crosses. Save the remnant seed of each full-sib cross (Figure 17.3).

**Season 2** Evaluate full-sib progenies in multiple location replicated trials. Select the promising half sibs (20–30).

**Season 3** Recombine the selected full sibs.



**Figure 17.3** Generalized steps in breeding by the full-sib method.



**Figure 17.4** Generalized steps in breeding based on  $S_1/S_2$  progeny performance.

*Procedure: cycle 1* This is the same as for cycle  $C_0$ .

*Genetic issues* The genetic gain per cycle is given by:

$$\Delta G_{FS} = i\sigma_A^2 / 2\sigma_{FS}$$

where  $\sigma_{FS}$  = phenotypic standard deviation of the full-sib families.

#### Selfed ( $S_1$ or $S_2$ ) family selection

An  $S_1$  is a selfed plant from the base population. The key features are the generation of  $S_1$  or  $S_2$  families, evaluating them in replicated multienvironment trials, followed by recombination of remnant seed from selected families (Figure 17.4).

**Applications** The  $S_1$  appears to be best suited for self-pollinated species (e.g., wheat, soybean). It has been used in maize breeding. One cycle is completed in three seasons in  $S_1$  and four seasons in  $S_2$ . A genetic gain per cycle of 3.3% has been recorded.

#### Procedure

- Season 1** Self-pollinate about 300 selected  $S_0$  plants. Harvest the selfed seed and keep the remnant seed of each  $S_1$ .
- Season 2** Evaluate  $S_1$  progeny rows to identify superior progenies.
- Season 3** Random mate selected  $S_1$  progenies to form a  $C_1$  cycle population.

**Genetic issues** The main reason for using this scheme is to increase the magnitude of additive genetic variance. In theory the genetic gain is given by:

$$\Delta G_{S1} = i\sigma_{A1}^2 / \sigma_{PS1}$$

where  $\sigma_{A1}^2$  = additive genetic variance among  $S_1$ , and  $\sigma_{PS1}$  = phenotypic standard deviation among  $S_1$  families. The additive genetic variation among  $S_2$  is two times that of  $S_1$ . The  $S_1$  and  $S_2$ , theoretically, have the highest expected genetic gain per cycle for intrapopulation improvement. However, various reports have indicated that, in practice, full-sib and testcross selections have produced greater genetic gain for both populations *per se* and the population crosses.

#### Family selection based on a testcross

The key feature of this approach to selection is that it is designed to improve both the population *per se* as well as its combining ability. The choice of the tester is most critical to the success of the schemes. Using a tester to aid in selection increases the duration of a cycle by 1 year (i.e., a 3-year cycle instead of a 2-year one as in phenotypic selection). The choice of a tester is critical to the success of a recurrent selection breeding program. The commonly used testers may be classified into two: (i) narrow genetic base testers (e.g., an inbred line); and (ii) broad genetic base testers (e.g., open-pollinated cultivars, synthetic cultivars, double-cross hybrid). Broad base testers are used for testing GCA in the population under improvement, whereas narrow genetic base testers are used to evaluate SCA and possibly GCA.

Generally, plants are selected from the source population and are selfed in year 1. Prior to intermating, the

selected plants are crossed as females to a tester in year 2. Intermating of selected plants occurs in year 3.

#### *Half-sib selection with progeny test*

**Half-sib** or **half-sib family selection** is so-called because only one parent in the cross is known. In 1899, C. G. Hopkins first used this procedure to alter the chemical composition of corn by growing progeny rows from corn ears picked from desirable plants. Superior rows were harvested and increased as a new cultivar. The method as applied to corn is called **ear-to-row breeding**.

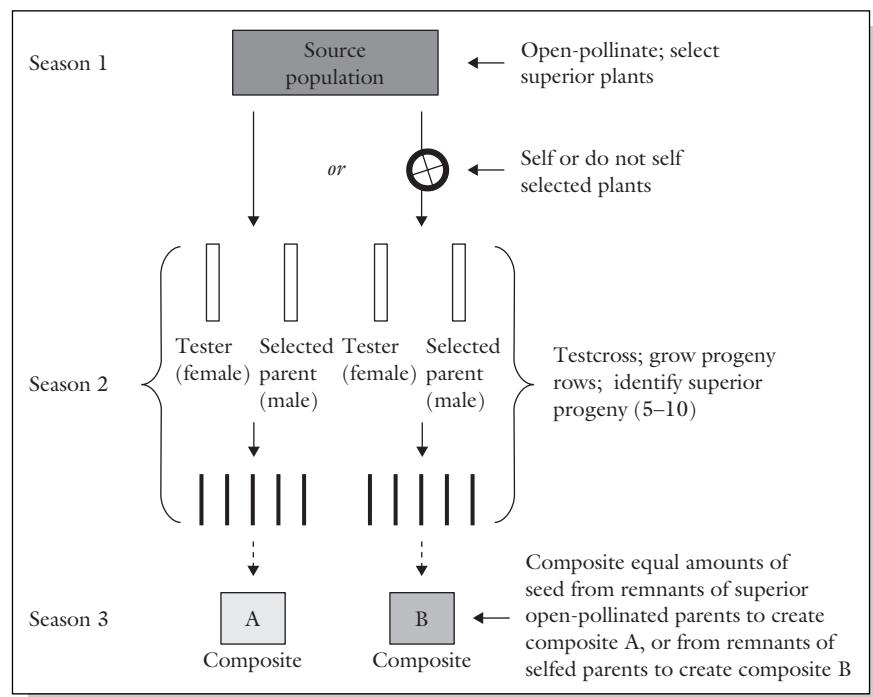
**Key features** There are various half-sib progeny tests, such as the topcross progeny test, open-pollinated progeny test, and polycross progeny test. A half sib is a plant (or family of plants) with a common parent or pollen source. Individuals in a half-sib selection are evaluated based on their half-sib progeny. Unlike mass selection, in which individuals are selected solely on phenotypic basis, the half sibs are selected based on the performance of their progenies. The specific identity of the pollen sources is not known.

**Applications** Recurrent half-sib selection has been used to improve agronomic traits as well as seed composition

traits in corn. It is suited for improving traits with high heritability, and in species that can produce sufficient seed per plant to grow a yield trial. Species with self-incompatibility (no self-fertilization) or some other constraint of sexual biology (e.g., male-sterile) are also suited to this method of breeding.

**Procedure** A typical cycle of half-sib selection entails three activities – crossing the plants to be evaluated to a common tester, evaluating the half-sib progeny from each plant, and intercrossing the selected individuals to form a new population. In the second season, each separate seed pack is used to plant a progeny row in an isolated area (Figure 17.5). The remnant seed is saved. In season 3, 5–10 superior progenies are selected, and the seed is harvested and composited; alternatively, the same is done with the remnant seed. The composites are grown in an isolation block for open-pollination. Seed is harvested as a new open-pollinated cultivar, or used to start a new population.

**Genetic issues** Like mass selection, half-sib selection is based on maternal plant selection without pollen control. Consequently, heritability estimates are reduced by 50%. Half-sib selection is hence less effective for changing traits with low heritability.



**Figure 17.5** Generalized steps in breeding by half-sib selection with a progeny test.

**Advantages and disadvantages** The major advantages and disadvantages of half-sib selection include the following.

#### *Advantages*

- 1 The procedure is rapid to conduct.
- 2 Progeny testing increases the success of selection, especially if quantitative gene action occurs or heritability is low.

#### *Disadvantages*

- 1 The trait of interest should have high heritability for success.
- 2 It is not readily applicable to species that cannot produce enough seed per plant to conduct a yield trial.
- 3 Lack of pollen control reduces heritability by half.

**Modifications** The modifications or alternative procedures for recurrent half-sib selection differ by: (i) the testers used, (ii) the selection of one or both parents; and (iii) the seed used for intercrossing. Some procedures use a population as a tester and half-sib seed for intercrossing. Selfed seed or clones may be used for intercrossing.

Where genetic male sterility is incorporated in the scheme, male-sterile plants may be tagged in the field at the time of flowering. After open-pollination, each head is harvested and threshed separately. Seed from one head (family) is evaluated as one entry in a yield trial, remnant seed is saved, superior entries are identified, and the remnant seed is bulked and planted for recombination to occur. Male-sterile plants are tagged again and harvested individually to form the next cycle of evaluation.

#### *Half-sib selection with a testcross*

Another way of evaluating genotypes to be composited is by conducting a testcross.

**Key features** This variation of half-sib selection allows the breeder to more precisely evaluate the genotype of the selected plant by choosing the most suitable testcross parent. The half-sib lines to be composited are selected based on a testcross evaluation not progeny performance. The tester may be inbred, in which case all the progeny lines will have a common parental gamete.

**Applications** Like half-sib selection with a progeny test, this procedure is applicable to cross-pollinated species in which sufficient seed can be produced by

crossing to grow a replicated testcross progeny trial. However, in procedures in which self-pollination is required, the method cannot be applied to species with self-incompatibility.

**Procedure** In season 1, the breeder selects 50–100 plants from the source population. A tester parent is pollinated with pollen from each of the selected plants (Figure 17.6). The crossed seed from the tester as well as the open-pollinated selected plants are harvested separately. In season 2, the testcross progenies are grown in replicated plots. In season 3, an equal amount of open-pollinated seed from 5–10 superior plants is composited and grown in isolation for open-pollination to occur.

**Advantages and disadvantages** The major advantages and disadvantages of half-sib selection with a testcross include the following.

#### *Advantages*

- 1 Control over the testcross parents permits a more precise evaluation of the genotype of the selected plant than would be obtained by open-pollination, as in half-sib selection with progeny test.
- 2 It is rapid to conduct.

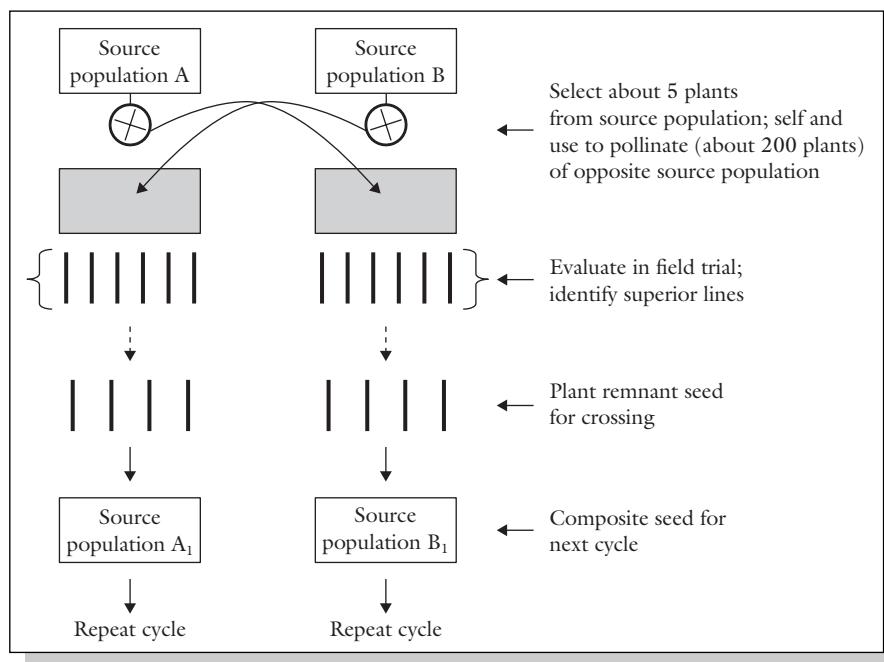
#### *Disadvantages*

- 1 This method of breeding is applicable to species that can produce sufficient seed by crossing for replicated testcross progeny trials.
- 2 When self-pollination is used, the method is applicable to species without self-incompatibility issues.

**Modifications** Pollen from each selected plant may be used to pollinate a tester plant and self-pollinate the selected plant. Also, in season 3, equal quantities of selfed seed may be composited and planted in isolation.

#### *Interpopulation improvement methods*

The purpose of this group of recurrent selection schemes is to improve the performance of a cross between two populations. To achieve this, interpopulation heterosis is exploited. The procedures are appropriate when the breeder's goal is hybrid production (i.e., the final product or cultivar is a hybrid). Developed by P. E. Comstock and his colleagues, the procedures allow the breeder to improve two genetically different populations for GCA and SCA, thereby improving their crossbred mean.



**Figure 17.6** Generalized steps in breeding by half-sib selection with a testcross.



## Industry highlights

*Developing a new cool-season perennial grass forage: interspecific hybrids of *Poa arachnifera* × *Poa secunda**

Bryan Kindiger  
USDA-ARS Grazinglands Research Laboratory, El Reno, OK 73036, USA

**Introduction**

An increasing awareness of the importance of sustainable grazing lands and the maintenance and enhancement of grasslands has led to a greater interest in the genetic improvement of native species. The *Poa* genus contains examples of productive cool-season perennial species that provide both turf and forage types. The genus is a variable one of about 300 species worldwide, of which several are indigenous to the United States (Huff 1992). The wide variation within the genus, along with its predominantly apomictic reproductive behavior, has made *Poa* a difficult genus with which to work. *P. arachnifera* Torr. (Texas bluegrass) is a vigorous, rhizomatous, sod-forming perennial, native to the southern plains of the USA (Figure 1a) (Silveus 1933). The species is dioecious, with male and female plants. To date, all of our native collections of *P. arachnifera* are completely sexual in their form of reproduction. However, due to the excessive amount of cobwebby hairs located at the base of the lemma, seed harvest and sowing of this species is difficult. *P. secunda* J. Presl. (big bluegrass) is also a native bluegrass (Figure 1b). It is an important component of the sagebrush grassland vegetation of the western United States (Silveus 1933). It is distributed from southeastern Alaska through the west and the Great Plains region of the USA. It is a long-lived perennial with a fibrous, bunch-type root system, which can occasionally develop short rhizomes. *P. secunda* is also a facultative apomictic species exhibiting aposporic embryo sac development (Kellogg 1987). The apomictic character of *P. secunda* is a second favorable characteristic that can be contributed to hybrids and would allow the long-term maintenance of any hybrid vigor resulting from the hybridization of these two species. Presumably, an extensive amount of genetic diversity exists in both *P. arachnifera* and *P. secunda* (Larson et al. 2001). By utilizing the dioecious *P. arachnifera* as the female parent in interspecific hybridizations, large numbers of interspecific hybrids can be rapidly generated. Such hybridization might be useful in developing diverse, native germplasm that could be widely adapted and biologically appropriate to the central and southern plains of the USA.



**Figure 1** Examples of (a) *Poa arachnifera* and (b) *P. secunda* individuals.

#### *Development of hybrids*

A mature female *P. arachnifera* individual ( $2n = 56$ ) collected near El Reno, Canadian County, Oklahoma, was pollinated by a mature clone of *P. secunda* var. Sherman ( $2n = 63$ ) (PI578850). Pollinations were achieved by placing the accessions adjacent to each other and agitating the *P. secunda* pollen parent between 7.30 and 10.00 am and allowing the pollen to fertilize the *P. arachnifera* females. One advantage in utilizing *P. arachnifera* in a *Poa* interspecific hybrid program relates to its dioecious nature. Crosses are greatly simplified because the typical necessity for emasculation of pollen-setting structures is not a necessary component of this type of hybridization scheme.

Seed heads were harvested about 25 days after pollination and were air dried for an additional 2 weeks. Due to the cobwebby nature of the *P. arachnifera* seed heads, the inflorescences were broken by rubbing the seed heads between two hands. The seed heads were sown in pots on top of a light potting soil mixture, then lightly covered with the same potting soil mixture. The pots were watered, covered, and placed in a germination chamber. As seedlings emerged, they were transplanted to separate pots. Seedlings were transplanted to the nursery and allowed to grow through the spring and winter. A vernalization period was required to initiate flowering.

Pollen viability in the hybrids was estimated by examining the pollen grains obtained from freshly extruded anthers with a 40 $\times$  hand lens. Pollen grains exhibiting a pearly white phenotype were considered viable. Pollen grains exhibiting a golden-yellow, crystalline phenotype were considered non-viable. Seed production was not estimated; however, the ability of individual plants to generate viable seed was assessed by examining the florets and looking for the presence of seed. If plump seeds were observed, additional seed heads from that individual were threshed by hand and any harvested seed was sown by the previously described method.

### Experimental results

#### Confirmation of hybrids with molecular markers

Following the generation of any wide cross hybrid, especially in polyploid species that have relatively unknown or poorly studied reproductive processes, it is worthwhile confirming the generation of true hybrids. Unique, but naturally occurring, events such as patrogenesis or androgenesis could give rise to unanticipated products. To identify hybrids, random amplified polymorphic DNA (RAPD) markers (Williams et al. 1990) or some other polymerase chain reaction (PCR)-based approach can be efficiently utilized (Weising et al. 1995). For this purpose, a series of *P. arachnifera* and *P. secunda* PCR-RAPD-based primers were utilized for evaluating genotypes. Standard DNA extraction and PCR-RAPD methods were utilized to evaluate the progeny of the hybrids.

#### Hybrid phenotypes

Of the 54 hybrids generated, 50 were true hybrids and four were androgenic haploids. Thirty-five flowered the first growing season and all flowered in their second season. A preliminary assessment of their phenotypic variability indicated that  $F_1$  individuals could possess glossy to waxy leaves, wide to narrow leaves, an erect or prostrate plant profile, variation in flowering time, differences in the shape and size of inflorescence, and variation in pollen viability. Twenty-four of the true  $F_1$  hybrids expressed a monoecious phenotype while 26 exhibited a dioecious phenotype. Of those exhibiting a dioecious phenotype, 12 were female and 14 were male. All androgenic haploids were monoecious in their phenotype. Chromosome counts of randomly selected hybrid individuals indicated that chromosome numbers ranged from  $2n = 54$  through to  $2n = 72$ . Chromosome counts of the androgenic haploids were: two with 16 chromosomes, one with 18 chromosomes, and one with 20. PCR analysis confirmed that the androgenic haploids did not possess any of the *P. arachnifera* genome and only a subset of the *P. secunda* genome and each possessed a fibrous, bunchgrass-type root system characteristic of the *P. secunda* parent. Each of the androgenic haploids was a smaller version of the monoecious *P. secunda* parent. The generation of haploids in these and additional interspecific hybrids of *Poa* spp. is not without precedent and has been reported elsewhere (Kiellander 1942; Kindiger 2004).

In the monoecious hybrid individuals, seed heads more closely resembled the *P. secunda* parent (Figure 2). In the dioecious individuals, most exhibited a phenotype resembling the *P. arachnifera* parent; however glume size varied considerably among the dioecious population. Stereoscopic examination of the florets of monoecious individuals indicated that the hybrids possessed fewer cobwebby hairs than the *P. arachnifera* parent or the female individuals of the dioecious group.

Many of the taxonomic characters commonly utilized to differentiate between the parents were clearly expressed in the hybrids. The segregation of monoecy and male or female dioecy was relatively equivalent among the  $F_1$  hybrids. All the  $F_1$



**Figure 2** Inflorescences from (a) *Poa arachnifera*, (b) *P. secunda*, and (c) a monoecious  $F_1$  interspecific hybrid.



**Figure 3** Range of seed development following pollination of *Poa arachnifera* × *P. secunda*. Not all the hybrid seeds are viable and they may require seed culture in nutritional media to generate plants.

nearby *P. arachnifera* males. For the dioecious females, self-pollination was found to vary among the hybrids. Upon examination of individuals representing both monoecious and dioecious hybrids, pollen fertility as estimated by I<sub>2</sub>KI staining ranged from 0% to as high as 90%. Field observations estimated pollen viability in the range of 0% to 80%. General comparisons of laboratory and field methods of estimating pollen viability in individual plants were within 5–10% of each other.

#### Pollen fertility

Pollen fertility – as judged by pollen grain size, the occurrence of starch-filled pollen grains, and the presence of an operculum – was found to vary among the hybrids. Upon examination of individuals representing both monoecious and dioecious hybrids, pollen fertility as estimated by I<sub>2</sub>KI staining ranged from 0% to as high as 90%. Field observations estimated pollen viability in the range of 0% to 80%. General comparisons of laboratory and field methods of estimating pollen viability in individual plants were within 5–10% of each other.

To determine if the seed had been generated by either a sexual or an apomictic mode of reproduction, seed was sown and 10 seedlings were submitted to PCR genotyping utilizing a series of informative PCR-based markers. If following electrophoresis of the PCR amplification products, band variation was observed among a set of siblings, sexual reproduction was indicated. If the band profile was uniform across siblings and identical to the maternal parent, an apomictic form of reproduction was implicated. Since no variation was observed, the analysis indicated that the seedlings were derived by an apomictic form of reproduction. As a consequence, the parent from which these seedlings were derived likely produced its offspring via an apomictic form of reproduction. Since *P. secunda* exhibits aposporic pseudogamous embryo sac development, this seed-fertile hybrid also likely expresses this form of apomixis.

#### Conclusion

Though more detailed cytogenetic studies are necessary, the observed 10–90% pollen fertility scored in a random sampling of the interspecific hybrids suggests that partial or high levels of genome affinity can exist between the *P. arachnifera* and *P. secunda* genomes. The combining of these two genomes resulted in both pollen-fertile and seed-fertile F<sub>1</sub> hybrids, of which at least one exhibited an apomictic form of reproduction. In addition, any heterotic response exhibited in the hybrids could potentially be maintained via apomixis. The transfer of apomixis from *P. secunda* to the hybrids indicates that sexuality is recessive to apomixis. Subsequent breeding of fertile interspecific hybrids and their utilization in a backcross program to *P. arachnifera* or *P. secunda*

hybrids possessed a rhizomatous root system, which is characteristic of the *P. arachnifera* parent, and suggests that genetic control of the fibrous, bunchgrass root system of *P. secunda* is recessive to the rhizomatous root system of *P. arachnifera*.

#### Seed development

Stereoscopic observations regarding the progression of seed maturity indicated a wide range in endosperm and mature seed development. In some instances, seed exhibited apparently normal endosperm development, while other seed exhibited complete abortion of the endosperm. Both the *P. arachnifera* and *P. secunda* parents are highly polyploid ( $2n = 56$  and  $2n = 63$ , respectively), which could cause their gametes, following microsporogenesis and megasporogenesis, to possess a wide range of chromosome numbers. Ploidy differences in interspecific crosses are known to account for endosperm developmental failure and such events often lead to embryo abortion (Bradshaw et al. 1995; Carputo 1997). It is assumed that the F<sub>1</sub> hybrids obtained in this study were obtained from seed exhibiting a non-abortive endosperm.

Evaluations regarding the seed production potential of the hybrids were determined by stereoscopic examination following pollination and seed maturation. All the hybrids were allowed to receive pollen through open-pollination by adjacent pollen-fertile sibs or

could be an effective approach toward the development of drought-tolerant forage or turf-type bluegrass. Continued hybridization, breeding, and selection of these hybrids may provide an indigenous, productive, and drought-tolerant cool-season perennial grass for pastures or rangelands.

### References

- Bradshaw, J.E., R.L. Wastie, H.E. Stewart, and G.R. Mackay. 1995. Breeding for resistance to late blight in Scotland. In: *Phytophthora infestans* 150 (Dowley, L.J., E. Bannon, L.R. Cooke, T. Keane, and E. O'sullivan, eds), pp. 246–254. EAPR Pathology Section Conference. Boole Press Ltd and Teagasc, Dublin, Ireland.
- Carputo, D. 1997. Ploidy and endosperm balance number (EBN) manipulations for germplasm introgression from 1EBN *Solanum commersonii* into 4EBN *S. tuberosum*. Ital. J. Agron. 1:123–128.
- Huff, D.R. 1992. Apomixis in *Poa*. In: Proceedings of the Apomixis Workshop, February 11–12, Atlanta, GA (Elgin, J., Jr., and J.P. Miksche, eds), pp. 20–25. USDA-ARS, Beltsville, MD.
- Kellogg, E.A. 1987. Apomixis in the *Poa secunda* complex. Am. J. Bot. 74:1431–1437.
- Kiellander, C.L. 1942. A subhaploid *Poa pratensis* L. with 18 chromosomes and its progeny. Svensk Botanisk Tidskr. 36:200–220.
- Kindiger, B. 2004. Generation of androgenic haploids from interspecific hybridization of *Poa arachnifera* × *Poa secunda*. Grassland Sci. 49:577–580.
- Larson, S.R., B. Waldron, S. Monson, L. St John, A.L. Palazzo, C. L. McCracken, and D. Harrison. 2001. AFLP variation in agamosperous and dioecious bluegrasses of western North America. Crop Sci. 41:1300–1305.
- Silveus, W.A. 1933. Texas grasses. Classification and description of grasses. Glegg Co., San Antonio, TX.
- Weising, K., H. Nybom, K. Wolff, and W. Meyer (eds). 1995. DNA fingerprinting in plants and fungi. CRC Press, Boca Raton, FL.
- Williams, J.G.K., A.R. Kubelik, K.J. Livbak, J.A. Rafalski, and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531–6535.

### Half-sib reciprocal recurrent selection

**Key features** The half-sib recurrent selection scheme involves the making of  $S_1$  plant testcrosses and evaluating them to identify and select superior progenies.

#### Procedure: cycle 0

- Season 1** Select and self-pollinate about 200–300  $S_1$  plants in each of two populations, A and B.
- Season 2** Grow 200–300 of the selected  $S_1$  progenies and produce half sibs of population A by crossing a number of plants with B as female, and vice versa. Self-pollinate the  $S_1$  plants used in making the testcrosses. Save  $S_1$  seed.
- Season 3** Evaluate about 100 half sibs in replicated trials. Select about 20 promising testcross families. This is done for both populations A and B.
- Season 4** Randomly mate the plants selected from  $S_1$  families within A and B to obtain new seed to initiate cycle 1.

**Procedure: cycle 1** Repeat cycle 0.

**Genetic issues** This scheme makes use of additive, dominance, and overdominance gene action. It is effec-

tive for selecting favorable epistatic gene combinations in the population. The change in the cross-bred mean may be calculated as follows:

$$\Delta G_{(A \times B)} = [i\sigma_{A(HSA)}^2]/4\sigma_{P(HSA)} + [i'\sigma_{A(HSB)}^2]/4\sigma_{P(HSB)}$$

where  $i$  and  $i'$  are the selection intensities in populations A and B, respectively; and  $\sigma_{A(HSA)}^2$  and  $\sigma_{A(HSB)}^2$  are the additive variances for populations A and B, respectively. Similarly,  $\sigma_{P(HSA)}$  and  $\sigma_{P(HSB)}$  are the phenotypic standard deviations among half sibs.

### Full-sib reciprocal recurrent selection

**Key features** Developed by Hallauer and Eberhart as modifications of the method by Comstock and colleagues, the full-sib method requires at least one of the populations to be prolific. The recombination units are half sibs (instead of  $S_1$  families). Developed for maize, full-sib families are produced by pairing plants from two populations, A and B. The top ear of a plant from population A is crossed with a plant from population B. The lower ear is selfed to be saved as remnant seed. The same is done for the reciprocal plant from population B, if they have two ears, otherwise they are selfed.

**Applications** The scheme has been used in crops such as maize and sunflower with reported genetic gains of 2.17% for the population *per se*, and 4.90% for the population hybrids.

#### Procedure: cycle 0

- Season 1** Plant population A as females (detassel) in an isolated block and population B as males in field 1. Plant population B as females and population A as males in field 2. The upper ears in each field are open-pollinated, while the lower ears are protected and pollinated manually. The result is that the upper ear is an interpopulation half-sib family, while the lower ear is an intrapopulation half-sib family.
- Season 2** Evaluate 100–200 A × B and B × A half sibs in replicated trials. Select the best half sibs from both sets of crosses.
- Season 3** Plant the remnant seed of the lower ears (selfed by hand pollination) that correspond to the best A × B half sibs in ear to rows as females (detassel). The males are the bulk remnant half-sib seed from population B corresponding to the best B × A crosses. They are randomly mated. The open-pollinated seed in populations A and B are harvested to initiate the next cycle.

**Genetic issues** An advantage of this method is that additive genetic variance of full-sib families is twice that of the half-sib families. The expected genetic gain is given by:

$$\Delta G = i\sigma_A^2 / 2\sigma_{PFS}$$

where  $\sigma_{PFS}$  = phenotypic standard deviation of the full-sib families.

#### Advantages

- 1 As compared to the half-sib method, one-half of the families are evaluated in each cycle because the evaluation of each full sib reflects the worth of two parental plants, one from each population.
- 2 Superior  $S_0 \times S_0$  crosses may be advanced in further generations and evaluated as  $S_1 \times S_1, S_2 \times S_2, \dots, S_n \times S_n$  to allow the breeder to simultaneously develop hybrids while improving the populations.

#### Optimizing gain from selection in population improvement

The goal of the breeder is to make systematic progress in the mean expression of the trait of interest from one

cycle to the next. Achieving progressive gains in yield depends on several factors.

- 1 **Genetic variance.** As previously indicated, it is critical to increase additive genetic variance per cycle. Additive genetic variance can be increased through increasing diversity in the entries used in population improvement.
- 2 **Selection intensity.** The rate of gain with selection is increased when selection intensity is increased. The number of individuals selected for recombination in each cycle should be limited to the best performers.
- 3 **Generations per cycle.** The breeder's choice of the breeding system to use in a breeding project is influenced by how rapidly each cycle of selection can be completed. When possible, using 2–3 generations per year can increase yield gains. Multiple generations per year is achieved by using off-season nurseries (winter nurseries), or planting in the dry season using irrigation.
- 4 **Field plot technique.** Breeders select in the field, often handling large numbers of plants. Heterogeneity in the field should be managed by using proper experimental designs to reduce random variation. Whenever possible, uniform fields should be selected for field evaluations. The cultural conditions should be optimized as much as possible (proper fertilization, irrigation, disease and pest control, weed control, etc.). This practice will reduce variation between replications. Other factors to consider are plot sizes, number of plants per plot, number of replications per trial, and number of locations. Implemented properly, these factors reduce random variations that complicate experimental results.

#### Development of synthetic cultivars

##### Synthetic cultivars versus germplasm composites

There are two basic types of open-pollinated populations of crops – those produced by population improvement, and synthetics. As previously discussed, population improvement methods can be categorized into two – those that depend on purely phenotypic selection (mass selection), and those that involve selection with progeny testing. A **synthetic cultivar** may be defined as an advanced generation of cross-fertilized (random mating in all combinations) seed mixture of parents that may be strains, clones, or hybrids. The parents are selected based on GCA. The primary distinction between the basic types of population mentioned in

this section is that population improvement cultivars can be propagated indefinitely as such. However, a synthetic cultivar is propagated for only a limited number of generations and then must be reconstituted from the parental stock. A synthetic population differs from a natural population by consisting of breeder-selected parental stocks.

**Germplasm composite** is a broad term used to refer to the mixing together of breeding materials on the basis of some agronomic trait (e.g., yield potential, maturity, disease resistance), followed by random mating. There are many ways to put a composite together. Germplasm composites are by nature genetically broad based and very complex. They can be used as for commercial cultivation over a broad range of agroecological environments. However, they can also be used as reservoirs of useful genes for use in breeding programs.

### Desirable features of a synthetic cultivar

K. J. Frey summarized three major desirable features of synthetic cultivars:

- 1 Yield reduction in advanced generations is less than with a single or double cross. For example, in maize, an estimated 15–30% reduction occurs between  $F_1$  and  $F_2$ , as compared to a reduction of only 5–15% from *syn-1* to *syn-2*. This slow rate of reduction in yield makes it unnecessary for producers to obtain new seed of the cultivar for planting in each season.
- 2 A synthetic cultivar may become better adapted to the local production environment over time, as it is produced in successive generations in the region.
- 3 A synthetic cultivar is genetically heterogeneous, a population structure that makes it perform stably over changing environmental conditions. Further, because of this heterogeneity, both natural and artificial selection can modify the genotypic structure of synthetic cultivars. That is, a breeder may achieve gain in performance by practicing selection in *syn-2* and subsequent generations.

### Key features

There are three primary steps in the development of a synthetic:

- 1 Assembly of parents.
- 2 Assessment of general combining ability.
- 3 Random mating to produce synthetic cultivars.

The parents used in synthetics may be clones (e.g., forage species) or inbred lines (e.g., corn, sugar beet).

Whereas forages can be increased indefinitely by clonal propagation, inbred lines are needed to perpetuate the genotypes used in hybrid production. The parental materials are reproducible and may be substituted with new genotypes as they become available, for some improvement in the synthetic cultivar. The parents are selected after progeny testing or GCA analysis using a testcross or topcross, but most frequently a polycross, for evaluation.

### Test for general combining ability

**Polycross test** A polycross test is generally preferred because it is simple and convenient to conduct and also, by nature, provides an efficient estimate of GCA, a desired attribute in synthetic production. Furthermore, it allows an adequate amount of seed to be obtained for more comprehensive testing using commercial standards. It provides a greater insurance to cultivars against genetic shifts that could arise during seed increase. However, any significant amount of selfing (especially if unequal among the component parents) or non-random cross-pollination could result in bias. The component clones may vary in self-fertility and other biological characteristics that impact fertilization. To minimize such deviations from a perfect polycross, the Latin square design (see Chapter 23) may be used to establish the polycross nursery. In theory, the polycross allows each clone in the nursery to be pollinated by the same pollen sources as a result of random pollination from all the entries in the same plot.

**Topcross test** Selected clones are grown in alternative rows with an open-pollinated cultivar as tester. The testcross seed includes both selfs and intercrosses among the clones being evaluated.

**Diallel cross test** A diallel cross entails achieving all possible single crosses, involving all the parents. This is laborious to conduct. It requires that each parent be grown in isolation. It provides information on both GCA and SCA.

### Applications

The synthetic method of breeding is suitable for improving cross-fertilized crops. It is widely used to breed forage species. Successful synthetic cultivars have been bred for corn, sugar beet, and other species. The suitability of forage species for this method of breeding

stems from several biological factors. Forages have perfect flowers, making it difficult to produce hybrid seed for commercial use. The use of male sterility may facilitate controlled cross-pollination, which is difficult to achieve in most forage species. In order to test individual plants for use in producing commercial seed, it is essential to obtain sufficient seed from these plants. The amount of seed obtained from single plants of these species is often inadequate for a progeny test. Furthermore, forage species often exhibit self-incompatibility, a condition that inhibits the production of selfed seed. Synthetic cultivars are also used as gene pools in breeding progeny. Synthetic cultivars are advantageous in agricultural production systems where farmers routinely save seed for planting. One of the well known and widely used synthetics is the Iowa stiff-stalk synthetic (BSSS) of maize.

### Procedure

A procedure for crops in which selections are clonally propagable is as follows.

**Year 1 The source nursery.** The source population consists of clones. The source nursery is established by planting several thousands (5,000–10,000) of plants assembled from many sources to provide a broad genetic base of the clonal lines for selection. The germplasm in the nursery is screened and evaluated to identify superior individuals according to the breeding objectives.

**Year 2 Clonal lines.** The breeder first selects 100–200 superior plants on a phenotypic basis to multiply clonally to produce clonal lines. A clonal line nursery is established, each line consisting of about 20–25 plants derived from the same parental line. The breeder may impose various biotic and abiotic selective pressures (e.g., drought, specific disease epidemic, severe clipping) to aid in identifying about the 25–50 most desirable clones.

**Year 3 Polycross nursery.** The selected clonal lines are planted in a polycross nursery to generate seed for progeny testing. Ideally, the layout of the polycross in the field should allow each clone to be pollinated by a random sample of pollen from all the other entries. A method of layout to achieve this objective is a square plot (e.g., 12 × 12) in which every clone occurs once in every row. Covering with a fine mesh tent or separating the plots by an adequate distance isolates

each square plot. The mesh is removed once the pollination period is over. A large number of replications (10 or more) of the single randomized clones should be used to achieve a highly mixed pollination. Seed from each clone is harvested separately. The polycross test is valid if the layout ensures random interpollination. Alternative methods of evaluating clones for quantitatively inherited traits are available. Self-fertilization may be used but it often yields only a small amount of seed. A diallel cross is cumbersome to conduct, especially for large entries. A topcross evaluates SCA. The polycross is used because it evaluates GCA.

**Year 4 Polycross progeny test.** Seed is harvested from the replicated clones and bulked for planting progeny rows for performance evaluation. The progeny test evaluates yield and other traits, according to the breeding objective. The top performing 5–10 clones are selected for inclusion in the synthetic cultivar.

**Year 5 *Syn-0* generation.** The selected clones are vegetatively propagated and randomly transplanted into an isolated field for cross-fertilization to produce *syn-0* seed. Leguminous species may be isolated in an insect-proof cage and cross-fertilized by using insects.

**Year 6 *Syn-1* generation.** The *syn-0* seed is increased by planting in isolation. Equal amounts of seed are obtained from each parent and mixed to ensure random mating in the field. Bulk seed is harvested from seed increased in the *syn-1* generation, which may be released as a commercial cultivar provided sufficient seed is produced.

**Year 7 Subsequent *syn* generations.** Frequently, the *syn-1* seed is not sufficient to release to farmers. Consequently, a more practical synthetic breeding scheme is to produce a *syn-2* generation by open-pollinated increase of seed from *syn-1*. The *syn-2* seed may be likened to a breeder seed. It is further increased to produce *syn-3* (foundation seed) and *syn-4* (certified seed). Commercial seed classes are discussed in detail in Chapter 24. The pattern of loss in vigor, progressively with the advancement of generations from *syn-1*, *syn-2*, to *syn-n*, is similar to that which occurs when hybrids are progressively selfed from  $F_1$ ,  $F_2$ , to  $F_n$  generations. It is important to maintain the original clones so that the synthetic can be reconstituted as needed. The steps described are only generalized and can be adapted and modified according to the species and the objectives of the breeder.

### Genetic issues

The highest yield performance is obtained in the *syn-1* generation, hybrid vigor declining with subsequent generations. It is generally estimated that a synthetic forage cultivar of cross-fertilized diploid or polyploid species will experience a maximum yield decline of 10–12% from the *syn-1* to *syn-2* generation, as previously stated. The yield decline is less in subsequent generations. Sewall Wright proposed a formula to predict the  $F_2$  yield of a group of inbred lines:

$$F_2 = F_1 - [(F_1 - P)/n]$$

where  $F_2$  = expected performance of the  $F_2$ ,  $F_1$  = mean  $F_1$  hybrid performance from combinations of inbred lines,  $P$  = average performance of inbred lines, and  $n$  = number of inbred lines. That is, one can increase the  $F_2$  yield by increasing the average  $F_1$  yield, increasing the yield of parental lines, or increasing the number of lines used to create the synthetic. This formula assumes that the species has diploid reproduction and that the parents are inbred. Hence, even though shown to be accurate for maize, it is not applicable to polyploid species and those that are obligate outcrossers.

The formula may also be written as:

$$syn-2 = syn-1 - [(syn-1 - syn-0)/n]$$

Studies involving inbred lines and diploid species have indicated that as the number of parental lines increase, the  $F_1$  performance increases. Parental lines with high combining ability will have a high  $F_1$  performance. In practice, it is a difficult task to find a large number of parents with very high combining ability. Furthermore, predicting yield performance of synthetic cultivars of cross-fertilized diploid and polyploid forage species is more complicated than is described by their relationship in the equation. Given a set of  $n$  inbred lines, the total number of synthetics,  $N$ , of size ranging from 2 to  $n$  is given by:

$$N = 2n - n - 1$$

As inbred lines increase, the number of possible synthetics increases rapidly, making it impractical to synthesize and evaluate all the possible synthetic cultivars.

The theoretical optimum number of parents to include in a synthetic is believed to be about 4–6. However, many breeders favoring yield stability over yield ability tend to use large numbers of parents

ranging from about 10 to 100 or more. Large numbers are especially advantageous when selecting for traits with low heritability.

Synthetics of autotetraploid species (e.g., alfalfa) are known to experience severe and widespread decline in vigor between *syn-1* and *syn-2*, which has been partly attributed to a reduction in triallelic and tetrallelic loci. Higher numbers of tetrallelic loci have been shown to be associated with higher agronomic performance (e.g., forage yield, seed yield, height) of alfalfa. The number of selfed generations is limited to one. Selfed seed from selected  $S_0$  plants are intermated to produce the synthetic population. The rationale is that  $S_0$  plants with high combining ability should contain many favorable genes and gene combinations. Selecting specific individuals from the segregating population to self could jeopardize these desirable combinations.

Additive gene action is considered more important than dominance genotypic variance for optimum performance of synthetic cultivars. In autotetraploids where intralocus and allelic interactions occur, high performing synthetic cultivars should include parents that have a high capacity to transfer their desirable performance to their offspring. Such high additive gene action coupled with a high capacity for intralocus or allelic interactions will likely result in higher performing synthetics.

Synthetic cultivars exploit the benefits of both heterozygosity and heterosis. J. W. Dudley demonstrated that yield was a function of heterozygosity by observing that, in alfalfa crosses, the  $F_1$  yields reduced as generations advanced. Further, he observed that allele distribution among parents used in a cross impacted heterozygosity. For example, a cross of duplex  $\times$  nulliplex ( $A_2 \times A_0$ ) always had a higher degree of heterozygosity than, say, a cross of simplex  $\times$  simplex ( $A_1 \times A_1$ ), regardless of the clonal generation used to make the cross.

Natural selection changes the genotypic composition of synthetics. The effect can have significant consequences when the cultivar is developed in one environment and used for production in a distinctly different environment. There can be noticeable shifts in physiological adaptation (e.g., winter hardiness) as well as morphological traits (e.g., plant height). For example, when growing alfalfa seed in the western US states (e.g., California) for use in the midwest, the cultivars may lose some degree of winter hardiness, a trait that is desired in the production region of the midwest. A way to reduce this adverse impact is to grow seed crop in the west using foundation seed from the midwest.

### Factors affecting the performance of synthetic cultivars

Three factors are key in determining the performance of a synthetic cultivar.

- 1 Number of parental lines used.** Synthetic cultivars are maintained by open-pollination. Consequently, the  $F_2$  (*syn-2*) yield should be high to make it a successful cultivar. Hardy-Weinberg equilibrium is reached in *syn-2* for each individual locus and hence should remain unchanged in subsequent generations. It follows then that the  $F_3$  (*syn-3*) should produce as well as *syn-2*. Some researchers have even shown that  $F_3$  and  $F_4$  generations yield as much or slightly better than  $F_2$  generations, provided the number of lines included in the synthetic cultivar is not small. With  $n = 2$ , the reduction in performance will be equal to 50% of the heterosis. Consequently,  $n$  has to be increased to an optimum number without sacrificing high GCA. When  $n$  is small, the yields of *syn-1* are high, but so is the decline in *syn-2* yields. On the other hand increasing  $n$  decreases *syn-1* and *syn-2* yields. A balance needs to be struck between the two effects. Some researchers estimate the optimum number of lines to include in a synthetic to be five or six.
- 2 Mean performance of the parental lines (in *syn-0*).** The lines used in synthetics should have high performance. A high value of parental lines reduces the reduction in performance of *syn-2* over *syn-1*. Preferably, the parents should be non-inbreds or have minimum inbreeding (e.g.,  $S_0$  or  $S_1$ ).
- 3 Mean *syn-1*.** In theory, the highest value of *syn-1* is produced by a single cross. However, alone, it will suffer from a higher reduction in performance. It is important for the mean  $F_1$  (*syn-1*) yield of all the component  $F_1$  crosses is high enough such that the *syn-2* yield remains high in spite of some decline.

### Advantages and disadvantages

The major advantages and disadvantages of the synthetic cultivar method include the following.

### Advantages

- 1 The method is relatively easy to implement.
- 2 It can be used to produce variability for hybrid breeding programs.
- 3 Advanced genotypes of synthetics show little yield reduction from *syn-1*, making it possible for farmers to save and use seed from the current season to plant in the next season.

### Disadvantages

- 1 Because inadequate seed is often produced in *syn-1*, the method fails to exploit to the maximum the effects of heterosis, as is the case in conventional  $F_1$  hybrid breeding. The method of synthetics is hence a compromise to the conventional means of exploiting heterosis.
- 2 Natural selection changes the genotypic composition of synthetics, which may be undesirable.

### Backcross breeding

The key concern in the application of the backcross technique to cross-pollinated species is the issue of inbreeding. Selfing cross-pollinated species leads to inbreeding depression (rapid loss of vigor). The use of a recurrent parent in backcrossing with cross-pollinated species is tantamount to inbreeding. To minimize the loss of vigor, large populations should be used to enable the breeder to sample and maintain the diversity of the cultivar and to ensure against the harmful effects of inbreeding.

Just like in self-pollinated species, it is relatively straightforward to improve a qualitative trait conditioned by a single dominant gene. The breeder simply selects and advances individuals expressing the trait. Where a recessive gene is being transferred, each backcross should be followed by one round of intercrossing to identify the recessive phenotype.

Improving inbred lines (to be used as parents in crosses) is equivalent to improving self-pollinated species. The key to success is for the breeder to maintain a broad gene base by using an adequate number (2–3) of backcrosses and a large segregating population.

## References and suggested reading

- Agrawal, R.L. 1998. Fundamentals of plant breeding and hybrid seed production. Science Publishers, Inc., Enfield, NH.
- Avey, D.P., H.W. Ohm, F.L. Patterson, and W.E. Nyquist. 1982. Three cycles of simple recurrent selection for early heading in winter wheat. *Crop Sci.* 15:908–912.
- Burton, J.W., and C.A. Brim. 1981. Recurrent selection in soybeans III: Selection for increased percent oil in seeds. *Crop Sci.* 21:31–34.
- Chahal, G.S., and S.S. Gosal. 2000. Principles and procedures of plant breeding: Biotechnological and conventional approaches. CRC Press, New York.
- Dudley, J.W. 1964. A genetic evaluation of methods of utilizing heterozygosis and dominance in autotetraploids. *Crop Sci.* 4:410–413.
- Dudley, J.W. (ed.). 1974. Seventy generations of selection for oil and protein in maize. Crop Science Society of America, Madison, WI.
- Dudley, J.W., and R.J. Lambert. 1992. Ninety generations of selection for oil and protein in maize. *Maydica* 37:81–87.
- Fehr, W.R. 1987. Principles of cultivar development. McMillan, New York.
- Feng, L., J.W. Burton, T.E. Carter Jr., and V.R. Pantalone. 2004. Recurrent half-sib selection with testcross evaluation for increased oil content in soybean. *Crop Sci.* 44:63–69.
- Hallauer, A.R. 1985. Compendium of recurrent selection methods and their applications. *Crit. Rev. Plant Sci.* 3:1–33.
- Hayward, M.D., N.O. Bosemark, and I. Romagosa (eds). 1993. Plant breeding: Principles and prospects. Chapman & Hall, London.
- Narro, L., S. Pandey, J. Crossa, C. DeLeon, and F. Salazar. 2003. Using line × tester interaction for the formation of yellow maize synthetics tolerant to acid soils. *Crop Sci.* 43:1718–1728.
- Poehlman, J.M., and D.H. Slepper. 1995. Breeding field crops. Iowa State University Press, Ames, IA.
- Sumarno, T., and W.R. Fehr. 1982. Response to recurrent selection for yield in soybean. *Crop Sci.* 22:295–299.
- Vogel, K.P., and J.F. Pedersen. 1993. Breeding systems for cross-pollinated perennial grasses. *Plant Breed. Rev.* 11:251–274.
- Wilcox, J.R., and I.F. Cavins. 1995. Backcrossing high seed protein to a soybean cultivar. *Crop Sci.* 35:1036–1041.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 The synthetic method of cultivar breeding is suitable for cross-pollinated species.
- 2 The highest yield performance is obtained in the *syn-1* generation.
- 3 A polycross evaluates specific combining ability.
- 4 The theoretical optimum of parents to include in a synthetic is about 10–15.
- 5 Additive gene action is considered more important than dominance genotypic variance for the optimum performance of synthetic cultivars.
- 6 The success of a recurrent selection program depends on the genetic nature of the base population.
- 7 The ear-to-row breeding in corn is equivalent to full-sib selection.
- 8 Combining ability is important in the half-sib selection method of breeding.
- 9 Half-sib selection is based on maternal plant selection.
- 10 In full-sib selection, both parents in a cross are known.

### Part B

Please answer the following questions:

- 1 Why is it more practical to release *syn-2* rather than *syn-1* seed to farmers?
- 2 Give a specific disadvantage of synthetic breeding.
- 3 Why is the Latin square design a preferred layout of a polycross nursery?
- 4 Define a synthetic cultivar.
- 5 Mass selection as applied to cross-pollinated species is also called .....
- 6 In recurrent selection the base population is designated cycle .....

- 7 Give a specific disadvantage of mass selection as applied to cross-pollinated species.
- 8 What is the technique of gridding in mass selection, and what is its effect on the selection process in breeding?

**Part C**

Please write a brief essay on each of the following topics:

- 1 Discuss the key features of a polycross nursery.
- 2 State and discuss the implications of Sewall Wright's proposed formula for predicting the  $F_2$  yield of a group of inbred lines.
- 3 Discuss the applications of synthetic breeding.
- 4 Compare and contrast the methods of half-sib selection with progeny test and half-sib selection with a testcross.
- 5 Discuss the issue of heritability in recurrent selection.
- 6 Describe the method of half-sib selection with a testcross.
- 7 Discuss the concept of recurrent selection.



# 18

# Breeding hybrid cultivars

---

## Purpose and expected outcomes

*The methods of breeding discussed so far that are preceded by crossing go beyond the  $F_1$  population. As previously indicated, the  $F_2$  is the most variable population following a cross. Consequently, selection often starts in this population. On the contrary, hybrid breeding ends with the  $F_1$ . The purpose of this chapter is to discuss the rationale of a hybrid and the genetics underlying the development of this type of cultivar. After studying this chapter, the student should be able to:*

- 1 Discuss the historical background of hybrid seed development.
  - 2 Discuss the concept of hybrid vigor and its role in hybrid development.
  - 3 Discuss the genetic basis of hybrid seed development.
  - 4 Present the steps in the procedure of hybrid breeding.
- 

### What is a hybrid cultivar?

A **hybrid cultivar**, by definition, is the  $F_1$  offspring of a planned cross between inbred lines, cultivars, clones, or populations. Depending on the breeding approach, the hybrid may comprise two or more parents. A critical requirement of hybrid production is that the parents be unidentical. As will be discussed next, it is this divergence that gives hybrids their superior performance. The outstanding yields of certain modern crops, notably corn, owe their success to the exploitation of the phenomenon of **heterosis** (hybrid vigor), which is high when parents are divergent. Much of what we know about hybrid breeding came from the discoveries and experiences of scientists engaged in corn hybrid cultivar development. However, commercial hybrids are now available for many crops, including self-pollinating species.

### Brief historical perspective

One of the earliest records on hybridization dates back to 1716 when American Cotton Mather observed the

effects of cross-fertilization in maize, attributing the multicolored kernels to wind-borne intermixture of different colored cultivars. However, it was the German T. G. Koelreuter who conducted the first systematic studies on plant hybridization in 1766. Even though previous observations had been made to the effect that offspring of crosses tended to exhibit superior performance over the parents, it was G. H. Shull who in 1909 first made clear scientific-based proposals for exploiting heterosis to produce uniform and high-yielding cultivars. Unfortunately, the idea was at that time impractical and potentially expensive to commercially exploit. In 1918, D. F. Jones proposed a more practical and cost-effective approach to producing hybrid cultivars by the method of the double cross. Double-cross hybrids produced significantly more economic yield than the single-cross hybrids originally proposed by Shull. Single-cross hybrid seed was then produced on weak and unproductive inbred parents, whereas double-cross seed was produced on vigorous and productive single-cross plants. The corn production industry was transformed by hybrids, starting in the 1930s.

Other notable advances in the breeding of hybrids were made by researchers including M. I. Jenkins in 1934 who devised a method (topcross performance) to evaluate the effectiveness of parents in a cross (i.e., combining ability). Through this screening process, breeders were able to select a few lines that were good combiners (productive in a cross) for use in hybrid breeding.

The next significant impact on hybrid production also came in the area of techniques of crossing. Because corn is outcrossed and bisexual, it is necessary to emasculate one of the parents (i.e., make one female) as part of the breeding process. In the early years of corn hybrid breeding, emasculation was accomplished by the labor-intensive method of mechanical detasseling (removal of the tassel). The discovery and application of **cytoplasmic male sterility (CMS)** to corn hybrid programs eliminated the need for emasculation by the late 1960s. Unfortunately, the success of CMS was derailed when the **Texas cytoplasm (T-cytoplasm)**, which was discovered in 1938 and was at that time the dominant form of male sterility used in corn breeding, succumbed to the southern leaf blight epidemic of 1970 and devastated the corn industry. It should be mentioned that mechanized detasslers (rather than CMS) are used by some major seed companies in hybrid seed production of corn today.

Realizing that the limited number of inbred lines used in hybrid programs did not embody the complete genetic potential of the source population, and with the need to develop new inbred lines, scientists embarked on cyclical recombination (by recurrent selection) to generate new variability and to improve parental lines. Breeders were able to develop outstanding inbred lines to make single-cross hybrids economical enough to replace double-cross hybrids by the 1960s. By this time, corn hybrid production programs had developed a set of standard practices consisting of the following, as observed by N. W. Simmonds:

- 1** Maintenance and improvement of source population by open-pollinated methods (recurrent selection).
- 2** Isolation of new inbreds and improvement of old ones (by backcrossing).
- 3** Successive improvement of single-cross hybrids by parental improvement.
- 4** CMS-based seed production.

The application of hybrid methodology in breeding has socioeconomic implications. The commercial seed industry has rights to its inventions, which generate royalties. More importantly, because heterosis is maxi-

mized in the F<sub>1</sub>, farmers are generally prohibited from saving seed from the current season's crop to plant the next year's crop. They must purchase seed from the seed suppliers each season. Unfortunately, poor producers in developing countries cannot afford annual seed purchase. Consequently, local and international (e.g., international agricultural research centers such as the Centro Internationale de Mejoramiento de Maiz y Trigo (CIMMYT) in Mexico) efforts continue to be largely devoted to producing propagable improved open-pollinated cultivars for developing countries.

The idea of commercializing hybrid seed production is traced to Henry A. Wallace, an Iowa farmer, who studied self-pollination and selfing of corn in 1913. His industry led to the founding of the Pioneer Hi-Bred Corn Company in Iowa, in 1925. In 1933, only about 0.1% of US corn production was devoted to hybrid seed. Today, hybrid seed is planted on almost all corn fields. Hybrids are also gradually being embraced by developing countries (see Chapter 26).

### Concepts of hybrid vigor and inbreeding depression

As previously stated, the hybrid production industry thrives on the phenomenon of hybrid vigor.

#### Hybrid vigor

**Hybrid vigor** may be defined as the increase in size, vigor, fertility, and overall productivity of a hybrid plant, over the midparent value (average performance of the two parents). It is calculated as the difference between the crossbred and inbred means.

$$\text{Hybrid vigor} = \{[F_1 - (P_1 + P_2)/2] / [(P_1 + P_2)/2]\}$$

The estimate is usually calculated as a percentage (i.e.,  $\times 100$ ).

The synonymous term, **heterosis**, was coined by G. H. Shull. It should be pointed out immediately that, as it stands, heterosis is of no value to the breeder (and hence farmer) if a hybrid will only exceed the midparent in performance. Such advantageous hybrid vigor is observed more frequently when breeders cross parents that are genetically diverse. The *practical* definition of heterosis is hybrid vigor that greatly exceeds the better or higher parent in a cross. Heterosis occurs when two inbred lines of outbred species are crossed, as much as when crosses are made between pure lines of inbreeders.



## Industry highlights

### *Pioneer Hi-Bred International, Inc.: bringing seed value to the grower*

Jerry Harrington

Pioneer Hi-Bred International, Des Moines, IA 50307, USA

#### ***Introduction***

Pioneer Hi-Bred International, Inc., is the world's largest seed company with annual sales of over US\$2 billion. With headquarters in Des Moines, Iowa, Pioneer, a subsidiary of DuPont, sells seed to growers in nearly 70 countries. The company has over 5,000 employees with over 70 seed conditioning plants and 120 plant research centers throughout the world.

Founded in 1926, Pioneer was the first company to develop, produce, and market hybrid seed corn. Before the advent of hybrid corn, farmers saved grain from one year's crop to use as seed for the next with annual yields averaging between 20 and 40 bushels per acre. With the new hybrids, yields improved dramatically, and corn was larger, stronger, and better able to stand up to the elements, resulting today in yields averaging 140 bushels per acre in the USA and commonly 200 bushels per acre throughout the US Corn Belt.

By combining seed research with programs to show the value of hybrid seed corn to growers, Pioneer Hi-Bred played a significant role in ushering in the modern age of farming.

The company looks quite a bit different than it did over 75 years ago. Pioneer has expanded beyond corn and now develops and markets seed for the soybean, sorghum, sunflower, canola, alfalfa, rice, millet, and soft red winter wheat markets. Rapid advances in biotechnology and genetics have dramatically changed research tools and procedures and the very products offered to growers.

In this box, I will concentrate only on corn/maize hybrid development, production, and sales.

#### ***Research***

The Pioneer Research and Product Development (R&D) Department develops hybrids of corn, sorghum, sunflower, and canola, and varieties of soybean, alfalfa, wheat, and canola for worldwide markets. Hybrids and varieties are developed at primary research locations and tested at thousands of other locations in order to be sure that products are adapted to a wide range of growing environments.

Like all successful businesses, Pioneer R&D has specific goals outlined for the marketplace. The goals of the maize research team are:

- 1 To develop hybrids with greater than 5% yield performance advantage.
- 2 To reduce crop losses, grower input costs, and risk through biotechnology that provides insect, disease, and herbicide resistance within maize.
- 3 To create more value and new uses for maize by improving the quality of the grain and forage produced.
- 4 To use available, appropriate technologies that result in improved products for customers.

Customers actually start the hybrid development process. Corn growers, processors, livestock producers, and commodity grain users, along with sales and marketing staff, identify specific traits they want in a hybrid. Then plant genetic researchers draw upon the universal gene pool, proprietary germplasm, and genetic technologies to develop new hybrids.

Lab and field researchers work together to develop products, and scientists apply the latest crop production technology all the way along the product development cycle. While scientists in the lab use technologies to test genes and proteins, scientists in the field evaluate germplasm combinations in numerous environments. Scientists developing Pioneer products lead the seed industry in the development and application of genomic tools. This information, gathered through the use of genomic tools, when used in conjunction with other technologies, helps researchers understand gene functions. This is critical because the information helps scientists better understand which genes determine important traits, gain valuable knowledge of how the genes work together, and get insight into genes that control complex traits such as drought resistance and maturity.

Researchers at Pioneer have already discovered some genes that impact important traits such as disease and insect resistance, drought tolerance, and grain traits, and are searching for more. Genomic tools also allow researchers to look across species for traits that are important for corn. Scientists use this information, in addition to an extensive library of elite genetics, to develop better products.

A series of additional tools within maize have included the introduction of transgenic traits that provide resistance to damaging insects and low-cost herbicides. These can lower production costs, enhance grower efficiency, and increase yields. Before intro-

ducing seed products with these traits onto the marketplace, Pioneer must make sure that these traits are registered within the country where the products are sold and follow all legal and regulatory guidelines within each market. This also includes following product stewardship guidelines within product development and educating its grower customers on these rules, so that the technology is preserved long into the future.

Another application of biotechnology tools is the gene mapping technique. This provides information for direct genetic selection of the gene combinations in breeding lines. Using this information, researchers make more accurate decisions about which lines to use in developing new hybrids.

Researchers at Pioneer spend 4–7 years evaluating products in the laboratory and in a wide variety of growing conditions before new hybrids are released for sale to farmers.

To some extent, plant breeding is a numbers game: the more genetic combinations that are developed and tested, the greater the odds of developing improved products faster. Every year, Pioneer maize researchers around the world evaluate about 130,000 new experimental hybrids. These hybrids enter a four- to five-generation testing cycle. One might think of these experimental products as “applicants” to college. The top 10%, 13,000 from the first season of testing, make up the “freshman college class”.

During each of the next four generations, the hybrids are tested at more locations and in a range of soil types, stresses, and climate conditions. At every stage of testing, researchers look for high, stable yield, standability, tolerance to stresses, and other agronomic characteristics. Only hybrids that meet Pioneer standards are advanced to the “sophomore” class. Typically, there are about 6,000 sophomores each season, less than half of the freshman class. Each year, approximately 200 Pioneer experimental hybrids make it to the “junior” level, and fewer than 100 hybrids make it to the “senior” level. And finally, from about 130,000 original candidates, only about 15–20 hybrids “graduate” to commercial status.

By the time a Pioneer® brand hybrid is offered for sale, it has been tested at more than 150 locations, and in more than 200 customers’ fields. This rigorous testing system helps Pioneer researchers develop leading edge, new genetics with a total package of traits that add value for customers.

### ***Supply management***

When the decision is made to sell a new hybrid, production of that seed is undertaken by Pioneer’s Supply Management Division. The mission of this group is to reliably provide the highest quality seed for Pioneer’s customers.

The process begins with a small number of seeds of the new product being handed off from R&D to Supply Management. From these seeds, parent inbred lines are produced and multiplied by the division’s parent seed operations, located in the US mainland, in Hawaii, and at locations around the globe. As the anticipated demand for the hybrid is identified, commercial volumes of seed are grown, conditioned, and stored to fill customer needs.

The volume of seed for a given product hybrid can number into the hundreds of thousands of units (80,000 seeds constitute a unit). However, regardless of volumes, Pioneer maintains high standards of quality – for germination rates, genetic purity, and physical purity – and regularly monitors seed quality throughout production for all products.

In North America, there are 21 commercial corn production facilities, allowing Pioneer to grow and condition products in a number of different environments to spread risk, while producing in the most cost-effective manner. In addition to a summer production cycle in North America, Pioneer manages a “winter” production cycle, using its facilities in South America, to accelerate the production of the latest products, i.e., those advanced to commercialization in the fall of one year for sale the next spring. This winter production option also allows for recovery of seed supplies in case of a reduced summer production of a given hybrid.

The Supply Management team makes use of state-of-the-art technology and science throughout production. Research agronomists provide the latest scientific information regarding optimum agronomic conditions to grow and harvest seed crops. Pioneer’s sophisticated physical facilities are designed by Pioneer engineers to ensure seed quality is maintained through harvesting, conditioning, and shipping. Despite the millions of units annually conditioned, the equipment is designed to carefully handle every seed with minimal physical damage. A wide range of technologies – global positioning systems, bar codes, automated sampling, and sophisticated inventory management systems – assure that Pioneer produces, ships, and tracks seed efficiently and accurately around the world.

The annual production cycle begins with seed fields being contracted from high quality growers around a production location. Pioneer production agronomists work closely with these growers to ensure each field meets Pioneer’s standards and is managed to optimum production. For example, to attain the very highest seed purity, seed fields are preferred to be on rotated ground – that is, the field had a crop other than corn during the prior growing season. Attention is also given to how well the seed crop can be isolated from other corn crops and corn pollen. The majority of Pioneer’s seed fields are irrigated, which assures adequate moisture is available during critical stages of the crop’s development.

Production agronomists help ensure that each field is planted with the correct ratio of individual rows of male or female inbred parent seed and that the seed is planted at the right soil depth and population densities. Field planters are thoroughly inspected before being filled to make sure there is no foreign grain present. Agronomists and contract seed growers work with neighboring farmers, whose fields adjoin Pioneer’s seed ground, to guarantee that Pioneer’s strict isolation requirements can be achieved.

As the plants emerge and grow, fields are regularly checked for rogues or off-type plants to ensure that only the intended corn plants are present. If any others are found, they are destroyed. All this is done to make sure that customers receive the superior genetics they expect.

Most modern hybrids are a single cross, meaning that the pollen comes from one male inbred and fertilizes – or crosses – with a different female inbred. To assure this cross – and only this cross – occurs, the female tassel is removed prior to shedding pollen so that it does not pollinate itself. Automated mechanical cutters and wheel puller-type detasslers are used to handle some of this work initially. Pioneer also employs thousands of teenagers and adults to complete female detasseling by hand. This work is finished only when Pioneer inspectors certify that at least 99.5% of the female tassels in a field have been pulled. Fields not meeting this standard are abandoned as far as seed production is concerned.

Harvest is an especially busy time of year at the production locations and generally begins when the seed corn moisture in the field reaches its physiological maturity level of 35–40%, this being much higher than the 15–20% moisture level for the harvest of grain. Pioneer research has shown that harvesting seed corn at this higher moisture level and drying it gradually in a controlled environment results in improved yield and quality. However, a freeze at this level of moisture could result in reduced germination levels. At this time of year, production locations run multiple shifts and longer workweeks to accommodate getting the crop in before the weather can adversely impact the crop.

Throughout its growth, the crop has been tracked with sophisticated systems. This careful supervision continues during conditioning. These systems will continue to keep the local production team informed of a wide range of product aspects – origin, amount, location, quality, conditioning activities underway, etc.

When it is ready, the crop is mechanically picked on the cob with the husk on – much like sweetcorn – rather than combined and shelled in the field as grain is harvested. “Husk-on” harvesting ensures gentler handling and helps protect the relatively soft corn seed during this harvest and transportation stage.

A given hybrid, and that hybrid alone, is brought to the unloading area at the production location where the ears are gently unloaded. Between hybrids, the area is thoroughly cleaned to assure there is no mixing of products. Next is the husking and sorting stage, where the husk is removed, and a visual inspection is made of every ear. Any that fail are discarded. Hybrid seed is then dried slowly at low temperatures to just under 13% moisture. The dried corn moves by conveyer to the sheller where seed kernels are carefully removed from the cob while avoiding damage to the living seed. Semifinished seed is then moved to special bulk storage bins, equipped with electronic sensors to monitor the seed temperature and which can quickly engage fans to keep the seed cool during storage.

From here, seed is gently moved by bucket elevators and belt conveyors to seed sizing. This stage physically categorizes seed by width and thickness, so that packaged seed will be uniform. Some farmers request uniform seed size to maintain proper plant populations.

All seed is treated with a fungicide to protect it from soil-borne fungi. Depending on customer requirements, other insecticide seed treatments may be subsequently applied to protect seed from insects once it is out in the farmers’ fields.

As the seed moves through the production location, quality tests are run to evaluate the seed’s status. Seed lots meeting minimum quality standards proceed, while seed lots failing the standard are held aside for further evaluation or discard. The overriding concern is maintaining purity and assuring seed quality.

The final stage of production is packaging. Seed can be packaged into bags of 80,000 kernels, packaged into larger containers, such as PROBOX (holding from 25–50 units) or handled in an unpackaged manner. Regardless of package type, however, every one contains a tag that gives the grower important information about the hybrid type and size, special traits, date of germination tests, origin of the seed, and other information.

Seed is then shipped to Pioneer sales agents around North America and other markets throughout the globe. Computer technology provides the tools needed to track the increasingly complex and vast inventory of seed from Pioneer, allowing the company to ship seed when and where it is needed.

At the end of the customers’ planting season, unused seed is returned to Pioneer’s Supply Management warehouses and stored in large coolers during the hot summer months to better assure that seed germination is preserved. Quality tests continue to be run on each seed lot even while in storage so that the company understands the seed vigor of all of its products. Seed lots that deteriorate beyond acceptable limits during this extended storage period are discarded. The company regularly checks its product supply against anticipated customer demand and lays out plans for growing new crops if needed to assure supplies meet what the customers need.

### **Sales**

The Pioneer sales organization is among the most well trained and equipped in the industry. In North America, Pioneer distributes its seed products through a network of more than 2,500 independent sales reps and several thousand retail dealer outlets in the south and west. More than 90% of the Pioneer sales reps are qualified as Certified Crop Advisors, requiring extensive study of plant agronomy, written exams, and follow-up courses.

When the Pioneer sales organization was created in the late 1920s and 1930s, many of the salesmen were farmers who saw the value of the hybrid corn, offering credible testimonials to their neighbors. This was key to getting farmers to accept the new concept of buying hybrid seed corn, rather than using grain produced in their fields. Initially, most of these salespeople were full-time farmers. Because the seed market has become increasingly complex, the modern Pioneer sales organization has evolved into a system of professional salespeople who are dedicated to more full-time selling, with extensive training programs and a demand

for computer technology tools and skills. This sales force of independent agents can connect to a Pioneer intranet site that provides them a convenient way to record and store customer information, as well to access product and agronomic information.

In North America alone, Pioneer offers about 300 different types of corn hybrids, with the complexity increasing each year. While each region sells only hybrids developed for that area, the sales force still needs to become familiar with a vast array of crop production knowledge to help each customer decide which set of hybrids works best on his or her farm.

In addition to the wide offerings of diverse genetics, corn hybrids are differentiated by agronomic characteristics, insect-resistant and herbicide-resistant traits, seed size, and insecticide seed treatments. In order for growers to get the specific hybrids they want, they are urged to order from Pioneer earlier and earlier each year. Often, the decision-making and seed ordering time for growers is at the end of the previous year's harvest. Farmers will evaluate next year's decisions at field days and through yield comparisons on their own farms and those throughout their area.

These independent sales agents in North America are supported by a staff of regional Pioneer agronomists who provide the sales force and customers with agronomic and production information to help get the greatest value from each bag of seed. The Pioneer agronomists, in turn, are supported by Pioneer research information, as well as information generated by Pioneer Agronomy and Nutritional Sciences, which conducts ongoing research on production and animal nutrition issues to support growers.

The Pioneer sales reps and dealers are also supported by a staff of full-time regional, area, and district salespeople. The sales effort is enhanced by national and regional advertising and a communications effort that reinforces the value of Pioneer for its customers and provides information to growers to help make their seed purchase decisions. This includes regional seed catalogs, a national magazine, and a website offered only to Pioneer customers, which provides valuable crop production information for growers.

All these efforts – within Pioneer research, supply management, and sales – have one goal in mind. It is to provide the greatest value to Pioneer customers year in and year out.

Heterosis, though widespread in the plant kingdom, is not uniformly manifested in all species and for all traits. It is manifested at a higher intensity in traits that have fitness value, and also more frequently among cross-pollinated species than self-pollinated species. All breeding methods that are preceded by crossing make use of heterosis to some extent. However, it is only in hybrid cultivar breeding and the breeding of clones that the breeder has the opportunity to exploit the phenomenon to full advantage.

Hybrids dramatically increase yields of non-hybrid cultivars. By the early 1930s (before extensive use of hybrids), maize yield in the USA averaged 1,250 kg/ha. By the early 1970s (following the adoption of hybrids), maize yields quadrupled to 4,850 kg/ha. The contribution of hybrids (genotype) to this increase was estimated at about 60% (the remainder being attributed to production practices).

### Inbreeding depression

Heterosis is opposite and complementary to **inbreeding depression** (reduction in fitness as a direct result of inbreeding). In theory, the heterosis observed on crossing is expected to be equal to the depression upon inbreeding, considering a large number of crosses between lines derived from a single base population. In practice, plant breeders are interested in heterosis

expressed by specific crosses between selected parents, or between populations that have no known common origin.

Reduction in fitness is usually manifested as a reduction in vigor, fertility, and productivity. The effect of inbreeding is more severe in the early generations (generations 5–8). Just like heterosis, inbreeding depression is not uniformly manifested in plants. Plants including onion, sunflower, cucurbits, and rye are more tolerant of inbreeding with minimal consequences of inbreeding depression. On the other hand plants such as alfalfa and carrot are highly intolerant of inbreeding.

### Genetic basis of heterosis

Two schools of thought have been advanced to explain the genetic basis for why fitness lost on inbreeding tends to be restored upon crossing. The two most commonly known are the **dominance theory** first proposed by C. G. Davenport in 1908 and later by I. M. Lerner, and the **overdominance theory** first proposed by Shull in 1908 and later by K. Mather and J. L. Jinks. A third theory, the mechanism of epistasis (non-allelic gene interactions) has also been proposed. A viable theory should account for both inbreeding depression in cross-pollinated species upon selfing, and increased vigor in F<sub>1</sub> organisms upon hybridization.

### Dominance theory

The dominance theory assumes that vigor in plants is conditioned by dominant alleles, recessive alleles being deleterious or neutral in effect. It follows then that a genotype with more dominant alleles will be more vigorous than one with few dominant alleles. Consequently, crossing two parents with complementary dominant alleles will concentrate more favorable alleles in the hybrid than either parent. The dominance theory is the more favored of the two theories by most scientists, even though neither is completely satisfactory. In practice, linkage and a large number of genes prevent the breeder from developing inbred lines that contain all homozygous dominant alleles. Too many deleterious alleles would be present to make it difficult to inbreed to recover sufficient loci with homozygous dominant alleles. Inbreeding depression occurs upon selfing because the deleterious recessive alleles that are protected in the heterozygous condition (heterozygous advantage), become homozygous and are expressed. It should be pointed out that highly productive inbred lines have continued to be produced for hybrid production, the reason why single-cross hybrids have returned to dominance in corn hybrid production.

To illustrate this theory, assume a quantitative trait is conditioned by four loci. Assume that each dominant genotype contributes two units to the phenotype, while a recessive genotype contributes one unit. A cross between two inbred parents could produce the following outcome:

P <sub>1</sub> (AAbbCCdd)	×	P <sub>2</sub> (aaBBccDD)
Phenotypic value 2 + 1 + 2 + 1 = 6	↓	1 + 2 + 1 + 2 = 6
	F <sub>1</sub> (AaBbCcDd)	
	2 + 2 + 2 + 2 = 8	

With dominance, each locus will contribute two units to the phenotype. The result is that the F<sub>1</sub> would be more productive than either parent.

D. L. Falconer developed a mathematical expression for the relationship between the parents in a cross that leads to heterosis as follows:

$$HF_1 = \sum dy^2$$

where HF<sub>1</sub> = deviation of the hybrid from the midparent value,  $d$  = degree of dominance, and  $y$  = difference in

gene frequency in the parents of the cross. From the expression, maximum midparent heterosis (HF<sub>1</sub>) will occur when the values of the two factors ( $d, y$ ) are each unity. That is, the populations to be crossed are fixed for opposite alleles ( $y = 1.0$ ) and there is complete dominance ( $d = 1.0$ ).

### Overdominance theory

The phenomenon of the heterozygote being superior to the homozygote is called **overdominance** (i.e., heterozygosity *per se* is assumed to be responsible for heterosis). The overdominance theory assumes that the alleles of a gene (e.g.,  $A, a$ ) are contrasting but each has a different favorable effect in the plant. Consequently, a heterozygous locus would have greater positive effect than a homozygous locus, and, by extrapolation, a genotype with more heterozygous loci would be more vigorous than one with less heterozygotes.

To illustrate this phenomenon, consider a quantitative trait conditioned by four loci. Assume that recessive, heterozygote, and homozygote dominants contribute one, two, and one and a half units to the phenotypic value, respectively.

P <sub>1</sub> (aabbCCDD)	×	P <sub>2</sub> (AABBccdd)
Phenotypic value 1 + 1 + 1½ + 1½ = 5	↓	1½ + 1½ + 1 + 1 = 5
	F <sub>1</sub> (AaBbCcDd)	
		2 + 2 + 2 + 2 = 8

Heterozygosity *per se* is most superior of the three genotypes.

### Biometrics of heterosis

Heterosis may be defined in two basic ways:

**1 Better-parent heterosis.** This is calculated as the degree by which the F<sub>1</sub> mean exceeds the better parent in the cross.

**2 Midparent heterosis.** This was previously defined as the superiority of the F<sub>1</sub> over the mean of the parents.

For breeding purposes, the breeder is most interested to know whether heterosis can be manipulated for crop improvement. To do this, the breeder needs to understand the types of gene action involved in the phenomenon as it operates in the breeding population

of interest. As Falconer indicated, in order for heterosis to manifest for the breeder to exploit, some level of dominance gene action must be present, in addition to the presence of relative differences in gene frequency in the two parents.

Assume two populations (A, B), in Hardy-Weinberg equilibrium, with genotypic values and frequencies for one locus with two alleles  $p$  and  $q$  for population A, and  $r$  and  $s$  for population B as follows:

Genotypes	Gene frequency	Genotypic values	Number of $A_1$ alleles
	Pop. A	Pop. B	
$A_1A_1$	$p^2$	$r^2$	+ $\alpha$
$A_1A_2$	$2pq$	$2rs$	$d$
$A_2A_2$	$q^2$	$s^2$	- $\alpha$
			0

After a cross ( $A \times B$ ) between the populations in Hardy-Weinberg equilibrium, the genotypic values and frequencies in the cross are:

Genotypes	Frequencies	Genotypic values
$A_1A_1$	$pr$	+2
$A_1A_2$	$ps + qr$	$d$
$A_2A_2$	$qs$	- $d$

where  $p$  and  $r$  are the frequencies of allele  $A_1$ , and  $q$  and  $s$  are the frequencies of allele  $A_2$ , in the two populations. Also,  $q = 1 - p$  and  $s = 1 - r$ . The mean values of the populations are  $P_A$  and  $P_B$ .

$$\begin{aligned} P_A &= [(p-q)\alpha] + 2pqd \\ &= [(2p-1)\alpha] + [2(p-q^2)d] \\ P_B &= (r-s)\alpha + 2rsd \\ &= (2r-1) + [2(r-r^2)d] \\ F &= [(pr-qs)\alpha] + [(ps+qr)d] \\ &= [(p+r-1)\alpha] + [(p+r-2pr)d] \end{aligned}$$

where  $F$  is the hybrid of  $P_A \times P_B$ . Calculating heterosis as a deviation from the midparent values is as follows:

$$\begin{aligned} H_{MP} &= F_1 - (P_1 + P_2)/2 \\ &= [(p+r-1)\alpha + (p+r-2pr)d] \\ &\quad - \frac{1}{2}[(2p-1)\alpha + 2(p-q^2)d + (2r-1)\alpha \\ &\quad + 2(r-1)\alpha + 2(r-r^2)d] \dots \\ &= (p-r)^2 d \end{aligned}$$

From the foregoing, if  $d = 0$  (no dominance), then heterosis is zero (i.e.,  $F = MP$ , the mean of the midparents). On the other hand, if in population A,  $p = 0$  or 1, and by the same token in population B,  $r = 0$  or 1 for the same locus, depending on whether the allele is in homozygous recessive or dominant state, there will be a

heterotic response. In the first generation, the heterotic response will be due to the loci where  $p = 1$  and  $r = 0$ , or vice versa. Consequently, the heterosis manifested will depend on the number of loci that have contrasting loci as well as the level of dominance at each locus. The highest heterosis will occur when one allele is fixed in one population and the other allele in the other. If gene action is completely additive, the average response would be equal to the midparent, and hence heterosis will be zero. On the other hand, if there is dominance and/or epistasis, heterosis will manifest.

Plant breeders develop cultivars that are homozygous (according to the nature of the method of reproduction). When there is complete or partial dominance, the best genotypes to develop are homozygotes or heterozygotes, where there could be opportunities to discover transgressive segregates. On the other hand, when non-allelic interaction is significant, the best genotype to breed would be a heterozygote.

Some recent views on heterosis have been published. Some maize researchers have provided evidence to the effect that the genetic basis of heterosis is partial dominance to complete dominance. A number of research data supporting overdominance suggest that it resulted from pseudo-overdominance arising from dominant alleles in repulsion phase linkage. Yet other workers in maize research have suggested epistasis between linked loci to explain the observance of heterosis.

### Concept of heterotic relationship

Genetic diversity in the germplasm used in a breeding program affects the potential genetic gain that can be achieved through selection. The most costly and time-consuming phase in a hybrid program is the identification of parental lines that would produce superior hybrids when crossed. Hybrid production exploits the phenomenon of heterosis, as already indicated. Genetic distance between parents plays a role in heterosis.

In general, heterosis is considered an expression of the genetic divergence among cultivars. When heterosis or some of its components are significant for all traits, it may be concluded that there is genetic divergence among the parental cultivars. Information on the genetic diversity and distance among the breeding lines, and the correlation between genetic distance and hybrid performance, are important for determining breeding strategies, classifying the parental lines, defining heterotic groups, and predicting future hybrid performance.

## Definition

A **heterotic group** may be defined as a group of related or unrelated genotypes from the same or different populations, which display a similar combining ability when crossed with genotypes from other germplasm groups. A **heterotic pattern**, on the other hand is a specific pair of heterotic groups, which may be populations or lines, that express in their crosses high heterosis and consequently high hybrid performance. Knowledge of the heterotic groups and patterns is helpful in plant breeding. It helps breeders to utilize their germplasm in a more efficient and consistent manner through exploitation of complementary lines for maximizing the outcomes of a hybrid breeding program. Breeders may use heterotic group information for cataloging diversity and directing the introgression of traits and creation of new heterotic groups.

The concept of heterotic groups was first developed by maize researchers who observed that inbred lines selected out of certain populations tended to produce superior performing hybrids when hybridized with inbreds from other groups. The existence of heterotic groups have been attributed to the possibility that populations of divergent backgrounds might have unique allelic diversity that could have originated from founder effects, genetic drift, or accumulation of unique diversity by mutation or selection. Interallelic interaction (overdominance) or repulsion phase linkage among loci showing dominance (pseudo-overdominance) could explain the observance of significantly greater heterosis following a cross between genetically divergent populations. Experimental evidence supports the concept of heterotic patterns. Such research has demonstrated that intergroup hybrids significantly outyielded intragroup hybrids. In maize, one study showed that intergroup hybrids between “Reid Yellow Dent” × “Lancaster Sure Crop2” outyielded intragroup hybrids by 21%.

D. Melchinger and R. R. Gumber noted that heterotic groups are the backbone of successful hybrid breeding and hence a decision about them should be made at the beginning of a hybrid crop improvement program. They further commented that once established and improved over a number of selection cycles, it is extremely difficult to develop new and competitive heterotic groups. This is because at an advanced stage, the gap in performance between improved breeding materials and unimproved source materials is often too large. However, the chance to develop new heterotic groups could be enhanced with a change in breeding objectives. Once developed, heterotic groups should be broadened continuously by

introgressing unique germplasm in order to sustain medium- and long-term gains from selection.

## Methods for developing heterotic groups

A number of procedures may be used by breeders to establish heterotic groups and patterns. These include pedigree analysis, geographic isolation inference, measurement of heterosis, and combining ability analysis. Some have used diallel analysis to obtain preliminary information on heterotic patterns. The procedure is recommended for use with small populations. The technology of molecular markers may be used to refine existing groups and patterns or for expediting the establishment of new ones, through the determination of genetic distances.

To establish a heterotic group and pattern, breeders make crosses between or within populations. Intergroup hybrids have been shown to be superior over intragroup hybrids in establishing heterotic relationships. In practice, most of the primary heterotic groups were not developed systematically but rather by relating the observed heterosis and hybrid performance with the origin of parents included in the crosses. One of the earliest contributions to knowledge in the areas of developing heterotic patterns was made in 1922. Comparing heterosis for yield in a large number of intervarietal crosses of maize, it was discovered that hybrids between varieties of different endosperm types produced a higher performance than among varieties with the same endosperm type. This discovery by F. D. Richey suggested that crosses between geographically or genetically distant parents expressed higher performance and hence increased heterosis. This information led to the development of the most widely used heterotic pattern in the US Corn Belt – the “Reid Yellow Dent” × “Lancaster Sure Crop”.

## Heterotic groups and patterns in crops

Heterotic patterns have been studied in various species. For certain crops, breeders have defined standard patterns that can guide the production of hybrids. As previously indicated in maize, for example, a widely used scheme for hybrid development in temperate maize is the “Reid” × “Lancaster” heterotic pattern. These heterotic populations were discovered from pedigree and geographic analysis of inbred lines used in the Corn Belt of the USA. In Europe, a common pattern for maize is the “European flint” × “Corn Belt Dent”, identified based on endosperm types. In France,  $F_2 \times F_6$  heterotic

patterns derived from the same open-pollinated cultivars were reported. Other patterns include "ETO-composite" × "Tuxpeno" and "Suwan 1" × "Tuxpeno" in tropical regions. Alternate heterotic patterns continue to be sought.

In rice, some research suggests two heterotic groups within *Oryza indica*, one including strains from south-eastern China, and another containing strains from South East Asia. In rye, the two most widely used germplasm groups are the "Petkus" and "Carsten", while in faba bean three major germplasm pools are available, namely, "Minor", "Major", and "Mediterranean".

Even though various approaches are used for the identification of heterotic patterns, they generally follow certain principles. The first step is to assemble a large number of germplasm sources and then make parent populations of crosses from among which the highest performing hybrids are selected as potential heterotic groups and patterns. If established heterotic patterns already exist, the performances of the putative patterns with the established ones are compared. Where the germplasm accession is too large to permit the practical use of a diallel cross, the germplasm may first be grouped based on genetic similarity. For these groups, representatives are selected for evaluation in a diallel cross.

According to Melchinger, the choice of a heterotic group or pattern in a breeding program should be based on the following criteria:

- 1 High mean performance and genetic variance in the hybrid population.
- 2 High *per se* performance and good adaptation of the parent population to the target region.
- 3 Low inbreeding of inbreds.

### Estimation of heterotic effects

Consider a cross between two inbred lines, A and B, with population means of  $\bar{X}_{P1}$  and  $\bar{X}_{P2}$ , respectively. The phenotypic variability of the  $F_1$  is generally less than the variability of either parent. This indicates that the heterozygotes are less subject to environmental influences than the homozygotes. The heterotic effect resulting from the crossing is roughly estimated as:

$$H_{F1} = \bar{X}_{F1} - \frac{1}{2}(\bar{X}_{P1} + \bar{X}_{P2})$$

This equation indicates the average excess in vigor exhibited by  $F_1$  hybrids over the midpoint (midparent) between the means of the inbred parents. K. R. Lamkey, and J. W. Edward coined the term **panmitic midparent**

**heterosis** to describe the deviation in performance between a population cross and its two parent populations in Hardy-Weinberg equilibrium. Heterosis in the  $F_2$  is 50% less than what is manifested in the  $F_1$ .

### Types of hybrids

As previously discussed, the commercial applications of hybrid breeding started with a cross of two inbred lines (a single cross: A × B) and later shifted to the more economic double cross [(A × B) × (C × D)] and then back to a single cross. Other parent combinations in hybrid development have been proposed, including the three-way cross [(A × B) × C] and modified versions of the single cross, in which closely related crosses showed that the single cross was superior in performance to the other two in terms of average yield. However, it was noted also that the genotype × environment interaction (hybrid × environment) mean sum of squares (from the ANOVA table; see Chapter 23) for the single cross was more than twice that for the double crosses, the mean sum of squares for the three-way cross being intermediate. This indicated that the single crosses were more sensitive or responsive to environmental conditions than the other crosses. Whereas high average yield is important to the producer, consistency in performance across years and locations (i.e., yield stability) is also important. As R. W. Allard and A. D. Bradshaw explained, there are two basic ways in which stability may be achieved in the field. Double and three-way crosses have a more genetically divergent population for achieving buffering. However, a population of single-cross genotypes that are less divergent can also achieve stability on the basis of individual buffering whereby individuals in the population are adapted to a wide range of environments.

Today, commercial hybrids are predominantly single crosses. Breeders continue to develop superior inbred lines. The key to using these materials in hybrid breeding is identifying pairs of inbreds with outstanding combining ability.

### Germplasm procurement and development for hybrid production

As previously indicated, the breeder needs to obtain germplasm from the appropriate heterotic groups, where available. It is critical that the source population has the genes needed in the hybrid. Plant breeders in ongoing breeding programs often have breeding lines in

storage or in nurseries from which potential parents could be selected for future programs. These materials should be evaluated for performance capabilities and especially for traits of interest in the proposed breeding program. Germplasm may be introduced from germplasm banks and other sources. Such material should also be evaluated as is done with local materials.

### **Development and maintenance of inbred lines**

An **inbred line** is a breeding material that is homozygous. It is developed and maintained by repeated selfing of selected plants. In principle, developing inbred lines from cross-pollinated species is not different from developing pure lines in self-pollinated species. About 5–7 generations of selfing and pedigree selection are required for developing an inbred line. As previously indicated, inbreeders tolerate inbreeding, whereas outbreeders experience varying degrees of inbreeding depression. Consequently, the extent of inbreeding in developing inbred lines varies with the species. Species such as alfalfa and red clover that are more intolerant of inbreeding may be selfed only a few times. Alternatively, sib mating may be used to maintain some level of heterozygosity in these sensitive species.

Hybrid breeding as previously stated exploits the phenomenon of heterosis. Heterosis will be highest when one allele is fixed in one parent to be used in a cross and the other allele fixed in the other parent.

#### *Inbred lines of inbreeding species*

Inbred lines in self-pollinated species were previously discussed. They are relatively easy to maintain. The breeder should be familiar with the material to be able to spot off-types that may arise from admixtures or outcrossing in the field. Off-types should be rogued out and discarded, unless they are interesting and warrant additional observation and evaluation. Physical mixtures occur at harvesting (e.g., due to equipment not cleaned properly before switching to another line), threshing, processing and handling, storage, and at planting. When maintaining certain lines, especially those developed from wild species, it may be necessary to be more vigilant and harvest promptly, or bag the inflorescence before complete maturity occurs to avoid losing seed to shattering.

#### *Inbred lines of cross-pollinated species*

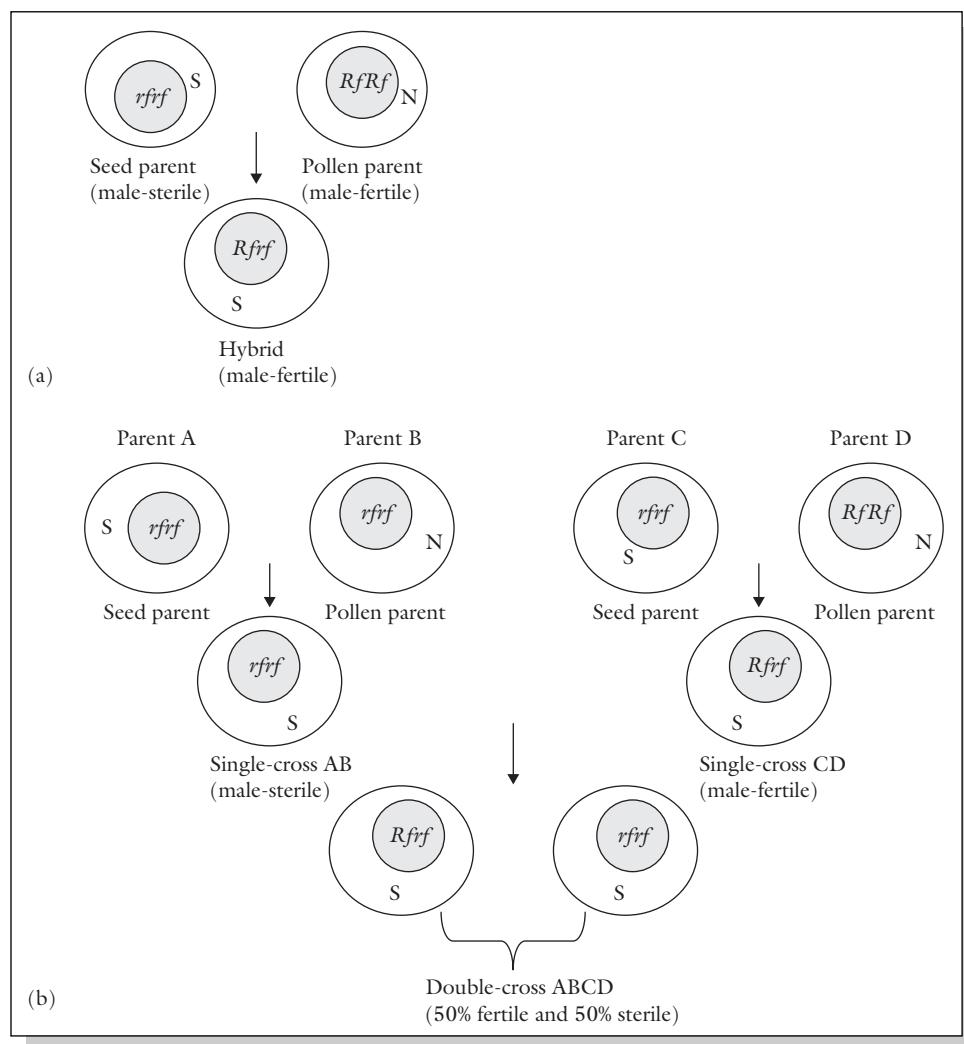
Because of the mode of reproduction, breeding lines from cross-pollinated species are more challenging to

develop and maintain. Inbred lines may be developed from heterozygous materials obtained from a natural population, or from  $F_2$  selected genotypes. Depending on the breeding procedure, parents for hybrid production may be developed in the conventional fashion, or non-conventional fashion.

**1 Conventional or normal inbreds.** Normal inbreds are developed by repeatedly self-pollinating selected plants, from  $S_0$ – $S_n$  (for materials drawn from natural populations) or from  $F_1$ – $F_n$  (for materials obtained from crossing), the latter being akin to the pedigree breeding method previously described for self-pollinated species. The  $S_1$  or  $F_2$  populations are heterogeneous, as are results of segregation of traits. Superior plants are selected and progeny-rowed to expose inferior genotypes. Superior individuals are selected for the next cycle of selfing. By  $S_3$ , the plants in the progeny should be fairly uniform. After about 6–8 generations of selfing, the negative effects of inbreeding ceases. The next step then is to compare different lines. The value of  $n$ , the number of generations of self-pollination, varies from about 5 to 8. The goal is to attain a level of homozygosity at which the inbred lines are uniform in characteristics and will remain so under continued selfing, with no further loss of vigor. At this stage, the inbred line may be maintained by self-pollination.

Inbred lines should be evaluated for performance and other general agronomic qualities (e.g., drought resistance, lodging resistance, disease resistance), especially those that are basic to the specific crop industry and the production region. This way, the final lines developed should have high desirability and productive potential. These materials are maintained by conventional selfing or sibbing procedures.

**2 Non-conventional inbred lines.** To facilitate hybrid production, cytoplasmic male sterility may be incorporated into lines to eliminate the need for mechanical emasculation. Three different inbred lines are required to implement a CMS breeding project. Different kinds of parent materials need to be developed and maintained when making use of a cytoplasmic-genetic male-sterility system in breeding. Two kinds of female parents are needed (Figure 18.1): an **A-line** (male-sterile, sterile cytoplasm (S), with non-restorer genes (*rfrf*) in the nucleus) and a **B-line** (male-fertile, fertile cytoplasm (N), with non-restorer genes (*rfrf*) in the nucleus). The A-line is the seed-producing parent. To develop an A-line, cross a B-line as male to a male-sterile female with sterile cytoplasm and fertility-restorer genes, followed by repeated back-crosses (5–7) to the B-line. The A-line and B-line are



**Figure 18.1** Breeding by CMS: (a) single cross and (b) double cross. N, normal cytoplasm; S, sterile cytoplasm.

hence **isogenic** (genetically different at only a specific locus).

To maintain the A-line, the breeder crosses sterile A-line (S, *rfrf*) with the B-line (N, *rfrf*) to obtain all male-sterile offspring. The B-lines (maintainer lines) are maintained by selfing, sibbing, or by open-pollination in isolated crossing blocks. Another inbred parent, the fertility-restorer parent, is required in a CMS breeding program. This parent (called the **R-line**) has the genotype *RfRf* and is the pollen parent in the hybrid. It is produced by first selecting a desirable line to be converted into an R-line. It is crossed to the restorer gene donor parent using backcrossing with the R-line as the recurrent parent. The process can be simplified if the donor has sterile cytoplasm (S, *rfrf*) and is used as the female parent in the cross.

This strategy will eliminate the need for a testcross or selfing after each backcross to distinguish between fertile and sterile individuals [(S, *Rfrf* × N, *rfrf*), S, *Rfrf* (fertile), or S, *rfrf* (sterile), which are phenotypically distinguishable]. Once produced the R-line is maintained like the B-line. It should be mentioned that inbred lines of self-pollinated species may also be developed to incorporate a male-sterility system for hybrid production.

**3 Genetically modified inbreds.** With the advent of genetic engineering, special lines with specific transgenes are developed for use in producing transgenic hybrids. The details of the techniques are discussed in Chapter 14. Inbred lines are transformed with these special genes to develop transgenic breeding stocks.

### Storage of seed

It is critical that germplasm be stored such that its viability is retained over the duration of storage. Seed germplasm should be stored at low seed moisture content in an environment in which the humidity, temperature, and oxygen content are low. To this end, seed to be stored is usually dried to about 10.0–12.5% moisture content, and stored at a temperature of less than 21°C. The specific requirements differ among species. The Harrington rule of thumb suggests that seed viability is retained for a longer time if the sum of the storage temperature (°F) and relative humidity (%) is less than 100°F. Relative humidity is more important in the storage of soybean. In corn, a sum of 60 is desired for long-term storage of corn. The rate of decline in seed viability in storage also varies among species. Storage in a household freezer may suffice for certain species, especially small-seeded legumes (e.g., alfalfa, clover). The oxygen level in the storage environment may be reduced by introducing gasses such as carbon dioxide, nitrogen, or argon. The seed may also be stored in a vacuum.

### Selection of parents (inbred lines)

The choice of parents to be used in a cross is the most critical step in a plant breeding program for the development of hybrids. The choice of parents depends on the specific objectives of the breeding program and what germplasm is available. Once the inbred lines have been developed, the breeder has the task of identifying a few lines with potential for use as parents in hybrid production. The number of inbred lines that would emerge from a random mating population in which a number of loci are segregating is given by  $2n$ . Hence, for  $n = 10$ , there will be 1,024 inbreds. First, the large pool of inbreds needs to be significantly reduced by phenotypic selection to identify a small number of high performing inbreds. This is effective for traits of high heritability. The next step is to subject the promising lines to a more rigorous test of their performance in crosses (combining ability test, see Chapter 8). Combining ability tests, as previously described, entail crossing each inbred with all other inbreds to be evaluated. Suppose 50 inbreds were selected, the cross combinations required in a combining ability test is given by  $n(n - 1) = 50(50 - 1) = 2,450$  crosses! To handle this large number, the practice is to use a common tester. As previously indicated, the breeder should select parents from different heterotic groups (interheterotic cross) rather than within the same group. A general combining ability (GCA) test

should be conducted first, to be followed by a specific combining ability (SCA) test to identify specific pairs of inbreds with exceptional performance in crosses. This sequence of activities is of practical and strategic importance in quickly reducing the large number of inbreds to a manageable size by the time of the more involved evaluations. Certain inbreds have high GCA, being able to produce high performing hybrids with a series of other inbreds. On the other hand, certain inbreds are able to “nick” with only a few in that set of inbreds tested. The key decision in combining ability testing is the type of tester to use. A tester can have a broad genetic base (e.g., open-pollinated cultivars) or a narrow genetic base (e.g., elite inbreds, related inbred lines).

Where a hybrid breeding program already exists, breeders may want to develop one or two new inbreds to replace those in the program that have been shown to have weaknesses. To replace an inbred in an established single cross, for example, the opposite inbred should be used as a tester. Substitute inbred lines may be developed by backcross procedures (so the inbred is least genetically reorganized), or by isolating new inbreds from the same genetic source. New inbreds may also be developed from completely new sources.

### Field establishment

Once a breeder has identified superior inbreds, these lines are used as parents for producing hybrid seed. Considerations for maximizing hybrid seed production in the field include the following.

- 1 **Field preparation.** The field should be properly prepared to obtain a seedbed suitable for the seed size. The field should be free from weeds (use pre- and postemergence weed control as appropriate). Competition from weeds will adversely affect crop establishment.
- 2 **Planting time.** It is important that the planting be timed such that the seed will germinate promptly for good establishment. Also, the time of pollination should coincide with good weather. In fact, the whole operation, from planting to harvesting, should occur within the growing season, making maximum use of the growing condition for optimum seed yield. The breeder may use **heat units** to calculate the best time for planting the parents (Table 18.1).
- 3 **Synchronization of flowering.** Because a hybrid depends on two different genotypes, the breeder should synchronize the flowering of these inbreds so that both male and female plants would be ready at that same time for effective pollination. A technique

**Table 18.1** Calculating heat units and degree-days.

Scientists use the correlation between temperature and plant growth to predict crop harvesting times and also to determine adaptability to a given area. To do this, plant growth is measured in **heat units** (the number of degrees the mean daily temperature exceeds a base minimum growth temperature):

$$\text{Heat units} = \frac{\text{Daily minimum temperature} + \text{daily maximum temperature}}{2}$$

The base temperature varies with the species (e.g., 4.5°C for small grains, 10°C for corn and soybean, 15.5°C for cotton). A modification of the heat units called **growing degree-days** (GDD) is used where, for the particular species, limits are placed on the daily minimum and maximum temperature (obtained from weather records). For example, if the critical maximum temperature is 30°C and the minimum is 10°C, the GDD for corn for that particular day is:

$$\text{GDD} = \frac{10 + 30}{2} - 10 = 10$$

The GDD are summed over the growing season to give the total needed for commercial production of these crops.

A generalized equation for estimating accumulated temperature above a threshold is:

$$\text{GDD} = T_{\text{mean}} - T_{\text{base}}$$

The application of this technology is in determining the staggering of planting dates of male and female parents in a hybrid production program in the field to ensure high cross-pollination. GDD is also used for predicting vegetative growth and other plant growth processes. Certain horticultural crops such as temperate fruits need to accumulate a certain critical amount of **chilling days** without which fruiting would not occur.

The estimation of chilling dates is the opposite of GDD. Field crops such as spring wheat are sown in the fall so that they receive cold treatment (**vernalization**) during the winter. This treatment triggers the reproductive phase of plant development in spring.

Some species are frost tolerant while others are frost sensitive. Frost damage is most critical when flower buds start to open.

The grower should pay attention to and use frost forecast information provided by the US Department of Agriculture.

Temperature is also an important requirement for germination. Some species (e.g., oat) can germinate in cool soils (3°C), while others (e.g., sorghum) prefer a warm temperature (5°C) for germination. Temperature is a key factor in the Hopkins Bioclimatic Law (crop production activities (e.g., planting, harvesting) and specific morphological developments are delayed 4 days for each 1° of latitude, 5° of longitude, and 122 m (400 ft), as one moves northward, eastward, and upward, respectively).

that is often used is **staggered planting**, whereby the inbred lines are planted at different times to ensure that pollen will be available for both early and later flowering females.

- 4 **Field layout.** Female and male inbred parents should be arranged in the field such that pollen distribution is effective and efficient. Effective layout patterns vary among species. Male and female lines may be planted in alternating rows, or a certain number of males can alternate with a certain number of females according to the pollen-producing ability of the pollinator.
- 5 **Plant density.** Hybrid seed is harvested from only the female parent (i.e., the space occupied by the male is not available for seed production). It is important to maximize plant populations in seed fields. The female plants may be increased relative to the pollen source.

## Maintenance

Once established, the crop should be properly maintained for optimal growth and development. Field care should include weed control, irrigation, proper fertilization, and disease and insect control. To enhance the development of the female line, and reduce the chance of contamination at harvesting, the male line may be removed after pollination is completed. This “thinning” of the field provides additional growth resources for the female line, as a result of reduced competition.

Where pollen control was implemented by mechanical means (as opposed to the use of the male-sterility system), workers need to walk through the field several times during hybrid seed production to remove any pollen sources that may have survived mechanical emasculation.

Further, it is good practice to walk through the field to rogue out any off-types prior to pollination.

### **Harvesting and processing**

The seed should be harvested at the proper maturity and moisture content. At physiological maturity in corn, for example, the kernel moisture is about 30–40%. Safe harvesting is done at a moisture content of 20% or less. The timing of this phase in hybrid seed production is critical because the seed is intended for use as a planting material and must be of the highest possible germination capacity. Mechanical damage, physiological immaturity, and improper seed moisture adversely impact seed quality, and reduce the germination capacity. Further, improper seed moisture may predispose the seed to rapid deterioration in storage.

The processing needed for the seed varies with the crop. In corn, for example, workers first clean the ears to remove diseased and discolored ears before shelling. All seed must be cleaned to remove weed seeds and debris as much as possible. It is required that the producer of the seed attach a label providing specific information including the seed analysis results.

### **Hybrid seed production of maize**

Commercial hybrid seed production in maize is used as an example because hybrid production in maize is one of the earliest and most successful exploitations of heterosis. Both CMS and mechanical detasseling are used in commercial seed production. In the USA, the single cross is used in maize hybrid breeding. The female plant is male sterile (the A-line). It is maintained by crossing with the B-line in isolation. The A-line is grown in alternating rows with the pollinator (the R-line) in a ratio of 1 : 2, 2 : 3, or 2 : 4. Some seed companies mechanically detassel their maize instead of using a CMS system (Figure 18.2).

### **Hybrids in horticulture**

A review by J. Janick (1996) indicates that hybrid seed is significantly used in horticultural production. A wide variety of mating systems are used in hybrid seed production of these species. These include hand emasculation (e.g., in sweet pepper, tomato, eggplant), CMS (e.g., in sugar beet, carrot, onion), self-incompatibility (e.g., in cauliflower, broccoli), and monoecy (e.g., in muskmelon, cucumber). The importance of hybrids is variable among species. The approximate percentages



**Figure 18.2** Using a mechanical detasseler to emasculate corn. (Courtesy of Pioneer Hi-Bred Seed Company.)

of hybrid seed in use in the commercial production of selected plants are: carrot (90% of fresh market and 40–60% of canning and freezing cultivars), broccoli (100%), cauliflower (40%), sugar beet (70%), spinach (90%), muskmelon (80–100%), sweet corn (99%), tomato (100% of fresh market), and onion (65%).

In the ornamental industry a similar picture prevails.  $F_1$  hybrid seed is used in begonia (100% by emasculation), impatiens (100% by CMS), petunia (100% by CMS), seed geranium (100% by genetic male sterility), carnation (80% by genetic male sterility), and dianthus (70% by genetic male sterility).

### **Exploiting hybrid vigor in asexually reproducing species**

#### **Plants with vegetative propagations**

Asexual (vegetative) reproduction is the propagation of plants using propagules other than seed. Many horticultural plants are vegetatively propagated. The economic parts of many important world food crops are non-seed, such as tubers (e.g., potato), stems (e.g., sugarcane), and roots (e.g., cassava). Heterosis can be effectively exploited in species that have the capacity to bear seed and yet be propagated vegetatively. In such species, the plant breeder only needs to create one superior genotype. There is no need for progeny testing. The hybrid vigor and other traits assembled in the  $F_1$  can be maintained indefinitely, as long as the genotype is propagated asexually thereafter.

In horticulture, the superior genotype may be propagated by using techniques such as micropropagation, grafting, budding, sectioning, and cutting (see Chapter 4). Successful hybrids have been developed in species

such as sugarcane (most commercial cultivars), turfgrass (e.g., "Tifway"), and forage crops.

### Apomixis

Apomixis is vegetative propagation through the seed. The seed in this instance is genetically identical to the female plant. Using the apomictic propagule is similar to reproducing the plant by other vegetative means as previously described, only more convenient. Hybrid vigor is fixed and expressed indefinitely through vegetative propagation. The genetics and mechanisms of apomixis are described in detail under breeding apomictic species (see Chapter 11).

### Monoecy and dioecy

The reproductive biology of monoecy and dioecy has been previously described. Monoecious species bear male and female flowers (imperfect) on the same plants but on different parts of the plant. Environmental conditions (e.g., photoperiod, temperature) can influence sex expression by making one plant more female or more male. In cucumber, short day and low night temperatures promote femaleness, while the application of gibberellic acid promotes maleness. Breeders may manipulate the environment to produce hybrid seed.

### Prerequisites for successful commercial hybrid seed production

As briefly reviewed, commercial hybrid seed production is undertaken for a wide variety of species including

field crops and horticultural species. Some species are more suited to commercial hybrid seed production than others. Generally, the following are needed for a successful commercial hybrid seed production venture.

- 1 High heterosis.** Just as plant breeding can not be conducted without variability, hybrid seed production is not meaningful without heterosis. The F<sub>1</sub> should exhibit superior performance over both parents. The degree of heterosis is higher in some species (e.g., corn) than others (e.g., wheat).
- 2 Pollen control and fertility-restoration system.** An efficient, effective, reliable, and economic system should exist for pollen control to exclude unwanted pollen from a cross. Some species have natural pollination control mechanisms (S<sub>1</sub>, male sterility) or reproductive behaviors (monoecy, dioecy) that can be exploited to facilitate the crossing program. A sterility system should have a fertility-restoration system to restore fertility to the commercial seed. In the absence of natural pollen control mechanisms, mechanical or hand emasculation should be feasible on a large scale.
- 3 High F<sub>1</sub> yield.** The F<sub>1</sub> seed is the commercial product. Species such as corn that bear a large number of seed per F<sub>1</sub> plant are more suited to hybrid seed production than species that produce small amounts of seed on an F<sub>1</sub> plant (e.g., wheat).
- 4 Economic seed production.** Hybrid seed production is more expensive overall than conventional seed production. The cost of seed production may be significantly higher when hand emasculation is the method used in the crossing process. In this latter scenario, hybrid seed production would be economic only in high priced crops (e.g., tomato), or where labor is cheap (e.g., cotton production in India).

### References and suggested reading

- Burton, G.W. 1983. Utilization of hybrid vigor. In: Crop breeding (Wood, D.R., ed.). American Society of Agronomy and Crop Science Society of America, Madison, WI.
- Crow, J.F. 1998. 90 years ago: The beginning of hybrid maize. *Genetics* 148:923–928.
- Hallauer, A.R., and J.B. Miranda. 1988. Quantitative genetics in maize breeding. Iowa State University Press, Ames, IA.
- Janick, J. 1996. Hybrids in horticultural crops. CSSA/ASHS Workshop on Heterosis, Indianapolis, IN, November 3, 1996. Department of Horticulture, Purdue University, West Lafayette, IN.
- Melchinger, A.E., and R.K. Gumber. 2000. Overview of heterosis and heterotic groups in agronomic crops. In: Concepts and breeding of heterosis in crop plants (K.R. Lamkey, and J.E. Staub, eds), pp. 29–44. Proceedings of the Plant Breeding Symposium sponsored by the CSSA and American Society of Horticultural Science, November 1996. Crop Science Society of America, Madison, WI.
- Norskog, C. 1995. Hybrid seed corn enterprises. A brief history. Curtis Norskog, 2901 15th Street SW, Willmar, MN.
- Riday, H., and E.C. Brummer. 2002. Heterosis of agronomic traits in alfalfa. *Crop Sci.* 42:1081–1087.
- Sinobas, J., and I. Monteagudo. 1996. Heterotic patterns among US Corn Belt and Spanish maize populations. *Maydica* 41:143–148.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Hybrid vigor is highest in a cross between two identical parents.
- 2 CMS may be used in hybrid breeding to eliminate emasculation.
- 3 The inbred A-line is male-sterile.
- 4 G. H. Shull proposed the dominance theory of heterosis.
- 5 A hybrid cultivar is the F<sub>1</sub> offspring of a cross between inbred lines.

### Part B

Please answer the following questions:

- 1 Define a hybrid cultivar.
- 2 What is hybrid vigor, and what is its importance in hybrid breeding?
- 3 What is an inbred line?
- 4 What is a heterotic group?
- 5 Explain the dominance of single-cross hybrids in modern corn hybrid production.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the dominance theory of heterosis.
- 2 Discuss the importance of synchronization of flowering in hybrid breeding.
- 3 Discuss inbred lines and their use in hybrid breeding.
- 4 Discuss the contributions of G. H. Shull and D. F. Jones in hybrid breeding.

# **Section 7**

## **Selected breeding objectives**

**Chapter 19 Breeding for physiological and morphological traits**

**Chapter 20 Breeding for resistance to diseases and insect pests**

**Chapter 21 Breeding for resistance to abiotic stresses**

**Chapter 22 Breeding compositional traits and added value**

---

This section is devoted to discussing the genetics and other issues associated with breeding for traits of major importance to crop production. The discussion focuses on the rationale or the economic importance of each trait, the genetics of the trait, and the germplasm resources available for breeding. Successes and challenges of breeding these traits are also addressed. However, specific breeding procedures are not discussed because the methods of breeding have been addressed in detail in previous chapters in the book. By understanding the genetics and other characteristics of the traits, the breeder can then select the most appropriate breeding approach for improving the trait. The traits have been grouped under general categories, but some could be discussed under other categories as well. Even though they have been discussed as if they are independent, some of these traits or factors are so interdependent that addressing one impacts the other, as pointed out in various sections.

---



# 19

# Breeding for physiological and morphological traits

---

## Purpose and expected outcomes

*Physiological processes are common to all plants. However, there are morphological and physiological differences among plants. Morphological and anatomical traits are products of physiological processes. Yield is the ultimate goal of a plant breeding program. It is the product of complex biochemical processes. Plant breeders rarely select solely on yield without regard to some morphological trait(s). After completing this chapter, the student should be able to:*

- 1 Define yield.
  - 2 Discuss the biological pathway to economic yield.
  - 3 Discuss the concept of yield potential.
  - 4 Discuss the concept of yield plateau.
  - 5 Discuss the concept of yield stability.
  - 6 Discuss breeding of lodging resistance.
  - 7 Discuss breeding of shattering resistance.
  - 8 Discuss breeding of plant stature.
  - 9 Discuss breeding for early maturity.
  - 10 Discuss breeding of photoperiod response.
- 

## Physiological traits

Plant growth and development depend on a complex interaction of biochemical and physiological processes. Plant physiological processes are under genetic control as well as under the influence of the environment. The genotype of the plant determines the total complement of enzymes in the cell and hence is a determining factor in plant growth. As D. C. Rasmusson and B. G. Gengenbach stated, physiological gene functions determine the manner and extent of the genotypic contribution to the phenotype of the plant. Physiological gene action also reflects gene differences that provide

the basis for selection of desirable genotypes in plant breeding.

The major physiological processes are photosynthesis, respiration, translocation, and transpiration. Crop yield and productivity depends on the proper functioning of these processes. These traits are quantitatively inherited. Physiological traits may be broadly defined to include the major physiological processes, and yield and its components. It also includes plant environmental responses (to photoperiod and environmental stresses). Some of the specific physiological traits that have been bred by plant breeders with varying degrees of success are photosynthetic rate, leaf angle, leaf area, stomatal frequency,

water utilization, photoperiod, harvest index, tolerance to environmental stresses (drought, cold, salt, heat), and mineral nutrition. Some of the significant achievements with breeding for physiological traits have resulted in the modification of plant architecture, specifically short stature (semidwarf) in cereals (e.g., rice, wheat), with all the advantages that such a plant architecture brings. Photoperiod response is discussed in this chapter because of its association with maturity and plant stature.

### What is yield?

**Yield** is a generic term used by crop producers to describe the amount of the part of a crop plant of interest that is harvested from a given area at the end of the cropping season or within a given period. The plant part of interest is that for which the crop producer grows the crop. It could be the leaves, fruits, stems, roots, or flowers, or any other morphological part. It could also be the chemical content of the plant such as oil, sugar, or latex. In certain industrial crops such as cotton, the plant part of economic interest to the producer is the fiber, while for the producer of tea or tobacco the part of interest is the leaf. It should be added that a producer might harvest multiple parts of the plant (e.g., grain and leaf) for use or sale (i.e., multiple economic parts). Plant breeders seldom select solely on yield basis, without some attention to other morphological features of the plants. Yield is the best measure of the integrated performance of a plant.

Biological yield may be measured by breeding for physiological and morphological traits. All crop production ultimately depends on photosynthesis (as well as other physiological processes, for example respiration and translocation). Over the years, various researchers have attempted to improve biological yield by: (i) increasing the photosynthetic capacity of the individual leaf; (ii) improving the light interception characteristics of plants; and (iii) reducing wasteful respiration. In addition to increasing plant biomass, the goals of breeding for physiological and morphological traits include the redistribution of assimilates to the economic products within the plant as well as alleviating or avoiding the effects of adverse environmental conditions.

The term **biomass** is used by scientists to describe the amount or mass of organic matter in a prescribed area at a given point in time. This measure of biological matter includes material formed above and below ground. Yields of liquid products (e.g., latex, syrup) are measured by quantifying the volume of the product harvested.

Depending on the type of product and the purpose of producing it, harvesting may be undertaken at various stages of maturity for various product quality preferences, as demanded by the targeted market. Plant breeders may breed certain crops for early harvesting (for the fresh market) and others for dry grain. The yields at various stages of harvesting will differ between premature and fully mature products. Sometimes, scientists eliminate the moisture factor by measuring the weight of the harvested product on a dry matter basis after drying the product in an oven prior to being weighed.

### Biological versus economic yield

Yield may be divided into two types:

**1 Biological yield.** This may be defined as the total dry matter produced per plant or per unit area (i.e., biomass). Researchers use this measurement of yield in agronomic, physiological, and plant breeding research to indicate dry matter accumulation by plants. All yield is at first biological yield.

**2 Economic yield.** The economic yield represents the total weight per unit area of a specified plant product that is of marketable value or other use to the producer. The producer determines the product of economic value. A producer of corn for grain is interested in the grain; a producer of corn for silage is interested in the young, fresh stems and leaves. All yield is biological yield, but all biological yield is not necessarily economic yield. For example, the above-ground parts of corn may be totally useful in one way or another (e.g., the grain for food or feed, and the remainder also for feed or crafts). The roots are of no practical or economic use. However, in certain root crops such as sugar beet, the total plant is of economic value (root for sugar extraction and the leafy tops for livestock feed).

Yield depends on biomass and how it is partitioned. To increase yield, the breeder may breed for increased biomass and efficient partitioning of assimilates. The potential biomass of a crop is determined by factors including genotype, local environment (soil, weather), and the agronomic practices used to grow it. N. W. Simmonds identified three strategies for enhancing biomass:

**1 Seasonal adaptation.** The objective of this strategy is to optimally exploit the growing season by sowing early and harvesting late to maximize biomass

accumulation. Of course, this will have to be done within reasonable agricultural limits, as dictated by weather and cropping sequence. Genotypes can be adapted to new growing conditions (e.g., cold tolerance to allow the farmer to plant earlier than normal).

- 2 Tolerance of adverse environmental factors.** Because of the vagaries of the weather and the presence of other inconsistencies or variation in the production environment (climate, product management, etc.), biomass can be enhanced by breeding for tolerance to these factors. Such breeding efforts may be directed at developing tolerance to abiotic stresses (e.g., drought, heat, cold). This would allow the cultivar to produce acceptable yields in the face of moderate to severe adverse environmental conditions.
- 3 Pest and disease resistance.** Diseases and pests can reduce biomass by killing plant tissue (or even an entire plant in extreme cases), and stunting or reducing the photosynthetic surface of the plant. Disease- and pest-resistance breeding will enhance the biomass potential of the crop. Breeding to control pests is one of the major undertakings in plant breeding.

### Ideotype concept

Plant breeders may be likened to plant structural and chemical engineers who manipulate the genetics of plants to create genotypes with new physical and biochemical attributes for high general worth. They manipulate plant morphology (shape, size, number of organs) to optimize the process of photosynthesis, which is responsible for creating the dry matter on which yield depends. Once created, dry matter is partitioned throughout the plant according to the capacity of meristems (growing points of the plant) to grow. **Partitioning** (pattern of carbon use) is influenced by both intrinsic (hormonal) and extrinsic (environmental) factors. Certain plant organs have the capacity to act as **sinks** (importers of substrates) while others are **sources** (exporters of substrates). However, an organ may be a source for one substrate at one point in time and then a sink at another time. For example, leaves are sinks for nutrients (e.g., nitrates) absorbed from the soil while they serve as sources for newly formed amino acids.

Plant genotypes differ in patterns for partitioning of dry matter. This means plant breeders can influence dry matter partitioning. Pole (indeterminate) cultivars of legumes differ in patterns of the partitioning of dry matter from bush (determinate) cultivars. Similarly, in cereal crops, tall cultivars differ from dwarf cultivars in the pattern of dry matter partitioning. The concept of

the **plant ideotype** was first proposed by C. M. Donald to describe a model of an ideal phenotype that represents optimum partitioning of dry matter according to the purpose for which the cultivars will be used. For example, dwarf (short statured) cultivars are designed to channel more dry matter into grain development whereas tall cultivars produce a lot of straw. Tall cultivars are preferred in cultures where straw is of economic value (e.g., for crafts or firewood). Consequently, ideotype development should target specific cultural conditions (e.g., monoculture, high density mechanized production, or production under high agronomic inputs). All breeders, consciously or unconsciously, have an ideotype in mind when they conduct selection within a segregating population.

Plant morphological and anatomical traits (e.g., plant height, leaf size) are relatively easy for the breeder to identify and quantify. They do not vary in the short term and also tend to be highly heritable. Consequently, these traits are most widely targeted for selection by breeders in these programs. The wheat ideotype defined by Donald comprised the following:

- 1 A relatively short and strong stem.
- 2 A single culm.
- 3 Erect leaves (near-vertical).
- 4 A large ear.
- 5 An erect ear.
- 6 Simple awns.

It is not practical to specify every trait in modeling an ideotype. The degree to which a plant model is specified is left to the discretion of the breeder. A more accurate ideotype can be modeled if the breeder has adequate information about the physiological basis of these traits. The group of traits used to define the ideotype presumably are those that would contribute the most to crop economic yield under the range of environmental and management conditions that the crop would encounter during its life. As previously indicated, physiological processes are common to all plants, but there is no universal ideotype in plant breeding. This is because of the vast morphological and physiological diversity of crops, and the wide range in their economic end products as well as the cultural environments.

The role of partitioning in determining crop yield depends on the species and the products of interest. In forages, the total above-ground vegetative material is harvested as the end product. The importance of partition in the economic yield in this instance is small. In other crops, the desired product is enhanced at the

expense of the rest of the plant. N. W. Simmonds has identified three outcomes of competition among plant parts for assimilates and their implications in plant breeding.

- 1 Vegetative growth is sacrificed for reproductive growth.** In this outcome, the breeder reduces vegetative growth by reducing the plant structure (breeding for dwarf cultivars, or breeding determinate cultivars). Other strategies include the reduction in foliage as in the okra leaf cotton phenotype. Dwarf cultivars have been developed for many major cereal crops of world importance (e.g., wheat, rice, sorghum, barley). Dwarf cultivars are environmentally responsive (i.e., respond to agronomic inputs – fertilizer, irrigation, etc.). The success of the Green Revolution was in part due to the use of dwarf cultivars of wheat and rice.
- 2 Reproductive growth is sacrificed for vegetative growth.** In crops in which the desired part is vegetative, flowering and seed set are either reduced, (e.g., yam, cassava, potato) in order to channel resources into the vegetative parts, or suppressed (e.g. sugar beet, carrot).
- 3 One vegetative growth is sacrificed for another desired vegetative growth.** The objective is to allocate dry matter to the harvestable underground vegetative structures (e.g., potato).

### Improving the efficiency of dry matter partitioning

Proposed by C. M. Donald, **harvest index** is the proportion of the plant that is of economic value. It is calculated as a ratio as follows:

$$\text{Harvest index} = (\text{economic yield}/\text{biological yield})$$

For cereals, for example, the ratio will be grain yield to total plant weight (grain + straw). The theoretical value of the harvest index ranges from 0.0 to 1.0 (the value may be converted to a percentage by multiplying by 100). The higher the value, the more efficient the plant is in directing assimilates to the plant parts of economic value. The harvest index is hence also referred to as the **coefficient of effectiveness**. The higher the harvest index, the more economically desirable the genotype, because it translocates more of the available assimilates into the economic parts of the plant.

Some researchers indicate that the dramatic increase in the grain yield of major world cereal crops is due

mainly to increases in the harvest index and to a lesser extent the biological yield. In maize, for example, the harvest index changed from 24% in 1950 to 43% in 1970, increasing yield from 3 to 8 metric tons/ha. Generally, tall cultivars have high biological yield and low harvest index, whereas semidwarf cultivars have high harvest index and high biological yield. On the other hand, full dwarf cultivars have low biological yield and low economic yield. The breeding question is how effectively can harvest index be selected to make it a breeding objective for increasing yield? One difficulty with selecting harvest index is that it is not a phenotypic trait that can be readily evaluated. It is calculated from data obtained from two separate weighings. Such data are problematic to obtain if experimental plants are harvested mechanically, as is the case in many large breeding programs.

The developmental pathway followed by the plant part or chemical component of economic value affects the harvest index. In cereal crops (e.g., corn, wheat), the economic part, the grain, fills in a linear fashion up to a definite point, and then ceases. The harvest index in these crops depends on the relative duration of the vegetative and reproductive phases of the plant life cycle. However, in crops such as sugar beet and Irish potato, the economic part follows a protracted developmental pathway. In these crops, harvest index depends more on genetics than environmental factors.

The harvest index can be decreased or increased by manipulating the cultural environment. For example, increasing plant density and drought or soil fertility (e.g., nitrogen application), are known to lower the harvest index in corn. However, planting early maturing cultivars under good management, increased the harvest index in rice in some studies. This happened because the plant was able to allocate assimilates to the seed sooner, thereby leading to reduced accumulation of reserves in the leaves.

Harvest index has also been increased in small grain cereals partly through decreasing plant stature (e.g., by using the *Rht* dwarfing gene in wheat or by selection technique).

### Harvest index as a selection criterion for yield

In spite of the role of the harvest index in increasing crop yield, using this trait as a selection criterion for grain yield is problematic for the breeder. This is largely due to the effect of the environment and cultural

conditions on the harvest index, as previously described. Sometimes, the breeder selects on the basis of single plants (e.g., in a space planted, segregating population in the early part of a breeding program) or on the basis of families, at some point. Also, sometimes, plants are evaluated in microplots and at other times in large field plots at commercial densities. The challenge is for the breeder to predict yielding ability from one plant arrangement (isolated plants) to another (field crops). A much more severe restriction to the practical use of the harvest index as a selection criterion is the fact that it is more tedious to measure than grain yield *per se*.

### Selecting for yield *per se*

As previously indicated, a plant breeder seeking to improve crop yield affects the trait through manipulating biomass or partitioning or both. Furthermore, because yield is a complex trait, an outcome or product of the interaction of numerous physiological processes, breeders seek effective and efficient ways of selecting superior genotypes in a breeding program. As also previously discussed, biomass and partition are tedious to estimate. The rationale of yield components as a basis for selecting for yield has not proven useful because of the occurrence of compensatory negative correlations (i.e., an increase in one component produces a decrease in another). Similarly, certain physiological parameters (e.g., photosynthetic rate, net assimilation) that have been proposed as possible indicators of improved biomass, have not materialized. Breeders have also resorted to a variety of statistical procedures to help the selection process to become more efficient and effective for yield. A notable application is the development of indices for selection (selection index). Other multivariate techniques such as path coefficients have been attempted with little success. Modern molecular technology is attempting to identify quantitative trait loci (QTLs) associated with complex traits. In view of the foregoing, little wonder that many breeders select for yield *per se* in their breeding programs. As previously stated, it is the best measure of the integrated performance of a plant or crop.

### Biological pathways to economic yield

Yield is a very complex trait. A good crop yield reflects a genotype with high yield potential, growing in a good environment. It reflects, also, proper growth and

development – processes that are very complex in themselves. In an effort to manipulate crop yield, plant breeders attempt to construct the path by which the reproductive, developmental, and morphological features of plants in a crop stand contribute to the yield of a specified product. The pathways to yield are collectively called **yield components**. In theory, the total yield can be increased by increasing one component while holding the others constant. By breaking down a complex trait into components, breeders hope to find selection criteria for improving it.

For grain yield, a model of yield components is:

$$\begin{aligned}\text{Yield/unit area} = & (\text{plants/unit area}) \times (\text{heads/plant}) \\ & \times (\text{mean number of seeds/head}) \\ & \times (\text{mean weight/seed})\end{aligned}$$

Where the plant produces tillers, the model may be modified as:

$$\begin{aligned}\text{Yield/unit area} = & (\text{plants/unit area}) \times (\text{mean}) \\ & \times (\text{number of tillers with ears/plant}) \\ & \times (\text{mean number of grains/ear}) \\ & \times (\text{mean grain weight})\end{aligned}$$

These plant characteristics describe yield. They all depend on energy in a fixed pool that is furnished through photosynthesis. Plant breeders seek to influence yield by manipulating its components to positively affect photosynthesis.

It is important to mention that in interpreting the correlation between yield and its components, the breeder should evaluate the components in terms of relative importance. The seasonal sequence of environmental conditions that affect plant development should also be considered. Growing conditions may be ideal in the early growth and development of the crop, leading to good initiation of reproductive features. However, if there is an onset of drought, few pods may complete their development and be filled with seed, leading to low correlation between yield and the number of seeds per pod.

Yield components vary from one species to another in terms of optimum value relative to other components. Further, yield components affect each other to varying degrees. For example, if increasing plant density drastically reduces the number of pods per plant, the number of seeds per pod may only be moderately affected whereas seed size may remain unchanged or only slightly affected. Whereas a balance among yield components has great adaptive advantage for the crop, the components are environmentally labile. High yield

usually results from one component with extreme value. Furthermore, yield components are determined sequentially. As such, they tend to exhibit **yield compensation**, the phenomenon whereby deficiency or low value for the first component in the sequence of developmental events is made up by high values for the subsequent components. The next effect is that yield is maintained at a certain level. However, yield compensation is not a perfect phenomenon. For example, it may occur over a wide range of plant densities in certain species. In beans, a reduction in pod number can be compensated for by an increase in seed number per pod and weight per seed.

### Concept of yield potential

A given crop has an inherent optimum capacity to perform under a given environment. This capacity is called its **yield potential**. It can be measured through yield trials. It is the maximum attainable crop yield from a specific soil–water regime under ideal production conditions (experimental conditions whereby there is no limit

on access to any needed production input). It is suggested that only about 20–40% of this yield potential can be attained economically in actual production on farmers' fields. When a farmer has reached an economic yield potential for the crop, attempts can be made to increase field yields in a variety of ways. The farmer can use production resources more efficiently; agronomic innovations that are more responsive to local needs and conditions can be introduced to the farmer by extension agents; and the government may institute incentive policies (e.g., credit) for farmers. However, improved cultivars constitute perhaps the most effective approach. To do this, a breeder would have to assemble appropriate variability (genotypes with complementary genes contributing to yield potential) and hybridize them to generate transgressive segregates with superior yield. Biotechnology can be used to develop new cultivars to cope with the constraints to the rise in field yields (abiotic and biotic stresses). Molecular markers may be used to assist the breeding of especially complex traits while recombinant DNA technology may be used to incorporate desirable unique genes from unrelated sources.

## Industry highlights

### *Bringing Roundup Ready® technology to wheat*

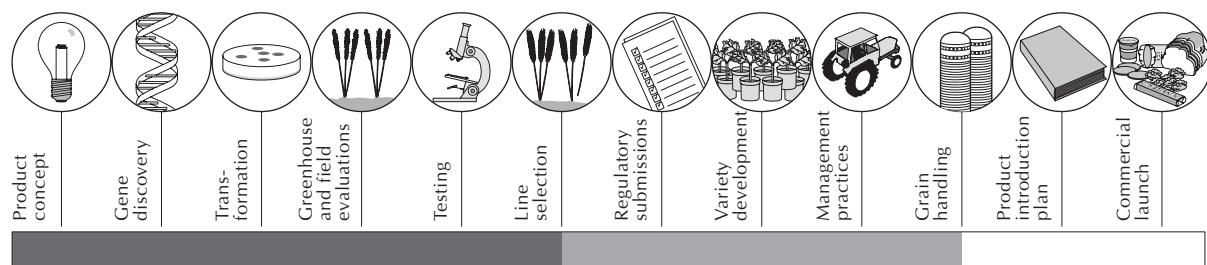


Sally Metz

Monsanto Corporation, 800 North Lindbergh Blvd, St Louis, MO 63167, USA

#### Introduction

Once a technology is proven successful in a crop (e.g. Roundup Ready® soybean; Padgett et al. 1996), researchers can theoretically transfer the technology to other crops. Industry refers to this as product extension. Given the complexity of the crop and trait, researchers determine how much additional optimization is required to achieve commercial success in subsequent crops. This development process can take 5–8 years and involves many different aspects of science. Costs can approach US\$40 million (Context Network 2004) over this period of time, requiring researchers to be strategic, focused, and precise. The development timeline and process described below occurred between 1997 and 2004 (Figure 1).



**Figure 1** The steps involved in the development and commercialization of Roundup Ready® wheat.

### Product concept

Understanding the needs of the customer, in this case the wheat growers, and then determining the technical criteria to meet those needs, are critical components in developing the product concept. In focus group sessions conducted in 2000, spring wheat growers identified numerous challenges with their current weed control options. These challenges included:

- 1 Incomplete control of many tough weeds.
- 2 Annual variability in product performance.
- 3 Crop damage/setback caused by herbicides.
- 4 The need for multiple products to broaden the spectrum of control.
- 5 Significant price variability based on the quantity and type of weeds.

The product concept for Roundup Ready® wheat is complete, dependable, cost-effective control of all weeds, often with only one herbicide application. Four years of field trials indicate that a single 26 oz/acre application of Roundup UltraMax® over the top of Roundup Ready® spring wheat provides 95–100% control of nearly all broadleaf and grassy weeds (Figure 2). Before a project is initiated, a market assessment is completed to understand the commercial potential for that trait/crop combination. If the results of that study are promising, then the project is initiated.

### The development team

Because there are so many aspects to bringing a new product to commercialization, a wide diversity of expertise is needed amongst team members. Early-stage development teams usually combine expertise in molecular biology, transformation, genetics, plant breeding, agronomy, and regulatory science. Bench scientists who have expertise in one of the team areas often lead early-stage teams. For instance, the early-stage Roundup Ready® wheat technical team was led by a molecular biologist. Mid-stage development teams evolve to incorporate additional expertise in regulatory affairs, seed production, industry affairs, public affairs, government affairs, and commercial development. People who specialize in managing product development typically lead these teams. Late-stage development teams evolve again – phasing out the expertise in molecular biology, transformation, and genetics, while adding expertise in marketing. Someone from the commercial arena usually leads these teams.

### Transformation

To be commercially successful, a crop not only has to be transformed, but the process has to be amenable to making hundreds of transgenic plants. Monsanto developed a process to transform wheat, using the trait of interest – glyphosate tolerance – as the selectable marker. The transformation method also must result in transgenic plants that are “clean” events – meaning that the DNA insertion is registerable with many regulatory bodies around the world. It has been our experience that *Agrobacterium* transformation results in more useful events than does particle gun transformation (Hu et al. 2003).



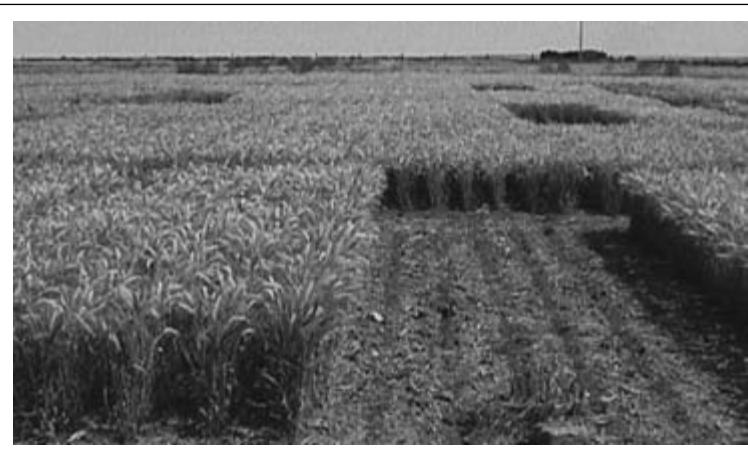
**Figure 2** Product concept demonstration: control of broadleaf and grassy weeds.

### *The gene and its expression in the plant*

In Roundup Ready® crops, the principal gene that has been used is the *CP4 EPSPS* gene, which was isolated from *Agrobacterium*. This is the gene present in Roundup Ready® wheat. The expression of this gene allows the plant to continue to produce aromatic amino acids after the application of one of the Roundup® family of herbicides. The challenge is enabling the gene to be expressed in the right tissues at the right time. The promoters control expression. Some promoters, like the cauliflower mosaic virus-enhanced 35S



**Figure 3** Field trials demonstrating gene efficacy.



**Figure 4** Selecting the commercial event.

ful for the team to set selection criteria before the data is analyzed. Selection criteria involves four areas: (i) molecular biology (single insertion, no vector backbone, intact insertion); (ii) genetics (trait inherited as a single dominant gene, expression maintained over generations); (iii) agronomics (yield, maturity, disease reactions – i.e., selection of plants that are not affected by the transformation/regeneration process); and (iv) trait performance (for herbicide tolerance, at least a 2x safety margin to the commercial rate and timing of application) (Figure 4). In the case of Roundup Ready® wheat, the development team invoked an additional selection criterion to select a transgenic event that was present on the A or B genome. Molecular breeding techniques were used to map the insertion of every potential commercial event. By selecting away from the D genome, research demonstrates that there is a high likelihood that crossing between wheat and its only North American wild relative, jointed goatgrass (*Aegilops cylindrica*), would not result in introgression of the trait into the wild jointed goatgrass population.

#### *Developing the trait*

In parallel, the push is often underway in four primary areas: (i) developing regulatory packages for the Environmental Protection

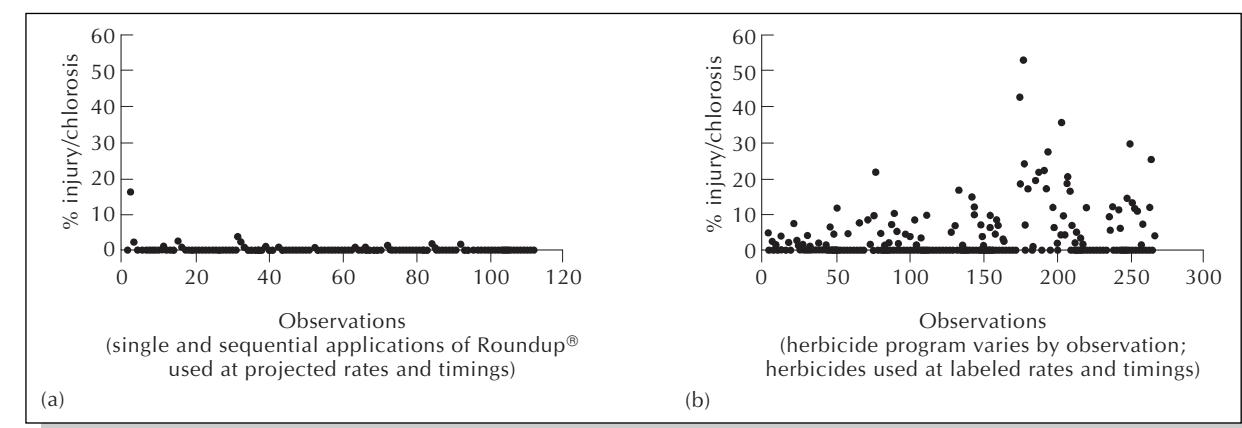
promoter (e35S), are expressed strongly in wheat leaves, but at lower levels in specific tissues critical for reproduction. Other promoters, like rice actin, are expressed at higher levels in these critical reproductive tissues but at lower levels during plant regeneration and in wheat leaves. In wheat, we used both the e35S and rice actin promoters to achieve plant selection and a commercial level of vegetative and reproductive tolerance to glyphosate.

Usually, many different constructs with different combinations of promoters, genes, and stop regions are tested to find the combination that provides the desired phenotypic result.

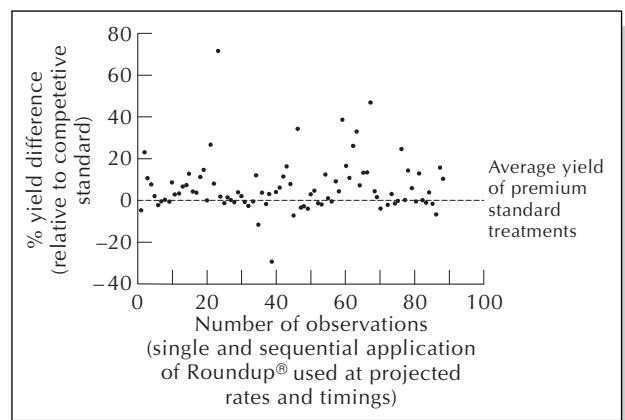
First, tests for the function of the introduced gene – for example, in the case of Roundup Ready® wheat, glyphosate spray tests – are conducted to select transgenic plants that express the trait. Seed of each event must be increased in greenhouses or growth chambers to provide seed for field testing. Any “environmental release” (field trial) is conducted under the rules and regulations of the US Department of Agriculture–Animal and Plant Health Inspection Service (USDA-APHIS) as well as individual state departments of agriculture. Similarly, the Canadian Food Inspection Agency (CFIA) controls field testing in Canada. For wheat, these regulations include specifications regarding minimum isolation distances, volunteer monitoring requirements, and crop rotation restrictions. Field trials are conducted over several locations and years to obtain enough performance data to select the commercial transgenic event (Figure 3) (Zhou et al. 2003).

#### *Selecting the commercial transgenic event*

In order to facilitate the selection of the eventual commercial event, it is most use-



**Figure 5** The crop safety of (a) the Roundup Ready® system versus (b) grower standard programs, showing that there is virtually no crop injury from the use of Roundup Ready® wheat. (Source: Monsanto and University 1999–2003 US field trials.)



**Figure 6** Results showing that Roundup Ready® wheat consistently yields better than wheat treated with the best commercial standard treatments. (Source: Monsanto and University 1999–2003 US field trials.)

eties in components important for human and animal nutrition, and that no unapproved food additives are present (Obert et al. 2004); and (ii) safety studies to ensure that Roundup Ready® wheat contains no increase in allergens, toxicants, or antinutrients when compared to standard wheat varieties (Goodman et al. 2003).

Additional marketplace-support studies demonstrated feeding equivalence for hogs (Hyun et al. 2005) and chickens (Kan & Hartnell 2004), as well as weed control and yield optimization (Blackshaw & Harker 2003). Between 2001 and 2004, Monsanto scientists or public scientists who are developing a basic database or evaluating the technology, published more than 39 scientific presentations, posters, and peer-reviewed journal articles.

Years of work and research demonstrates the successful transfer of the Roundup Ready® technology to wheat – meeting or exceeding all selection criteria and resulting in a product that provides 95–100% dependable, cost-effective control of all weeds, often with only a single 26 oz/acre application of Roundup UltraMax®. In addition, Roundup Ready® wheat optimizes the yield potential by incurring less stress due to herbicide injury (Figure 5) and by decreasing the quantity of weeds capable of competing with the wheat crop for available water and nutrients. This has resulted in a consistent 5–15% yield increase over conventional wheat treated with conventional herbicides (Figure 6).

Agency (EPA), Food and Drug Administration (FDA), USDA, and other regulatory agencies in Canada, Europe, Japan, etc.; (ii) developing commercial varieties that contain the new trait; (iii) developing agronomic and stewardship practices for the trait; and (iv) developing benefits data for growers, consumers, and the internal marketing and salespeople who will promote the product.

Questions are asked such as: When do you apply the herbicide? At what rate? Are split applications beneficial? How does rate and timing affect residue levels in grain or forage? How do you control volunteers? To what extent is outcrossing an issue? Are there any data published in the literature? Has end-use quality been affected? Can you optimize yields in the system? What about crop rotations? What is the most productive way to use this technology in a total cropping system? Has the feed efficiency of livestock been affected?

Monsanto conducted more than 48 formal regulatory studies, which demonstrated the substantial equivalence and safety of Roundup Ready® wheat compared to conventional wheat. These studies fall into two categories: (i) compositional analysis to ensure that Roundup Ready® wheat is substantially equivalent to standard wheat varieties

### References

- Blackshaw, R.E., and K.N. Harker. 2003. Selective weed control with glyphosate in glyphosate-resistant spring wheat (*Triticum aestivum*). *Weed Technol.* 16(4):885–892.
- Context Network. 2004. Biotech traits commercialized, 2004. CD ROM.
- Goodman, R.E., Bardina, L., Nemeth, M.A., et al. 2003. Comparison of IgE binding to water-soluble proteins in genetically modified and traditionally bred varieties of hard red spring wheat. In: World Allergy Organization Congress – Vancouver, September 2003. Abstract. World Allergy Organization.
- Hu, T., S. Metz, C. Chay, et al. 2003. Agrobacterium-mediated large-scale transformation of wheat (*Triticum aestivum* L.) using glyphosate selection. *Plant Cell Rep.* 21:1010–1019.
- Hyun, Y., G.E. Bressner, R.L. Fischer, et al. 2005. Performance of growing-finishing pigs fed diets containing YieldGard Rootworm corn (MON 863), a nontransgenic genetically similar corn, or conventional corn hybrids. *J. Animal Sci.* 83:1581–1590.
- Kan, C.A., and G.F. Hartnell. 2004. Evaluation of broiler performance when fed Roundup Ready® wheat (event MON 71800), control, and commercial wheat varieties. *Poultry Sci.* 83:1325–1334.
- Obert, J., W. Ridley, R. Schneider, et al. 2004. The composition of grain and forage from glyphosate tolerant wheat, MON 71800, is equivalent to that of conventional wheat (*Triticum aestivum* L.). *J. Agric. Food Chem.* 52(5):1375–1384.
- Padgette, S.R., D.B. Re, G.F. Barry, et al. 1996. New weed control opportunities: Development of soybeans with a Roundup Ready gene. In: Herbicide-resistant crops (Duke, S.O., ed.), pp. 54–84. CRC Press, Boca Raton, FL.
- Zhou, H., J.D. Berg, S.E. Blank, et al. 2003. Field efficacy assessment. *Crop Sci.* 43:1072–1075.

### Concept of yield plateau

It is important for food production to keep pace with population growth. Even though global total crop yields are continually rising, the rate of yield growth is slowing. This trend is termed **yield stagnation** or **yield plateau**, and has been observed in many of the crops that feed the world, especially cereals. For example, yield growth rates for wheat declined from 2.92% (between 1961 and 1979) to 1.78% (between 1980 and 1997). The decline in corn yield for the same periods was from 2.8% to 1.29%.

Several key factors may be responsible, at least in part, for yield stagnation. From the research perspective, many agricultural research programs do not focus on yield *per se* increases but rather on improving specific traits such as drought tolerance, insect resistance, and disease resistance. A major factor is a shifting away from cereal production to the production of more profitable crops, with the decline in world cereal prices. There are no incentives for producers to pump production inputs into an enterprise that could raise field yields, only to take a loss at harvest time. Another factor is crop intensification whereby multiple crops are being grown where previously only one was grown. This practice is suspected to cause a rapid decline in soil fertility.

The genetic potential of many important crops remains untapped. For example, the global average yield of wheat is 2 metric tons/ha compared to the record yield of 14 metric tons/ha, and even a possible 21 metric

tons/ha. Plant breeding is needed to increase crop yields by developing high-yielding cultivars for various ecotypes.

### Yield stability

Plant breeders are not only interested in developing high-yielding cultivars. They are interested in developing cultivars with sustained or stable high performance over seasons and years (**yield stability**). One of the key decisions made by the plant breeder at the end of the breeding program is the genotype to release as a cultivar. This decision is arrived at after yield trials conducted over locations, seasons, and years, as applicable. When different genotypes exhibit differential responses to different sets of environmental conditions, a genotype  $\times$  environment ( $G \times E$ ) interaction is said to occur. This subject is best fully discussed elsewhere (see Chapter 23). Genotypes that are more responsive to the environment are less stable in performance, doing well in a good production environment and poorly under less optimal production conditions. On the contrary, less responsive genotypes perform generally well across varied environments.

Stability is rather difficult to determine or breed for in a breeding program. While it might be desirable to set as an objective to breed for either more or less environmental responsiveness, it is more practical and realistic to exploit whatever turns up during yield trials. Each

kind of adaptation has its value. A cultivar that is responsive to the environment may be released for a narrowly defined area of adaptation, whereas another that has a low  $G \times E$  interaction may be suitable for release for use over a wider region of production.

### Lodging resistance

**Lodging resistance** may be defined as the resistance to leaning, bending, or breaking of the plant prior to harvesting.

#### Nature, types, and economic importance

There are two basic types of **lodging** that may be caused by biotic or abiotic factors. Lodging may originate at the root level (**root lodging**) or at the stem or stalk level (**stalk lodging**). Soil-borne disease and insect pests may destroy plant roots, causing the plant to lean over starting at the root level. Disease and insect attack can also cause the stem or stalk to lodge. For example, the European corn borer weakens the stalk and predisposes it to bending. Strong winds and other weather factors such as hail, ice, or snow, are common causes of lodging in plants.

Certain plant characteristics make plants susceptible to lodging. These include tall plant height, thin stems, excessive vegetative growth, and succulence. In addition, the crop cultural environment may encourage lodging. For example, high nitrogen fertilization promotes luxuriant vegetative growth, and hence excessive amounts produce top-heavy and succulent plants that are also prone to disease and insect attack. Stems that have been attacked by pests are weak and lodge easily.

Breeding for lodging resistance is important because lodging results in yield reduction. Lodging prior to pod filling results in partial fruit or seed development. Lodging at maturity may also make pods or cobs inaccessible to mechanized harvesters (combines) and hence left unharvested. Lodged plants are exposed to disease infestation. For example, in cereal crops such as wheat, contamination by mycotoxins produced by *Fusarium* species may increase.

#### Genetics and breeding

Breeding for lodging resistance is challenging, partly because it is a quantitative trait and hence conditioned by many genes. Furthermore, its expression is significantly impacted by the environment. It is difficult to score for lodging resistance on a phenotypic basis in the field

because the factors that cause lodging may occur at different stages in plant development, or never at all. Over the years, researchers have tried to find above-ground morphological traits that strongly correlate with lodging resistance in various species. Even though culm parameters (e.g., weight, diameter) have been reported to correlate significantly with lodging resistance in wheat, such an observation is not universal. Consequently, developing a reliable morphological index for field selection in a breeding program is difficult.

Nonetheless, resistance to lodging may be improved by targeting a combination of the following traits (depending on species) – short stature, stronger stalk, sturdy stem, thick rind, stiff culm, strong root system, resistance to stalk or stem diseases and insect pests, and other traits. It should be pointed out that no single trait or group of traits has proven to be universally reliable as an index of lodging resistance.

To improve selection, QTLs for lodging resistance have been identified and used as molecular markers in some breeding programs using marker-assisted selection. Furthermore, dwarfing genes have been used to breed short-statured cultivars in crops such as rice and wheat, resulting in increased lodging resistance. In corn, for example, recurrent selection has been used to improve lodging resistance. The success of the Green Revolution is attributed to the development of semi-dwarf cultivars that were high yielding and responsive to the environment without lodging.

### Shattering resistance

Dry fruits that split open upon maturity to discharge their seeds are called **dehiscent fruits**. Whereas shattering is advantageous in the wild, it is undesirable in modern crop production.

#### Nature, types, and economic importance

Some fruits split along only one side (called a **follicle**), while others split along two sides (called a **legume**), or multiple sides (called a **capsule**). This natural mechanism of seed dispersal has adaptive value to plants in the wild. In crop production, the splitting of dry fruits to release their seeds prior to harvesting is called **grain shattering**. In serious cases, some cultivars can lose over 90% of their seed to shattering if harvesting is delayed by just a few days. Whereas shattering is often identified with pod-bearing species (e.g., soybean, pea), it also occurs in cereal crops such as wheat and rice.

Shatter-sensitive cultivars are susceptible to high losses during harvesting. The physical contact of the harvesting equipment with the plant may be enough pressure to trigger shattering. However, most susceptible cultivars spontaneously shatter their seeds when the environmental condition is right (dry, sunny, and windy). Shattered seeds are not only lost but also become a nuisance when they germinate as volunteer plants in the next year's crop. Being weeds, the volunteer crops are controlled at additional production cost.

### Breeding grain-shattering resistance

Like lodging resistance, grain-shattering resistance is a complex trait. There is a large variation in the degree of shattering in existing rice gene pools, ranging from extremely shattering to extremely hard to thresh. Furthermore, researchers have identified at least five genes that condition shattering in rice, including *sh1*, *sh2*, and *sh4*.

## Reduced plant height

Modern production of certain cereal crops is dominated by semidwarfs or dwarf cultivars (e.g., rice, wheat, sorghum). These cultivars have advantages in mechanized agriculture and high input production systems.

### Nature, types, and economic importance

Reduced plant height is associated with, or promotes, lodging resistance. Similarly, early maturity also reduces plant internode length. Producers desire crop cultivars with reduced plant stature because they are easier to harvest mechanically. They produce less straw after the economic product has been harvested. However, in certain cultures, the straw is used for crafts or firewood, and hence tall cultivars are preferred.

Reduced stature also increases the harvest index. Dwarf cultivars can be more closely spaced in the field for increased crop yield. These cultivars are also environmentally more responsive, responding to agronomic inputs, especially fertilizers, for increased productivity.

### Genetics and germplasm resources

A variety of height-reducing genes have been discovered in various species. These genes reduce plant height when in the recessive form. In castor bean, the dwarf internode gene, *di*, reduces plant height by 25–50%. This reduction in height makes the plant sturdier and

more lodging resistant. It reduces the swaying of the pods in the wind thereby reducing shattering. In wheat, the reduced stature plants are called **semidwarfs**. The dwarfing genes in wheat are designated *Rht* (reduce height), of which about 20 have been discovered. These genes differ in their effects on the plant. For example, *Rht8*, discovered in wheat cultivars from Yugoslavia, is widely used because of its less adverse effect on kernel density and weight. Furthermore, the wheat dwarfing genes increase grain yield by increasing plant tillering and number of seeds per plant. Sometimes, wheat breeders intensify the height reduction by including two different *Rht* genes in a cultivar. Such cultivars, called **double dwarfs**, are shorter and tiller more than single dwarfs. Using monosomic analysis, scientists have associated *Rht8* with chromosome 4A and *Rht2* with chromosome 4D.

Dwarfing genes have been used in rice cultivar development. One set of genes designated *d* has been known to reduce kernel size and grain yield along with reducing internode length, and hence is not used in commercial cultivar development. Instead, the *sd1* gene, a spontaneous mutant discovered in a Taiwanese cultivar, "Deegeo-woo-gen", has been used to develop successful cultivars of rice. The gene has been induced by mutagenesis and used in cultivar development. There are several major dwarfing genes in oats, of which, so far, the dominant *Dw6* and semidwarf *Dw7* are readily available for use in breeding. However, the use of these genes in improving lodging resistance has been limited because their use results in reduced grain quality in many environments.

Sorghum breeding has also benefitted significantly from the discovery and use of dwarfing genes. Four dwarfing genes – *dw1*, *dw2*, *dw3*, and *dw4*, have been discovered. These genes produce a type of dwarfism described as **brachytic dwarfism**, which reduces plant stature without significantly affecting leaf number, leaf size, plant maturity, or yield. The gene *dw3* is observed to be mutable (to *Dw3*), resulting in one tall mutant occurring out of about every 600–1,200 plants in a field of dwarf plants. These rogues must be removed before harvesting to keep the seed as pure as possible. Furthermore, the reduction in plant stature is dependent upon the genotype with respect to the dwarfing loci. Cultivars may have one dwarfing gene (e.g., *Dw1*, *Dw2*, *Dw3*, *dw4*), two dwarfing genes (e.g., *Dw1*, *Dw2*, *dw3*, *dw4*), three dwarfing genes (e.g., *dw1*, *dw2*, *Dw3*, *dw4*), or four dwarfing genes (*dw1*, *dw2*, *Dw3*, *dw4*). Most sorghum cultivars in the USA are three-dwarfs (have three dwarfing genes). They are superior in yield

to four-dwarfs. It is also known that the manifestation of the dwarf trait is influenced by modifying genes such that cultivars with an identical set of dwarfing genes vary in plant stature.

Plants that have dwarfing genes are unable to respond to gibberellin or gibberellic acid, the plant hormone that promotes stem elongation. Using transposon tagging, scientists have cloned a gene, *GAI*, in *Arabidopsis* that is responsible for dwarfism in that species. DNA sequence analysis of the gene followed by bioinformatics has revealed that other species have equivalent genes. This indicates the possibility of using rDNA technology to transfer the *GAI* gene into other species.

### Breeding determinacy

Another way in which plant form is manipulated is through determinacy.

#### Nature, types, and economic importance

Plant growth form may be **indeterminate** or **determinate**. Some species such as corn and wheat are determinate in growth form, whereas soybean has both determinate and indeterminate types. In indeterminate cultivars, new leaves continue to be initiated even after flowering has begun. Flowering occurs in axillary racemes. On the other hand, determinate cultivars or species do not initiate new leaves after flowering has begun. Also, flowering occurs in both axillary and terminal racemes.

Generally, determinate plant types tend to have stiffer and shorter stems (called **bush types**). In soybean, both determinate and indeterminate commercial cultivars are used in production. Determinate cultivars are used in areas with long growing seasons (i.e., southern latitudes), while indeterminate cultivars are grown in areas with short growing season (i.e., northern latitudes).

#### Genetics and germplasm resources

Determinacy in soybean is conditioned by a recessive gene, designated *dt<sub>1</sub>*. This gene has been used to breed determinate cultivars for the northern growing regions. Semideterminacy is conditioned by the *Dt<sub>2</sub>* allele, while the *S* allele conditions short plant internodes.

### Photoperiod response

Plants exhibit environmentally determined developmental switches from the initiation of leaves (vegetative

phase) to flowering (reproductive phase). The two developmental switches that plant breeders pay attention to are **photoperiod** and **vernalization**. The key environmental factors are temperature and day length. In some plants, flowering is not promoted by temperature and day length but occurs regardless of the conditions (called facultative plants), whereas in others, flowering will not occur without the appropriate temperature–day length combination (called obligate plants). Photoperiodism is a photomorphogenic responsive to day length (actually, plants track or measure the duration of the dark period rather than the duration of day light). Three categories of responses are known.

- 1 Long-day (short-night) plants.** These plants require a light period longer than a certain critical length in order to flower. They will flower under continuous light. Cool season species (e.g., wheat, barley, alfalfa, sugar beet) are examples of long-day plants.
- 2 Short-day (long-night) plants.** These plants will not flower under continuous light, requiring a photoperiod of less than a certain critical value in a 24-hour daily cycle. Examples include corn, rice, soybean, peanut, and sugarcane.
- 3 Day-neutral (photoperiod-insensitive) plants.** Photoperiod-insensitive plants will flower regardless of duration of day length. This trait is very desirable, enabling producers to grow the crops in a broad range of latitudes. Examples include tomato, cucumber, cotton, and sunflower.

Plant breeders need plants to flower at the appropriate time for hybridization in a breeding program, and also to influence the cultivars they develop for different growing areas. Photoperiod influences the duration of the vegetative phase versus reproductive phase and hence crop yield at different latitudes. Soybean, for example, has 13 recognized adaptation zones, ranging from 000 in the northern latitudes to IX in the southern latitudes. Day length increases as one goes north in summer in the northern hemisphere. Consequently, a cultivar developed for the southern latitudes may not be as productive in the northern latitudes where reproductive growth is not initiated until the fall season when day length is short.

Vernalization is the process by which floral induction in some plants is promoted by exposing plants to chilling temperatures for a certain period of time. Plant breeders of crops such as wheat either sow in the fall so the plant goes through a natural vernalization in the winter, or they place trays of seedlings in a cold chamber for the same purpose, prior to transplanting.

### Nature, types, and economic importance

Photoperiod and temperature are two major environmental factors that influence crop adaptation through their effect on days to flowering. Photoperiod is also known to affect photosynthate partitioning in some species such as peanut, in which researchers found a reduction in the partitioning of dry matter to pods, in certain genotypes, under long photoperiods. Decreasing partitioning to grain favors partitioning to vegetative parts of the plant, resulting in increased leaf area and dry matter production. Crop cultivars that are developed for high altitudes should mature before the arrival of winter as well as be less sensitive to long photoperiods so that seed yield is high.

### Genetics and germplasm resources

In soybean, a number of maturity genes that influence flowering under long day length have been reported. Of these, the  $E_3$  locus has the most significant effect on flowering under long day length. Plants with genotypes of  $e_3e_3$  have insensitivity to fluorescent long day length. The genetic control of photoperiod insensitivity in wheat is variable among cultivars, including two genes with major effects, one dominant gene, or one recessive gene controlling the condition. In sorghum, four genes ( $Ma_1$ ,  $Ma_2$ ,  $Ma_3$ ,  $Ma_4$ ) affect plant maturity and response to long day length. As J. R. Quimby observed, the recessive alleles of these genes condition a degree of sensitivity to longer day length. The genotype did not affect the time of initiation of flower primordial under short days (10 hours). However, under long days (14 hours), the genotype  $Ma_1Ma_2ma_3Ma_4$  in sorghum induced flowering in 35 days, while  $Ma_1ma_2Ma_3Ma_4$  flowered in 44 days. The genotype  $Ma_1Ma_2Ma_3Ma_4$  flowered in 70 days.

### Early maturity

Crop maturity in general is affected by a variety of factors in the production environment, including photoperiod, temperature, altitude, moisture, soil fertility, and plant genotype. Early maturity could be used to address some environmental stresses in crop production such as drought and temperature. Maturity impacts both crop yield and product quality. Sometimes, the producer desires the crop to attain its maximum dry matter under the production conditions, before harvesting is done. In some crops, premature harvesting produces the product quality for premium prices.

### Nature, types, and economic importance

There are two basic types of maturity – **physiological** and **harvest (market) maturity**. Physiological maturity is that stage at which the plant cannot benefit from additional production inputs (i.e., inputs such as fertilizer and irrigation will not translate into additional dry matter or gain in economic product) because it has attained its maximum dry matter. In certain crops, the product is harvested before physiological maturity to meet market demands. This stage of maturity is called harvest maturity. For example, green beans are harvested before physiological maturity to avoid the product becoming “stringy” or fibrous.

It is desirable for a producer to grow a cultivar that fully exploits the growing season for optimal productivity. However, under certain production conditions, it is advantageous for the cultivar to mature early (i.e., exploit only part of the growing season). Early maturity may allow a cultivar to escape environmental stresses (e.g., disease, insects, early frost, early fall rain storms, drought) that may occur later in the season. Also, early maturing cultivars are suitable for use in multiple cropping systems, allowing more than one crop to be grown in a production season. Early maturity has made it possible to extend the production of some crops to regions in higher altitudes and with shorter summers, as well as with low rainfall.

Early maturity has its disadvantages. Because the plant only partially exploits the growing season, economic yield may be significantly reduced in species including corn, soybean, wheat, and rice. In cotton, earliness is negatively correlated with traits such as fiber length.

### Genetics and germplasm resources

Research reports indicate earliness to be controlled by either dominant or recessive genes. Modifier genes of major genes have also been reported. The inheritance appears to differ among species. Researchers have conducted QTL analysis for improvement of earliness. Furthermore, a flowering promoting factor (*fpp1* gene) has been cloned in *Arabidopsis*. This gene is being used to transform other species to create early maturity in those species. In sorghum, four gene loci,  $Ma_1$ ,  $Ma_2$ ,  $Ma_3$ , and  $Ma_4$ , influence time to maturity. These are the same genes that are implicated in photoperiod sensitivity. Tropical sorghum species are dominant at the  $Ma_1$  locus, and do not flower during the long summer photoperiods in the USA.

## References and suggested reading

- Frey, K.J. 1959. Yield compensation of oats II: Effect of nitrogen fertilization. *Agron. J.* 51:605–608.
- Lee, E.A., A. Ahmadzadeh, and M. Tollenaar. 2005. Quantitative genetic analysis of the physiological processes underlying maize grain yield. *Crop Sci.* 45:981–987.
- Milach, S.C.K., and L.C. Federizzi. 2001. Dwarfing genes in plant improvement. *Adv. Agron.* 73:35–65.
- Rasmussen, D.C., and B.G. Gengenbach. 1983. Breeding for physiological traits. In: *Crop breeding* (Wood, D.R., ed.), pp. 231–254. American Society of Agronomy, Madison, WI.
- Rosielle, A.A., and J. Hamblin. 1981. Theoretical aspects of selection for yield in stress and non-stress environments. *Crop Sci.* 21:943–946.
- Simmonds, N.W. 1979. *Principles of crop improvement*. Longmans, New York.
- Worland, A.J. 1996. The influence of flowering time genes on environmental adaptability in European wheats. *Euphytica* 89:49–57.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 All yield is biological, but all yield is not economic.
- 2 Sinks are exporters of substances.
- 3 Biomass is the yield per unit area of the grain.
- 4 The higher the harvest index, the more efficient the plant is in directing assimilates to the part of the plant of economic value.
- 5 Lodging may occur at either the root or stem levels.
- 6 Grain shattering occurs only in legumes.
- 7 Most sorghum cultivars in the USA are two-dwarfs.
- 8 The dwarfing gene in wheat is designated *Rht*.
- 9 A bush-type cultivar is indeterminate in growth form.

### Part B

Please answer the following questions:

- 1 There are two types of lodging: ..... or .....
- 2 Give the genotype of three-dwarf sorghum.
- 3 Distinguish between single-dwarf and double-dwarf sorghum cultivars.
- 4 ..... is the proportion of the plant that is of economic value.
- 5 What is yield?
- 6 Distinguish between economic yield and biological yield.
- 7 What is yield potential?
- 8 Define harvest index.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the genetics of dwarfing in sorghum.
- 2 Distinguish between determinate and indeterminate growth in plants.
- 3 Discuss the importance of lodging resistance in crop production.
- 4 Discuss the concept of plant ideotype.
- 5 Discuss the concept of yield components.
- 6 Discuss the importance of yield stability in cultivar development.
- 7 Discuss the concept of yield component composition.



# Breeding for resistance to diseases and insect pests

---

## Purpose and expected outcomes

*Plants are plagued by a host of insects and pathogens that often must be controlled or managed in crop production. To control a parasite effectively requires an understanding of its biology, epidemics, spread, and the damage it causes. A variety of methods are used in pest control, each with advantages and disadvantages. These methods are chemical, biological, cultural, legislative, and physical controls. A specific tactic in the method of biological pest control is the use of disease-resistant cultivars in crop production. Breeding for disease and insect resistance is one of the primary objectives in plant breeding programs. After completing this chapter the student should be able to:*

- 1 Discuss the economic importance of plant pests.
  - 2 Discuss the genetic basis of disease resistance.
  - 3 Compare and contrast resistance breeding strategies.
  - 4 Discuss the role of wild germplasm in disease- and insect-resistance breeding.
  - 5 Discuss specific applications of biotechnology in plant breeding to control pests.
  - 6 Discuss disease epidemics and their breeding implications.
- 

## Groups of pests targeted by plant breeders

Plant diseases are caused by pathogens that vary in nature and may be microscopic or readily visible (e.g., virus, plant, animal). These pathogens may be airborne or soil-borne. Six general groups of causal agents of disease, which represent six general approaches to breeding for pest resistance, may be identified as: **airborne fungi**, **soil-borne fungi**, **bacteria**, **viruses**, **nematodes**, and **insects**. Through an understanding of the biology, epidemics, spread, and damage caused by these organisms in each category, breeders have developed certain strategies and methods for breeding cultivars to resist certain types of biotic stress in plant production.

It should be pointed out that plant breeders have devoted varying amounts of resources to breeding for resistance in these categories with varying degrees of success. Plant species vary in their susceptibility to

diseases caused by pathogens or pests in each group. Some crop production pests are conspicuously absent from the list because they are relatively unimportant (e.g., mites) or are not practical to breed against (e.g., birds).

Cereal crops tend to have significant airborne fungal disease problems, while solanaceous species tend to experience viral attacks. Diseases that afflict crops of world importance and cause major economic losses and are readily transmitted across geographic boundaries receive funding from major donors. Breeding for resistance to fungi, especially airborne fungi, is the most prominent resistance breeding activity. N. W. Simmonds has suggested that the relative importance of the six groups of pathogens of importance to plant breeders, might be something like this: **airborne fungi > soil-borne fungi > viruses > bacteria = nematodes = insects**.

Weed control is a major activity in crop production. However, crop tolerance to weeds is seldom, if ever, a breeding objective, and hence is not discussed in this chapter.

### **Biological and economic effects of plant pests**

Disease is basically a change from the state of metabolism necessary for the normal development and functioning of an organism. An abnormal growth and development of the plant will cause a reduction in biological yield and invariably in economic yield, hence, the need to control pathogens and insect pests in crop production. There are four basic ways by which diseases and insect pests adversely affect plant yield and general performance, and, eventually, reduce economic value.

- 1 Complete plant death.** Certain parasites sooner or later will completely kill the afflicted plant. When this occurs early in crop production, gaps are left in the crop stand. Where such gaps cannot be compensated for by plants in the vicinity, additional nutrients may become available to the existing plants due to reduced competition. However, the reduced plant density results in reduced biomass. Crop stand-reducing pests include those that cause mildews, vascular wilts, and insects such as cutworms that cause a seedling to fall over and die.
- 2 Stunted growth.** Viruses are known to reduce the metabolic performance of plants without killing them outright. Afflicted plants grow to only a fraction of their normal size, and usually cause severely reduced economic yield of plants.
- 3 Partial plant death.** Some diseases that afflict adult plants do not completely kill them. Rather, only certain parts of the plant (e.g., branches) are killed (e.g., as observed in fungal diebacks).
- 4 Direct product damage.** The three effects of disease discussed above indirectly affect the economic or harvested product if it is the fruit, seed, or modified organs (tubers, bulbs). Some pests directly injure these products completely (e.g., by causing rotting of tissue) or reducing quality (e.g., by causing blemishes, holes).

It is a principle in pest control that a method of control (especially a chemical method) is warranted only when economic loss is eminent. It does not make economic sense to spend \$1,000 in pest control to save \$100 of harvestable product.

### **Overview of the methods of control of plant parasites**

Four strategies are available for controlling plant pathogens and pests, as summarized here:

- 1 Exclusion of pathogen from the host.** This strategy may use methods such as legislation (plant quarantine, crop inspection) or crop isolation to prevent the pathogen or pest from making initial contact with the host plant. Quarantine laws help to prevent the introduction of new pathogens into a production region. The laws are commonly enforced at the points of entry for people and goods into a country.
- 2 Reduction or elimination of the pathogen's inoculum.** In the event that a pathogen gains access into a production area, various methods may be used to remove it or reduce the inoculum to contain it. A method such as crop rotation reduces disease buildup in the field, while observance of sanitation (e.g., removing diseased plants and burning them) also reduces the spread of the pathogen. The producer may also implement management practices that discourage the growth and spread of parasites (e.g., soil drainage, weeding, soil sterilization, seed treatment).
- 3 Improvement of host resistance.** This is the strategy of most concern to plant breeders. It entails breeding to introduce genetic resistance into adapted cultivars. This is the primary subject of this chapter.
- 4 Protection of the host.** Economic plants may be protected from parasites by using chemicals (pesticides). While this is widely used, the method is environmentally intrusive and expensive.

### **Concept of resistance in breeding**

Breeding for yield and other morphological traits, as well as breeding for resistance to abiotic stress, are conceptually different from breeding for resistance to biotic stresses. Breeding in the former cases entails the manipulation of one genetic system – plants. Breeding for resistance to biotic stresses on the other hand involves the manipulation of two genetic systems – one for plants (host) and the other for the organism (parasite) – not independently, but with regard to the interaction between the two systems. The breeder needs to understand the interrelationships between plants and their parasites that have persisted through coevolution and coexistence.

**Resistance** is a response to a cause. There are degrees of resistance. In terms of disease, resistance is always

relative to a benchmark. A genotype that is immune to a disease is completely or totally resistant to it. In practice, plant breeders arbitrarily set the susceptibility–resistance scale. A resistant cultivar has *less disease* than the standard cultivar – it could be a little or a lot less diseased. True disease resistance has a genetic basis and hence is amenable to plant breeding methodologies. It is manifested in two basic forms: inhibition of infection or inhibition of subsequent growth of the pathogen, the former being the more common form of resistance. Resistance can be a qualitative or quantitative trait.

As A. L. Hooker observed, breeding for pest resistance is an integral part of any system of crop pest control. If yield and other desirable traits are maintained, resistance (host resistance) is an ideal method of breeding to control plant pests. The development and use of resistant cultivars has several advantages. It is relatively inexpensive to deploy (versus use of other pest control measures) and has no adverse environmental consequences. Also, as Hooker indicated, it is convenient for farmers to use, requiring no additional production costs or decision-making during the production season. However, when host resistance is not satisfactory, farmers may need to supplement host resistance with other pest management practices (e.g., use of pesticides).

## Concepts of pathogen and host

### The pathogen

The **pathogen** is a living organism that is capable of inflicting a distinct disease or disorder in another organism (the host). The capacity of the pathogen to cause disease or disorder in a member of a host species is called its **pathogenicity**, while the extent of disorder or disease development it causes is its **virulence**. The pathogenicity and virulence of a pathogen vary among pathogen types (**races** or **pathotypes**). Races or pathotypes that fail to cause disease symptoms or successfully attack a given host are said to be **avirulent**. It should be pointed out that the presence of a pathogen on a susceptible host is not enough to cause disease symptoms to occur. A third factor – favorable environment – is needed, the trio (pathogen plus susceptible host plus favorable environment) referred to as the **disease triangle**. Pathotypes or races of pathogens may also be described in terms of **aggressiveness** or **non-aggressiveness** in relation to the rate at which they produce disease symptoms.

### The host

The host (genotype, plant) is the organism in which a pathogen may produce disease symptoms. A susceptible host is one in which a pathotype or race can manifest a disease symptom. A host may employ one of several mechanisms (defense mechanisms) to resist pathogens. These mechanisms may be pre-existing or induced upon infection.

- 1 **Pre-existing defense mechanisms.** These include morphological features that pose as barriers to the penetration of the pathogen into the plant (e.g., presence of lignin, cork layer, callose layers), or secondary metabolites (phenols, alkaloids, glycosides) that have antimicrobial properties.
- 2 **Infection-induced defense mechanisms.** Upon infection, the host quickly produces chemical products (e.g., peroxidases, hydrolases, phytoalexins, etc.) to combat the infection.

## Mechanisms of defense in plants against pests

Plants exhibit a wide variety of strategies and mechanisms of defense against pathogens and insects pest that may be classified into three major groups – **avoidance**, **resistance**, and **tolerance**. The mechanisms are given different names as pertain to insect pests, as will be discussed.

### Avoidance

Also described as **escape**, avoidance is a mechanism that reduces the probability of contact between pathogens or insect pests and the plant. That is, it operates prior to the establishment of any intimate host–pest/pathogen contact. This mechanism is rare in pathogens and hence is primarily applicable to insect pests where entomologists describe it as **antixenosis** (or **non-preference**). A plant using this mechanism has characteristics (morphological or chemical) that make it unattractive to insects for feeding, oviposition, or shelter. In the cabbage, a variant leaf color makes the plant unattractive to the cabbage aphid (*Brevicoryne brassicae*), while pubescence in cotton prevents oviposition in the cotton boll weevil (*Heliothis zea*). Also, in cotton, a repulsive odor deters the boll weevil from feeding on the plant.

Avoidance or non-preference may not be effective if there are no alternative hosts. Non-preference

conditioned by allelochemicals is not as effective in modern agriculture where monocropping and breakdown of resistance are common. However, morphological traits that interfere with feeding and breeding of insects are important breeding goals as a first line of defense against insect pests. Crop producers may help plants escape pest attack by timing the planting of the crop (early or late planting) so that the susceptible growth stage escapes the time when the insect pest is most abundant.

### Resistance

The mechanism of resistance manifests after a host has been attacked by a pathogen or insect pest. The mechanism operates to curtail the invasion or to reduce the growth and/or development of the pathogen. The nature of the mechanism may be biochemical, physiological, anatomical, or morphological. The equivalent term used to describe this mechanism as it relates to insect pests is **antibiosis**. As previously discussed, resistance is a relative term for the genetic-based capacity of a host to reduce the adverse effect of a pathogenic attack. It does not imply complete overcoming of the pathogen and may be quantitative in its expression (i.e., some genotypes are more resistant than others in the face of equal amount of initial inoculum). Further, host-resistance mechanisms may be pre-existing (called **passive resistance**). For example, onion is resistant to smudge disease (caused by *Colectotrichum circinaus*) because of the presence of various chemicals (e.g., catechol) in its outer scales. The resistance may also be induced or activated (called **active resistance**) in response to an infection.

Resistance reactions may also be described as hypersensitive, overdevelopment, or underdevelopment.

**1 Hypersensitive reaction.** This mechanism acts to prevent pathogen establishment. An infection evokes a rapid localized reaction whereby the cells immediately surrounding the point of attack die so that the infection is contained. The pathogen eventually dies, leaving a necrotic spot. This infection inhibition response is the most common and is evoked in diseases such as leaf spots, blights, cankers, and decays.

**2 Overdevelopment of tissue.** Meristematic activity may be stimulated, resulting in excessive abnormal tissue growth as exemplified by galls and also leaf curling. In some cases, a layer of suberized cells form around the invaded tissue to curtail the spread of the pathogen.

**3 Underdevelopment of tissue.** Afflicted plants become stunted in growth, or organs become only partially developed. Viral infections produce this kind of effect.

### Tolerance

Unlike avoidance and resistance mechanisms that operate to reduce the levels of infection by the pathogen or pest, tolerance (or endurance) operates to reduce the extent of damage inflicted. The afflicted host attempts to perform normally in spite of the biotic stress. A host may be highly susceptible (i.e., support a high population of pathogens or pests) and yet exhibit little reduction in economic yield (i.e., the plant host is tolerant of the parasite). Because tolerance is measured in terms of economic yield, it is not applicable to diseases or insect pests that directly attack the economic part of the plant (grain or fruit). The mechanism may be applied to a situation in which a plant recovers quickly following a pest attack such as grazing. From the point of view of virologists, plants with few or no symptoms are described as tolerant.

The exact mechanism(s) of tolerance is not understood. However, some attribute it to plant vigor, compensatory growth, and changes in photosynthate partitioning, among other factors. Because it is difficult to determine the yield reduction per unit of infection, it is very difficult to determine the difference in tolerance between genotypes.

### Specificity of defense mechanisms

Almost all plant species are non-hosts for almost all parasites. Most avoidance mechanisms are non-specific (general or broad) in that they are capable of avoiding groups of parasites with similar ecological requirements. For example, the closed flowering habit of some barley cultivars that excludes pathogens that invade through the open flower, such as *Ustilago nuda* (causes loose smut), also excludes *Claviceps purpurea* (causes ergot). However, some specificity may occur in the case of certain repellents. Resistance mechanisms may also be specific or non-specific. Specificity may not only be at the pathogen level but even at the race level (cultivar × race interaction). Race non-specific resistance can be highly pathogen-specific. Plant breeders are primarily interested in resistance in the host-pathogen system that has the race-specific and race-non-specific effects.

## Types of genetic resistance

The complexity of host-pathogen interaction makes it difficult to categorize resistance into finite types. A large number of host-pathogen interaction systems occur at various stages of coevolution. Nonetheless, some rough categorization, albeit simplistic, is widely acceptable for its convenience and practical plant breeding applications. Resistance reactions may be generally categorized into two major kinds – **vertical** or **horizontal** – based on their epidemiological status and stability of resistance. Vertical (specific) or horizontal (non-specific) effects are expressed in terms of population dynamics, not in terms of genes.

### Vertical resistance

Vertical resistance is known by other terms including **hypersensitivity**, **race- or pathotype-specific resistance**, **major gene** or **oligogenic resistance**, **non-durable** or **qualitative resistance**, and **non-uniform resistance**. This reaction is said to occur when a race of a pathogen produces disease symptoms on some cultivars of a host but fails to do so on others. That is, there is a complete display of resistance in some genotypes while others are susceptible. Vertical resistance is a differential resistance, or a host × isolate interaction of cultivars against different races of a pathogen.

However, there are examples of resistance genes that the pathogen apparently fails to match (e.g., dominant resistance in *Periconia cincinata*, the causal agent of the milo disease of sorghum). Major gene resistance is often absolute, conferring immunity against the appropriate pathotype (but not necessarily so). It delays the start of an epidemic, and disappears when the pathotype virulent to the vertical resistance gene is prevalent. This type of resistance is relatively easy to breed because the major genes are easy to identify and transfer through simple crosses. These genes control specific races or genotypes of pests and hence do not protect against new races of the pests. Cultivars equipped with this narrow-based genetic protection are often vulnerable when a genetically variable population of a pathogen occurs in the environment. They may not succumb immediately to the new strains of the pathogen but under certain cultural systems (e.g., monoculture without crop rotation), the populations of the new strains could build up high enough to cause economic injury to the crop. This form of protection is also described as **non-durable resistance**. Disease or insect resistance may be conditioned

by dominant or recessive genes. Dominant genes occur in resistance to, for example, rusts, viruses, mildews, and nematodes.

### Horizontal resistance

Horizontal resistance is also known as **slow rusting**, **partial resistance**, **field resistance**, **race-non-specific resistance**, **minor gene reactions**, or **polygenic inheritance**. There is a lack of differential reaction of isolates on different cultivars. The resistance is effective against all genotypes of the parasite species without cultivar × isolate interaction (i.e., race-non-specific). Following the initial establishment of the pathogen, the plant may resist its spread and reproduction such that the disease develops at a slower rate. Unlike vertical resistance, evolution of new races due to selection pressure against any specific race is absent under horizontal resistance, making this type of resistance more stable.

Horizontal resistance is controlled by polygenes. Each of the genes that condition the disease contributes toward the level of resistance, and hence resistance is also called **minor gene resistance**. Polygenic resistance is widespread and provides protection against a wide range of races of a pest and hence is less susceptible to being overcome by a new race. Polygenic resistance protects plants by slowing down the spread of disease and the development of epidemics in the field. Breeding polygenic resistance is more challenging. The many minor genes cannot be individually identified and consequently cannot be transferred through crossing in a predictable fashion. However, the heritability for polygenic resistance is high.

### Durability of resistance

An aspect of resistance of concern to plant breeders is the **durability**. Resistance that is durable is one that remains effective in a widely grown cultivar for a long time in an environment that favors the parasite. Horizontal resistance is said to be durable whereas vertical resistance is not. Adaptations of the parasite to newly introduced resistant cultivars (i.e., breakdown of resistance) occurs frequently in pathogens but less so in insect pests. Such cultivars lack durable resistance. Durability of resistance is variable even within pathosystems. For example in one study involving resistance to stripe rust in wheat, durability ranged from 1 to 18 years, depending on the cultivars.

## Genetics of host-pathogen reactions

### Genetics of resistance

R. H. Biffen is credited with providing the first report on the genetics of resistance. Working on stripe rust (*Puccinia striiformis*), he reported that resistance to disease was controlled by a single Mendelian gene. Since then many reports have supported this finding that much host resistance to pathogens is simply inherited. Dominance gene action is very common for resistance to pests and pathogens (especially hypersensitive resistance). Monogenic recessive resistance is less common. However, it is known that resistance may be controlled by any number of genes whose effects may be large or small. Further, the genes may interact epistatically or additively. Commonly, resistance and virulence genes usually operate on a gene-for-gene basis as discussed next. Also, genes for resistance are often reported to be clustered in linkage groups. Vertical resistance or race-specific resistance is oligogenic. Other examples are the *Sr* (for stem rust), *Lr* (for leaf rust), and *Pr* (for powdery mildew) genes of wheat (cause hypersensitive reaction for resistance), and the *Dm* (for downy mildew) genes of lettuce.

The inheritance of horizontal resistance is more complex than for vertical resistance. Resistance is often polygenic, such as the partial resistance in potato to *Phytophthora infestans* and in maize to *Puccinia sorghi*. Any number of genes could be involved in horizontal resistance. It is a stable resistance because a pathogen has to change for a number of genes to overcome the defense mechanism in that one host (non-specific). Horizontal resistance may arise when: (i) the host genes do not operate in a gene-for-gene fashion with the pathogen genes (no differential interactions are possible); or (ii) when several to many host genes with small effects operate on a gene-for-gene basis with an equivalent number of genes in the pathogen population (differential effects are too small to be detectable and appear to be horizontal resistance).

Vertical resistance and horizontal resistance may be viewed as extremes of a continuum, if one assumes that gene-for-gene relationships are common in host-specific pathogen systems. When few genes with large effects occur, differential interactions are readily identified and the result is vertical resistance. However, when more genes with smaller effects are involved in the reaction, differential interactions are less readily identified, resulting in a mixture of vertical and horizontal resistance. When numerous genes of small effect are involved,

differential reactions cannot be discerned, and the resistance is primarily horizontal resistance.

### Gene-for-gene reactions (genetics of specificity)

Working on flax rust (caused by *Melampsora lini*), H. H. Flor discovered that the major genes for resistance in the host interacted specifically with major genes for avirulence in the fungus. He proposed the gene-for-gene concept that states that for each gene conditioning resistance in the host, there is a specific gene conditioning virulence in the parasite. In other words, a resistance gene can act only if a locus in the pathogen carries a matching allele for avirulence. By themselves, resistance genes are ineffective if the pathogen carries the appropriate allele for virulence (usually a recessive allele). Most researchers agree that the specific interaction occurs between dominant resistance alleles and the dominant avirulence alleles (Figure 20.1).

The original concept proposed in 1950 by H. H. Flor suggests that the virulence of a pathogen and the resistance of a host have a genetic basis. For each gene that confers resistance in the host, there is a corresponding gene in the pathogen that confers virulence to the pathogen, and vice versa. This host-pest interaction is called the **gene-for-gene hypothesis**. J. A. Browning simplified the hypothesis by comparing the interaction between the genetic systems of the host and pathogen to a set of locks and keys. The host has the locks – each locus that conditions resistance to a pathogen is a potential for a lock. The locks are equivalent to dominant alleles that inhibit pathogen establishment in the host (plant). To succeed, the pathogen must have the appropriate key (a virulent allele) to open the lock, or the host must have no lock for the key. For example, a genotype *AABBcc* in a plant requires the pathogen to have the genotype *aabb-* to be susceptible. The third allele can be *C* or *c*, since there is no lock in the host.

Host	Pathogen			
	<i>AB</i>	<i>Ab</i>	<i>Ab</i>	<i>ab</i>
<i>rrss</i>	+	+	+	+
<i>RRss</i>	-	+	- <sup>a</sup>	+
<i>rrSS</i>	-	-	+	+
<i>RRSS</i>	-	-	-	+

‘a,’ differential interaction

**Figure 20.1** Demonstration of the concept of gene-for-gene interaction in disease resistance.

## General considerations for breeding resistance to parasites

In spite of the advantages of host resistance as a plant breeding approach to controlling pests, the approach is not always the best choice, for several reasons.

- 1 Breeding is an expensive and long-duration undertaking that makes it only justifiable (from at least the economic standpoint) for major pests that impact crops that are widely produced or have significant benefit to society.
- 2 Natural resistance is not available for all pests. Sometimes, the resistance is available in unadapted gene pools, requiring additional costs of prebreeding.
- 3 Breeding for resistance varies in ease and level of success from one pest to another. Resistance to vascular pathogens, viruses, smuts, rusts, and mildews is relatively easier to breed than breeding against pathogens that cause rots (root rot, crown rot, storage rot) and ectoparasitic nematodes. Similarly, it is relatively easier to have success with breeding resistance to aphids, green bugs, and hoppers, than to breed resistance to root-chewing or grain-storage pests.
- 4 Instability of pest resistance is a key consideration in breeding for pest control because diseases and insect pests continue to change. New pathogenic races may arise, or the cultural environment may modify the resistance of the cultivar.
- 5 The techniques of biotechnology may be effective in addressing some breeding problems more readily than traditional methods.
- 6 After being satisfied that breeding for disease resistance is economical, the breeder should select the defense mechanism that would be most effective for the crop, taking into account the market demands. For example, horticultural products and produce for export usually require that the product to be free from blemish. In these cases, breeding for major gene resistance with complete expression is desirable. It is also easy to breed for this type of reaction. However, the breeders should note that this resistance is not durable.
- 7 When a crop is grown for food or feed, breeding for mechanisms that increase the levels of chemical toxins in plant tissues is not suitable.

## Challenges of breeding for pest resistance

Breeding for pest resistance differs fundamentally from breeding for other traits because, as A. L. Hooker

observed, the breeding objectives may involve a change in a relationship with an evolving and variable pest or pathogen population. Further, the genes for resistance cannot be identified unless the plant containing the genes is interacting with the pathogen or insect pest in an environment in which susceptible plants normally would be injured or succumb to disease. Plant breeders breed or, as on other occasions, must develop a segregating population with adequate diversity to include the desired combination of genes of interest. In disease breeding, the challenge is identifying and selecting a desired genotype in a form that would be genetically stable for eventual release as a cultivar.

The breeder must utilize reliable methods for detecting differences in pest reaction among segregants, in order to avoid selecting and advancing escapes. Whereas natural infection may be used, artificial inoculation is frequently more reliable.

Hooker also recognized that a major problem in breeding for resistance to disease and insect pests is the dynamic nature of the problem. Changes in the cultural environment of the crop as well as in the genotype (crop and pathogen) are responsible for such shifts. Some pathogens evolve continuously, making it necessary for breeders to breed pest resistance to ensure stability of crop production, by preventing the development of destructive epiphytotics and infestations, and to reduce the annual loss of products from pests.

## Resistance breeding strategies

Disease breeding is a major objective for plant breeders. It is estimated that 95–98% of cultivars of small grains grown in the USA have at least one gene for disease resistance. Further, an estimated 75% of cultivated land in US crop production grows a crop with at least one disease-resistance gene. It should be pointed out that a combination of traits rather than just one trait, makes a cultivar desirable. Yield and resistance to disease are top considerations in breeding programs, the former trait being usually the most important breeding objective.

## Specificity in the parasite

Another distinction between breeding for pest resistance and breeding for other traits is that pest resistance is influenced by genetic variability of the pest population, especially, in diseases. When pathogen genotypes share a group of cultivars to which they are virulent, they are said to belong to a **physiological race**. The

equivalent term applied to viruses is **strain**, while it is **biotype** in insects and **pathotype** in nematodes. It should be pointed out that these terms have other usage in other contexts. Physiological races of pathogens occur in rusts, powdery mildew, and some insects; the physiological races may be identified by using **differential cultivars** (contain known genes for disease reaction). Breeders use a series of differentials to determine what genes would be most effective to incorporate into a cultivar. The concept of differentials stems from the ability of a cultivar to differentiate between races of a parasite on the basis of disease reaction. If a cultivar has resistance to one race but is susceptible to another, it has differential properties to identify two races of a pathogen and hence is called a **differential**. In the case of two categories of reaction (R = resistant, S = susceptible),  $n$  differentials may be used to differentiate  $2^n$  races of a pathogen (where  $n = 2$ ). In that case, four ( $2^2$ ) races of a pathogen can be differentiated as follows:

	Races				
	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	
Cultivars	C <sub>1</sub>	-	-	+	+
	C <sub>2</sub>	-	+	-	+

The ideal set of differential cultivars is one in which each cultivar carries a gene for resistance to only one race.

It should be pointed out that physiological races are an abstraction, since they are not pure biotypes of an organism, but simply groups of genotypes that express the same reaction upon inoculation over a set of differential cultivars determined by experimentation. The differential cultivars provide some information on the virulence characteristics present in resistance to which the pathogen population carries avirulent genes, and, similarly, the genes for resistance of the host that would fail because the pathogen possesses the necessary genes for virulence.

### Planned release of resistance genes

Strategy is important in breeding for resistance to parasites. However, sometimes, breeders just simply focus on developing cultivars with good resistance to a disease of economic importance for a specific growing region. The result is that all cultivars developed have the same resistance genes because it is the best available source to breeders. It is recommended to have a planned release (consecutive release of different resistance genes) of resistance genes so that only one or a few are used in agricultural production at one time. The virulence gene

composition of the pathogen population should be monitored annually using a differential series of host genotypes that carry different resistance genes either one at a time or in some combination. Once a current cultivar succumbs to a new race of a pathogen (i.e., a new race that is virulent with the resistance gene in use), breeders then release new cultivars that carry another effective gene. This way, plant breeding stays ahead of the pathogen.

### Application of gene pyramiding

The concept of transferring several specific genes into one plant is called **gene pyramiding**. Because there are different races of pathogens, plant breeders may want to transfer a number of genes for conferring resistance to different races of a disease into a cultivar. Three major genes conditioning resistance to blast in rice, *Pi-1*, *Pi-2*, and *Pi-3*, were pyramided through pairwise crosses of the isogenic lines carrying the genes. Similarly, genes for resistance to biotype L of the Hessian fly (*Motyella destructor*) of wheat were successfully pyramided into the crop. This strategy is applicable to regions where plant breeding is centrally coordinated, and where the production region using the cultivar with the multiple resistance genes is isolated from other areas not using the system. Simultaneously releasing cultivars with single resistance genes along with the multiple gene cultivars would reduce the success of the latter approach.

### Using multilines

The rationale of a multiline is that a host population that is heterogeneous for resistance genes would provide a buffering system against destruction from diseases.

### Breeding multilines

Multiline breeding procedures were discussed in Chapter 16.

### Breeding for oligogenic and polygenic resistance

The breeder should guard against breeding highly resistant cultivars that have no economic worth. A good strategy is to breed for **middling resistance** with high yield. To this end breeding for polygenic horizontal resistance is the most desirable strategy since it accounts

for most middling resistance. It should be pointed out that some single-gene resistance events are also effective and do not confer immunity on the cultivar.

### Breeding for horizontal resistance

As previously discussed, horizontal (general) resistance is pathotype-non-specific, polygenic, and not hypersensitive. Its expression is quantitative and stable. It is suitable for both annuals and perennials and is applicable to all pathogens. Breeding for general resistance is more challenging because many genes with minor effects are involved. It is laborious to develop breeding stocks with horizontal resistance. However, it is easy to improve on the very low level of horizontal resistance that normally underlies a failed vertical resistance. Such improvements may be accomplished by using recurrent selection methods.

### Breeding for vertical resistance and the strategy of gene deployment

Vertical resistance is pathotype-specific and easy to breed. It is susceptible to **boom and bust** cycles, the cyclical failure of vertical resistance genes to deter newly evolved pathotypes. Boom and bust cycles arise when major genes for vertical resistance against a major economic race of a pathogen are used in cultivar development for a region, leading to widespread adoption of the resistant cultivars by most producers in the region (the **boom phase**). Selection pressure on the races of the pathogen present in the cultivars reduces the virulent ones. However, the less virulent one against which the cultivars carry no major genes continues to increase until it becomes epidemic in the vast region of production of the crop (the **bust phase**). Vertical resistance is suitable for annuals but not perennials. It is very effective against immobile or localized pathogens (e.g., soil-borne pathogens). However it must be deployed strategically to be useful against mobile airborne pathogens such as epidemic fungi of annual plants. Five general strategies may be used to make vertical resistance a success.

- 1 Temporal deployment.** A strategy for enhancing the success of vertical resistance is to develop and release several cultivars in successive or cyclical fashion. The rationale is that temporal deployment of cultivars will prevent the pathogen population from becoming completely adapted to the cultivar of the growing season.
- 2 Geographic deployment.** The application is of limited practical use because it requires a special circumstance

that one crop population protects another by acting as a filter to delay the advance of disease.

- 3 Spatial deployment.** In a situation where virulent pathogens are spatially localized, a cultivar diversification strategy whereby the fields of different farmers are planted to different cultivars containing two or more virulent genes, would slow down disease epidemics
- 4 Multiline deployment.** A multiline consisting of genotypes carrying different major genes would also put a damping effect on epidemics just like in the case of spatial deployment.
- 5 Mixture deployment.** A mixture of distinct cultivars with complementary vertical resistance genes can be deployed as in the case of multilines.

### Combining vertical and horizontal resistance

As previously indicated, vertical resistance is not suited to perennials. Tree breeding is a much longer process than breeding species of annuals. Errors are costly to correct. Should vertical resistance be overcome, it cannot be quickly replaced with a new genotype with new resistance genes. It is tempting to think that combining vertical resistance and horizontal resistance will provide the best of two worlds in the protection of plants. Unfortunately, this is wishful thinking simply because it is not possible to select horizontal resistance in the presence of vertical resistance.

The erosion of horizontal resistance while breeding for race-specific vertical resistance is called the **vertifolia effect** (after the potato cultivar "Vertifolia") in which the major gene is so strong that while the breeder focuses on vertical resistance, no evaluation and selection for horizontal resistance is possible, eventually leading to the loss of horizontal resistance. However, the vertifolia effect is not of universal occurrence. Some researchers have reported race-specific resistance in addition to high-level polygenic resistance to leaf rust in barley. To reduce the incidence of vertifolia effect, some suggest that breeders select and discard susceptible plants in segregating populations, rather than selecting highly resistant genotypes. Also, others suggest to first breed for a high level of horizontal resistance in a genotype then cross it with one that has high vertical resistance.

### Role of wild germplasm in disease- and pest-resistance breeding

Wild relatives of cultivated crops have been a source of genes for solving a variety of plant breeding problems,

especially disease resistance. The success and effectiveness of introgression of disease-resistance genes into crop species from wild relatives varies by crop. Factors to consider include the amount of diversity within the crop species, ease of hybridization with wild relatives, and the complexity of the genetic control of the trait. Some crop breeders (e.g., tomato breeders) use genes from wild relatives more frequently than other breeders, such as sorghum breeders, who appear to find their needs in adapted species. In tomato, wild relatives provided genes for enhancing the nutritional value (vitamin C and beta-carotene) and solids content, significantly boosting the commercial value of the crop. The impact of introgression of genes from the wild into adapted cultivars has been dramatic in some crops. For example, the resistance to the devastating late blight of potato was found in a wild species. Similarly, resistance to the root knot nematode in peanut was obtained from three wild species. A wild relative of rice, *Oryza nivara*, growing in the wild in Uttar Pradesh was found to have one single gene for resistance to the grassy stunt virus, a disease that devastated the crop in South and South East Asia in the 1970s.

### **Screening techniques in disease- and pest-resistance breeding**

One of the critical activities in breeding for resistance to diseases and insect pests is screening or testing for resistance. Various facilities, techniques, and approaches are used, depending on the parasite and host characteristics.

#### **Facilities**

Disease, as previously indicated, depends on the interaction among three factors – pathogen, host, and environment. Whereas field screening has the advantage of representing the conditions under which the resistant cultivar would be produced, it has its limitations. The weather (or the environmental component of the disease triangle) is unpredictable, making it difficult to have the uniformity and consistency of the parasite population. In some years, the weather may not favor an adequate pathogen population for an effective evaluation of plants. Controlled environmental tests provide reliable, uniform, and consistent evaluation of disease, but it has less field correspondence. Screening for resistance to mobile pests is challenging, requiring special provisions to confine the parasites. The distribution of the pest on the plants is often uniform.

### **Factors affecting expression of disease and insect resistance**

Certain specific factors may complicate breeding for resistance that may be environmental or biological in nature.

#### *Environmental factors*

- 1 **Temperature.** Low or high temperature over a period of time may cause loss of resistance.
- 2 **Light.** Light intensity affects the chemical composition of plants that is related to pest resistance (e.g., glycoside in potato).
- 3 **Soil fertility.** High soil fertility makes plants more succulent and more susceptible to disease development.

#### *Biological factors*

- 1 **Age.** The response of a plant to a pathogen or insect pest may vary with age. Some diseases are more intense at the early stage in plant growth than others.
- 2 **New pathotypes or biotypes.** New variants of the parasite that overcome the current resistance in the host may exhibit a different kind of disease expression.

### **Breeding procedures**

The first step in breeding for resistance to pathogens or insect pests is to assemble and maintain resistance genes. The sources of resistance genes include commercial cultivars, landraces, wild progenitors, related species and genera, mutagenesis, and biotechnology. As indicated elsewhere in the book, obsolete and current commercial cultivars are most preferred because they have minimum undesirable traits. Once a desirable source has been found, the backcross method of breeding is commonly used to transfer resistance genes into adapted cultivars. As previously discussed, there should be an effective and efficient screening technique for disease-resistance breeding. For cross-pollinated species, recurrent selection is effective for increasing the level of resistance in a population of genetically heterogeneous population.

### **Applications of biotechnology in pest-resistance breeding**

One of the successful applications of agricultural biotechnology is in pest-resistance breeding. The first disease-resistance gene, *Pto* (binds with products of the

pathogen to give resistance), was cloned in 1993 by Greg Martin.

### Engineering insect resistance

There are two basic approaches to genetic engineering of insect resistance in plants:

- 1 The use of protein toxins of bacterial origin.
- 2 The use of insecticidal proteins of plant origin.

#### *Protein toxins from Bacillus thuringiensis (Bt)*

The *Bacillus thuringiensis* (*Bt*) endotoxin is a crystalline protein. It was first identified in 1911 when it was observed to kill larvae of the flour moth. It was registered as a biopesticide in the USA in 1961. *Bt* is very selective in action, that is, one strain of the bacterium kills only certain insects. Formulations of whole sporulated bacteria are widely used as biopesticide sprays for biological pest control in organic farming. There are several major varieties of the species that produce spores for certain target pests: *B. thuringiensis* var. *kurstaki* (for controlling lepidopteran pests of forests and agriculture), var. *brliner* (wax moth), and var. *israelensis* (dipteran vectors of human disease). The most commercially important type of the crystalline proteinaceous inclusion bodies are called **δ-endotoxins**. To become toxic, these endotoxins, which are predominantly protoxins, need to be proteolytically activated in the midgut of the susceptible insect to become toxic to the insect. These endotoxins act by collapsing the cells of the lining of the gut regions.

*Bt* resistance development has been targeted especially at the European corn borer, which causes significant losses to corn in production. Previous efforts developed resistance in tobacco, cotton, tomato, and other crops. The effort in corn was more challenging because it required the use of synthetic versions of the gene (rather than microbial *Bt* *per se*) to be created.

Two genes, *cryB1* and *cryB2*, were isolated from *B. thuringiensis* subsp. *kurstaki* HD-1. These genes were cloned and sequenced. The genes differed in toxin specificities, the *cryB1* gene product being toxic to both dipteran (*Aedes aegypti*) and lepidopteran (*Manduca sexta*) larvae, while *cryB2* affects only the latter. The *Bt* toxin is believed to be environmentally safe as an insecticide. In engineering *Bt* resistance in plants, scientists basically link the toxin to a constitutive (unregulated) promoter that will express the toxin systemically (i.e., in all tissues).

Transgenic plants expressing the δ-endotoxin gene have been developed. The first attempt involved the fusion of the *Bt* endotoxin to a gene for kanamycin resistance to aid in selecting plants (conducted by a Belgian biotechnology company, Plant Genetic Systems, in 1987). Later, Monsanto Company researchers expressed a truncated *Bt* gene in tomato directly by using the CaMV 35S promoter. Agracetus Company followed with the expression of the *Bt* endotoxin in tobacco with the CaMV 35S promoter linked to an alfalfa mosaic virus (AMV) leader sequence. Since these initial attempts were made, modifications to the protocols have increased expression of the toxin in transgenic plants by several hundred-fold. Transformation for expressing the chimeric *Bt* genes was *Agrobacterium*-mediated, using the TR2' promoter. This promoter directs the expression of manopine synthase in plant cells transformed with the TR-DNA of plasmid pTiA6.

The original *Bt* coding sequence has since been modified to achieve insecticidal efficacy. The complete genes failed to be fully expressed. Consequently, truncated (comprising the toxic parts) genes of *Bt* var. *kurstaki* HD-1 (*cry1A[b]*) and HD-73 (*cry1A[c]*) were expressed in cotton against lepidopteran pests. In truncating the gene, the N-terminal half of the protein was kept intact. For improved expression, various promoters, fusion proteins, and leader sequences have been used. The toxin protein usually accounts for about 0.1% of the total protein of any tissue, but this concentration is all that is needed to confer resistance against the insect pest.

Genetically engineered *Bt* resistance for field application is variable. For example, Ciba Seed Company has developed three versions of synthetic *Bt* genes capable of selective expression in plants. One is expressed only in pollen, another in green tissue, and the third in other parts of the plant. This selectivity is desirable for several reasons. The European corn borer infestation is unpredictable from year to year. The life cycle of the insect impacts the specific control tactic used. The insect attack occurs in broods or generations. The *Bt* genes with specific switches (pollen and green tissue) produce the *Bt* endotoxin in the plant parts that are targets of attack at specific times (i.e., first and second broods). This way, the expression of the endotoxin in the seed and other parts of the plant where protection is not critical is minimized. Monsanto's "YieldGard" corn produces *Bt* endotoxin throughout the plant, and protects against both first and second broods of the pest. The commercially available *Bt* corn cultivars were developed by different transformation events, each with a different promoter.



## Industry highlights

### *Genetic improvement of cassava through biotechnology*

Nigel J. Taylor

International Laboratory for Tropical Agricultural Biotechnology (ILTAB), Donald Danforth Plant Science Center, St Louis, MO 63132, USA

#### **Introduction**

The vast majority of the world's farmers reside in developing countries where they cultivate crops on a hectare or less of land for on-farm consumption and for sale in the local marketplace. Delivering enhanced plant varieties to these farmers is central to improving their food security, health, and economic well-being. The starchy root crop cassava (*Manihot esculenta*), otherwise known as tapioca, manioc, yucca, and mandioca, is an important component of agricultural systems throughout much of the world's tropical regions. After rice and maize, cassava is the most important source of dietary calories in the tropics, and worldwide is cultivated over an area only 7% less than that of the *Solanum* potato (FAO 2004). Cassava is grown for its large swollen storage roots, which develop to store large amounts of starch within specially developed xylem parenchyma. Resistance to drought, tolerance to poor soils, and a flexible harvest window – the roots can be dug up when needed any time between 9 and 24 months after propagation from woody stem cuttings – makes the crop attractive to small-scale and subsistence farmers. As a result cassava is widely cultivated in Africa, where it is the most important staple crop after maize, and by resource-poor farmers in Latin America and tropical Asia. In addition, cassava is grown on a commercial scale for animal feed and as a source of starch for industrial applications and the food processing industries in Asia, southern India, and increasingly in South America.

#### **Why use biotechnology for the genetic improvement of cassava?**

Conventional breeding in cassava is complicated by strong heterozygosity and inbreeding depression. Breeding programs consist of bringing together preferred parents, crossing these and screening the first generation of resulting offspring for desired traits. Preferred individuals are then propagated vegetatively to generate sufficient clones to allow for further screening and multilocation testing (Ceballos et al. 2004). Large numbers of crosses must be carried out and thousands of offspring examined in order to identify candidate lines. Most importantly, the inability to perform backcrossing greatly hinders the capacity to introgress multiple desired traits into a single genetic background. Breeding programs in Africa and South America have been successful in developing and releasing cassava varieties with an enhanced harvest index and resistance to disease. However, there is general acceptance that transgenic technologies could hold the key to realizing the full genetic potential of cassava. Modern biotechnology allows the direct integration of beneficial traits into a plant genome. In the case of cassava, this could be utilized to work hand in hand with breeding programs to integrate traits missing from otherwise high performing germplasm, or employed to directly improve existing varieties and landraces already favored by farmers. An example of the latter approach is provided by ongoing programs utilizing transgenic technologies to generate resistance to cassava mosaic disease in East African cassava; the aim being to increase yields in highly susceptible landraces and in some cases prevent such germplasm from being abandoned by farmers. Transgenic technologies also offer the possibility to integrate beneficial traits that do not exist within the species or its sexually compatible relatives, into cassava. Examples of such genes include those imparting resistance to herbicides and the *Bt* family of genes that provide resistance to insect pests.

#### **Development of a transgenic capacity in cassava**

The use of genetic transformation to improve a crop plant requires two technologies. Firstly, molecular biologists identify genes with the potential to impart desired traits and then generate "gene constructs" that will allow effective expression of this genetic material in the target species. Most often this involves fusing the coding sequence to a suitable promoter sequence that will direct expression of the gene(s) in the appropriate tissues. For example, because cassava is a root crop, significant resources have been directed at identifying promoters that drive gene expression in this organ and to the discovery of new promoter sequences, free from existing intellectual property rights, which are capable of carrying out this function (Verdageur et al. 1998; Zhang et al. 2003). Secondly, technical capacity must be developed to deliver the new genetic material through the cell wall and facilitate its integration into the plant's genome. The latter requires that tissues be manipulated in culture to generate totipotent cell lines. It is these totipotent cells that act as the target for transgene insertion, and which after selection for the successful integration events, can be regenerated to produce whole genetically transformed plants.

Efforts to develop transgenic technologies in cassava were initiated in the late 1980s. However, it was not until 1996 that the first transgenic cassava plants were reported (Li et al. 1996; Schöpke et al. 1996). This delay resulted from an inability to produce suitable totipotent target tissues for transgene insertion. Although techniques for producing somatic embryos existed before 1996, these tissues proved to be ineffective for the recovery of transgenic cassava plants via microparticle bombardment or *Agrobacterium tumefaciens*. The multicellular origin and highly organized nature of these embryogenic structures meant that only chimeric embryos and plants could be recovered when this target tissue was used for genetic transformation. Two breakthroughs concerning the *in vitro* manipulation of cassava somatic embryos were instrumental in allowing the first transgenic cassava plants

to be recovered. In the first, it was discovered that when the organized embryogenic structures (OES) were transferred from a Murashige and Skoog-based medium to one containing the basal salts of Greshoff and Doy, the OES would break down into a highly friable embryogenic callus (FEC) from which somatic embryos could be regenerated and plants recovered upon subculture in a medium devoid of auxin (Taylor et al. 1996). This FEC is composed of tens of thousands of submillimeter-sized pre-embryogenic units that proliferate via disorganized divisions from single cells at their surface. As such they act as excellent target tissue for transgene insertion and recovery of genetically transformed plants, by both microparticle bombardment (Schöpke et al. 1996) and *Agrobacterium* (Gonzalez-de Schöpke et al. 1996). In the second system, it was discovered that if cotyledons were allowed to develop from the OES, shoots could be induced to develop via an organogenic process from the cut surface of these organs in the presence of an appropriate cytokinin. This system resulted in the recovery of transgenic cassava plants after co-culture of the cotyledons with *Agrobacterium* (Li et al. 1996). Subsequent research has also proven that it is possible to recover transgenic cassava plants directly from leaf explants via *Agrobacterium*-mediated gene transfer (Siritunga & Sayre 2003). For a more detailed description of genetic transformation systems in cassava, the reader is referred to a recent review by Taylor et al. (2004).

The production of transgenic cassava plants is now routine in five laboratories around the world. Based in Europe, the USA, and at the Centro Internacional de Agricultura Tropical (CIAT) in Colombia, four of the five utilize the FEC system as the target tissue of choice and all now employ *Agrobacterium*-mediated gene insertion for the production of transgenic cassava plants. As for other crops, it has been found that in cassava, *Agrobacterium* more reliably generates plants with single copy insertions of the transgene – an issue considered important with regard to the stability of transgene expression and future deregulation of genetically modified plants for release to farmers. The traits presently being targeted within cassava transgenic programs reflect the interests of the specific laboratories and include: resistance to virus disease, herbicide tolerance, increased resistance to insect pests, and enhanced starch quality, elevated protein, and reduced cyanogenic content of the storage roots (Taylor et al. 2004).

#### *Steps to produce a transgenic crop plant in the tropics using virus-resistant cassava as an example*

As an example of the steps and processes involved in producing a transgenic crop plant for delivery to farmers in the tropics, we will examine in more detail the program and strategies employed at the Donald Danforth Plant Science Center (DDPSC) to produce cassava with resistance to cassava mosaic disease (CMD). CMD impacts cassava production throughout sub-Saharan Africa and the Indian subcontinent and is caused by at least eight species of whitefly-transmitted geminiviruses. The geminiviruses that infect cassava are single-stranded DNA viruses possessing two circular genomic components, each approximately 2.6 kb in size. The disease rarely kills infected plants but can significantly reduce the production of storage roots in susceptible varieties. The disease is most severe when plants are simultaneously infected with more than one species of geminivirus, when the two pathogens act synergistically to overcome the plant's defense mechanisms. It is estimated that CMD is responsible for 30–40% yield reduction in Africa, which is equivalent to as much as 25 million metric tons of food each year. Workers at the DDPSC are employing transgenic technologies to impart resistance to CMD in varieties already preferred by farmers and consumers in Africa.

Transgenic programs start with an idea based on existing knowledge about how integration and expression of a specific gene or genes might impart beneficial agronomic characteristics within the crop of interest. Before commencing work in the laboratory, it is essential to determine whether the strategy in question can pass the rigorous food and environmental safety testing that all transgenic plants must satisfy before they can be grown by farmers. For example, a gene that codes for a toxin from scorpions, or a known human allergen, might provide very effective resistance against a given insect pest when expressed transgenically in a crop plant. However, it can be predicted that such a product will not be acceptable to the regulatory agencies (nor the public at large) and can therefore never be made available to farmers. Proceeding with such a program may be justified on a purely scientific basis, but if the intention is to bring a new product to market, it would be a waste of resources and the program would be terminated before it began. Likewise, before significant investment is made in the research and development phases, scientists must work with lawyers experienced in the field of intellectual property rights. Many genes, genetic sequences, transformation protocols, and other tools required to develop a given transgenic plant are protected through patents and licensing agreements. Failure to obtain the "freedom to operate" for commercial release early in the development program for all the technologies needed to develop the final product, can cause significant problems later in the product delivery process. It is at that time that the owners of such technologies may assert their legal rights and demand royalties or otherwise block commercial release. A well-known example of such problems arose with "Golden Rice", where unresolved intellectual property right issues were not addressed until late in the development process. Significant resources were required to resolve outstanding issues and further product development was delayed.

Workers at the DDPSC have adopted a pathogen-derived resistance strategy to combat the effects of CMD. In such approaches, genetic sequences from the viral pathogen are cloned, fused to an appropriate expression cassette, and integrated into the plant's genome. In this specific case the *AC1* (or replication associated) gene from African cassava mosaic virus (a causal agent of CMD) was chosen as the candidate gene. *AC1* and the use of replication-associated genes to impart resistance to geminiviruses resides in the public domain and therefore its use is not restricted due to issues of intellectual property. The promoter sequence used to drive expression of this gene in cassava was the cassava vein mosaic virus (CsVMV) promoter, previously isolated and characterized at the International Laboratory for Tropical Agricultural Biotechnology (ILTAB) (Verdaguer et al. 1998). Although licensed to a commercial entity, it is available through special agreement for use and release within transgenic products to farmers in developing countries. A history of safe use for pathogen-derived resistance technologies deployed to farmers in transgenic papaya, potato, and squash, generated confidence that this approach in cassava would not encounter problems of food and environmental safety, should a successful product be developed.

Initial proof of concept for the ability of transgenically expressed *AC1* to impart resistance to the geminiviruses causing CMD was achieved using tobacco plants. The production of transgenic tobacco (*Nicotiana benthamiana*) is straightforward and inexpensive. In comparison, recovering transgenic cassava is labor intensive, requires specially trained staff, and requires 5–6 months to recover plantlets. *N. benthamiana* is susceptible to infection with a range of virus species, making it a highly effective model species for such studies. Research by workers at the DDPSC and in the UK demonstrated that, indeed, when the *AC1* gene was integrated into tobacco, the resulting plants were significantly more resistant to infection by cassava geminiviruses than the non-transgenic control plants (Hong & Stanley 1996; Sangaré et al. 1999). Armed with this information, workers could progress with some confidence to the next stage and develop this strategy in cassava itself. Technologies described above were utilized to regenerate approximately 20 transgenic cassava plants from independent integration events in the CMD-susceptible, West African variety 60444. Presence of the transgenes was confirmed by polymerase chain reaction (PCR) amplification for both the *AC1* and the *nptII* selectable marker genes in the FEC lines and subsequently in regenerated plantlets.

Production of transgenic plants containing a desired transgene does not guarantee that the desired trait will be expressed at the level needed to generate a product of the quality required for release to farmers. In order to test for efficacy of the *AC1* gene in cassava, transgenic cassava plants were challenged with infectious clones of three geminivirus species. Although whiteflies are the natural vector for CMD, maintaining virifilous populations of the specific biotype responsible for spreading the disease in Africa is highly problematic in the temperate regions. As an alternative, viral DNA was coated onto micron-sized gold particles and shot into young cassava plants using Biolistic® microparticle bombardment technology. When cassava plants were inoculated in this manner, four lines were found to exhibit significantly enhanced resistance to the pathogens (Chellappan et al. 2004). Most interestingly, these plants were resistant not only to African cassava mosaic virus (ACMV) but also to East African cassava mosaic virus (EACMV), Sri Lankan cassava mosaic virus, and to simultaneous, dual challenge with ACMV and EACMV.

Cross-protection imparted by the *AC1* transgene from ACMV was not predicted at the beginning of the project, but has important implications for deploying such a technology to the field in Africa. As described above, there are at least six species of geminiviruses infecting cassava in sub-Saharan Africa. An effective defense strategy against CMD, whether developed through conventional breeding or biotechnology, must provide protection against most or all of these, and also against the synergistic effects of dual infection with more than one species of viral pathogen. Small interfering RNAs (siRNAs) specific to the *AC1* gene were detected in the two plant lines showing highest resistance to CMD prior to infection with virus (Chellappan et al. 2004), indicating that the *AC1* transgene was triggering gene silencing mechanisms against this sequence within the plant. The *AC1* gene is well conserved across different species of cassava-infecting geminiviruses, and it is therefore hypothesized that such siRNAs are capable of targeting the mRNA of this sequence for degradation immediately the transgenic plant becomes infected with a geminivirus, thereby inhibiting viral replication and imparting resistance.

To summarize the above it can be stated that by integrating one transgene (two if the *nptII* selectable marker gene is counted) we were able to generate very resistant plant lines from an otherwise highly CMD-susceptible cassava cultivar. The caveat to this proof of efficacy was that resistance was demonstrated in the greenhouse in the US midwest and plants were inoculated with the pathogens in an artificial manner. Important questions remain as to whether the plants will perform well when grown under field conditions in Africa where the viruses are transmitted by whiteflies. Testing the plants under such conditions is a critically important step in demonstrating the usefulness of this technology. Success at this stage is critical to deciding whether to proceed along the delivery pathway.

Establishing field trials of transgenic plants is straightforward in North America where many thousands of such tests have been completed over the last decade. The situation in Africa is dramatically different. As of late 2004, outside of South Africa, only Kenya, Zimbabwe, and Burkina Faso have carried out field trials, with the total combined number being less than 10 trials. The remaining countries either do not have the required regulations, biosafety infrastructure, and trained personnel in place, or have never exercised their existing regulatory procedures to establish field trials of transgenic plants. This situation obviously presents challenges to those wishing to develop and deploy a transgenically enhanced crop in such regions. The DDPSC has established collaboration with the Kenyan Agricultural Research Institute (KARI) to carry out trials of the cassava plants described above. After consideration by the Kenyan National Biosafety Committee it was recommended that a contained screenhouse trial should be performed in a cassava-growing region of western Kenya – the first test of transgenic cassava in Africa. Plantlets were exported from the DDPSC to Kenya and planted in pots in a biosafety level 2 screenhouse constructed for the trial. At the time of writing, whitefly inoculation of transgenic and control plants is ongoing within this structure. If the transgenic plants demonstrate acceptable levels of resistance to the geminiviruses infecting cassava in western Kenya, it is planned to carry out multilocational, confined field trials in this and other regions of that country.

The product development and delivery pathway contains a loop back from the field trial phase to the laboratory. There are two reasons as to why this takes place. Firstly, the technology in question could perform well in the field, but not well enough to proceed in its initial form to further product development and delivery. In this case more work is required in the laboratory, most likely to improve transgene express and/or to direct this expression to desired tissues in a more effective manner. Further testing in the greenhouse and field will then be required to confirm any improvements. For crops such as cassava, even if initial field trials are successful, it will be necessary to return to the laboratory in order to generate more genetically transformed plants. This is because the variety in which the initial transgene integrations take place is most likely not in the background most suitable to farmers' needs. In sexually propagated crops it is possible to take the best performing transgenic plant produced and backcross this with preferred parents within established breeding programs. In this way the beneficial transgenic trait is introgressed into a

variety (or range of varieties) suitable for deployment to farmers. As described above, this is not a good option for cassava and the other heterozygous, vegetatively propagated crops. Here the desire is to directly enhance the performance of existing farmer-preferred varieties, landraces, and elite breeding lines, without changing their other beneficial characteristics, which may be lost due to unpredictable segregation when sexual crossing is carried out. It is thus necessary to genetically transform each desired cassava variety individually. This presents important challenges, as one has to correctly identify which varieties to target for such investment and then develop the technical capacity to manipulate this germplasm within the tissue culture and genetic transformation systems established for the crop. Greenhouse and field testing must then follow for each new transgenic variety. With support from the United States Agency for International Development, the DDPSC is currently engaged in this process for preferred cassava varieties from East Africa.

### **Future directions for transgenic cassava**

Biotechnology can have a significant impact on the genetic improvement of vegetatively propagated crops, including cassava, if sufficient resources are committed to such efforts. However, with only five research institutes capable of producing transgenic cassava plants, it is obvious that the funding being committed to this and other crops important in developing countries such as plantains, sorghum, millet, and sweet potato are not in proportion to their importance as sources of food security and economic well-being for people in the tropical regions. Nevertheless, important progress has been made over the last 5–10 years in developing transgenic programs for cassava. Genes of agronomic interest have been integrated into the crop and field trials – a critical step in the process towards product development – that are being initiated in Africa and at CIAT, Colombia. Future programs should benefit from mapping and other genomic-based research. BAC (bacterial artificial chromosome) libraries are being developed for the crop and being used to identify and isolate genes responsible for resistance to CMD and bacterial blight disease. Once available these can be integrated into farmer-preferred varieties using the transgenic technologies described above. It is also hoped to access beneficial genes present in the wild relatives of cassava. Within this germplasm can be found traits for enhanced protein accumulation in the roots, longer shelf-life of harvested roots, and disease and pest resistance. Cassava has an inherent capacity for high rates of photosynthesis and the ability to accumulate large amounts of starch within its storage roots. It has already been shown possible to shut down starch accumulation in cassava through the application of transgenic technologies. Can this unused energy be diverted towards the accumulation of biosynthetic plastics? Could genes for apomixis be introduced to cassava to enable true seeds to be produced, thereby revolutionizing the propagation of disease-free propagules? Reaching such goals though traditional breeding alone is not possible. They are only feasible if robust biotechnology programs are developed for cassava.

### **References**

- Ceballos, H., C.A. Iglesias, J.C. Perez, and A.G.O. Dixon. 2004. Cassava breeding: Opportunities and challenges. *Plant Mol. Breeding* (available at <http://www.kluweronline.com/issn/0167-4412/current>).
- Chellappan, P., M.V. Masona, V. Ramachandran, N.J. Taylor, and C.M. Fauquet. 2004. Broad spectrum resistance to ssDNA viruses associated with transgene-induced gene silencing in cassava. *Plant Mol. Biology* (available at <http://www.kluweronline.com/issn/0167-4412/current>).
- FAO. 2004. FAOSTAT statistical database, agriculture data. Available at <http://apps.fao.org>.
- Gonzalez-de Schöpke, A.E., C. Schöpke, N.J. Taylor, R.N. Beachy, and C.M. Fauquet. 1998. Regeneration of transgenic plants (*Manihot esculenta* Crantz) through *Agrobacterium*-mediated transformation of embryogenic suspension cultures. *Plant Cell Rep.* 17:827–831.
- Hong, Y., and J. Stanley. 1996. Virus resistance in *Nicotiana benthamiana* conferred by African cassava mosaic virus replication-associated protein (AC1) transgene. *Mol. Plant Microbe Interact.* 9:219–225.
- Li, H.Q., C. Sautter, I. Potrykus, and J. Pounzi-Kaerlas. 1996. Genetic transformation of cassava (*Manihot esculenta* Crantz). *Nature Biotechnol.* 14:736–740.
- Sangaré, A., D. Deng, C.M. Fauquet, and R.N. Beachy. 1999. Resistance to African cassava mosaic virus conferred by mutant of the putative NTP-binding domain of the Rep gene (AC1) in *Nicotiana benthamiana*. *Mol. Biol. Rep.* 5:95–102.
- Schöpke, C., N.J. Taylor, R. Carcamo, et al. 1996. Regeneration of transgenic cassava plants (*Manihot esculenta* Crantz) from microbombarded embryogenic suspension cultures. *Nature Biotechnol.* 14:731–735.
- Siritunga, D., and R.T. Sayre. 2003. Generation of cyanogen-free transgenic cassava. *Planta* 217:367–373.
- Taylor, N., P. Chavarriaga, K. Raemakers, D. Siritunga, and P. Zhang. 2004. Development and application of transgenic technologies in cassava. *Plant Mol. Biology* (available at <http://www.kluweronline.com/issn/0167-4412/current>).
- Taylor, N.J., M. Edwards, R.J. Kiernan, C. Davey, D. Blakesley, and G.G. Henshaw. 1996. Development of friable embryogenic callus and embryogenic suspension cultures in cassava (*Manihot esculenta* Crantz). *Nature Biotechnol.* 14:726–730.
- Verdaguer, B., A. de Kochko, C.I. Fux, R.N. Beachy, and C.M. Fauquet. 1998. Functional organisation of the cassava vein mosaic virus (CsVMV) promoter. *Plant Mol. Biol.* 37:1055–1067.
- Zhang, P., S. Bohl-Zenger, J. Pounzi-Kaerlas, I. Potrykus, and W. Gruisse. 2003. Two cassava promoters related to vascular expression and storage root formation. *Planta* 218:192–203.

*Bt* cotton is another widely grown bioengineered crop. The pest resistance conferred by the *Bt* gene has led to a dramatic reduction in pesticide use and has consequently reduced adverse impact on the environment from agropesticides. As indicated previously, *Bt* sprays are widely used in organic farming for pest control. However, such application is ineffective if the insect bores into the plant. Further, *Bt* sprays have short duration activity.

### Engineering viral resistance

Even though viruses may utilize DNA or RNA as hereditary material, most of the viruses that infect plants are RNA viruses. One of the most important plant viruses in biotechnology is the cauliflower mosaic virus (CaMV) from which the widely used 35S promoter was derived (CaMV 35S promoter). A virus is essentially nucleic acid encased in a protein coat. The primary method of control of viral infections is through the breeding of resistance cultivars. Also, plants can be protected against viral infection by a strategy that works like inoculation in animals. Plants may be protected against certain viral infections upon being infected with a mild strain of that virus. This strategy, called **cross-protection**, provides protection to the plant against future, more severe, infections.

Engineering transgenic plants with resistance to viral pathogens is accomplished by the method called **coat protein-mediated resistance**. First, the viral gene is reverse transcribed (being RNA) into DNA from which a double-stranded DNA is then produced. The product is cloned into a plasmid and sequenced to identify the genes in the viral genome. A chimeric gene is constructed to consist of the open reading frame for the coat protein to which a strong promoter is attached for high level of expression in the host. This gene construct is transferred into plants to produce transgenic plants.

Successes with this strategy have been reported in summer squash (the first product developed by this approach), and for resistance to papaya ring spot virus (a lethal disease of papaya), among others.

### Engineering herbicide resistance

Herbicides constitute one of the most widely used agrochemicals in crop production. Organisms genetically engineered for herbicide resistance are among the major applications of biotechnology in plant food biotechnology.

#### Why engineer herbicide-resistant crops?

A successful herbicide should destroy weeds only, leaving the economic plant unharmed. Broad-spectrum herbicides (non-selective) are attractive but their use in crop production can be problematic, especially in the production of broadleaf crops such as soybean and cotton. There is a general lack of herbicides that will discriminate between dicot weeds and crop plants. Preplant applications may be practical to implement. However, once the crop is established and too tall for the safe use of machinery, chemical pest management becomes impractical. Grass crops (e.g., wheat, corn) may tolerate broadleaf herbicides better than the reverse situation. Consequently, when cereal crops and broadleaf crops are grown in rotation or adjacent fields, the broadleaf plants are prone to damage from residual herbicides in the soil, or drift from herbicides applied to grasses. When a crop field is infested by weed species that are closely related to the crop (e.g., red rice in a rice crop or nightshade in a potato crop), herbicides lack enough sensitivity to distinguish between the plants.

To address these problems, one of two approaches may be pursued: (i) the development of new selective post-emergent herbicides; or (ii) the genetic development of herbicide resistance in crops to existing broad-spectrum herbicides. The latter strategy would be advantageous to the agrochemical industry (increased market) and farmers (safer alternative to pesticides that are already in use). New herbicides are expensive to develop and take time.

#### Modes of action and herbicide-resistance mechanisms

Most herbicides are designed to kill target plants by interrupting a metabolic stage in photosynthesis. Because all higher plants photosynthesize, most herbicides will kill both weeds and desirable plants. Plants resist phytotoxic compounds via one of several mechanisms.

- 1 The plant or cell does not take up toxic molecules because of external barriers such as cuticles.
- 2 Toxic molecules are taken up but sequestered in a subcellular compartment away from the target (e.g., protein) compounds the toxin was designed to attack.
- 3 The plant or cell detoxifies the toxic compound by enzymatic processes, into harmless compounds.
- 4 The plant or cell equipped with resistance genes against the toxin may produce a modified target compound that is insensitive to the herbicide.

- 5 The plant or cell overproduces the target compound for the phytotoxin in large amounts such that it would take a high concentration of the herbicide to overcome it.

### Epidemics and plant breeding

Modern agriculture is drastically different to that done by early plant domesticators. Plant breeding, in adapting wild species for cultivation, often removes natural means of crop protection needed for survival in the wild, but undesirable in modern production. Starting with the early domesticators, mass production of a few desirable genotypes has become the norm of crop production. The monocultures of modern agriculture are extreme versions of mass production, in which genetic diversity is further restricted. It is not difficult to see how such breeding and production practices predispose crops to widespread disease and insect pest damage. With the arrival of a devastating pest attack, breeders counter with the production of resistant cultivars, thus setting the stage for sequential cycles of pest resistance and pest susceptibility of crop plants.

The patterns of disease incidence vary from one class of pathogen to another. Soil-borne pathogens tend to be localized and persistent in the soil, season after season. Airborne pathogens have a different biology and spread pattern, as well as the potential to become epidemic. An **epidemic** may be defined as an outbreak of disease characterized by an infection that starts from a low level and then progresses to a high one. Airborne pathogens tend to have annual epidemic cycles. These cycles depend on seasonal weather patterns and the presence or absence of a susceptible host (crop). Soil-borne pathogens usually stabilize in a new location after several years, and persist in the chronic phase.

Each type of disease pathogen and the epidemic it causes is characterized by its own ecology and spatial and temporal dynamics. Plant fungal diseases are especially significantly impacted by weather conditions and climatic patterns. Epidemics in agricultural crop production impact US agriculture by affecting the economic value, quantity, and quality of food and fiber produced.

J. E. Van der Plank provided the underlying scientific principles of disease epidemics. An epidemic occurs in phases. It starts with a small inoculum of the pathogen (that may be from overwintering spores, or freshly introduced into the area from outside). The initial rate of infection is proportional to the infection level. As time proceeds, the infection rate slows down as uninfected tissue decreases. Van der Plank expressed the initial infection rate by this mathematical relationship:

$$\frac{dI}{dt} = rI \text{ or } I = I_0 \exp^{rt}$$

where  $I$  = infection at time  $t$ ,  $I_0$  = initial infection, and  $r$  = multiplication rate of the pathogen. A general and more realistic expression is:

$$\begin{aligned} dI &= rI(1 - I) \\ I &= 1/[1 + (\exp^{-rt})] \text{ or } \log_e(1/I - 1) = rt + \log_e(1/C) \end{aligned}$$

where  $C$  is a constant that depends upon  $I_0$  but is not equal to it.

In plant breeding, the breeder's goal in countering a disease epidemic is to reduce  $r$  (the multiplication rate of the pathogen) to a level such that, at the time of crop maturity, the final attack by the pathogen can inflict only a small and acceptable yield loss. It may be best if  $r = 0$  (i.e., immunity), but this is not necessary. In practice, breeders do not aim for immunity but for reducing  $r$  below that of the current commercial cultivar that is susceptible to the pathogen.

### References and suggested reading

- Alam, S.N., and M.B. Cohen. 1998. Detection and analysis of QTLs for resistance to the brown planthopper, *Nilaparvata lugens*, in a doubled haploid rice population. *Theor. Appl. Genet.* 97:1370–1379.
- Bogdanove, A.J. 2002. Protein–protein interactions in pathogen recognition in plants. *Plant Mol. Biol.* 50:981–989.
- Flor, H.H. 1956. The complementary genic system in flax and flax rust. *Adv. Genet.* 8:29–54.
- Flor, H.H. 1971. Current status of the gene-for-gene concept. *Ann. Rev. Phytopathol.* 9:275–296.
- Hill, C.B., Y. Li, and G.L. Hartman. 2004. Resistance to the soybean aphid in soybean germplasm. *Crop Sci.* 44:98–106.
- Hooker, A.L. 1980. Breeding to control pests. In: *Crop breeding* (Wood, D.R., ed.). American Society of Agronomy/Crop Science Society of America, Madison, WI.
- Khush, G.S., and D.S. Brar. 1991. Genetics of resistance to insects in crop plants. *Adv. Agron.* 45:223–274.
- Niks, R.E., P.R. Ellis, and J.E. Parlevliet. 1993. Resistance to parasites. In: *Plant breeding: Principles and prospects*

- (Hayward, M.D., N.O. Bosemark, and I. Ramagosa, eds). Chapman & Hall, London.
- Parlevliet, J.E. 1981. Disease resistance in plants and its consequences for plant breeding. In: Plant breeding II (Frey, K.J., ed.). Iowa State University Press, Ames, IA.
- Powles S.B., C. Preston, I.B. Bryan, and A.R. Jutsum. 1997. Herbicide resistance: Impact and management. *Adv. Agron.* 58:57–93.
- Simmonds, N.W. 1985. A plant breeder's perspective of durable resistance. *FAO Plant Protein Bull.* 33:13–17.
- Simmonds, N.W. 1991. Genetics of horizontal resistance to diseases of crops. *Biol. Rev.* 66:189–241.
- Simmonds, N.W. 1993. Introgression and incorporation strategies for the use of crop genetic resources. *Biol. Rev.* 68:539–562.
- Smith, G.M. 1989. Plant resistance to insects: A fundamental approach. Wiley, New York.

---

### Outcomes assessment

#### Part A

Please answer the following questions true or false:

- 1 Major gene resistance is durable.
- 2 Most of the viruses that affect plants are RNA viruses.
- 3 The genes for *Bt* products were obtained from bacteria.
- 4 Vertical resistance is conditioned by numerous minor genes.

#### Part B

Please answer the following questions:

- 1 ..... is the concept of transferring several specific genes into one plant.
- 2 What is durable resistance?
- 3 Explain the strategy of cross-protection in viral-resistance engineering.
- 4 Give two specific reasons for engineering herbicide resistance.
- 5 Describe the mechanism of hypersensitive reaction used by plants in response to the invasion of certain pathogens.

#### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the method of coat protein-mediated resistance in plant breeding.
- 2 Discuss the concept of gene-for-gene hypothesis in disease genetics.
- 3 Discuss the mechanism plants use to respond to pathogen invasion.
- 4 Discuss the mechanisms commonly used by plants to respond to insect pest attacks.
- 5 Discuss the approach to breeding for vertical resistance.
- 6 Discuss the breeding of *Bt* resistance in plants.



# Breeding for resistance to abiotic stresses

---

## Purpose and expected outcomes

*Climate is the variation in meteorological factors over a large area formed over many years. It determines plant adaptation to a growing region. Weather on the other hand is the environmental condition described by the short-term variations in meteorological factors within a local area. It determines crop development and productivity. Weather affects the choice of specific crop production activities (i.e., how the crop is produced in the field). Plant breeders usually develop breeding programs to produce cultivars for specific production regions. Crop production is subject to the vagaries of the weather. Unpredictable weather can drastically reduce crop yield. Crop varieties used in regions that are prone to adverse weather during production need to have the capacity to resist or tolerate environmental stresses to an extent that they produce acceptable economic yield. After completing this chapter, the student should be able to:*

- 1 Describe the types of environmental stresses that reduce plant performance in crop production.
  - 2 Discuss breeding for drought resistance.
  - 3 Discuss breeding for cold tolerance.
  - 4 Discuss breeding for salt tolerance.
  - 5 Discuss breeding for resistance to heat stress.
  - 6 Discuss breeding for tolerance to aluminum toxicity.
  - 7 Discuss breeding for tolerance to oxidative stress.
  - 8 Discuss breeding for resistance to water-logging.
- 

## Importance of breeding for resistance to abiotic stresses

Only about 30% of the earth is land. Of this, about 50% is not suitable for economic crop production, mainly because of constraints of temperature, moisture, and topography. Of the remaining portion of arable land, optimum production is further limited by a variety of environmental stresses, requiring mineral and moisture supplementation for economic crop production. As the world population increases more food will have to be produced by increasing the productivity of existing farmlands as well as bringing new lands into production.

This means, marginal lands will have to be considered. Plant breeders will have to develop cultivars that are adapted to specific environmental stresses. It is estimated that abiotic environmental stresses are responsible for about 70% of yield reduction of crops in production.

In modern agricultural production, producers (especially those in developed countries) are able to alleviate moisture stress by providing supplemental irrigation. Also, fertilizers are used to alleviate nutritional stress, while soil reaction (pH) can be modified by applying soil amendments such as lime or sulfur. Excessive soil moisture can be removed by installing drainage systems. However, in poor regions of the world where such

technologies are not readily available or affordable (which, incidentally, are the regions in most need of food increases), the alternative is to breed cultivars that are able to resist these environmental stresses enough to produce acceptable crop yields.

Each species has natural limits of adaptability. There are tropical plants and temperate plants. Breeders are able, within limits, to adapt certain tropical plants to temperate production, and vice versa. To achieve this, plants are genetically modified to resist environmental stresses in their new environment. Sometimes, in modern crop production, breeders develop cultivars to resist certain environmental stresses to make better use of the production environment. For example, cold resistance enables producers to plant early in the season while the soil is still too cold for normal planting. This may be done to extend the growing season for higher productivity or for some other reasons.

### **Resistance to abiotic stress and yield potential**

Plant growth and development is the product of the interaction between the genotype (genetic potential) and the environment in which the plant grows. Any stress in the environment will adversely impact growth and development. Plants perform well in environments to which they are well adapted. Yield potential was previously defined as the highest yield attainable by a genotype growing in an environment to which it is adapted and in which there is no environmental stress (i.e., optimum growing conditions). Except, perhaps, under controlled environments, it is very difficult to find a production environment (especially on a large scale) in which some environmental stress of some sort does not occur. Stress resistance is an inherent part of all cultivar development programs. Prior to releasing a cultivar, genotypes of high potential are evaluated at different locations and over several years to determine adaptedness.

High yield potential has been achieved in cereal crops through enhanced harvest indexes, whereby dry matter is redistributed into the grain (by reducing its distribution to vegetative parts – roots, stems, leaves). However, this pattern of dry matter distribution also reduces the ability of these genotypes with high yield potential to cope with environmental stresses. Studies on genotype  $\times$  environment ( $G \times E$ ) interactions have shown cross-over effects whereby, under conditions of severe stress, genotypes of high yield potential perform poorly. Fur-

ther, as the abiotic constraints intensify (e.g., drought, low temperature, high temperature), a report by the Centro Internationale de Mejoramiento de Maiz y Trigo (CIMMYT) indicates that it becomes more difficult to improve genetic and agronomic yield of crops.

R. A. Fisher and R. Maurer partitioned stress effects on yield into parameters that measure sensitivity to stress ( $S$ ) and the extent of the stress ( $D$ ) and yield potential ( $\Upsilon_p$ ):

$$\Upsilon = \Upsilon_p(1 - S \times D)$$

where  $D = (1 - \bar{X}/\bar{X}_p)$ , and  $\bar{X}$  and  $\bar{X}_p$  are the mean yield of all cultivars under stressed and optimal conditions, respectively. This relationship may be manipulated algebraically to obtain:

$$\begin{aligned} S &= (1 - \Upsilon/\Upsilon_p)/D \\ &= (\Upsilon_p - \Upsilon)(\Upsilon_p \times D) \end{aligned}$$

where  $D$  = a constant for a particular trial. Hence,  $S$  is a measure of the yield decrease due to the stress relative to the potential yield, with a low value of  $S$  being desirable.

The problem with using  $S$  as a measure of adaptation to stress is that there are cases where  $S$  has been positively correlated with  $\Upsilon_p$  (i.e., cultivars whose yield was affected little by the stress also had very low yield potential). In other words, cultivars with low  $S$  also may have low stress resistance ( $\Upsilon$ ) and would not have been useful to the producer in the first place. The correlation between  $S$  and  $\Upsilon_p$  indicates that it may not be possible (or it would be challenging) to combine the desirable traits associated with a low  $S$  and high yield potential.

### **Types of abiotic environmental stresses**

J. Lewitt (1972) observed that plants respond to stress by strain reactions, which take the form that may be classified as either plastic or elastic, and manifest as  $G \times E$  interactions. Plastic responses produce a permanent change in the phenotype, whereas elastic responses are flexible and permit the normal state of the plant to return. Plants respond physiologically to stress by changing this reaction norm. Plant growth and development depend on biochemical processes (e.g., photosynthesis, respiration) that in turn depend on factors in the environment in order to proceed optimally. As previously indicated, when conditions in the environment are less than optimum, the plant experiences stress, which adversely affects its growth and development, and

ultimately its productivity and economic value. The common stresses that plants may be exposed to in agroecological systems include the following:

- 1 Drought.** This is the environmental condition caused by lack of rainfall.
- 2 Heat.** Heat stress occurs when temperatures are high enough to cause irreversible damage to plant function.
- 3 Cold.** Cold stress manifests when plants are exposed to low temperatures that cause physiological disruptions that may be irreversible.
- 4 Salinity.** Stress from salinity occurs when the dissolved salts accumulate in the soil solution to an extent that plant growth is inhibited.
- 5 Mineral toxicity.** Mineral toxicity occurs when an element in the soil solution is present at a concentration such that plants are physiologically impaired.
- 6 Oxidation.** Oxygen-free radicals (or activated oxygen) are known to cause degenerative conditions in plant cells.
- 7 Water-logging.** Excessive soil moisture as a result of prolonged rainfall can cause anoxic soil conditions, leading to roots suffering from lack of oxygen.
- 8 Mineral deficiency.** Inadequate amounts of essential soil minerals for plant growth cause crop injury.

Over the years, plant breeders have devoted attention and resources to addressing these environmental stresses to varying degrees and with varying success. The most widely studied stresses will be discussed in more detail in the next sections.

### Tolerance to stress or resistance to stress?

The terms **tolerance** and **resistance** are used in the literature to describe the mechanisms by which a plant responds to stress. Often, they are used as though they were interchangeable. According to J. Lewitt, from a physiological standpoint, a plant's response to stress may be characterized as "avoidance" (i.e., the environmental factor is excluded from the plant tissue) or "tolerance" (i.e., the factor penetrates the tissue but the tissue survives). The term resistance, from a physiological standpoint, is mechanism-neutral (implying neither tolerance nor exclusion). When the term is applied to bacteria, the development of resistance to an antibiotic has evolutionary stages. Full antibiotic resistance is not necessarily conferred by an immediate change in the bacterial genome. It is preceded by tolerance. Because bacteria can survive in the presence of an antibiotic, they have the opportunity to develop resistance.

Researchers who use resistance to describe plant response to a stress (cause) appear to view resistance as a generic term for describing a number of mechanisms of which tolerance is one. In breeding for response to a stress, the ultimate goal of the breeder is to transfer genes to the cultivar that would enable it to perform to a desirable degree in spite of the stress. In this chapter, the term that is most widely associated with a particular stress in the literature is used.

### Screening for stress resistance

Because of the complexity of environmental stress as previously discussed, simple, practical, and effective tests that can readily be used by breeders as selection aids are not widely available. A. Acevedo and E. Fereres (1993) summarized the criteria for the development and use of screening tests as follows:

- 1 Genetic variation should occur in the germplasm pool for the trait under consideration.
- 2 The heritability for the trait should be greater than the heritability for yield *per se*.
- 3 The trait should be correlated with a yield-based stress-resistance index.
- 4 It is desirable for the trait to be causally related to yield.
- 5 The screening test should be easy, rapid, and economical to apply.

Some of the traits associated with resistance to economic stress for which genetic variation has been found in wheat include osmotic adjustment, proline accumulation, leaf area per plant, epicuticular wax content, organ pubescence, tolerance in translocation, root growth, and many more. However, few of these traits can be readily applied as selection aids by plant breeders.

### Drought stress

Water is the most limiting factor in crop production. In tropical regions of the world, moisture extremes are prevalent. There is either too much of it when the rain falls, or there is little or lack of rainfall. Drought is responsible for severe food shortages and famine in developing countries.

#### What is drought stress?

**Drought** occurs both above ground (**atmospheric drought**) and below ground (**soil drought**). The

duration of drought is variable, sometimes lasting for a short time and without severe adverse physiological impact, sometimes lasting throughout an entire growing season or even years, resulting in complete devastation of crops. The efforts of breeders are directed at short-duration drought that often is experienced when crop production is rainfed. Under rainfed conditions, rainfall is often erratic in frequency, quantity, and distribution. To avoid disruption in growth and development processes, and consequently in crop performance (yield), plants need to maintain a certain level of physiological activity during the adverse period.

The effect of drought varies among species and also depends on the stage of plant growth and development at which the moisture stress occurs. Drought at flowering may cause significant flower drop and low fruit set. Similarly, when drought occurs at fruiting, there will be fruit drop and/or partially filled or shriveled fruits. In the end, both quality and product yield will be decreased.

### Overview of drought stress concepts

Scientists have developed crop simulation models that a researcher may use to estimate and quantify the impact of specific drought stress conditions on crop productivity. Models are available for cereals (e.g., maize, wheat, rice), grain legumes (e.g., soybean, dry bean, peanut), root crops (e.g., cassava, potato), and other crops (e.g., tomato, sugarcane).

As the demand for water exceeds supply, a plant develops what is called **plant water deficit**. The supply of water is determined by the amount of water held in the soil to the depth of the crop root system. The demand for water is determined by plant transpiration rate or crop evapotranspiration (both plant transpiration and soil evaporation). The rate of transpiration is influenced by solar radiation, ambient air temperature, relative humidity, and wind. These factors control transpiration at the single leaf level. The most important factor that controls transpiration at the whole-plant or crop level is total leaf area.

It is difficult to sense and estimate plant water stress at whole-crop levels because of the need to integrate an estimate based on the whole canopy. Various plant-based methods are used to obtain direct measurements of water status, stress status, and other physiological consequences of water deficit. These include leaf water potential, relative water content, and osmotic potential.

Soil moisture is measured in a variety of ways. The soil moisture content (volume) is measured by gravimetric

methods (soil is weighed before and after drying to determine water content). The soil water status is measured in terms of potential and tension. The amount of water a given crop can extract from the soil at a given water potential and depth is called the extractable soil moisture. Most crops usually extract between 50% and 80% of the extractable soil moisture before crop transpiration is reduced and symptoms of water deficit occur.

The relative importance of the major mediators of cellular response to water deficit are not exactly known. Plant hormones (e.g., abscisic acid) are believed to be a key factor in water stress response. Numerous water stress responsive genes have been identified. At the whole-plant or crop level, water deficit effects manifest in various ways – phenology, phasic development, growth, carbon accumulation, assimilate partitioning, and reproductive processes. These manifestations are primarily responsible for the variations in crop yield attributable to drought stress. Water deficit causes reduced cell expansion, reduced plant water use, and reduced plant productivity. Reduced cell expansion also adversely impacts meristematic development of yield components (e.g., inflorescence or tiller initials in cereals), leading to potentially small reproductive organs and hence reduced yield. Damage to the meristem is usually irreversible and cannot be corrected by irrigation. Water deficit also causes advanced or delayed flowering, depending on the species. This alteration in phasic development is known to be critical to maize yield under stress. Water deficit can cause reproductive failure whereby the pollen may desiccate, creating an effect similar to male sterility, and leading to reduced grain set. The duration of grain filling is reduced under stress. The root : shoot dry weight ratio increases as plants slip into water stress.

### Managing drought stress

Irrigation is the primary means of addressing drought in crop production, provided this approach is economical. Crop production may be designed for irrigated or dryland (rainfed) environments. In irrigated production, the common practice is to implement a supplemental irrigation regime, whereby irrigation is applied when rainfall is inadequate. Various soil and water conservation practices may be adopted to conserve soil moisture from one season to another. Practices such as fallow and the use of ground cover are recommended practices for soil and water conservation.

## Breeding for drought resistance

Drought resistance is highly cropping region specific. Drought resistance in crops is a major breeding objective in dryland farming systems. In more advanced agricultural production systems, the goal of combating drought is to obtain economic plant production in spite of the stress; plant survival is not the goal. However, in less advanced agriculture economies, plant survival is critical to producing some yield. The yield, albeit small, may mean the difference between famine and livelihood.

### Underlying principles

In order to formulate proper and effective breeding objectives in a drought breeding program, the breeder should understand the nature of the trait to be manipulated. Earlier plant breeding efforts were directed at developing a genotype that had water-saving capacity. However, this concept has proven to be inadequate in addressing modern crop development. Researchers (J. B. Passioura and C. T. de Wit) have produced a model to describe the relationship between crop yield and water use:

$$\text{Yield} = T \times \text{WUE} \times \text{HI}$$

where  $T$  = total seasonal crop transpiration,  $\text{WUE}$  = crop water use efficiency, and  $\text{HI}$  = crop harvest index. This relationship clearly indicates that the focus in breeding for drought resistance should be on water use (efficient water use), rather than water saving. In order for a crop to sustain yield, it should be able to use water under stress. Plants cannot live without water. Drought resistance is a finite trait. What is desirable in modern crop production is a plant with an ability to use water efficiently when this resource is limited by drought.

A successful drought breeding program depends on the formulation of an appropriate and relevant ideotype for a drought target environment. This task is very challenging, requiring the breeder to put together a conceptual description of detailed morphological, physiological, and developmental attributes of the ideal genotype.

### Characterization of the drought environment

Generally, constraints to crop production manifest as a combination of physiological stress factors (rather than one), even though one may predominate. For example, drought spells may be associated with low or high

temperature, depending on the region. In irrigated production, drought is often associated with salinity stress. Further, the soil may introduce additional factors that limit crop production (e.g., unfavorable pH, toxic levels of elements such as aluminum, and deficiencies of other such micronutrients). To complicate the stress system, the predominant factors may be unpredictable in occurrence, intensity, and duration. Two drought conditions may have significantly different composite properties. For example a drought event may occur with either a low vapor pressure deficit or high vapor pressure deficit.

Characterization of the drought environment is needed for several reasons, including the following.

- 1 Drought resistance is highly location-specific. This is because various factors (edaphic, biotic, agronomic) that are often location-specific are involved.
- 2 The drought-resistant genotypes to be developed need to be evaluated under a specific set of environmental conditions. Genotypes that are adapted to a given set of drought conditions may not be equally adapted to a different set of drought conditions.

It is clear then that breeders will have to characterize the target production region to know the most effective combination of adaptation features in genotypes best suited for successful production. Environmental characterization is a challenging undertaking. Consequently, breeders often limit the activity to only assembling qualitative meteorological data and soil descriptions. These data may be obtained at the regional level, or may even be as specific as farm level, and collected over seasons and years. Breeders engaged in breeding for resistance to drought should work closely with experts including agronomists, soil scientists, and meteorologists, or at least be familiar with techniques for computing soil moisture variations and rainfall probabilities. The modern technology of GIS (geographic information system) is helpful in detailed characterization of environments.

### Plant traits affecting drought response

Researchers have identified numerous stress response genes. However, these genes are not necessarily stress adaptive in terms of drought resistance. Nonetheless, both adaptive and non-adaptive genes are important in plant performance under drought stress. A good example of this fact is the role of genes that condition early flowering. Whereas these genes are expressed in any environment (i.e., not drought induced), early flowering may play a role in plant response to, and

performance under drought stress. Major plant traits that play a role in plant drought-resistance response include phenology, development and size, root, plant surface, non-senescence, stem reserve utilization, photosynthetic systems, and water use efficiency.

### *Phenology*

**Phenology** is the study of the times of recurring natural phenomena, especially as it relates to climate. Botanical phenology is the study of timing of vegetative activities, flowering, fruiting, and their relationship to environmental factors. For the plant breeder, developing a cultivar with short growth duration (i.e., early flowering) may help the cultivar escape drought that may occur in the late season of crop production. However, it is known that a longer growth duration is associated with high crop yield. Hence, using earliness for drought escape may reduce crop yield. Optimizing phenology is easier when there is a predictable environment. If the environment is highly unpredictable, using an early flowering indeterminate cultivar for crop production may be disadvantageous because stress would be unavoidable. Under such conditions, cultivars that are either determinate or indeterminate but have longer growth duration (late flowering) may have a chance to rebound after a drought episode to resume growth. However, late maturing cultivars may face late season stress from disease and frost.

Research indicates a genetic linkage between growth duration and leaf number, and often leaf size. Consequently, early maturing genotypes have a small leaf area index (LAI) and reduced evapotranspiration over most of the growth stages. Corn breeders make use of a phenological trait called **anthesis-to-silk interval** (ASI). A short ASI is desirable, whereas a longer ASI results in poor pollination. Crop plants differ in their phenological response to drought stress, some (e.g., wheat) advancing flowering, whereas others (e.g., rice) delay flowering.

### *Plant development and size*

Water use by plants is significantly influenced by single plant leaf area or LAI. Genotypes with small size and reduced leaf area are generally conducive to low productivity, while they limit water use. These genotypes may resist drought but their growth rate and biomass accumulation are severely slowed. Modern plant breeders tend to select for genotypes with moderate productivity and moderate size for use in dryland production.

### *Plant root characteristics*

Root characteristics are critical to ideotype development for combating drought. The most important control of plant water status is at the root. The primary root characteristics of importance are root depth and root length density, of which root depth is more important to breeding drought resistance. In cereals where tillering occurs, deeper root extension occurs when fewer tillers are produced. Unfortunately, many drought-prone regions of the world have shallow soils. Under such conditions, breeders may focus on other factors such as shoot developmental characteristics and osmotic adjustment. Also, development of lateral roots at very shallow soil depth may be advantageous in capturing the small amounts of rainfall that occur.

### *Plant surface*

Plants interact with the environment through surface structures. The form and composition of such structures determine the nature of the interactions. Dehydration avoidance is positively correlated with yield under stress. The reflective properties of leaves and resistance to transpiration depend on plant surface structures. The stomatal activity primarily determines the resistance of plant leaves to transpiration, but cuticular properties (e.g., wax load) also play a role. In sorghum, genotypes with lower epicuticular wax load (*bm*) had greater total leaf transpiration than genotypes with higher wax load (*Bm*). Another leaf surface feature with implication in drought resistance is pubescence. In soybean, high leaf pubescence genotypes had higher water use efficiency stemming from lower net radiation and transpiration with sustained photosynthesis. In wheat and barley, the role of leaf color in drought resistance has been noted. Yellow-leaf cultivars (with about one-third less leaf chlorophyll) tend to perform better than cultivars with normal green color, under drought stress.

### *Non-senescence*

Environmental factors such as drought, heat, and nitrogen deficiency are known to accelerate plant senescence. Certain genotypes, called “non-senescence”, “delayed senescence”, or “stay-green”, have been identified to have the capacity to delay or slow senescence. These genotypes have high chlorophyll content and leaf reflectance. The expression of stay-green properties varies among species. In sorghum, the condition is better expressed when plants have been exposed to

drought. Consequently, phenotypic selection for the trait may be more efficient under post-flowering drought stress.

#### *Stem reservation utilization*

Small grains and cereals store carbohydrates in their stems. The size of storage depends on favorable growing conditions before anthesis, and the genotype. The potential stem storage is determined by stem length and stem weight density. It has been found that the presence of *Rht<sub>1</sub>* and *Rht<sub>2</sub>* dwarfing genes in wheat limits the potential reserve storage by about one-half, as a result of the reduction in stem length. This stem reduction may be partly responsible for the observed greater drought susceptibility of the dwarf high-yielding wheat cultivars. The percentage of grain yield accounted for by stem reserves varies according to the environment and the genotype, and ranges between 9% and 100%. Stem reserve mobilization is a major source of carbon for grain filling under any stress. High reserve utilization, however, accelerates shoot senescence – a consequence of the export of stored carbohydrates into the grain. This finding suggests that a breeder may not be as successful in selecting for both traits (delayed senescence and high reserve mobilization) as a strategy for developing a cultivar with high grain filling under stress.

#### **Mechanisms of drought resistance**

Plant species differ in the stages at which they are most susceptible to drought stress. Some species are most prone to stress damage during the early vegetative stage, while others are most susceptible during the pre- or postanthesis stage, with others in between. Four general mechanisms may be identified by which plants resist drought:

- 1 **Escape.** Using early maturing cultivars may allow the crop to complete its life cycle (or at least the critical growth stage) before the onset of drought later in the season. The plants use the optimal conditions at the beginning of the season to develop vigor.
- 2 **Avoidance.** Some plants avoid drought stress by decreasing water loss, for example by having cuticular wax or by having the capacity to extract soil moisture efficiently.
- 3 **Tolerance.** In species such as cereals in which grain filling is found to depend on both actual photosynthesis during the stage, as well as dry matter distribution from carbohydrates in pre-anthesis, terminal drought significantly reduces photosynthesis. This

shifts the burden of grain filling to stored carbohydrate as the source of dry matter for the purpose. Consequently, such species may be more tolerant of postanthesis drought, being able to produce appreciable yield under the stress.

- 4 **Recovery.** Because drought varies in duration, some species are able to rebound (recover) after a brief drought episode. Traits that enhance recovery from drought include vegetative vigor, tillering, and long growth duration.

#### **Approaches for breeding drought resistance**

Plant breeders have two basic approaches for breeding for drought resistance – indirect breeding and direct breeding.

#### **Indirect breeding**

In this strategy, the breeder exposes genotypes to an environmental stress, even though they are not being directly evaluated for environmental stress. Indirect selection pressure is applied to these genotypes by conducting performance trials at locations where stress conditions exist. This approach is not advisable if the cultivars to be released are not intended for cultivation in the location where the evaluation was conducted. Under such atypical conditions, it is possible the cultivars might exhibit susceptibility to other stresses in their area of production.

#### **Direct breeding**

Direct selection for drought is best conducted under conditions where the stress factor occurs uniformly and predictably. Temperature and moisture are highly variable from one location to another and hence are difficult to predict. There are several methods used for direct breeding.

#### *Field selection*

Field selection is often problematic in drought-resistance breeding. Water requirement is variable from year to year, and may be sufficiently severe in one year to cause a loss of breeding materials. Further, drought in different seasons can occur at different growth stages. Without special management for stress, inconsistencies in the field may result in inconsistent selection pressure from one cycle to the next. The ideal field selection site would

be the desert, where irrigation may be used to dispense water at desired rates.

#### *Selection under managed stress environments*

Most drought breeding research is conducted under managed stress environments. A common facility for such research is the rainout shelter. This is essentially a mobile roof that protects selected plots from rain.

#### *Selection based on yield per se*

The concept of genotype  $\times$  environment ( $G \times E$ ) is used to evaluate genotypes at the end of a breeding program. The desired genotype is one with minimal  $G \times E$  interaction, and stable yield in its target environment and across other environments. Selection for yield under stress is generally an inefficient approach. The generally low heritability of yield becomes even lower under stress. To get around this problem, some breeders resort to the tedious and expensive approach of screening very large populations. Some also use molecular markers to tag and select certain yield-related quantitative trait loci (QTLs) to aid in the selection process.

#### *Selection based on developmental traits*

Most breeders tend to select for drought resistance on the basis of certain developmental traits, major ones including root size and development, and stem reserve utilization. Root studies are tedious to perform. The most widely used technologies for detailed studies of roots are the **rhizotrons** and **lysimeters**. These pieces of equipment are available in various forms and provide various kinds of information. The **minirhizotron** even has a minute video camera included in the rig to record roots as they appear on the external surface of the tubes in which the plant is growing. Some researchers grow plants in soil-filled tubes (polyethylene tubes) and remove and wash away the soil, usually at flowering time, to measure root length. To select by using stem reserve utilization for grain filling, researchers completely inhibit the photosynthetic source (e.g., spraying plants with oxidizing chemicals such as potassium iodide) and measuring grain filling without photosynthesis, to compare with normal plants.

#### *Selection based on assessment of plant water status and plant function*

Methods used in this category of selection approaches include assessing stress symptoms (e.g., leaf rolling, leaf

desiccation, leaf tip burning). An infrared thermometer is also used to measure canopy temperature, whereas infrared photography is used to measure spectral reflection from leaves. Selection based on plant function includes the measurement of cell membrane stability and chlorophyll fluorescence.

### **Breeding methods**

In breeding for drought resistance, breeders may select for early maturity for use in avoiding the stress. Also, genotypes with proven drought resistance may be used in hybridization programs to transfer the resistance genes into superior cultivars. The selection methods discussed and the various breeding methods previously discussed are used in the selection process.

## **Cold stress**

In the USA, the US Department of Agriculture (USDA) has developed a plant adaptation map (plant hardiness zones) based on temperature requirements, to guide in the selection of plants for use in various parts of the country. Tropical and semitropical plants are cold intolerant.

#### **Overview of cold stress concepts**

Plants use a variety of adaptive mechanisms to combat environmental stress caused by cold temperature. Seed dormancy is a physiological condition that delays seed germination until the embryo has gone through an after-ripening period, during which certain biochemical and enzymatic processes occur for the seed to attain full maturity. In many temperate zone trees and shrubs, buds undergo a dormant stage, starting in late summer or early fall, ending only when the buds have been exposed to an extended period of cold or increasing day length in spring. Most winter annuals and biennials have a cold temperature requirement (vernalization) before they will flower. When plants are exposed to gradually decreasing temperatures below a certain threshold, they acclimatize (low-temperature acclimation) to the stress, a process called **cold hardening**.

In spite of various adaptations to cold, plants may be injured through exposure to cold temperatures in a variety of ways, depending on the temperature range. One type of injury, called **chilling injury**, occurs at exposure to temperatures between 20 and 0°C. Some injuries are irreversible. Common chilling injuries include



## Industry highlights

### *Discovering genes for drought adaptation in sorghum*

Andrew Borrell<sup>1</sup>, David Jordan<sup>1</sup>, John Mullet<sup>2</sup>, Patricia Klein<sup>2</sup>, Robert Klein<sup>3</sup>, Henry Nguyen<sup>4</sup>, Darrell Rosenow<sup>5</sup>, Graeme Hammer<sup>6</sup>, and Bob Henzell<sup>1</sup>

<sup>1</sup>Department of Primary Industries and Fisheries, Hermitage Research Station, Warwick, Queensland 4370, Australia;

<sup>2</sup>Institute for Plant Genomics and Biotechnology, Texas A&M University, College Station, TX 77843, USA; <sup>3</sup>USDA-ARS Southern Agricultural Research Station, College Station, TX 77843, USA; <sup>4</sup>Plant Sciences Unit and National Center for Soybean Biotechnology, University of Missouri, Columbia, MO 65211, USA (previously Texas Tech University, Lubbock, USA); <sup>5</sup>Texas A&M Agricultural Research and Extension Center, Lubbock, TX 79403, USA; <sup>6</sup>School of Land and Food, University of Queensland, Queensland 4072, Australia

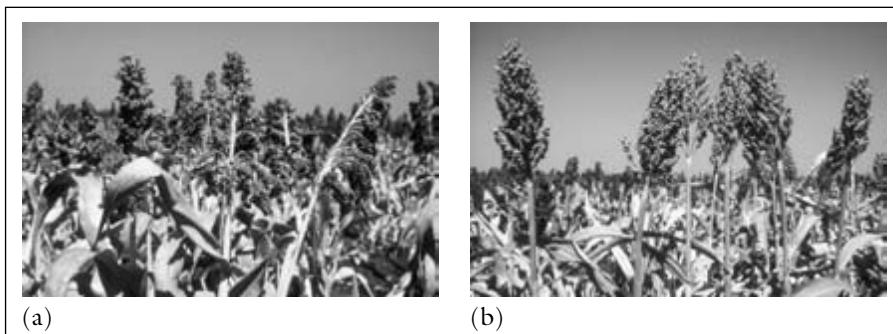


**Figure 1** Harvesting the sorghum experiment in southern India at the end of the dry season. The B35 variety displayed remarkable resilience under extremely dry conditions.

*Looking up from my notebook, I was stunned to see a plot of sorghum with a number of large green leaves and strong green stems (Figure 2). Balancing on the end of these stems were large panicles yielding about three times as much grain as the other plots. How could this be? Our field plan revealed that this particular line was B35, a stay-green line from Ethiopia that was first documented by Dr Darrell Rosenow, a plant breeder at Texas A&M University. B35 is derived from a durra landrace, an ancient*

*The heat shimmered above the dry earth, engulfing all in its wake. The black clay soil was deeply cracked, and like a cosmic vacuum cleaner, the sun sucked the last hint of moisture from its depths. Yet this desolate landscape was not without life. A crop of sorghum stood defiantly, thrusting its red grain into the copper sky. It was the dry season in southern India and no rain had fallen for many weeks. The crop was almost ready to harvest, but the yield would be low due to the severe drought (Figure 1). I walked through my experimental plots at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) on the outskirts of Hyderabad. A young Indian scientist accompanied me on this stroll through our 280 plots. We stopped briefly at each plot, noting if any leaves still remained alive. We were examining a population of sorghum lines that varied in the stay-green drought-resistance trait.*

*Plants containing the stay-green trait maintain more green leaf and stem under drought compared with senescent (non-stay-green) plants, resulting in stronger stems and higher grain yield. In most plots we found all the leaves had died. It was not unusual to find whole plots laying on the ground, their stems greatly weakened by the drought. I will never forget what happened next.*



**Figure 2** Photographs of (a) senescent and (b) stay-green sorghum taken on the same day just prior to harvesting in the field experiment. The B35 stay-green variety is derived from a durra landrace from Ethiopia.

*type of sorghum that was eaten in Egypt over 4,000 years ago. I knelt in the dust to examine the soil, making sure this plot wasn't receiving any additional water. Like the other plots, deep cracks snaked through the soil indicating the severity of drought. Now it was time to check the other replicates. A critical component of experimental design is the existence of replicates to verify outcomes. Excitedly, we made our way over to the next replicate of B35. Before we even arrived I could see it standing out from its lifeless neighbours. Yes! Quickly we headed to the third and fourth replicates. The same again! I was amazed at the resilience of B35, and determined to find out what drought adaptation mechanisms contributed to its remarkable survival. For me, the adventure of cracking the stay-green phenomenon had just begun.*

**Andrew Borrell**

### ***Introduction***

Producing more grain with less water is one of the greatest challenges facing crop scientists in the 21st century. Globally, the availability of fresh water per capita has declined 37% since 1970 as population growth and degradation of water supplies has surpassed the capacity to develop new sources (Downer 2000). Governments all over the world are choosing carefully how they allocate water between agricultural, urban, and industrial uses. In a contest between these three, agriculture is often the loser because water used for irrigation generally produces a smaller economic return than water diverted to industry (Dupont 2000), with urban requirements being even more important for many governments. Yet in the face of diminishing water resources, the world is expected to consume twice as much food in the next 50 years as it has in the past 10,000 years. To meet this demand, world grain production will have to increase 40% by 2020 (Dupont 2000).

The case study above describes how a multidisciplinary team of Australian and US scientists are collaborating to discover genes for drought adaptation in sorghum. The potential to utilize these genes in the world's other major cereals is also discussed. Sorghum is a repository of drought-resistance mechanisms, and has developed biochemical, physiological, and morphological characteristics such as C<sub>4</sub> photosynthesis, deep roots, and thick leaf wax that enable growth in hot and dry environments. Sorghum is the dietary staple of more than 500 million people in over 30 countries, making it the world's fifth most important crop for human consumption after rice, wheat, maize, and potatoes (Miller 1996).

### ***Multidisciplinary approach***

In many areas of human endeavor, it is often the integration of fields of knowledge that proves to be the fertile ground for innovation. So it is with "gene discovery" in the world's most important cereal crops. The pursuit of drought-resistance genes in sorghum is a multidisciplinary effort involving plant breeders, crop physiologists, molecular biologists, biometricalians, functional genomists, and simulation modelers. Scientists from Australia and the USA are collaborating in the search for genes (*Stg1*, *Stg2*, *Stg3*, and *Stg4*) associated with the "stay-green" trait in grain sorghum. Keeping leaves alive for longer is a fundamental strategy for increasing crop production, particularly under water-limited conditions. During postanthesis drought, genotypes possessing the stay-green trait maintain more photosynthetically active leaves than genotypes not possessing the trait. The broad objective of this research is to identify and understand the function of the genes and gene networks that contribute to improved plant drought resistance under water-limited conditions.

### ***Approaches to gene discovery***

There are two general approaches to identifying and isolating genes involved in drought resistance (Mullet et al. 2001). First, genes are targeted that show relatively rapid changes in expression at the RNA level in response to water limitation. Second, sorghum genes involved in drought adaptation are identified and isolated using map-based gene discovery. The current stay-green project primarily utilizes map-based gene discovery undertaken by scientists at Texas A&M University, although microarray analysis is being used simultaneously to assist in gene discovery.

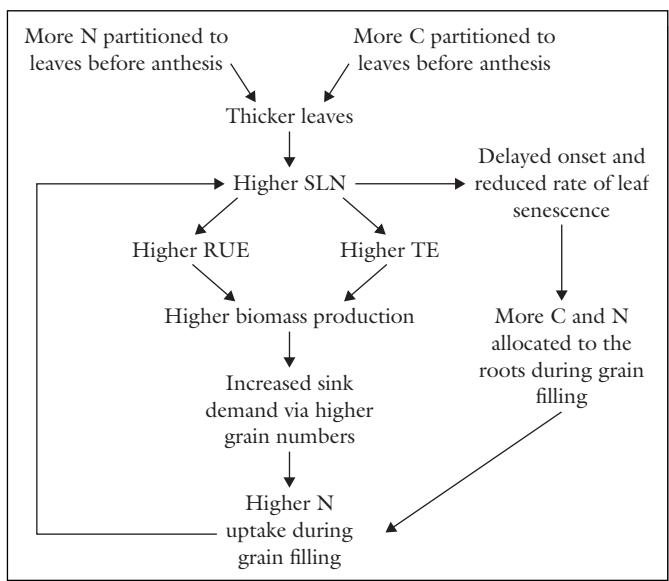
### ***Phenotyping, genotyping, and physiological characterization***

#### **Phenotyping driving genotyping**

Map-based cloning requires the accurate screening of the phenotype and genotype of large segregating populations (Tanksley et al. 1995), highlighting the need for collaboration between plant breeders, crop physiologists, and molecular biologists. Typically, plant breeders develop a range of populations for mapping (e.g., recombinant inbred lines), fine mapping (e.g., segregating populations with breakpoints across the loci of interest), and physiological dissection (e.g., near-isogenic lines). Such populations are systematically phenotyped and genotyped by crop physiologists and molecular biologists, respectively, resulting in the identification of regions of genomes (trait loci) that modulate the expression of traits such as stay-green.

#### **Genotyping driving phenotyping**

Following the mapping of drought-resistance loci, efficient map-based cloning requires the availability of a high-resolution integrated genetic and physical map, large populations, and careful phenotyping (Mullet et al. 2001). The construction of an integrated



**Figure 3** Nitrogen (N) dynamics and the stay-green phenomenon in grain sorghum (Borrell et al. 2003). C, carbon; RUE, radiation use efficiency; SLN, specific leaf nitrogen; TE, transpiration efficiency.

Stay-green can be viewed as a consequence of the balance between nitrogen demand by the grain and nitrogen supply during grain filling at a whole-plant level (Figure 3). Matching nitrogen supply from age-related senescence and nitrogen uptake during grain filling with grain nitrogen demand found that the shortfall in nitrogen supply for grain filling was greater in the senescent than the stay-green hybrids, resulting in more accelerated leaf senescence in the former (Borrell et al. 2001). Preliminary simulation modeling to assess the value of stay-green in a range of environments found that improved nitrogen dynamics alone could not explain the observed yield increases under drought; enhanced transpiration efficiency (TE) was also required (Chapman et al. 2003).

#### Identifying candidate genes

Initial mapping activities can generally map targeted loci to 1–5 cM regions of the sorghum genetic map (e.g., Klein et al. 2001). Analysis of large segregating populations (c. 1,000 plants) is usually required to provide sufficient genetic resolution for efficient map-based cloning. Fine mapping can then reduce the target locus in euchromatic regions to less than 100 kbp, a size that can be readily sequenced using standard BAC-based shotgun sequencing approaches. Interestingly, c. 100 kbp of sorghum DNA, on average, will encode c. 10 sorghum genes. There are several ways to identify genes within the target genomic interval. First, if the targeted region is less than 500 kbp, shotgun sequencing of BAC DNAs spanning the region followed by BLASTX analysis can be used to identify sorghum genes that are related to other known protein coding genes. Second, the sorghum sequence can be compared to the sorghum expressed sequence tag (EST) database to identify the transcribed portions of the BAC sequence. Third, other genes encoded by the sorghum BAC sequence can be identified by aligning the sorghum sequence with orthologous rice or maize sequences. Finally, gene prediction programs such as FGENESH (<http://www.softberry.com/>) and riceGAAS (<http://ricegaas.dna.affrc.go.jp/>) can be used to identify regions of the sorghum sequence that may encode genes.

#### Proof of gene function

Identification of a candidate gene(s) is not the end point. Direct validation that a candidate gene causes variation in the trait under investigation is still required. Currently, this step is difficult in sorghum because sorghum gene transformation technology is time-consuming, throughput is low, and not all genotypes are easily transformed (Mullet et al. 2001). Other technologies such as RNAi are now under development in order to accelerate candidate gene validation (Holzberg et al. 2002).

#### Assessing gene function across environments via simulation modeling

Crop modeling contributes to the genetic regulation of plant performance and improvement in a number of ways (Hammer et al. 2002). For the purpose of this discussion, however, it is worth noting the role of simulation modeling as a means of both

sorghum genome map is well underway. A genetic map with about 3,000 points has been constructed with AFLPs (amplified fragment length polymorphisms), SSRs (simple sequence repeats), and RFLPs (restriction fragment length polymorphisms) using a *Sorghum bicolor* recombinant inbred population (Menz et al. 2002). In addition, physical maps of the sorghum genome are being constructed using BAC (bacterial artificial chromosome) libraries that provide about 20x coverage of the sorghum genome (Klein et al. 2003). The resulting integrated genome maps are being aligned to the rice genome sequence (e.g., Klein et al. 2003).

#### Physiological characterization

The aim is to dissect complex traits such as stay-green into their functional components. The characteristics of such complex traits can be viewed as emergent consequences of the interactions between underlying determinants and the prevailing environmental conditions (Hammer 1998). Integration of knowledge from gene to cropping system is also necessary. For example, stay-green can be viewed at a cell, leaf, whole-plant, crop, or system level. Understanding how gene networks respond to water deficits across these levels is critical to capturing traits like stay-green in plant breeding programs.

determining and assessing gene function *in silico*: the complexity arising from genes interacting with each other and their environments (both natural and managed) requires a mechanism to assimilate the many and varied combinations. Simulation modeling provides such a mechanism.

### A blue future

As a result of this research, we will know the identity and function of the different genes involved in stay-green, enabling sorghum breeders to develop a new generation of crops adapted to our increasingly water-scarce world. Ultimately, we hope to transfer the stay-green genes to grain crops less well adapted to dry environments (i.e., rice, wheat, maize), or to identify similar drought-resistant mechanisms in those crops. The global conservation of gene order among sorghum, rice, and other grass species is clear from RFLP mapping information (Wilson et al. 1999). Utilizing stay-green genes in rice could revolutionize rain-fed rice production.

The 1960s and 1970s were dominated by the “Green Revolution”, when semidwarf varieties of wheat and rice were grown with high water and nitrogen inputs to produce record yields of grain. The early 21st century needs to be dominated by a “Blue Revolution” in which genetic, agronomic, and management solutions are integrated to ensure food security in the face of global water shortages.

### Acknowledgments

The Grains Research and Development Corporation and the Queensland Department of Primary Industries and Fisheries funded the Australian aspects of this research. Components of this research undertaken in the United States were funded by the National Science Foundation.

### References

- Borrell A.K., G.L. Hammer, and E. Van Oosterom. 2001. Stay-green: A consequence of the balance between supply and demand for nitrogen during grain filling? *Ann. Appl. Biol.* 138:91–95.
- Borrell, A.K., E. Van Oosterom, G. Hammer, D. Jordan, and A. Douglas. 2003. Proceedings of the Australian Agronomy Conference. Australian Society of Agronomy, Gosford, Australia.
- Chapman, S.C., M. Cooper, D. Podlich, and G.L. Hammer. 2003. Evaluating plant breeding strategies by simulating gene action and dryland environment effects. *Agron. J.* 95:99–113.
- Downer, A. 2000. Sustenance and security: Australia’s multi-layered approach. In: Proceedings of a conference on food, water and war: Security in a world of conflict. Crawford Fund for International Agricultural Research, Parliament House, Canberra, 15 August 2000. ACIAR Monograph No. 73, pp. 9–15.
- Dupont, A. 2000. Food, water and security: What are the connections? In: Proceedings of a conference on food, water and war: Security in a world of conflict. Crawford Fund for International Agricultural Research, Parliament House, Canberra, 15 August 2000. ACIAR Monograph No. 73, pp. 39–59.
- Hammer, G.L. 1998. Crop modelling: Current status and opportunities to advance. *Acta Horticulturae* 456:27–36.
- Hammer, G., M.J. Kropff, T.R. Sinclair, and J.R. Porter. 2002. Future contributions of crop modelling – from heuristics and supporting decision-making to understanding genetic regulation and aiding crop improvement. *Eur. J. Agron.* 18:15–31.
- Holzberg, S., P. Brosio, C. Gross, and G.P. Pogue. 2002. Barley stripe mosaic virus-induced gene silencing in a monocot plant. *Plant J.* 30:315–327.
- Klein, P.E., R.R. Klein, J. Vrebalov, and J.E. Mullet. 2003. Sequence-based alignment of sorghum chromosome 3 and rice chromosome 1 reveals extensive conservation of gene order and one major chromosomal rearrangement. *Plant J.* 34:605–621.
- Klein, R.R., P.E. Klein, A.K. Chhabra, et al. 2001. Molecular mapping of the *rf1* gene for pollen fertility restoration in sorghum (*Sorghum bicolor* L.). *Theor. Appl. Genet.* 102:1206–1212.
- Menz, M.A., R.R. Klein, J.E. Mullet, J.A. Obert, N.C. Unruh, and P.E. Klein. 2002. A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP®, RFLP and SSR markers. *Plant Mol. Biol.* 48:483–499.
- Miller, F. 1996. Sorghum. In: Lost crops of Africa, Vol. I. Grains (National Academy of Sciences, ed.), pp. 127–143. National Academy Press, Washington, DC.
- Mullet, J.E., R.R. Klein, and P.E. Klein. 2001. *Sorghum bicolor* – an important species for comparative grass genomics and a source of beneficial genes for agriculture. *Curr. Opinion Plant Biol.* 5:118–121.
- Tanksley, S.D., M.W. Ganal, and G.B. Martin. 1995. Chromosome landing: A paradigm for map-based gene cloning in plants with large genomes. *Trends Genet.* 11:63–68.
- Wilson, A.W., S.E. Harrington, W.L. Woodman, M. Lee, M.E. Sorrells, and S.R. McCouch. 1999. Inferences on the genome structure of progenitor maize through comparative analysis of rice, maize and the domesticated panicoids. *Genetics* 153:453–473.

interruption in normal germination, flowering, and fruit development, which eventually adversely impact crop yield. Stored products may also suffer chilling injury. A more severe low-temperature injury is **freezing injury**, which occurs when temperatures drop below the freezing point of water. Sometimes, ice crystals form in the protoplasm of cells, resulting in cell death and possibly plant death.

Plants may be classified into three groups according to their tolerance to low temperatures. **Frost-tender** plants are intolerant of ice in their tissues and are hence sensitive to chilling injury. The plant (e.g., beans, corn, tomato) can be killed when temperatures fall just below 0°C. **Frost-resistant** plants can tolerate some ice in their cells and can survive cold temperatures of up to -40°C. **Cold-hardy** plants are predominantly temperate woody species. They can survive temperatures of up to -196°C.

Most crops that originate in the tropics and subtropics are sensitive to chilling temperatures. However, some temperate fruits are also susceptible to chilling injury. The temperature at which chilling injury starts varies among species and depends on where they originate. Temperate fruits exhibit chilling injury starting at 0–4°C, whereas the starting temperature is 8°C for subtropical fruits and 12°C for tropical fruits. Grain such as corn and rice suffer chilling injury at temperatures below 10°C. When chilling temperatures occur at the seedling stage, susceptible crops suffer stand loss. Also, crop maturity is delayed while yield is reduced.

### Genetic basis of low-temperature stress tolerance

The capacity of a genotype to tolerate low temperatures has been extensively studied. Whereas it is agreed that low-temperature tolerance is a complex trait (quantitative), researchers are not unanimous on the mode of gene action governing the expression of the trait. Reports indicate recessive, additive, partial dominance, and overdominance as the modes of action that occur in nature for cold stress. The inconsistency in the results is partly blamed on the way research is often conducted. Some workers use controlled freeze tests while others use field tests. Further, various reports indicate the role of cytoplasmic factors and non-additive gene effects, even though such effects are generally believed to be minor. Genes that condition varying levels of low-temperature tolerance occur within and among species. This genetic variability has been exploited to a degree in cultivar development within production regions.

A large amount of low-temperature tolerance research has been conducted in wheat. Low-temperature tolerance in cereals depends on a highly integrated system of structural, regulatory, and developmental genes. Several vernalization genes have been identified (e.g., *vrn*<sub>1</sub>, *vrn*<sub>4</sub>). The *vrn*<sub>1</sub> is homeoallelic to the locus *Sh*<sub>2</sub> in barley and *Sp*<sub>1</sub> in rye. These two genes have been linked to genetic differences in low-temperature tolerance. Winter cereals also produce several proteins in response to low-temperature stress, for example the dehydrin families of genes (*dh*<sub>5</sub>, *Wcs120*).

### Mechanisms of resistance to low temperature

Like drought resistance, certain physiological or morphological adaptations can make plants either avoid or tolerate stress due to low temperatures. Plants are described as cold hardy (concept of cold hardiness) when they have the capacity to withstand freezing temperatures. On the other hand winter-hardy (winter hardiness) species are able to avoid or tolerate a variety of weather-related effects associated with winter (e.g., freezing, heaving, desiccation, frost resistance, etc.).

The mechanism of low-temperature resistance may be grouped into two.

#### Chilling resistance

The factors that confer resistance to chilling are believed to operate at the cell membrane level where they influence membrane fluidity. Chilling-resistant seeds are known to imbibe moisture slowly. The presence of phenols in the seed coat of legumes is implicated in conferring chilling resistance.

#### Freezing resistance

Several mechanisms are used by plants to resist freezing, including the following:

**1 Escape.** Like drought, cultural practices may be adapted by producers to prevent the vulnerable stage of growth coinciding with the presence of the stress factor.

**2 Avoidance.** One of the injuries of low temperature results from the intracellular formation of ice following nucleation of ice in the tissue. Water may remain supercooled without forming ice crystals. Certain compounds that are capable of promoting ice nucleation are active at low temperatures. Bacteria such as *Pseudomonas syringae* are capable of producing

ice-nucleating proteins. The first field test of a bio-engineered organism was the testing of “**ice minus**”, a microbe genetically engineered to be incapable of producing the bacterial protein that causes ice nucleation. This was intended to be an approach for helping frost-sensitive plants survive frost.

- 3 Tolerance.** Freezing tolerance occurs when a plant is able to withstand both intracellular and extracellular ice formation.

### Selection for low-temperature tolerance

As previously stated, field survival trials have proven to be inefficient for selecting genotypes with low-temperature tolerance. Because low-temperature stress brings about many changes in plants (including morphological, biochemical, and physiological changes), researchers are pursuing these avenues in search of selection aids. Some factors that have shown promise in predicting low-temperature tolerance include plant erectness in winter (for cereals), tissue water content, and cell size. Unfortunately, these tests are not effective in discriminating among small differences that are of practical breeding importance.

Researchers commonly use controlled freeze tests conducted in artificial environments to measure low-temperature tolerance. For example the field survival index developed by Fowler and Gusta is often used. The researchers found that the crown and leaf water content of field-acclimated plants was a good indication of field survivability. Molecular marker technology is being pursued in the quest for QTLs associated with low-temperature tolerance. Other biotechnology tools are being explored to help transfer low-temperature tolerance genes into cultivars. In spite of not being as efficient and as desirable as controlled tests, field testing remains a widely used screening approach in low-temperature tolerance breeding. When other selection approaches are used, field testing is used as a final measure of plant winter survival. Researchers can take various precautions to improve the efficiency of field tests.

### Breeding for tolerance to low-temperature stress

Whereas the genetics of low-temperature tolerance have been studied to a reasonable degree, breeders have only had minimal success in applying research results to practical breeding. Breeding superhardy cultivars remains a challenge. Several reasons have been proposed by D. B. Fowler and A. E. Limin to be causal:

- 1 Exploitable genetic variability for low-temperature tolerance has been largely exhausted within the existing gene pools of most species.
- 2 A large number of genes with small effects and complex interaction are assumed to determine the phenotypic expansion of low-temperature tolerance, making selection difficult.
- 3 Current methodologies for measuring low-temperature tolerance give poor resolution of small phenotypic differences.
- 4 Measures of low-temperature tolerance lack the precision for single-plant analysis and many are destructive, making selection procedures complicated.
- 5 Poor expression of low-temperature tolerance in alien genetic backgrounds has prevented the expansion of gene pools through interspecific and intergeneric transfers (e.g., the superior low-temperature tolerance of rye is suppressed in the wheat background).

These researchers further observe that the ability to acclimate or avoid low-temperature stress varies among species and stages of crop growth. Hence, it is impossible to develop a breeding approach that will be applicable to all low-temperature tolerance breeding programs. Each breeder would essentially have to design a unique approach to breeding for cold stress.

### Salinity stress

Soil salinity constraints to crop production occur in an estimated 95% million hectares worldwide. **Salinity** is the accumulation of dissolved salts in the soil solution to a degree that inhibits plant growth and development.

#### Overview of salinity stress concepts

Soils with salinity problems are described as **salt affected**. When the salt concentration measured in terms of electrical conductivity ( $EC_e$ ) is more than 4 dS/m and the pH is less than 8.5, the soil is called a **saline soil**. When the  $EC_e$  value is less than 4 dS/m and the pH is more than 8.5, the soil is a **sodic soil**. Sodic soils are high in  $Na^+$  but low in other soluble salts. Semiarid regions have saline/sodic soils, whereby salts accumulate in subsoils because of low permeability of the subsoil.

Salinity may have a natural origin (called **primary salinity**) as a result of weathering of parent materials that are rich in soluble salts. Human-aided salinity (called **secondary salinity**) occurs as a result of agricultural activities, especially irrigation with impure (salt-rich) water. Salinity is often caused by a rising watertable.

**Table 21.1** Relative salt tolerance in plants.

Tolerant	Moderately tolerant	Moderately sensitive	Sensitive
Barley (grain)	Barley (forage)	Alfalfa	Strawberry
Cotton	Sugar beet	Peanut	Bean
Bermudagrass	Wheat	Corn	Potato
	Oat	Rice (paddy)	Tomato
	Soybean	Sweet clover	Pineapple
	Sorghum	Sweet potato	Onion

Plant growth is inhibited in salt-affected soils because the high salt concentration in the soil solution inhibits the process of water absorption by osmosis (osmotic stress). When excessive amounts of salt enter the transpiration stream, plant cells may be injured. Plants that are tolerant of high soil salt concentration are called **halophytes**. Wheat is one of the most salt-tolerant crops, while rice is one of the most salt-sensitive crops, although both are only moderately tolerant. Maize is also moderately tolerant of salts in the soil solution (Table 21.1).

### Breeding for salt tolerance

A common approach to breeding salt tolerance starts with assembling and screening germplasm for salinity tolerance. The selected genotypes are used as parents to transfer the trait to desired cultivars, followed by selecting desirable recombinants from the segregating population. This approach has yielded some success in species such as rice, wheat, and lucerne. The challenge in breeding for salt tolerance is how to measure salinity tolerance. Screening is commonly based on the growth of plants under salt stress. Two distinct mechanisms exist for salinity tolerance:

- 1 Tolerance to the osmotic effect of the saline solution (the osmotic effect makes it harder for plants to extract water from the soil).
- 2 Tolerance to the salt-specific nature of the soil saline solution (a high sodium concentration makes it difficult for the plant to exclude NaCl while taking up other ions).

Screening for salinity tolerance is a long process and requires a large amount of space to screen progeny from crosses. Screening for specific traits is quicker and more effective (they are less influenced by the environment than growth rates are). The most successful traits for assessing the salt-specific effect in salt-tolerance breeding is the rate of  $\text{Na}^+$  or  $\text{Cl}^-$  accumulation in leaves.

This is measured as the increase in salt in a given leaf over a period of time. Selection for these ions was used in breeding rice and lucerne cultivars with high salt tolerance. Traits for osmotic effects are related to growth (e.g., leaf elongation, root elongation, shoot biomass, leaf area expansion), the latter two being the most effective indices. Molecular marker technology and genetic engineering techniques are being used in salt-tolerance breeding efforts. Salinity tolerance has been found in the wild species of crops such as tomato, pigeon pea, and common bean.

### Heat stress

**Heat stress** may be defined as the occurrence of temperatures hot enough for a sufficient time to cause irreversible damage to plant function or development. A heat-resistant genotype is one that is more productive than another genotype in environments where heat stress occurs. Heat tolerance is the relative performance of a plant or plant process under heat compared with the performance under optimal temperatures.

### Overview of heat stress concepts

Heat stress occurs to varying degrees in different climatic zones. High temperatures can occur during the day or night. Also, temperature effects can be atmospheric or in the soil, with air temperature varying considerably during the day and night. Annual crop species may be classified into two categories according to maximum threshold temperatures as either cool season annuals or warm season annuals (Table 21.2). Cool season species are more sensitive to hot weather than warm season species.

High night temperatures have detrimental effects on the reproductive function of plants. It has been shown that there is a distinct period during the 24-hour day cycle when pollen development is most sensitive to high

**Table 21.2** Examples of warm and cool season crops.

Cool season plants	Warm season plants
Sugar beet	Okra
Cabbage	Eggplant
Apple	Corn
Wheat	Cotton
Barley	Sugarcane
Cauliflower	Peanut
	Sunflower
	Sorghum

Note: some species have wide adaptation with varieties that are adapted to both cool and warm growing regions.

night temperatures. In cowpea, plants that were exposed to high temperature during the last 6 hours of the night showed a significant decrease in pollen viability and pod set. Further, this damage was more pronounced in long days than short days. Other researchers also show that the stage of floral development most sensitive to high night temperature was between 7 and 9 days before anthesis.

Excessive heat in the soil affects the emergence of seedlings of both cool and warm season crops causing reduced crop stands. High temperatures tend to accelerate reproductive development. This may be part of the reason why the potential grain yields of warm season crops (e.g., rice, cowpea) are usually higher in the subtropics than tropics.

### Breeding for resistance to heat stress

Breeding for resistance to heat stress has not been as widely addressed as other environmental stresses that plants face in crop production. Heat resistance is more beneficial to the producer than heat tolerance. Some plant breeders use a direct measure of heat resistance in an approach to breeding whereby advanced lines are grown in a hot target production environment. Genotypes with greater yield than current cultivars are selected as superior. This breeding approach is more applicable for species that can be efficiently yield-tested in small pots (e.g., wheat) than for those that require larger plots or are more difficult to harvest. Breeders may also use this approach in environments where heat is the only major stress. When other stresses occur, the evaluation of heat damage is less conclusive (e.g., insect pests can cause damage to developing flower buds, similar to that which would occur under heat stress).

An approach to breeding heat resistance that is deemed by some to be more efficient is to select for specific traits that confer heat tolerance during reproductive development. To do this, genotypes with heat tolerance have to be discovered. This involves screening large accessions from germplasm collections. These genotypes can then be crossed with desirable cultivars if they lack the yield and other plant attributes desired.

The use of a controlled environment (hot greenhouse) has the advantage of providing a stable high night-time temperature and stable air temperature from day to day and over a longer period of time. It is conducive to screening for reproductive-stage heat tolerance. However, the facility can handle only a limited number of plants, compared to thousands of plants in a field evaluation. Selection aids (e.g., leaf electrolyte leakage) have been used by some researchers to identify genotypes with heat tolerance.

### Mineral toxicity stress

Plants obtain most of their nutrient requirements from the soil, largely from the products of weathering of mineral rocks or the decomposition of organic matter. Uptake in improper amounts may lead to toxic consequences to plants.

### Soil nutrient elements

Metals occur naturally in soils, some of which are beneficial and essential for plant growth and development, while others are toxic. About 16–20 elements have been identified as essential to plant nutrition. These may be broadly classified into two groups based on the amounts taken up by plants as **major (macro) nutrient elements** (these are required in large amounts) and **minor (micro) nutrient elements** (required in very small amounts) (Table 21.3). Each element has an optimal pH at which it is most available in the soil for plant uptake. However, at extreme conditions of soil reaction, excessive amounts of some elements become available. Some micronutrients are required in only trace amounts; their presence in large quantities in the soil solution may be toxic to plants. Some of the known toxicities of metallic elements occur at low pH (high acidity) and include iron and aluminum toxicities.

### Aluminum toxicity

Aluminum (Al) is one of the most abundant elements in the earth's crust. One of the most important metal

**Table 21.3** Summary of selected essential mineral nutrients for plants and their roles.

<b>Macronutrients</b>	
Nitrogen (N)	Used in synthesis of amino acids and proteins; component of chlorophyll and enzymes
Phosphorus (P)	Found in proteins and nucleic acids; critical in energy transfer process (adenosine triphosphate)
Potassium (K)	Catalyst for enzyme reactions; important in protein synthesis, translocation, and storage of starch
<b>Micronutrients</b>	
Calcium (Ca)	Important in cell growth, cell division, and cell wall formation
Magnesium (Mg)	Central atom of chlorophyll molecule; essential in formation of fats and sugars
Sulfur (S)	An ingredient in vitamins and amino acids
Boron (B)	Role in flowering, fruiting, cell division, and water relations
Iron (Fe)	Component of many enzymes; catalyst in synthesis of chlorophyll
Molybdenum (Mo)	Role in protein synthesis and some enzymes
Manganese (Mn)	Role in phosphorylation, activation of enzymes, and carbohydrate metabolism
Zinc (Zn)	Role in enzyme activation
Copper (Cu)	Catalyst for respiration and carbohydrate and protein metabolism

toxicities of economic importance to crop production is that which occurs when aluminum concentrations are greater than 2–3 ppm. At acid pH, Al<sup>3+</sup> ions predominate in the soil. Aluminum is not an essential nutrient for plants. At a pH of 5 or less, aluminum inhibits plant growth by interfering with cell division in root tips and lateral roots, increasing cell wall rigidity, reducing DNA replication, decreasing respiration, and other effects. In some cases, excess aluminum induces iron deficiency in some crops (e.g., rice, sorghum, wheat). A visual symptom of aluminum toxicity is so-called **root pruning**, whereby root growth is severely inhibited. Stunting of roots leads to chronic drought and nutrient stress in afflicted plants.

### Breeding for aluminum tolerance

Aluminum-tolerant genotypes have been identified. Based on the patterns of aluminum accumulation in plant tissue, three groups of aluminum-tolerant plants may be identified: (i) those with an apparent exclusion mechanism allowing lower accumulation of aluminum in their roots than aluminum-sensitive plants (e.g., wheat, barley, soybean); (ii) those with less aluminum in the shoot but more in the roots (e.g., wheat, barley, potato); and (iii) those with high aluminum accumulation in the shoot (e.g., pine trees). Research in wheat suggests the possibility of more than one aluminum-tolerance gene and more than one aluminum-tolerance mechanism. In one piece of research, two QTLs associated with aluminum tolerance were identified in the F<sub>2</sub> population of diploid alfalfa and confirmed in the back-cross population. Breeding aluminum tolerance helps to expand crop productivity to acidic soils.

### Mineral deficiency stress

#### Concepts associated with mineral deficiency

Mineral deficiencies or toxicities are widespread. A report from the Centro Internacional de Agricultura Tropical (CIAT) in Peru estimates that about 60% of the soils in the common bean production regions of the world have some soil mineral problem. Soils that are high in calcareous minerals tend to have high amounts of basic elements (e.g., Ca, Mg, K) that tend to raise soil pH. A high soil pH in turn causes mineral deficiency problems (e.g., Fe, Zn, P). Common mineral deficiency symptoms are summarized in Table 21.4. Zinc deficiency in common bean has been reported in production areas such as southern Idaho and Michigan.

#### Breeding efforts

Cultivars vary in their sensitivity to zinc deficiency. Sensitive cultivars take up and store less zinc in various plant parts and the seed than resistant cultivars. Researchers in common bean identified a zinc deficiency-resistant cultivar, “Matterhorn”, and subsequently determined that a single dominant gene, *Znd*, conditioned resistance to soil zinc deficiency.

### Oxidative stress

#### Concepts associated with oxidative stress

Oxygen free radicals (or activated oxygen) have been implicated in a variety of environmental stresses in

**Table 21.4** Common deficiency symptoms of selected essential nutrient elements for plants.

Mineral	Deficiency symptom
Nitrogen (N)	Chlorosis or yellowing of leaves; stunted growth
Phosphorus (P)	Dark green leaves; purpling of plant parts
Potassium (K)	Marginal necrosis; weak stem and lodging; leaf curling
Calcium (Ca)	Terminal bud growth ceases or is defective
Magnesium (Mg)	Chorosis of older leaves
Sulfur (S)	Chlorosis; weak stems
Boron (B)	Death of terminal bud producing growth called witches' broom
Iron (Fe)	Interveinal chlorosis of young leaves
Molybdenum (Mo)	Whip tail growth in coniferous species
Manganese (Mn)	Interveinal chlorosis
Zinc (Zn)	Mottled leaves
Copper (Cu)	Stunting; interveinal chlorosis

plants. They are involved in many degenerative conditions in eukaryotic cells (e.g., peroxidation of lipids, cross-linking and inactivation of protein, and mutation in DNA). However, the biosynthesis of some complex organic molecules, detoxification of xenobiotic chemicals, polymerization of cell wall constituents, and defense against pathogens are examples of essential cellular activities that depend on oxygen free radicals. Hence, the issue is not preventing their formation but how to control and manage the potential reactions of activated oxygen. Plants have a system of complex scavenging or activated oxygen that is highly conserved among plants.

Numerous sites of oxygen activation occur in the plant cell. These sites are highly controlled and coupled to prevent the release of intermediate products. It is presumed that such a control or coupling breaks down when a plant is under stress, resulting in leaking of activated oxygen. Injuries to the plant occur when the production of activated oxygen exceeds the plant's capacity to detoxify it. Symptoms of oxidative stress include loss of osmotic responsiveness, wilting, and necrosis.

There are two forms of activated oxygen that are produced via distinctly different mechanisms. Most

biological systems produce activated oxygen via reduction of oxygen to form superoxide, hydrogen peroxide, and hydroxyl radicals. In photosynthetic plants, a singlet oxygen form is also produced by photosynthesis.

### Applications and breeding efforts

Several herbicides are designed to function by the involvement of activated oxygen. These herbicides promote the accumulation of metabolic intermediates and the energy used to create singlet oxygen, which kills the plant. These herbicides are described as photo-bleaching (e.g., *p*-nitrodiphenyl ethers). Other herbicides that depend on light and chlorophyll are paraquat and diquat (both bipyridylium herbicides). So far, few plants have been selected for tolerance to oxygen free radicals.

### Flood stress (water logging)

Whereas some plants are adapted to water-logged conditions (e.g., flooded rice culture), most plants need well-drained soils to grow properly.

### Concepts associated with water-logging stress

In soybean, stress due to water logging can reduce crop yield by 17–43% when it occurs at the vegetative stage, and by about 50–56% if the stress occurs at the reproductive stage. Floods are often caused by excessive rainfall due to a prolonged seasonal rainfall. The excessive amount of water quickly creates anoxic (oxygen-deficient) soil conditions causing flood-sensitive plants to suffer anoxia or hypoxia. Fermentation occurs in plant roots under such conditions. The photosynthetic capacity of plants is significantly inhibited. Flood-tolerant species have certain adaptive mechanisms, such as the formation of aerenchyma and adventitious roots. Some studies indicate that root tissue survival under hypoxia depends on the fermentation rate and sufficient sugar supply to maintain cell energy and membrane function.

### Breeding efforts

Tolerance to water-logging appears to be quantitatively inherited. QTLs for tolerance to water-logging have been reported in rice and soybean.

## References and suggested reading

- Acevedo, E., and E. Fereres. 1993. Resistance to abiotic stresses. In: Plant breeding: Principles and prospects (Hayward, M.D., N.O. Bosemark, and I. Romagosa, eds). Chapman & Hall, London.
- Chinnusamy, Y., A. Jagendorf, and J.-K., Zhn. 2005. Understanding and improving salt tolerance in plants. *Crop Sci.* 45:437–448.
- Cinchy, K.A., S. Foster, K.F. Grafton, and G.L. Hosfield. 2005. Inheritance of seed zinc accumulation in navy bean. *Crop Sci.* 45:864–870.
- Fischer, R.A., and R. Maurer. 1978. Drought resistance in spring wheat cultivars I. Grain yield response. *Aust. J. Agric. Res.* 29: 897–912.
- Ibrahim, A.M.H., and J.S. Quick. 2001. Heritability of heat tolerance in winter and spring wheat. *Crop Sci.* 41:1401–1405.
- Lee, E.A., M.A. Staebler, and M. Tollenaar. 2002. Genetic variation in physiological discriminators of cold tolerance – early autotrophic phase of maize development. *Crop Sci.* 42:1919–1929.
- Lewitt, J. 1972. Responses of plants to environmental stresses. Academic Press, New York.
- McKersie, B.D., and Y.Y. Leshem. 1994. Stress and coping in cultivated plants. Kluwer Academic Publishers, Dordrecht.
- Voigt, P.W., and T.E. Stanley. 2004. Selection of aluminum and acid-soil resistance in white clover. *Crop Sci.* 44:38–48.
- Yordanov, I., V. Velikova, and T. Tsonev. 2000. Plant responses to drought, acclimation, and stress tolerance. *Photosynthetica* 38:171–186.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Sodic soils have a high accumulation of potassium.
- 2 Harvest maturity is determined by crop producers.
- 3 Early maturity is recessive to late maturity.
- 4 Most irrigated farms in the world have salinity problems.

### Part B

Please answer the following questions:

- 1 Distinguish between physiological maturity and marker maturity.
- 2 Distinguish between weather and climate.
- 3 Distinguish between drought tolerance and drought avoidance.
- 4 Briefly explain the role of plant growth regulators in drought tolerance.
- 5 Give four examples of crop plants that are susceptible to salts in the soil.
- 6 What is drought resistance?

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the advantages and disadvantages of early maturity.
- 2 Discuss the breeding of early maturity in crop plants.
- 3 Discuss the rationale for breeding drought resistance.
- 4 Discuss the mechanisms used by plants to avoid the effects of drought.



# Breeding compositional traits and added value

---

## Purpose and expected outcomes

The market value and utilization of a plant product is affected by a variety of factors. For example, for grain crops, the factors that affect grain quality and form the basis of crop quality improvement programs, include market quality, milling quality, cooking and processing qualities, and nutritional quality. The specific breeding goals with respect to each of these factors differ among crop species and agricultural production regions. For example, for the same crop, the quality aspect of importance in developed economies may be very different and even opposite to that of developing economies. Some cultures prefer white, non-scented, or non-sticky rice, whereas other cultures prefer colored, scented, or sticky rice. After completing this chapter, the student should be able to:

- 1 Discuss breeding for improved protein content.
  - 2 Describe breeding for improved fatty acid content.
  - 3 Discuss breeding for seedlessness in fruits.
  - 4 Discuss breeding for delayed ripening.
  - 5 Discuss breeding for novel traits in plant using biotechnology.
- 

## Concept of quality

Quality means different things to different people. The terms used to describe quality vary from crop production to food consumption, and include terms for appearance, storage quality, processing quality, and nutritional worth or quality. Of these, most attention has been devoted to processing quality for the major crops, such as milling quality and baking quality of wheat, canning quality of beans, chipping or baking quality of potatoes, malting quality of barley, and fermenting quality of grapes. These crops and others are of such high economic value that special labs operating at the private, state, and national levels have been set up to research and develop standards for these specialized processes for the benefit of plant breeders and industry.

Plant breeders should be very familiar with the market quality standards for their crops. These standards are

based on a complex interaction of social, economic, and biological factors, and are highly crop-specific.

## Nutritional quality of food crops

Plant parts used for food differ in nutritional quality. Different species and cultivars of the same species may differ significantly in total protein as well as in the nutritional value of the protein. The amino acid profiles of cereal grains and legumes differ according to certain patterns. Cereals tend to be low in lysine while legumes tend to be deficient in tryptophan (Table 22.1).

Three of the main crops that feed the world are cereals (corn, wheat, rice). Other important species are roots or tubers. Cereals and tubers are generally low in protein content. Rice averages about 8% protein, corn 10%, and potato 2%, versus 38–42% in soybean and 26%

**Table 22.1** Essential amino acids that are low in selected major world food crops.

Crop	Deficient amino acid
Corn	Tryptophan Lysine
Wheat	Lysine
Rye	Tryptophan Lysine
Rice (polished)	Lysine Threonine
Millet	Lysine
Soybean	Methionine Cystine Valine
Lima bean	Methionine Cystine
Peanut	Lysine Methionine Cystine Threonine
Pigeon pea	Tryptophan
Common bean	Tryptophan
Potato	Methionine Cystine

in peanut. Protein augmentation is a major breeding objective in many major world crops.

### Brief history of breeding for improved nutritional quality of crops

Breeding for high protein content in crop plants is perhaps the highest priority in improving the nutritional quality of plants because about 70% of the protein supply of human consumption is of plant origin. Further, cereals are deficient in some essential amino acids and low in total protein. Maize was one of the first crops on which formal nutritional augmentation work was done. In 1896, C. G. Hopkins initiated a project to breed for high protein and oil content at the Illinois Agricultural Experimental Station. Work by T. B. Osborne in the early 1900s resulted in the fractionation and classification of proteins according to solubility properties. He and his colleague discovered **zein** (the prolamin or alcohol soluble fraction) as comprising the bulk of the protein of maize endosperm. Later work in the mid 1900s by K. J. Frey demonstrated that breeding for protein augmentation primarily increased the zein content. There was a need to find a way to enhance the useful part of the protein. E. J. Mertz in 1964 discussed the

nutritional effects of the *opaque-2* gene in maize. The mutant gene increased the lysine content, called **high lysine** (discussed below). High lysine research has since been conducted in sorghum. Another cereal food of world importance is rice. However, it has significant nutritional problems, being low in protein as well as completely lacking vitamin A. Rice nutritional augmentation was initiated in 1966 at the International Rice Research Institute (IRRI) in the Philippines. The vitamin A deficiency is being addressed using genetic engineering (see below).

### Breeding for improved protein content

The key components of food that impact nutrition are carbohydrates, fats, proteins, minerals, water, vitamins, and fiber. The first three components provide caloric energy, while proteins, minerals, and water play a role in body tissue and structure. The roles of regulation and utilization are played by proteins, minerals, water, vitamins, and fiber. After satisfying caloric energy needs, proteins are the next most important nutritional component of a diet. Twenty-two amino acids are generally recognized in human nutrition, of which eight are essential for monogastric animals (Table 22.2). The utilization efficiency of the entire protein is diminished if the diet is deficient in any of the essential amino acids.

### Breeding high lysine content of grain

Breeders using conventional methods of ear-to-row selection were able to increase the total protein content of corn kernels from 10.9% to 26.6%. Unfortunately, because the protein of corn is about 80% zein, and hence nutritionally inadequate, the high increase in total protein was nutritionally unprofitable to non-ruminant animals. The zein fraction of the total protein is deficient in lysine and tryptophan. This deficiency was corrected in

**Table 22.2** Important amino acids in animal and human nutrition; those in bold are essential in human adults.

Isoleucine	Alanine	Serine
Leucine	Arginine	Tyrosine
Lysine	Cysteine	Asparagine
<b>Methionine</b>	Glutamic acid	Glutamine
Phenylalanine	Glycine	Cystine
Threonine	Histidine	Hydroxyglutamic acid
<b>Tryptophan</b>	Proline	Norleucine
<b>Valine</b>		

1964 when researchers at Purdue University discovered mutant genes, called *opaque-2* and *floury-2*, which increased the lysine content of the kernel. The patterns of expression of the mutant genes differ slightly. The *opaque-2* gene has a recessive gene action, whereas the *floury-2* gene exhibits a dosage effect. The resulting corn is called **high lysine corn**, and has a characteristic soft and starchy endosperm. Consequently, the softer endosperm predisposes high lysine kernels to breakage, cracking, and rot. Generally, high lysine cultivars have lower yields than their conventional counterparts. Cross-pollination with normal dent corn reverses the soft endosperm to normal dent endosperm. High lysine corn production must be done in isolated fields. The *opaque-2* recessive gene increased the lysine content of the kernel from about 0.26–0.30% to about 0.34–0.37%. High lysine has also been transferred into sorghum.

### Quality protein maize

**Quality protein maize (QPM)** may be described as an extension of the improvement of high lysine maize. It is

a high lysine product because it uses the *opaque-2* gene. However, it is unlike the traditional high lysine maize because it lacks all the undesirable attributes of high lysine products (i.e., low yields, chalky-looking grain, and susceptibility to diseases and insect pests). It looks like regular maize but has about twice the levels of lysine and tryptophan. QPM was developed by two researchers, K. V. Vasal and E. Villegas over about three decades. They used conventional breeding methods to incorporate modifier genes to eliminate the undesirable effects of the lysine gene. The two scientists were rewarded with the World Food Prize in 2001 for their efforts.

QPM has less of the indigestible prolamine-type amino acids that predominate in the protein of normal maize. Instead, QPM cultivars have about 40% of the more digestible glutelins and a balanced leucine : isoleucine ratio for enhanced niacin production upon ingestion. Research also indicates that QPM has better food and feed efficiency ratings (grain food intake/grain weight gain) following feeding tests with animal (e.g., pigs, poultry). QPM cultivars have been released for production in over 20 developing countries since 1997.



Twumasi Afriyie

## Industry highlights

### *QPM: enhancing protein nutrition in sub-Saharan Africa*

International Maize and Wheat Improvement Center (CIMMYT), PO Box 5689, Addis Ababa, Ethiopia

#### **Introduction**

Maize is a major staple in sub-Saharan Africa and also constitutes an important source of food for children in particular. For example, children in Ghana grow well during the first 6 months of life but thereafter when breast milk ceases to be sufficient to sustain their rapid growth, malnutrition becomes normal. This nutritional trend is explained by the pervasive use of a thin gruel porridge made from maize or millet as the first weaning food fed to children. Few mothers supplement such cereal diets with other sources of protein such as beans, fish, or milk due to ignorance about proper nutrition, high cost, or lack of time. The cereals alone do not provide a balanced diet because they are low in lysine and tryptophan, essential amino acids, which cannot be synthesized by monogastric animals including humans (National Research Council 1988). Normal maize, for example, has approximately 10% protein but the full amount is not utilizable by monogastric animals because the protein is low in lysine and tryptophan. When children are fed normal maize without any better-balanced protein supplement, they become malnourished and develop the protein deficiency disease called kwashiorkor.

In 1963, Mertz and his coworkers at the University of Purdue discovered a recessive mutant maize gene, *opaque-2*, which resulted in grain protein with approximately twice the quantities of lysine and tryptophan, the two limiting amino acids in ordinary maize (Mertz et al. 1964). There was an immediate upsurge of worldwide interest to develop nutritionally improved maize varieties. However, it was soon discovered that the gene conferring the improved nutritional quality also resulted in several undesirable agronomic characters including low grain yield potential, unacceptable chalky grain type, high moisture at harvest, and high susceptibility to insects and disease attacks (National Research Council 1988). When farmers rejected the early "high lysine" hybrids quickly released to them, the research in *opaque-2* or high lysine maize waned markedly worldwide. However, unrelenting research continued for some 30 years at the CIMMYT, Mexico, resulting in the development of maize germplasm combining

better protein quality with desirable grain yield potential and agronomic characteristics similar to normal maize (Bjarnason 1990). This new source of nutritionally improved maize was the result of the accumulation of modifier genes in a rather complex breeding technique supported by a strong laboratory for analysis of protein quality (Vasal et al. 1993). The new material had normal-looking hard endosperm grain type and was designated quality protein maize (QPM). At the time QPM germplasm became available, there were still doubts about the usefulness of QPM for farmer production and since researchers – especially those in developing countries – continued to show little interest in QPM, therefore in 1991, CIMMYT closed its research on QPM. It was at this time, however, that the Ghana Grains Development Project (GGDP) within the Crops Research Institute, Kumasi, Ghana, assigned one fulltime breeder to initiate a QPM development project for Ghana. The government of Ghana, the Canadian International Development Agency (CIDA), and Sasakawa Global 2000 (SG2000) provided research funding. The main objectives of QPM research in Ghana were to develop high and stable yielding QPM varieties with comparable performance to their normal counterparts, to demonstrate their nutritional advantages, and to promote their production, marketing, and utilization.

### *Breeding approaches*

The QPM germplasm used to initiate QPM breeding in Ghana was collected from CIMMYT, Mexico in 1991. The germplasm included open-pollinated experimental varieties and early generation inbred lines from CIMMYT populations 62 (white flint) and 63 (white dent). The maize streak virus disease was a major problem in Ghana at the onset of the QPM program. Therefore, the QPM germplasm was converted to streak virus disease resistance by backcrossing the susceptible materials to resistant sources obtained from the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. This was followed by growing the lines under artificial streak pressure and self-pollinating resistant lines to produce  $S_1$  lines. Half-sib (HS) and  $S_1$  recurrent improvement schemes were employed to develop the first QPM variety in Ghana. In the HS procedure, about 300 families were grown in an isolation block. Three rows were detasseled before anthesis to constitute female rows and these were grown in alternation with one male row (not detasseled), which served as the pollen source. Equal numbers of seed from each family were composited to plant the male rows, which served as the sole source of pollen. At harvest, HS families with desired agronomic traits were selected. These families were planted again as ear-to-row and selfed to produce  $S_1$  lines. Twenty seeds of selected  $S_1$  lines were sent for protein quality analysis (basically measuring tryptophan levels in total protein) at CIMMYT, Mexico. At the same time, light tables were used to select modified grains in the  $S_1$  lines. On the basis of visual selection of HS lines in the field, and the results of analysis of protein quality (tryptophan level) in the laboratory and grain modification under light tables,  $S_1$  lines were selected and recombined in isolation to constitute the next cycle of improvement. A streak-resistant QPM variety named "Obatanpa2" (literally "good nursing mother") was released in 1992. This was a white dent, medium maturity (105 days) variety.

We also initiated a hybrid development program. We developed several inbred lines through inbreeding by self-pollination, and conducted early generation testing and topcross evaluation. Diallel cross evaluations were used to determine the combining abilities of advanced inbred lines. Several single and three-way hybrids were developed. Open-pollinated varieties and hybrids developed in the project were first tested at the research stations located in the major agroecological zones in Ghana. The Ghanaian QPM hybrids performed as well as or better than local check hybrids. Consequently, three of the hybrids were released in Ghana and one was released in South Africa for production in southern African countries. QPM germplasm development was supported by biochemical analysis of the grain at CIMMYT, Mexico and Ghana.

### *QPM nutritional studies*

Despite the progress obtained in QPM germplasm development in Ghana, widespread doubts persisted about the usefulness of QPM technology. We conducted several collaborative animal-feeding studies on pigs, chickens, and rats to ascertain the nutritional advantages of QPM when used as human or animal food. The collaborative institutions included the Crops Research Institute (CRI), SG2000, and the Animal Science Department of the Kwame Nkrumah University of Science and Technology and the Health and Nutrition Department of the Ministry of Health.

#### *Feed ingredient for pigs*

Fourteen starter pigs from two litters (8.4 kg average weights) were divided into two equal groups (each containing three females and four males) and were fed similar diets (*ad libitum*) for 16 weeks (Osei et al. 1994). The group 1 diet contained 91% QPM ("Obatanpa") and the group 2 diet contained 91% normal maize (NM, "Okomasa"). The balance of both diets comprised equal quantities of mineral and vitamin supplement. The average growth rate of pigs fed the QPM diet was faster (13.9 kg) than pigs fed normal maize (5.9 kg). The corresponding daily gains were 124.9 g and 52.7 g for the QPM and NM pigs, respectively, showing that QPM-fed pigs had a growth rate 2.30 times that of pigs on an NM diet. The feed conversion efficiency of QPM pigs was also greater than NM pigs (Okai et al. 1994; Osei et al. 1994).

#### *Feed ingredient for broiler chickens*

A series of three feeding trials was conducted with broiler chickens to assess the commercial viability of "Obatanpa" in poultry feed (Osei et al. 1994). The initial studies focused on using QPM or normal maize as the sole source of protein and energy.

Subsequent studies investigated the feasibility of reducing the levels of fishmeal in commercial diets when "Obatanpa" was the source of maize. Results of the studies indicated that the use of "Obatanpa" allowed the level of fishmeal (an expensive high protein ingredient in Ghana) to be reduced from 19.5% to 13.5%, whilst still maintaining good performance. The use of "Obatanpa" in broiler diets resulted in a significant economic advantage because of the reduced use of fishmeal – due mainly to the huge price disparity between QPM and fishmeal.

#### Processed food from normal maize and QPM

*Kenkey* is a popular local food in Ghana. It is made from fermented maize meal. We studied the effect of processing and cooking on the nutritional quality of *kenkey* made from normal maize or QPM (Ahenkora et al. 1995). The QPM, "Obatanpa", and the normal maize, "Okomasa", were processed into *kenkey*. Weaning rats were fed *ad libitum* on *kenkey*-based diets, which served as the sole source of protein and amino acid, for 28 days. Analysis of samples of the *kenkey* revealed that processing and cooking raw grains into *kenkey* reduced the lysine content by 13% and the tryptophan content by 22% (Ahenkora et al. 1995). However, *kenkey* from QPM contained 51% more lysine and 63% more tryptophan than *kenkey* from normal maize. The individual average gain by rats fed on a QPM *kenkey* diet was 37.2 g compared with 16.2 g for NM *kenkey* – a 2.3-fold difference. Rats fed the QPM diet had a better feed conversion ratio and higher protein efficiency ratio values than their counterparts fed a normal maize *kenkey* diet.

#### Agricultural technology/nutrition impact study

A series of studies investigated the impact of QPM utilization on community-based agricultural technology interventions in the Ejura-Sekodumasi District, Ashanti Region, Ghana. The study was done through the Ministry of Health Nutrition Division with collaboration from other agricultural and health institutions in Ghana particularly the CRI, Ministry of Food and Agriculture (MOFA), and SG2000. The results showed that QPM enhanced growth relative to normal maize when fed to children.

#### *Breeding challenges related to QPM*

A series of experiments were conducted to dispel some of the doubts, myths, and fallacies concerning QPM (Twumasi-Afriyie et al. 1996).

- 1 QPM produces a lower grain yield than NM counterparts.** In Ghana, it was shown that QPM varieties could produce better yields than their normal counterparts.
- 2 Lysine and tryptophan are "heat labile" and would be destroyed during processing, thus QPM will lose its nutritional advantage during processing into local dishes.** We demonstrated that the nutritional advantage was maintained when QPM was processed into the most popular local dishes (Ahenkora et al. 1995).
- 3 QPM is conferred by a recessive gene and thus will lose its nutritional advantage in farmers' production plots, which are normally planted on small areas.** We conducted an experiment in which we surrounded a 1-acre field of QPM with a yellow endosperm NM with the same maturity, and allowed the two to cross freely. Results from 2 years of data at several locations showed a maximum of 10% contamination by the NM. The contamination was most pronounced within 12 m of the QPM field nearest the NM, and was most serious at the southwestern sector of the field due to the prevailing wind. The nutritional quality of the bulked grain from the most contaminated lot was still not significantly different from the non-contaminated QPM, based on a rat-feeding study.
- 4 QPM will not store well at the farm level.** From our study, when weevils were introduced into grains of NM or QPM grain, there was no difference in the extent of damage incurred. All samples were equally damaged in a short time period. Moreover, it was detected that, in general, post-harvest handling was very poor in Ghana, and that available improved technology if followed could enable farmers to store both NM and QPM with minimal problems.
- 5 Marketing will be difficult because QPM lacks visible identifiers that could facilitate sale at a higher price to offset additional costs of production.** In fact, there was no additional cost of production of "Obatanpa" as it produced higher yields than its normal counterparts under identical recommended practices. Agronomic performance *per se* became a driving force behind the adoption of this variety. Special marketing channels developed for "Obatanpa". Private purchasing agents began to market "Obatanpa" to satisfy the demands of commercial users such as food and feed processors and relief agencies. The private purchasing agents linked with producers and guaranteed the quality of "Obatanpa" to the users.

#### *Current efforts in QPM development in Africa*

The CIMMYT recommenced QPM development in 1997 partly due to the success achieved in Ghana. Current efforts in QPM development in sub-Saharan Africa involve several national agricultural research institutes and increasing numbers of private seed companies. QPM development and deployment largely follows the Ghana model involving multidisciplinary and multi-institutional approaches in germplasm development, nutritional studies, variety releases, seed production, and agricultural



**Figure 1** Farmer (left) and researcher (center) admiring a bag of quality protein maize being offered for sale in a store in Ghana, West Africa.

extension (Figure 1). To date, QPM varieties have been released in about 15 sub-Saharan African countries. Current efforts led by CIMMYT and IITA seek to incorporate QPM into elite and local cultivars through a process of conversion of NM germplasm to QPM. Conversion involves the use of backcrossing of normal maize to QPM using donor QPM populations or inbred lines. After one or two backcrosses, the plants are selfed or sib-pollinated, and segregating *opaque-2* phenotype grains, which appear partly opaque on light tables, are selected for further backcrossing. Two to three backcrosses are often enough to recover the recurrent parent genotype and phenotype in addition to the modified *opaque-2* grain character. However, in some cases, additional cycles of improvement may be required to accumulate enough modifier genes to recover the normal maize endosperm phenotype. Marker-assisted selection has been employed to speed up the backcrossing. In this procedure, a DNA marker that is very closely linked to the *opaque-2* gene is used – as a replacement of the phenotypic test – to select progeny carrying the desired allele based on the analysis of leaf samples from young plants.

### References

- Ahenkora, K., S. Twumasi-Afriyie, W. Haag, and B.D. Dzah. 1995. Ghanaian *kenkey* from normal maize and quality protein maize: Comparative chemical composition and rat growth trials. *Cereals Communic.* 23:299–304.
- Mertz E.T., L.S. Bates, and O.E. Nelson. 1964. Mutant genes that changes protein composition and increases lysine content of maize endosperm. *Science* 145:279–280.
- National Research Council. 1988. Quality-protein maize. Academy Press, Washington, DC.
- Okai, D.B., S.A. Osei, A.K. Tua, et al. 1994. The usefulness of Obatanpa, a quality protein maize variety in the feeding of pigs in Ghana. *Proc. Ghana Anim. Sci. Symp.* 22:37–43.
- Osei, S.A., C.C. Atuahene, A. Donkoh, et al. 1994. Further studies on the use of quality protein maize as a feed ingredient for broiler chickens. *Proc. Ghana Anim. Sci. Symp.* 22:51–55.
- Twumasi-Afriyie S., B.D. Dzah, and K. Ahenkora. 1996. Why QPM moved in Ghana. In: Maize productivity gains through research and productivity dissemination (Ransom, J.K., B.T. Palmer, Z.O. Mduruma, S.R. Waddington, K.V. Pixley, and D.C. Jewel, eds), pp. 28–31. Proceedings of the Fifth Eastern and Southern Africa Maize Conference, Arusha, Tanzania. CIMMYT, Mexico.
- Vasal, S.K., G. Srivisan, S. Pandey, F. Gonzalez, J. Crossa, and D.L. Beck. 1993. Heterosis and combining ability of CIMMYT's protein maize germplasm: Lowland tropical. *Crop Sci.* 33:46–51.

### Improving protein content by genetic engineering

Nutritional quality augmentation through the addition of new quality traits, removing or reducing undesirable traits, or other manipulations, is an important goal in the bioengineering of food crops. Crops that feed the world are primarily cereals, roots and tubers, and legumes. Unfortunately, they are nutritionally inadequate in providing certain amino acids required for proper growth and development of humans and monogastric animals. For example, cereals are generally deficient in lysine and

threonine, whereas legumes are generally deficient in sulfur amino acids. In some species (e.g., rice) in which the amino acid balance is relatively appropriate, the overall protein quantities are low.

Molecular genetic approaches are being adopted for genetically engineering seed protein. They may be categorized as:

- 1 Altering the amino acid profile of the seed.
- 2 Selective enhancement of expression of existing genes.
- 3 Designing and producing biomolecules for nutritional quality.

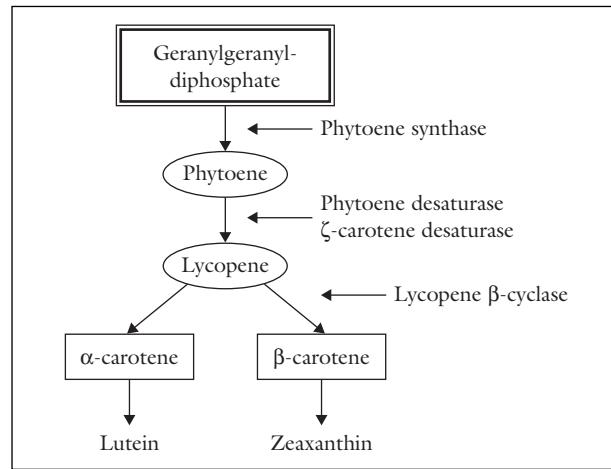
### The making of “Golden Rice”

“Golden Rice” is so-called because it has been genetically engineered to produce  $\beta$ -carotene (responsible for the yellow color in certain plant parts like carrot roots) in its endosperm. This rice produces  $\beta$ -carotene or pro-vitamin A, the precursor of vitamin A, which does not occur in the endosperm of rice.

An estimated 3 billion people of the world depend on rice as their staple food. Of this number, about 10% are at risk for vitamin A deficiency and the associated health problems that includes blindness and deficiency of other micronutrients such as iron and iodine. The effort to create such golden rice was led by Dr Ingo Potrykus, a professor of plant science at the Swiss Federal Institute of Technology. In 1990, Garry Toennissen, the director of food security for the Rockefeller Foundation recommended the use of the sophisticated tools of biotechnology to address the problem of lack of vitamin A in rice. Later, at a Rockefeller-sponsored meeting, Potrykus met Peter Beyer of the University of Freiburg in Germany, an expert on the  $\beta$ -carotene pathway in daffodils. In 1993, and with seed money of US\$100,000 from the Rockefeller Foundation, the two embarked upon an ambitious project to create a transgenic plant in a manner unlike any before. After 7 years, the duo announced to the world their outstanding achievement, “Golden Rice”, at a cost of \$2.6 million. The bill was partly footed by the Swiss government and the European Union.

The scientific feat accomplished in engineering  $\beta$ -carotene into rice is that it marks the first time a complete metabolic pathway has been engineered into an organism. Rice lacks the metabolic pathway to make  $\beta$ -carotene in its endosperm. Potrykus and Beyer had to engineer a metabolic pathway consisting of four enzymes into rice (Figure 22.1). Immature rice endosperm produces geranylgeranyl-diphosphate (GGPP), an early precursor of  $\beta$ -carotene. The first enzyme engineered was phytoene synthase, which converts GGPP to phytoene (a colorless product). Enzyme number 2, called phytoene desaturase, and enzyme number 3, called  $\zeta$ -carotene desaturase, each catalyzes the introduction of two double bonds into the phytoene molecule to make lycopene (with a red color). Enzyme number 4, called lycopene  $\beta$ -cyclase converts lycopene into  $\beta$ -carotene. A unit of transgenic construct (called an expression cassette) was designed for each gene for each enzyme. These expression cassettes were linked in series or “stacked” in the final construct.

The source of genes for enzymes 1 and 4 was the daffodil, while genes for enzymes 2 and 3 were derived



**Figure 22.1** The key enzymes involved in the development of “Golden Rice”.

from the bacterium *Erwinia uredovora*. Three different gene constructs were created, the first and most complex combining enzyme 1 with enzymes 2 and 3, together with an antibiotic resistance marker gene that encodes hygromycin resistance, along with its promoter, CaMV 35S. The second gene construct was like the first, except that it lacked any antibiotic resistance marker gene. The third gene construct contained the expression cassette for enzyme 4, plus the antibiotic marker. By separating the genes for the enzymes and antibiotic resistance marker into two different constructs, the scientists reduced the chance of structural instability following transformation (the more cassettes that are stacked, the more unstable the construct).

These gene constructs were transformed into rice via *Agrobacterium*-mediated gene transfer in two transformation experiments. In experiment 1, the scientists inoculated 800 immature rice embryos in tissue culture with *Agrobacterium* containing the first transgenic system. They isolated 50 transgenic plants following selection by hygromycin marker. In the second experiment, they used 500 immature rice embryos, inoculating them with a mixture of *Agrobacterium* (T-DNA) vectors carrying both the second and third constructs. This experiment yielded 60 transgenic plants. The second experiment was the one expected to yield the anticipated results of a golden endosperm. This was so because it had all the four enzymes required for the newly created metabolic pathway. However, the scientists also recovered transgenic plants with yellow endosperm from experiment 1. Subsequent chemical analysis confirmed

the presence of  $\beta$ -carotene, but no lycopene. This finding suggests that enzyme 4 may be present in rice endosperm naturally, or that it could be induced by lycopene to turn lycopene into  $\beta$ -carotene. Analysis also showed the presence of lutein and zeaxanthin, both products derived from lycopene. None of the above was found in the control (non-engineered) plants.

### Matters arising from the development of "Golden Rice"

The initial golden rice lines produced 1.6–2.0  $\mu\text{g}$  of  $\beta$ -carotene per gram of grain. The recommended daily allowance (RDA) set by health agencies for children is 0.3 mg/day. Estimates of the bioavailability of  $\beta$ -carotene have been put at less than 10% in some cases. The scientists intend to refine their invention to make it produce 3–5-fold its present level of  $\beta$ -carotene.

Upon an international intellectual property rights (IPR) audit commissioned by the Rockefeller Foundation through the International Service for the Acquisition of Agri-Biotech Applications, Potrykus and his team realized their invention utilized 70 IPRs and TPRs (technical property rights) owned by 32 different companies and universities. Because of the humanitarian goal of the project, the development team negotiated with owners of these patents to allow the use of their inventions under the "freedom to operate" clause. Further, because of public pressure and the need for big business to tone down their profit-oriented public image, the key companies (e.g., Monsanto) offered free licenses for their IPRs involved with "Golden Rice".

It might take an estimated additional 5–10 years before the ordinary person for whom it is intended to benefit can produce golden rice. The characters of the present genotype must be bred into as many locally adapted varieties and ecotypes in as many rice-growing countries as quickly as possible. It is also important that such breeding efforts be organized such that all rules and regulations concerning the handling and use of genetically modified organisms be strictly followed to avoid stirring up additional controversies. Agronomic and other studies will have to be conducted to determine how well such golden rice yields, its palatability and digestibility, and public acceptance.

The next phase requires additional funds. Because of public protests against the product for a variety of reasons, public funding began to quickly dwindle. In January of 2001, a new effort was launched in the Philippines for a comprehensive set of tests to determine the efficiency, safety, and usefulness of "Golden Rice"

for people in the developing world. The joint effort includes the Philippines-based IRRI, Syngenta, and the Rockefeller Foundation. Further, the IRRI has set up a humanitarian board to oversee this project and to ensure that the highest standards for testing, safety, and support are achieved. The board includes several public and private organizations, such as the World Bank, Cornell University, the Indo-Swiss Collaboration in Biotechnology, and the Rockefeller Foundation.

### Breeding improved oil quality

Oil quality improvement is a major breeding objective for major oil crops such as soybean and rape. Soybean oil accounts for about 22% of the world's total edible oil production.

#### Chemical composition of seed oil

By chemical composition, soybean oil consists primarily of palmitic (C16:0), stearic (C18:0), oleic (C18:1), linoleic (C18:2), and linolenic (C18:3) acids, of which palmitic acid is one of the two major components of saturated fatty acids. Consequently, the physical, chemical, and nutritional quality of soybean depends significantly on its palmitic acid content. Breeders are interested in lowering the palmitic acid content of soybean oil so as to lower its total saturated fatty acid content for higher desirability for human consumption. On average, soybean seed oil contains about 120 g/kg of palmitic acid.

On the other hand, seed oil of modern oil seed rape contains about 60% oleate (C18:1), 20% linoleate (C18:2), 10% linolenate (C18:3), and only 4% palmitate (C16:0) and 2% stearate (C18:0). The major breeding goal for oil seed rape quality improvement is to increase the oleate content, which decreases the amounts of polyunsaturated fatty acids, linoleate and linolenate. This would enhance its nutritional quality and increase its potential value as an industrial oil.

Generally, food products with less than 7% of total saturated fatty acids (C16:0 + C18:0 + C20:0 + C22:0) qualifies to be labeled "low in saturated fatty acids" in the USA.

#### Approaches to breeding oil quality

Conventional breeding approaches have been successfully applied to improve the seed quality of various oil crops. Soybean lines with reduced palmitic acid content

have been developed using traditional approaches of hybridization, recurrent selection, and chemical mutagenesis. Studies have shown that reduced palmitic acid is conditioned by at least two loci, without maternal effects. Some of these genes have been designated *fap1*, *fap2*, and *fap3*. Genes modifying the major palmitic acid loci have also been found to influence the trait by about 2–23 g/kg. Molecular markers, including RFLPs (restriction fragment length polymorphisms), SSRs (simple sequence repeats), and QTLs (quantitative trait loci) associated with palmitic acid reduction have been cloned and mapped.

Germplasm containing between about 75% and 90% of oleate have been developed in soybean, sunflower, and rape, among other species. Transgenic high oleate soybean has been developed. Soybean is also a major source of protein for animal feed. Reduced linolenate soybean is conditioned by three independent mutants – *fan1(A5)*, *fan 2*, and *fan 3*.

Research has shown that seed protein and oil content are negatively correlated. Consequently, developing high protein and high oil seed has limited success. Alternatively, breeders have devoted efforts to breeding cultivars with high protein and low oil, and those with high oil but low protein.

In sunflower, a breeding objective is to increase the stearic acid (C18:0) content of the seed oil. Mutagenesis was used to develop different lines with a higher C18:0 content of about 50 g/kg. The results of one genetic analysis indicated the inheritance of high stearic content to be under the control of one locus (*Es<sub>1</sub>*, *es<sub>1</sub>*) and partial dominance. Further studies indicated the presence of a second locus (*es<sub>2</sub>*, *es<sub>2</sub>*). Generally, fatty acid composition is controlled by the genotype of the developing embryo. Hence, selection for fatty acid may be conducted at the single-seed level, by non-destructive techniques such as gas-liquid chromatography.

### Breeding low phytate cultivars

Human and monogastric animals produce small amounts of the phytate enzyme needed to utilize phytate phosphorus. Soybean, for example, contains about 4.3 g/kg of phytate phosphorus and only 0.7 g/kg inorganic phosphorus. In humans, diets high in phytic acid decrease the absorption of essential minerals such as calcium, iron, and zinc. It would be desirable to remove phytate from cereals and oil seeds.

In soybean, low phytate mutants with about 1.9 g/kg phytate phosphorus and 3.1 g/kg inorganic phospho-

rus have been discovered. The trait is conditioned by recessive alleles designated *pha<sub>1</sub>* and *pha<sub>2</sub>* at two independent loci that exhibit duplicate dominant epistasis. Both of the alleles must be homozygous in order for low phytate seed to be obtained. In wheat, low phytic acid mutants *lpa<sub>1</sub>* and *lpa<sub>2</sub>* have been identified.

### Breeding end-use quality

Breeders set breeding objectives to meet the needs of producers as well as consumers. The crop producer needs to focus on traits that facilitate crop production and increase crop yield (e.g., pest resistance, maturity, high yield, lodging resistance, drought resistance). Consumers are more concerned about nutritional quality traits (taste, protein content, appearance). Another group of consumer needs that is not nutritional but concerned with how they are used or stored is end-use quality. Certain cultivars are bred for specific industrial quality traits (e.g., for processing, cooking).

### Extended shelf-life

Plant products that are harvested and used fresh (e.g., fruits, vegetables) are perishable and highly susceptible to spoilage soon after harvesting. In cases where production is far from marketing centers, the fresh produce has to be transported over long distances, and hence requires protection from bruising and rotting in transit. Fully vine- or plant-ripened fruits, even though desirable for their superior taste, are more susceptible to damage under such conditions than unripened fruits. Grocery stores need to display their produce for a period of time in good condition while waiting to make a sale. Extended shelf-life is hence an important plant trait from the point of view of producers, wholesalers, and consumers.

### Delayed ripening

Delayed ripening is desired in crops such as tomato and banana. Biotechnology has been successfully used to develop this quality in some crops. Certain fruits exhibit elevated respiration during ripening with concomitant evolution of high levels of ethylene. Called **climacteric fruits** (e.g., apple, banana, tomato), the ripening process of these fruits involves a series of biochemical changes leading to fruit softening. Chlorophyll, starch,

and cell walls are degraded. There is an accumulation of lycopene (red pigment in tomato), sugars, and various organic acids. Ripening is a complex process that includes fruit color change and softening.

Ripening in tomato has received great attention because it is one of the most widely grown and eaten fruits in the world. Ethylene plays a key role in tomato ripening. When biosynthesis of ethylene is inhibited, fruits fail to ripen, indicating that ethylene regulates fruit ripening in tomato. The biosynthesis of ethylene is a two-step process in which *s*-adenosyl methionine (SAM) is metabolized into aminocyclopropane-1 carboxylic acid (ACC), which in turn is converted to ethylene. Knowing the pathway of ethylene biosynthesis, scientists can manipulate the ripening process by either reducing the synthesis of ethylene or reducing the effects of ethylene (i.e., plant response).

In reducing ethylene biosynthesis, one successful strategy by Agritope of Oregon has been the cloning of a gene that hydrolyzes SAM, called SAM hydrolase, from a bacterial virus. After bioengineering the gene to include, among other factors, a promoter that initiates expression of the gene in mature green fruits, *Agrobacterium*-mediated transformation was used to produce transgenic plants. The effect of the chimeric gene was to remove (divert) SAM from the metabolic pathway of ethylene biosynthesis. The approach adopted by researchers was to prevent the ACC from being converted to ethylene. A gene for ACC synthase was isolated from a bacterium and used to create a chimeric gene as in the Agritope case.

The technology of antisense has been successfully used to develop a commercial tomato that expresses the antisense RNA for ACC synthase and ACC oxidase. United States Department of Agriculture (USDA) scientists pioneered the ACC synthase work, while scientists from England in collaboration with Zeneca pioneered the ACC oxidase work. Because transgenic tomatoes with an incapacitated ethylene biosynthetic pathway produced no ethylene, they failed to ripen on their own, unless exposed to artificial ethylene sources in ripening chambers. The technology needs to be perfected so that fruits can produce some minimum amount of ethylene for autocatalytic production for ripening over a protracted period.

#### *The "FlavrSavr" tomato*

Another application of antisense technology is in preventing an associated event in the ripening process, fruit softening, from occurring rapidly. Vine-ripened fruits

are tastier than green-harvested and forced-ripened fruits. However, when fruits vine ripen before harvesting, they are prone to rotting during shipping or have a short shelf-life in the store. It is desirable to have fruits ripen slowly. In this regard, the target for genetic engineering is the enzyme **polygalacturonase (PG)**. This enzyme accumulates as the fruit softens, along with cellulases that breakdown cell wall cellulose and pectin methylesterase that together with PG break the pectic cross-linking molecules in the cell wall. Two pleiotropic mutants of tomato were isolated and studied. One mutant, never ripe (*Nr*), was observed to soften slowly and had reduced accumulation of PG, while the second mutant, ripening inhibitor (*rin*), had very little accumulation of PG throughout the ripening process. This and other research evidence strongly suggested a strong association between PG and fruit ripening. PG is biosynthesized in the plant and has three isoenzymes (PG1, PG2, PG3).

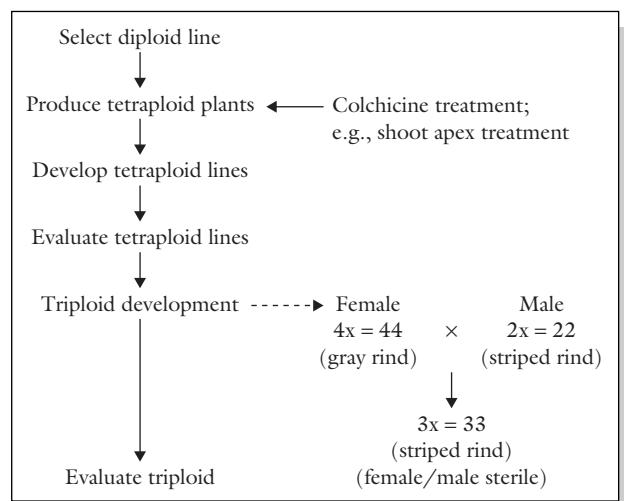
This technology was first successfully used by Calgene to produce the "**FlavrSavr**" tomato, the first bioengineered food crop, in 1985. The protocol has been previously described. This pioneering effort by Calgene flopped for several reason, among which was the poor decision to market a product intended for tomato processing as a fresh market variety.

#### Breeding cooking and processing qualities

Cooking food changes its texture, color, taste (palatability), and digestibility, among other changes. The heat treatment applied during cooking breaks down some toxic compounds in food, where applicable. What is considered to be a good cooking or processing quality depends on the product and the culture in which the product is used. As previously cited, for example, some cultures prefer sticky rice, others non-sticky rice for certain food preparations. Similarly, some potato cultivars are suitable for frying, others for baking, and yet others for cooking.

Similarly, canning or processing quality is an important breeding objective in crops that are grown for that purpose. It is desirable for canned produce to retain its texture and color to an appreciable degree. Some cultivars remain firm and of good color, whereas others crack or become mushy after canning.

Other products are crushed, ground, or milled during processing. In corn, for example, milling may be dry or wet. For dry milling white endosperm and semihard kernel is preferred, while wet milling (for starch and oil) requires softer kernels.



**Figure 22.2** Generalized steps in the breeding of seedless watermelons.

### Breeding seedlessness

Fresh fruits without seeds are more convenient to eat, because there are no seeds to spit out. Common fresh fruits in which seedless cultivars exist include watermelon, grape, orange, and strawberry.

The conventional way of producing seedless fruits is the use of triploid hybrids. A tetraploid ( $4x$ ) parent that is crossed with a diploid ( $2x$ ) line is exemplified by seedless watermelon breeding (Figure 22.2). In the watermelon,  $4x = 44$  and  $2x = 22$ . The tetraploid is always the female parent; the reciprocal cross (with the diploid as male parent) does not produce seed. The resulting triploid ( $3x = 33$ ) is female-sterile and hence the fruit is seedless. Furthermore, because the triploid is also male-sterile, producers of seedless watermelon must plant rows of diploid lines as pollinators for stimulation of fruit formation. In commercial production fields, growers usually plant a ratio of three triploid rows to one diploid row. It is important that the rows be marked to enable harvesters pick fruits from only triploid lines.

The breeding of seedless watermelon will be described to illustrate the conventional production of seedless fruits. The process involves a number of steps:

- 1 **Selection of diploid lines for developing tetraploids.** Many producers use tetraploids with gray-colored rind and diploids with striped rind. This way, it is easy to identify and discard selfed tetraploids (gray rinds) in the progeny.

**2 Tetraploid induction.** Chromosome doubling to produce tetraploids from the diploid line is accomplished by using colchicine. Other methods are also available. Colchicine is applied to the shoot apex of diploid plant seedlings, just as the cotyledons first emerge from the soil. This causes the chromosomes at the shoot apex to double resulting in the tetraploid seedling. The mutagen is applied at a concentration ranging between 1% and 2%, the lower concentration for small-seeded cultivars. Similarly, one or two drops of colchicine are then applied, one drop to smaller seedlings and two drops to larger seedlings. The success rate of chromosome doubling is low (about 1%). Detection of tetraploids is by chromosome count or other methods such as a count of the number of chloroplasts in each side of the guard cell (diploids have 5–6 per side, whereas tetraploids have 10–14 per side). Morphologically, tetraploids tend to have thicker leaves, shorter stems, and slower growth than diploids.

**3 Development of tetraploid lines.** The putative tetraploids from the mutagenesis ( $T_1$  plants) are evaluated in the next generation ( $T_2$ ) to authenticate their polyploidy. Selected plants are selfed or sibbed repeatedly in isolation to stabilize the new genotypes and increase fertility for high seed yield to produce sufficient seed for evaluation of the tetraploid lines.

**4 Evaluation of tetraploid lines.** The tetraploid lines in the  $T_4$  generation are evaluated especially for performance in a hybrid by making selected crosses to produce triploids. These triploids are evaluated to identify tetraploids with high potential as parents. Successful parents produce hybrids with high yield, good rind color (gray), and that lack empty seed coats.

**5 Development and evaluation of triploid lines.** Successful tetraploid parents are used in producing triploids by hand pollination (expensive) or insect pollination in isolation blocks. Hand pollination enables the breeder to exercise control over the pollination process to ensure a high success rate of desirable crosses. Pistilate flowers from the female parent may be tagged. When developing triploids by insect pollination, the two parents should have different rind colors for easy identification of hybrids. As indicated previously, the tetraploid female usually has a gray rind whereas the diploid is striped. The triploid produced by hybridization will be striped.

### Breeding for industrial uses

Some crops have multiple uses – food, feed, and industrial uses. Corn, for example, may be used for flour

milling, extraction of oil, starch production, or sugar production. Breeders have developed special purpose hybrids for these uses. For example, by inserting the endosperm gene, sugary (*su*), and another mutant gene, shrunken (*sh2*), in the same genotype, the resulting hybrid has increased sugar content (called supersweet or extrasweet corn). Similarly, waxy corn is developed for use in the production of adhesives, gums, and puddings, because of its high amylopectin content.

Edible oils that are high in polyunsaturated fatty acids are considered more healthy because they have the capacity to lower blood cholesterol. However, such oils are unstable for high-temperature cooking because they oxidize and break down readily at high temperatures. The hydrogenation process that is used to artificially stabilize oils converts the polyunsaturates back to monounsaturates. Oil plants such as soybean and cotton have genes that naturally convert oleic acid, a monounsaturated fatty acid, into polyunsaturated fatty acid. In cotton, about 25% of the seed oil comprises two fatty acids – palmitate and stearate. Conventional cotton seed oil contains primarily palmitate, which is thought to raise blood cholesterol more than stearate. Scientists at the Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia, turned off the genes for polyunsaturates to produce cotton seed with high levels of oleic acid (monounsaturates). This high oleic oil is more stable for high-temperature cooking and can be used in place of hydrogenated oils. Further, the scientists modified the plants to produce more stearate than palmitate, making the genetically modified (GM) product healthier for human use.

Similarly, in soybean, other scientists used the technology of gene silencing to develop high oleic cultivars. Using biotic techniques, a second copy of the *fad2* gene (fatty acid desaturase gene) that encodes the enzyme

delta-12 desaturase, which is involved in fatty acid synthesis, was introduced into the genome of soybean. This event switched off the desaturase gene, causing the accumulation of oleic acid in the seed only (normal fatty acid biosynthesis occurs in other plant parts). Consequently, only small amounts of the polyunsaturated fatty acids, linoleic and linolenic, are produced in the seed. The high oleic soybean GM cultivars such as G94-1 contain about 80% more oleic acid than conventional seed, and higher levels than in olive oil and rapeseed oil.

### Breeding plants for novel traits

An application of genetic engineering to breed novel traits is the use of organisms as bioreactors to produce pharmaceuticals. One of the earliest applications of this technology was the commercial production of human insulin in microbial systems. Similarly, certain pharmaceuticals are commercially produced in mammalian milk of sheep, goats, and rabbits. The application is being applied to plants to produce selected chemical compounds. Plant-made vaccines are currently under development for protection against cholera, diarrhea (Norwalk virus), and hepatitis B. The most common plants that are being used in plant-made pharmaceuticals are corn, tobacco, and rice. Other crops being investigated include alfalfa, potato, safflower, soybean, sugarcane, and tomato. To be usable, the plant should be readily amenable to genetic engineering and capable of producing high levels of protein. Further, there should be an efficient method for extracting the protein products from the plant tissues. Another example of a plant-manufactured pharmaceutical is taxol, a secondary product derived from the Pacific yew tree. This product has been found to be effective against certain cancers.

### References and suggested reading

- Bjarnason, M. 1990. CIMMYT's Quality Maize Program: Present status and future strategies. Paper presented at the 18th Congresso Nacional de Milho e Sorgo, Vitoria, Brazil, 29 July–3 August.
- Helms, T.C., and J.H. Orf. 1998. Protein, oil, and yield in soybean lines selected for increased protein. *Crop Sci.* 38:707–711.
- Hulke, B.S., W.R. Fehr, and G.A. Welke. 2004. Agronomic and seed characteristics of soybean with reduced phytate and palmate. *Crop Sci.* 44:2027–2031.
- Mertz E.T., L.S. Bates, and O.E. Nelson. 1964. Mutant gene that changes protein composition and increases lysine content of maize endosperm. *Science* 145:279–280.
- National Research Council. 1988. Quality-protein maize. Academy Press, Washington, DC.
- Wang, T.L., C. Domoney, C.L. Hedley, R. Casey, and M.A. Grusak. 2003. Can we improve the nutritional quality of legume seeds? *Plant Physiol.* 131:886–891.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 “Golden Rice” is rice bred for high protein content.
- 2 Triploid hybrid fruits are seedless.
- 3 The “FlavrSavr” tomato has high protein content.
- 4 Genes were obtained from the daffodil in the development of “Golden Rice”.
- 5 Cereals tend to be high in lysine.

### Part B

Please answer the following questions:

- 1 Quality protein maize is bred by incorporating the ..... gene.
- 2 Give a specific problem encountered in the development of “Golden Rice”.

### Part C

Please write a brief essay on each of the following topics:

- 1 Describe the breeding of seedless watermelon.
- 2 Discuss quality protein maize.
- 3 Discuss the breeding of supersweet corn.
- 4 Discuss high lysine corn.

# **Section 8**

## **Cultivar release and commercial seed production**

- 
- Chapter 23 Performance evaluation for crop cultivar release**
  - Chapter 24 Seed certification and commercial seed multiplication**
  - Chapter 25 International plant breeding efforts**
  - Chapter 26 Emerging concepts in plant breeding**
- 

Plant breeding is often a very lengthy process, lasting over a decade in many cases. After the plant breeder assembles and organizes variability into a new genetic matrix, the next and very crucial stage is to decide which of the many genotypes to release as a cultivar for use by crop producers. This step involves the evaluation of candidate genotypes under various conditions, including those under which the crop will be produced. The plant breeder resorts to field plot techniques to assist in experimental design and conduct, as well as for the evaluation of genotypes to determine the most promising for release. Computers are often needed to handle the complex analysis that considers not only the final yield, but also yield stability. After identifying the most promising genotype, the material is submitted for seed increase, certification according to criteria specific to the particular crop, and then release for sale.

Whereas seed production and cultivar development are well organized and highly commercialized enterprises in developed economies, farmers in developing countries often depend on low-yielding farmer-developed cultivars for planting. Unable to afford the high cost of improved seed marketed by multinational companies, international agricultural centers are a major source of improved seed for farmers in poor countries. Further, plant breeders are regularly seeking new ways of cultivar development for increasing the productivity of farmers and meeting consumer demands. Emerging concepts in plant breeding and the role of international agricultural centers in global cultivar development are also discussed in this section.

---

# 23

# Performance evaluation for crop cultivar release



## Purpose and expected outcomes

*The ultimate goal of the plant breeder is to be able to identify a superior genotype that can be released as a new cultivar to farmers for commercial production. To arrive at this goal, many experimental genotypes of high genetic potential are evaluated for performance capabilities under various environmental conditions, over several seasons and years, and at different locations. Detailed records are compiled and analyzed to help in the decision process. The breeding materials are evaluated by using appropriate statistical tools that entail design of the trials, collection of data, analysis, and interpretation of results. This calls for an understanding of how genotypes interact with the environment and of the field plot technique. After studying this chapter, the student should be able to:*

- 1 Discuss the concept and role of genotype × environment interaction in plant breeding.
- 2 Describe the field plot technique in plant breeding.
- 3 Discuss the process of crop registration.

## Purpose of performance trials

As indicated in all the breeding schemes discussed in this book, the breeder conducts performance or field trials of the advanced generation of the materials developed in a breeding program, primarily to identify a genotype to be released as a cultivar to producers. In a sense, these tests or field trials are designed to forecast the performance of the genotype to be released as a cultivar. These trials are usually called **yield trials** because yield is usually the most important trait in a breeding program. Breeding programs are frequently undertaken to address “secondary” traits (disease resistance, early maturity, high seed protein, etc.). If these traits are successfully transferred but the genotype has poor yield, it would not be released as a cultivar. The field trial also enables the breeder to collect data about the characteristics of the potential cultivar for other uses (e.g., registration of the cultivar). It is the primary source of information for the breeder to use in the decision-making process of cultivar release.

The breeder is not only concerned about the level of yield of the cultivar but also about the stability of yield. Consequently, stability analysis is often part of the performance evaluation of genotypes prior to cultivar release. The concepts of adaptation and yield stability analysis are further discussed later in this chapter.

## Kinds of field trials

There are two basic kinds of field trials conducted in plant breeding – **breeders’ trials** and **official trials**.

### Breeders’ trials

These trials are conducted for the primary purpose of evaluating the performance of the final set of genotypes (advanced generations in a breeding program) to allow the breeder to make a decision as to which genotype to release as a cultivar. Some breeders conduct these trials

in two stages. The first stage, called the **preliminary yield trial (PYT)**, starts at an earlier generation (e.g., F<sub>6</sub>, depending on the objectives and method of breeding), and consists of a larger number of entries (genotypes). Further, these entries may be planted in fewer rows per plot (e.g., two rows without borders) and fewer replications (2–3) than would be used in the final trial, the **advanced yield trial (AYT)**. Superior genotypes are identified for more detailed evaluations. The PYT is designed to be a quick evaluation of the breeding efforts. The PYT consists of a fewer number of promising genotypes (10–20), depending on resources. It is conducted for several years at different locations, using more replications and plots with more rows and with border rows. It is also subjected to more detailed statistical analysis.

Breeders' trials vary in scope, according to the crop, its distribution and importance, and resources available to the breeder. Some breeders (especially in the public sector) limit their evaluations to within the state or mandate region. Commercial breeders may conduct regional, national, and even international trials through established networks. Public breeders may have wide networks for trials (e.g., INTSOY – the International Soybean Program in the USA). In terms of management, breeders' trials may also be conducted in one of two ways – research managed or farmer managed.

- 1 Research managed.** This is a trial conducted at a research station or experimental farm under the supervision of researchers. These are usually replicated, full-scale, and self-contained trials designed to collect data that can be published in a scientific publication. Extensive data are often collected.
- 2 Farmer managed.** The trials are conducted on farmers' fields (see participatory breeding in Chapter 26). Often, it is a scaled down version of what is conducted at a research station (fewer replications, fewer plots, etc.), so that the farmer is not be overburdened. The breeder may use some creative analysis to obtain valid data from these trials. For example, different farmers in the same location may be considered as a block in a randomized complete block design.

### Official trial

After a genotype has been identified as a potential cultivar, the breeder may seek legal protection by applying for protection under the Plant Variety Protection (PVP) Act (see Chapter 15) and/or registration of the cultivar with an official seed agency. This trial is more detailed than the performance trial for yield and provides information

needed to establish legal identity for the cultivar, showing its distinctness from existing ones. Data must be collected to also indicate its uniformity and stability (i.e., the genotype breeds true from year to year).

### Designing field trials

One of the phases of plant breeding in which statistical analysis is used extensively is the design and conduct of performance evaluations. The key considerations in the design of a field trial are as follow:

- 1 Number of genotypes to evaluate.** As previously indicated, PYTs have more entries than AYTs. Whereas research-managed trials have the full complement of genotypes, farmer-managed trials may be reduced to a small number.
- 2 Where to conduct the trials (locations).** Breeders usually conduct trials at multiple locations. These locations, ideally, should be representative of the target regions for which the cultivars are to be released. In practice, test locations are seldom randomly selected. Breeders are limited to sites where they have collaborators (e.g., institutes, research stations, universities), or farmers who have an interest in participating in the project. Where possible, the breeder should endeavor to test at both research stations (where an optimal selection environment can be obtained), as well as at sites that reflect the major cropping areas and farming practices. Even when trials are conducted at research-managed sites, efforts should be made to replicate the actual production conditions in the farmers' fields (e.g., crop management practices). Research institutions often strategically locate a few research farms in target regions that represent the climatic and soil conditions of the area. The total number of sites is variable (about 5–10), but it depends on the extent of variability in the target region. Areas of major production should have more sites than those with less production of the crop.
- 3 What statistical design to use for field layout.** Randomized complete block designs are commonly used in breeders' trials. Research-managed trials may adopt more sophisticated designs, but farmer-managed trials should be as simple as possible. The former should have more replications than the latter, as previously stated.
- 4 What data to collect.** Researchers at experimental stations may use equipment and machinery designed for research (e.g., plot planter and combines). They have the time and the expertise to collect a wide

variety of data in addition to yield. Farmer-managed trials should be designed to permit farmers to utilize existing equipment and machinery already at their farms. Also, data collection, if it is to be done by the farmer, should be easy to collect and as minimal as possible. Some breeders sometimes request a farmer to make available land for the trial. Planting, management, and data collection are then done by the researcher.

- 5 Number of seasons/years to conduct the trial.** For effective evaluation of genotype  $\times$  location ( $G \times E$ ) interaction, at least 2 years of testing (more for repeatability) are needed for annual crops.
- 6 How to analyze and interpret the analysis to draw valid conclusions.** The breeder may use more efficient designs based on incomplete blocks. There may be unbalanced data (e.g., missing plots). Analyzing data over seasons, years, and locations is complex. The breeder should have the software for these and other analyses (e.g., stability analysis) and be familiar with the statistical methods (or at least have assistance to correctly analyze and interpret the results).

Before describing the steps involved in conducting field trials, some key concepts that are critical to the design and analysis of such experiments will be discussed.

### Role of the environment in field trials

The terms site and location are used interchangeably to indicate spatial variation. The term environment is used to represent the conditions under which plants grow and includes locations, years, and management practice adopted. A location/year constitutes one environment. The nature and effect of the environment has implications in the design and conduct of field trials. Test environments may be artificial (e.g., different levels of fertilizer) or natural (e.g., seasons, location) or both.

#### Types of environmental variables

The environmental variables that plant breeders face during genotype evaluation may be divided into two general categories – predictable and unpredictable factors.

- 1 Predictable factors.** Predictable environmental factors are those that occur in a systematic fashion or can be controlled and manipulated by the breeder. These include natural variables such as soil type (e.g., clay, sandy, organic), which are immutable over a short range of time, and breeder-imposed variations (e.g.,

planting dates, intra- and inter-row spacing, rates of fertilizer or irrigation application). Breeders may design studies to evaluate each of the imposed factors separately or several simultaneously. Variations in the soil are managed through the way plots are oriented, their shape, and sizes, as well as how the breeder-imposed variations (called **treatments**) are allocated to plots in the field.

- 2 Unpredictable factors.** Unpredictable factors of the environment are those that vary erratically over short or long periods of time. The local weather (the short-term meteorological characteristics of a place) is more fickle and relatively unpredictable than climate (the long-term patterns in meteorological characteristics of a region). As previously indicated, climate is the basis of crop adaptation; weather is the basis for crop production. Key unpredictable environmental factors of interest in genotype evaluation include rainfall, temperature, and relative humidity. To evaluate the effect of these factors, breeders test their materials at different locations (genotype  $\times$  location) or in different years (genotype  $\times$  years) or a combination of these factors (genotype  $\times$  location  $\times$  years).

#### Scale

Another way in which environments are categorized is according to scale:

- 1 Microenvironment.** This is the immediate environment often pertaining to the organism (plant). This includes soil and meteorological factors (e.g., light, moisture, temperature), and biotic factors in a limited space, intimately associated with the organism.
- 2 Macroenvironment.** This refers to the abiotic and biotic factors on a larger scale (location, region) at a particular period of time.

### Genotype $\times$ environment ( $G \times E$ ) interactions

**Genotype  $\times$  environment ( $G \times E$ ) interaction** is said to occur when two or more genotypes are compared across different environments and their relative performance (responses to the environment) are found to differ. That is, one cultivar may have the highest performance in one environment but perform poorly in others. Another way of stating this is that, over different environments, the relative performance of genotypes is inconsistent.  $G \times E$  is a differential genotypic expression across multiple environments. The effect of this interaction is that the association between phenotype and

genotype is reduced. This raises the important issue of adaptation because a breeder's selection in one environment of superior performers may not hold true in another environment. By measuring the  $G \times E$  interaction, the breeder will be better equipped to determine the best breeding strategy to use to develop the genotype that is most adapted to the target region.

### Classification of $G \times E$ interactions

The type of  $G \times E$  interaction influences the nature of the cultivar the breeder eventually releases for the production region. The environment, as previously described, can be complex, and so can the genotype of the plant. Consequently, the biological basis of  $G \times E$  interactions is complex by nature. Environmental factors are constantly changing. The interaction between the genotype (cultivar) and the environment is ongoing. As the number of environments ( $n$ ) and number of genotypes ( $m$ ) increase, the number of possible  $G \times E$  interactions is given by  $mn! / m!n!$ . Of this, there is theoretically only one genotype that is the best performer under all environments, odds that make a search for it futile. Allard and Bradshaw classified this interaction into three common patterns using two genotypes (A, B) and two environments ( $E_1, E_2$ ) for a graphic illustration of the concept of  $G \times E$  interaction. In statistical terms, a  $G \times E$  interaction has occurred when the difference in performance between the two genotypes is inconsistent over the environment:

$$A_1 - B_1 \neq A_2 - B_2 \text{ [or } A_1 - B_1 - (A_2 - B_2) \neq 0]$$

A  $G \times E$  interaction exists when:

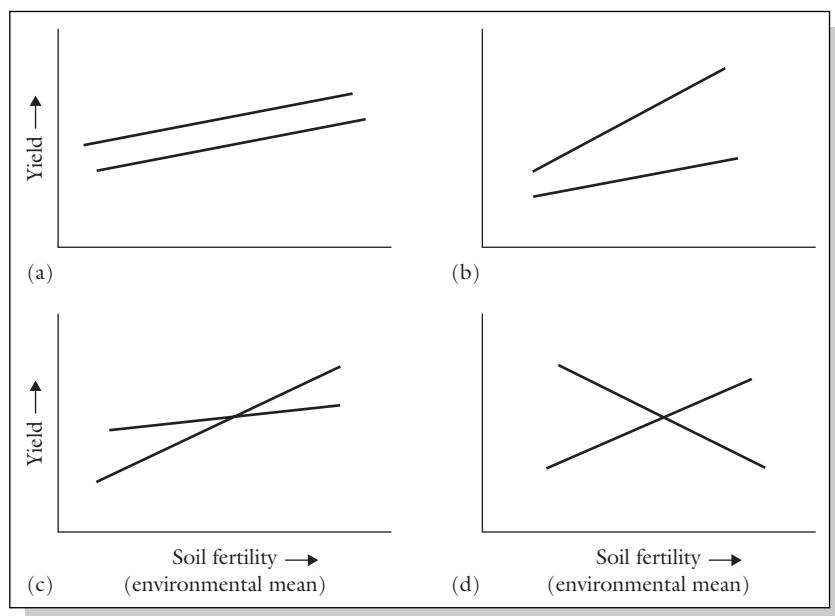
$$A_1 - B_1 - A_2 + B_2 \neq 0$$

Three basic types of  $G \times E$  interaction – **no interaction**, **non-crossover interaction** (quantitative interaction), and **crossover interaction** (qualitative interaction) – are recognized. A numerical example can be used to distinguish these classes of interactions (Table 23.1). A graphic illustration may also be used to demonstrate the nature of these interactions (Figure 23.1). Consider two genotypes, A and B, in a field trial analysis.

- 1 No  $G \times E$  interaction.** A no  $G \times E$  interaction occurs when one genotype (e.g., A) consistently performs better than the other genotype (B) by about the same amount across all the environments included in the test:
- 2 A non-crossover  $G \times E$  interaction.** A non-crossover  $G \times E$  interaction is said to occur when a genotype (A) consistently outperforms genotype B, across the entire test environment. However, the differential performance is not the same across the environment. That is, whereas there is no change in rank, genotype A may exceed genotype B by 20 units in one environment and 60 units in another.
- 3 A crossover  $G \times E$  interaction.** This is the most important  $G \times E$  interaction to plant breeders. A crossover  $G \times E$  interaction occurs when a genotype (A) is more productive in one environment, but a

**Table 23.1** Demonstration of  $G \times E$  interaction.

No interaction	Environment 1	Environment 2	Difference
Genotype A	10	14	+4
Genotype B	16	20	+4
Difference	+6	+6	
Non-crossover interaction	Environment 1	Environment 2	Difference
Genotype A	10	14	+4
Genotype B	16	24	+8
Difference	+6	+10	
Crossover interaction	Environment 1	Environment 2	Difference
Genotype A	16	14	-2
Genotype B	10	20	+10
Difference	-6	+6	



**Figure 23.1** Graphical presentation of: (a) genotype  $\times$  environment ( $G \times E$ ) interactions, (b) heterogeneity, (c) crossover interactions, and (d) combined interactions.

different genotype (B) is more productive in another environment. The basic test for crossover interaction (also called qualitative interaction) is to compare the performance of two genotypes in two environments and to determine if the difference in performance is significantly less than zero in one environment and significantly greater than zero in the other.

**4 Combined  $G \times E$  interaction.** The first three interactions previously described are the ones commonly discussed. If one of the factors considered on the axes increases for one genotype and reduces for the other genotype, there is a combined  $G \times E$  interaction.

The axes in the graph may be for any relevant factor of interest to the breeder. For example, the  $x$ -axis may be rainfall, while the vertical axis ( $y$ ) may be grain yield. In spite of the complexity of the environment, sometimes one factor may predominate to characterize the environment (or may be imposed by design). It should be pointed out these four graphs are only a selected unique few of the numerous patterns that may occur in reality. The breeder is most interested in repeatable  $G \times E$  interactions.

#### Measurement of $G \times E$ interactions

Interactions occur at various biological levels, such as genotypic, QTL (quantitative trait locus), and phenotypic

levels – the first two requiring genetic analysis.  $G \times E$  interaction at the phenotypic level requires observations at the plant or crop level. The  $G \times E$  interactions can also be partitioned into linear trends (e.g.,  $G \times$  location,  $G \times$  year,  $G \times$  time). Statistical methods are used to assess  $G \times E$  interactions. Consequently, the proper field plot design and analysis are required for an effective assessment of the interactions. These methods include both parametric and non-parametric procedures – partitioning of variance, regression analysis, non-parametric methods, and multivariate techniques.

#### Analysis of variance (ANOVA)

To ascertain the presence of a  $G \times E$  interaction, breeders conduct a network of comparative trials, as previously described, in which prospective cultivars are compared with standard cultivars at multiple locations or agroecological regions. The premise for such trials, according to Mather and others, is expressed by a linear equation:

$$X = \mu + g + e + ge$$

where  $X$  = yield of some other quantitative traits,  $\mu$  = mean value of the population (trial),  $g$  = value of the genotype (cultivar),  $e$  = value of the environmental effect, and  $ge$  = genotype  $\times$  environment interaction.

Different models of ANOVA are used for partitioning variance. The genotypes are never chosen at random since they are deliberately selected by the breeder as prospective cultivars. Similarly, locations are often not randomly chosen as previously discussed. However, they may be considered random if there are many of them spread over a large region. The genotypes are

evaluated over several years. The effects are random since the environment is not controlled. Also, where the genotypes are a random sample from a large population, their effects are random.

For a two-factor mixed model (fixed genotypes + random environment) the ANOVA table is as follows:

Source	df	MS	EMS
Replications ( $r$ )	$ly(r-1)$		
Years ( $y$ )	$y-1$		
Location ( $l$ )	$l-1$		
$y \times l$	$(y-1)(l-1)$		
Genotypes ( $G$ )	$g-1$		
$G \times y$	$(g-1)(y-1)$	$M_1$	$\sigma_e^2 + r\sigma_{gy}^2 + rl\sigma_{gy}^2 + ry\sigma_{gl}^2 + rly\sigma_g^2$
$G \times l$	$(g-1)(l-1)$	$M_2$	$\sigma_e^2 + r\sigma_{gl}^2 + ry\sigma_{gl}^2 + rly\sigma_g^2$
$G \times y \times l$	$(g-1)(y-1)(l-1)$	$M_3$	$\sigma_e^2 + r\sigma_{gl}^2 + ry\sigma_{gl}^2$
Error ( $G \times r$ )	$(g-1)(r-1)ly$	$M_4$	$\sigma_e^2 + r\sigma_{gl}^2$
		$M_5$	$\sigma_e^2$

Variances are calculated as follows:

$$\begin{aligned}\sigma_e^2 &= M_5 \\ \sigma_{gy}^2 &= (M_4 - M_5)/r \\ \sigma_{gy}^2 &= (M_2 - M_4)/rl \\ \sigma_{gl}^2 &= (M_3 - M_4)/ry \\ \sigma_g^2 &= (M_1 - M_2 - M_3 + M_4)/rly\end{aligned}$$

#### Breeding implications of ANOVA results

Lack of significant  $G \times E$  interactions for genotype  $\times$  location or location  $\times$  years indicates that the breeder may be able to select a superior genotype for release for use throughout a specified production region, following genotype evaluations in just 1 year. Crossover  $G \times E$  interactions are those that require careful interpretation by breeders. In this instance decisions are based on the practical significance of the results of the analysis. Breeders need to include in their decision-making process factors such as the magnitude of the change in rank. Consequently, it is not uncommon for different conclusions to be drawn by different breeders examining the same results. General interpretations of  $G \times E$  interactions resulting from unpredictable causes are as follow:

- If significant genotype  $\times$  location effects are observed and the rankings fluctuate by wide margins, the results indicate that the breeder should consider establishing separate breeding programs for the different locations (i.e., develop different cultivars for different locations). However, before making a decision, it is wise to examine the data to see what specific factors are

responsible for the variation. If stable factors such as soil are the source of variation, separate breeding efforts may be warranted.

- A significant genotype  $\times$  year interaction is similar in effect to genotype  $\times$  location. However, because the breeder cannot develop programs for different years, a good decision would be to conduct tests over several years and select the genotype with superior average performance over the years for release. Because conducting one trial per year for more years will prolong the breeding program, the breeder may include more locations and decrease the number of years.
- The breeding implications for a complex interaction like genotype  $\times$  years  $\times$  location is for the breeder to select genotypes with superior average performance across locations and over years, for release as new cultivars for the production region. Farmers will benefit from growing more than one cultivar each cropping season. This strategy will reduce the effects of the fluctuations attributed to genotype  $\times$  year interactions.
- The magnitude of a  $G \times E$  interaction is influenced by the genetic structure of the genotype. Genotypes with less heterogeneity (e.g., pure lines, single-cross hybrids, clones) or heterozygosity (e.g., pure lines) generally interact more with the environment than open-pollinated genotypes or mixtures, because of lower amounts of adaptive genes.
- Also, it is widely known that only  $G \times$  location interactions (rather than all the kinds of  $G \times E$  interactions) are useful for depicting adaptation patterns. This is because they are the only interactions that can be exploited by selecting for specific adaptation or by growing specifically adapted genotypes. For

example, a significant  $G \times$  year interaction cannot be exploited because the climatic conditions that generate year-to-year environmental variation are not known in advance. Consequently, the analysis of multiple environment yield trials should focus primarily on  $G \times$  location interactions, the other interaction being considered in terms of yield stability.

### Stability analysis

An ANOVA identifies statistically significant and specific interactions without telling the breeder anything about them. If a  $G \times E$  interaction is significant and the environmental variations are unpredictable, the breeder needs to know which genotypes in the trials are stable. A **stability analysis** is used to answer this question. **Field stability** may be defined as the ability of a genotype to perform consistently, whether at a high or low level, across a wide range of environments. Stability may be static or dynamic. **Static stability** is analogous to the biological concept of homeostasis (i.e., a stable genotype tends to maintain a constant yield across environments). **Dynamic stability** is when a stable genotype has a yield response in each environment such that it is always parallel to the mean response of the genotypes evaluated in the trials (i.e.,  $G \times E = 0$ ). Static stability is believed to be more useful than dynamic stability in a wide range of situations, and especially in developing countries. The genotype produces a better response in unfavorable environments or years. Whenever the  $G \times E$  interaction variation is wide, breeding for high-yield stability is justifiable.

As previously stated, ANOVA only detects the existence of  $G \times E$  effects. Breeders will benefit from additional information that indicates the stability of genotype performance under different environmental conditions. The stability of cultivar performance across environments depends on the genotype of individual plants and the genetic relationship among them. Generally, heterozygous individuals (e.g.,  $F_1$  hybrids) are more stable in their performance than their homozygous inbred parents.

A variety of methods have been proposed for genotype stability analysis. Examples are **regression analysis** and the **method of means**.

#### Regression analysis

This method of simple linear regression (also called **joint regression analysis**) stability analysis was developed by K. W. Finlay and G. N. Wilkinson (1963) and later by S. A. Eberhart and W. A. Russell (1966).

It is preceded by an ANOVA to assess the significance of  $G \times E$ . The breeder proceeds to the next step of regression analysis only if the  $G \times E$  interaction is significant.

Statistically, the observed performance ( $\gamma_{ij}$ ) of the  $i$ th genotype ( $i = 1, \dots, 5$ ) in the  $j$ th environment ( $j = 1, \dots, e$ ), may be expressed as:

$$\gamma_{ij} = \mu + g_i + e_j + ge_{ij} + \varepsilon_{ij}$$

where  $\mu$  = grand mean over all genotypes and environments,  $g_i$  = additive contribution of the  $i$ th genotype (calculated as the deviation from  $\mu$  of the mean of the  $i$ th genotype averaged over all environments),  $e_j$  = additive environmental contribution of the  $j$ th environment (calculated as the deviation  $\mu$  of the mean of the  $j$ th environment averaged over all genotypes),  $ge_{ij}$  =  $G \times E$  interaction of the  $i$ th genotype in the  $j$ th environment, and  $\varepsilon_{ij}$  = error term attached to the  $i$ th genotype in the  $j$ th environment. The regression coefficient for a specific genotype is obtained by regressing its observed  $\gamma_{ij}$  value against the corresponding mean of the  $j$ th environment.

If the yield was conducted over a wide range of environments, and hence a wide range of yields obtained, it is reasonable to assume that the individual trial means sufficiently summarize the effects of the environments. The mean performance of each genotype over all the test environments constitutes the **environmental index** (Table 23.2). In effect, this method of analysis produces a scale of environmental quality. The results for each genotype are plotted, trial by trial, against trial means, to obtain a regression line. According to Eberhart and Russell, an average performing genotype will have a regression coefficient of 1.0 and deviations from regression of 0.0, since it will tend to agree with the means. However the genotypes that were responsible for the  $G \times E$  interactions detected in the ANOVA, will have slopes that are unequal to unity. Furthermore, a genotype that is unresponsive to environments (i.e., a stable performer) will have a low slope ( $b < 1$ ), while a genotype that is responsive to environments (i.e., an unstable performer) will have a steep slope ( $b > 1$ ).

The regression analysis technique has certain limitations, both physiological and statistical, that can result in inaccurate interpretation and recommendations for cultivar release. For example, the use of mean yield in each environment as an environmental index flouts the statistical requirement of an independent variable. Also, when regression lines do not adequately represent data, the point of intersection of these lines has little biological meaning, unless considered with

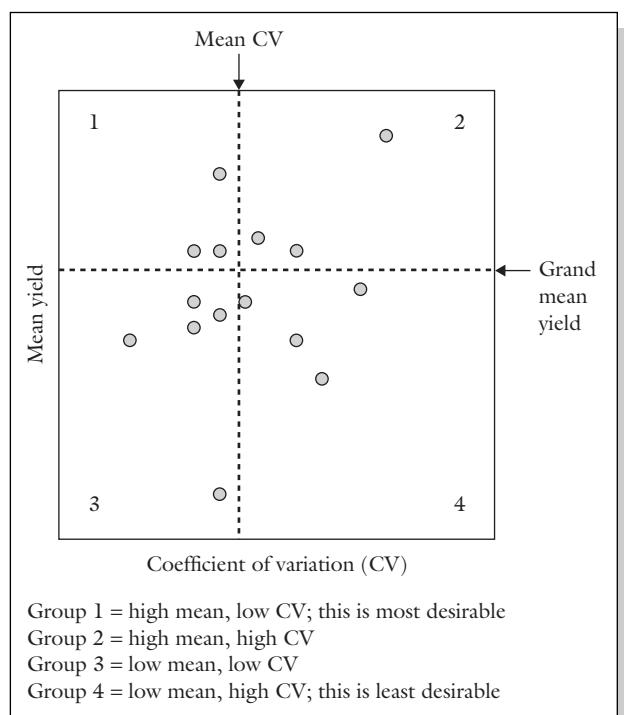
**Table 23.2** Stability analysis.

ANOVA					
Source	df	SS	MS	F	Probability
Environment ( $E$ )	4	27,103.87	6,775.96	104.03	0.0001
Reps ( $R$ ) $\times E$	15	22,580.65	1,505.37	23.11	0.0001
Genotype ( $G$ )	9	4,595.65	510.65	7.84	0.0001
$E \times G$	36	6,068.63	168.57	2.59	0.0001
Pooled error	135	8,792.85	65.13		

Regression analysis			Environmental index	
Genotype	Mean	Regression coefficient	Genotype	Index
N1	105.2	0.906	N1	93.47
N2	103.3	0.759	N2	96.3
N3	108.3	1.741	N3	112.4
N4	99.9	0.972	N4	123.6
N5	101.9	0.559	N5	96.7
N6	108.5	1.141		
N7	106.7	0.926		
N8	103.7	0.999		
N9	94.8	0.968		
N10	112.9	1.028		

Genotypes N2 and N5 are stable performers; N3 and N6 are responsive to the environment and unstable in performance; N8 and N10 are average performers.

**Figure 23.2**  $G \times E$  interaction based on the coefficient of variation (CV).

their associated statistical error – something often ignored by many breeders. Limitations notwithstanding, the regression technique is simple and has biological significance. Complex interactions are reduced to linear responses.

#### Plot of means versus coefficient of variation

Proposed by Francis and Kannenburg in 1978, this method entails calculating for each variety, the overall mean and the coefficient of variations (CVs) across the environments. A plot of means versus CVs yields a scattergram that can be divided into four sections by transecting the average CV and the grand mean yield (Figure 23.2). The most desirable genotype will be found in group 1 (high yield, low CV) while the least desirable (low yield, high CV) will occur in group 4.

#### Non-parametric methods

Multivariate procedures used to analyze  $G \times E$  interactions include clustering, principal component analysis (PCA), and factor analysis. These procedures perform uniformly across environments. A recent addition to these techniques is the additive mean effects multiplicative



Russell Freed

## Industry highlights

### *MSTAT: a software program for plant breeders*

Department of Crop and Soil Science, Michigan State University, East Lansing, MI 48824, USA

Data management is crucial for a successful plant breeding program. Plant breeders must design and analyze experiments, print labels and field books, and manage pedigree and other information about different plant populations. As computers came into use in the early 1980s, there was a need for a standard management and statistical package for plant breeders. Russell Freed at Michigan State University along with Oivind Nissen from Norway developed a basic program that incorporated the principles of good plot design and appropriate statistical analysis. As advances in data management and computer technology occurred, MSTAT became what it is today: a tool to help plant breeders develop successful programs through better data management and analysis. Over the years, MSTAT has been used by thousands of researchers to facilitate the development and identification of new plant cultivars in over 110 different countries (<http://www.msu.edu/~freed/mstatc.htm>).

MSTAT helps plant breeders design, manage, and analyze simple and complex experiments. MSTAT is capable of generating random block designs with one to five factors (0–4 splits), and square or rectangular lattices; and also prints field books and labels. It enables plant breeders to quickly and accurately determine the biological and economic significance of their experiments. It can analyze one-way, two-way, factorial, lattice, hierarchical, and non-orthogonal data sets. Figure 1 is a sample output from the ANOVA-2 program. MSTAT can also calculate correlations, regression, multiple regression, frequency tables, and basic statistics. MSTAT can analyze data over multiple years and locations with the STABIL program option. This is very important for breeders to identify the yield stability of promising lines. The ease and speed of data analysis provides timely results and thereby facilitates the generation of new plant cultivars.

The BRSERIES program keeps pedigree information for population development and will automatically update generation information. To prepare next year's planting book, the BRSERIES will ask how many rows need to be planted from the current year's book and then automatically updates the generation information and creates a book with the new row numbers. It can print field books, maps, and labels. MSTAT can also be used to help select parents for the crossing program. Figure 2 is a sample print-out from the BRSERIES.

Title: BEAN EXPERIMENT					
Function: ANOVA-2					
Data case 1 to 120					
Two-way Analysis of Variance over variable 2 (REPLICATION) variable 1 (ENTRY) Variable 4: YIELD					
A N A L Y S I S   O F   V A R I A N C E   T A B L E					
Source	Degrees of Freedom	Sum of Squares	Mean Square	F-value	Prob
REPLICATION	3	5.09	1.697	2.20	0.0939
ENTRY	29	75.34	2.598	3.37	0.0000
Error	87	67.16	0.772		
Non-additivity	1	0.36	0.358	0.46	
Residual	86	66.80	0.777		
Total	119	147.59			
Grand Mean = 3.558 Grand Sum = 427.000 Total Count = 120					
Coefficient of Variation = 24.69%					
The means of the 4 replications and the 30 varieties are also listed.					

**Figure 1** MSTAT output from analysis of variance for 30 cultivars and four replications using the ANOVA-2 program.

Seed Source	Exp. Number 84HT01 Cross Gen Parent Lines	2004 WHEAT HEADROWS Code Plot No	
	Check variety = Redcoat	*****	
83GH01	4 T830001 F1 FRANKENMUTH/AUGUSTA	PMHFLR	101
83GH01	4		102
83GH01	6 T830002 F1 CALDWELL/AUGUSTA//HOUSER	LRPM	103
83GH01	6		104
			105

**Figure 2** Output of a plant breeding book generated with BRSERIES in MSTAT.

The use of MSTAT strengthens plant breeding programs in several ways: (i) by quickly and accurately designing yield trials and other experiments; (ii) by enabling breeders to maintain better records; (iii) by increasing the size of the breeding programs because of the ease of information handling; (iv) by creating books with information collected over several seasons; and (v) by analyzing experimental data over multiple locations and years. MSTAT helps to improve the efficiency, timeliness, and confidence of conclusions that are all very important concerns for plant breeders.

interaction (AMMI). These procedures were discussed in Chapter 7.

### Adaptation

According to P. M. A. Tigerstedt, in the context of evolutionary biology, **adaptation** is a process, **adaptiveness** is the level of adaptation of a plant material to an environment, while **adaptability** is the ability to show adaptiveness in a wide range of environments. However, according to Cooper and Byth, in a plant breeding context, adaptation and adaptiveness relate to a condition rather than a process. They indicate the ability of a plant material to be high yielding with respect to a specific environment, one to which it is adapted. One of the purposes of a  $G \times E$  analysis is to help breeders decide whether to breed for narrow or wide adaptation or adaptability. Breeding for narrow adaptability, the breeder's goal is to release a cultivar for a specific part of the target region (with a unique set of conditions), whereas in breeding for wide adaptability, the breeder focuses on releasing a cultivar with high performance across all environments.

As previously stated, only a  $G \times$  location interaction is useful for determining genotype adaptation patterns in an ANOVA analysis. Further, only repeatable  $G \times$  location interactions are of practical importance. However, if the  $G \times$  location variance is small in magnitude compared to other sources of variance, even though significant, it minimizes the specific advantage for breeding for narrow adaptation.

### Field plot technique in plant breeding

The subject of **field plot technique** deals with decision-making processes and other activities conducted by the plant breeder to fairly and effectively evaluate genotypes from a breeding program, to select cultivars for release. Field trials are designed to evaluate genotypes in a natural environment that is similar to that which farmers operate in. To evaluate the genetic potential of genotypes, they must be grown in an identical environment so that any observed differences can be attributed to the genotype not chance variations. Unfortunately, outside a controlled environment condition such as a greenhouse, it is impossible to find a truly homogeneous environment in the field. One of the goals of the field plot technique is to evaluate genotypes error free (or, in practice, with a minimum error). It is important, therefore, to identify the source of error in a field study.

In order to collect data in the proper manner for analysis and decision-making, the breeder should follow a set of statistically based rules for laying out an experiment in the field, called the **experimental design**. Such a design not only allows the breeder to estimate random error in the field but the choice of a good design can minimize this error.

### Sources of experimental error

In an experiment, a researcher may deliberately impose conditions or factors that would generate variance in outcomes; such a factor is called a **treatment**. The unit to which a treatment is applied is the **experimental**

**unit:** in the field, these are plots (of rows of plants); in the greenhouse, a unit could be a pot. For example, applying different levels of fertilizer, planting at different spacing, or evaluating different genotypes (as in breeding), are all treatments that elicit variable responses from plants. However, in addition to the expected variation from a treatment, there is always some variation that is unintended and cannot be accounted for, and is known in research as **experimental error**. By definition, this is the variation among plots that are treated alike. Major sources of error are those due to soil heterogeneity, the operator's inability to conduct the experiment uniformly as intended, and interplant competition.

#### *Soil (site) variability*

Soils are naturally heterogeneous, some more so than others. Natural variation may originate from differences in soil minerals, soil moisture, organic matter, or topography. The tops of slopes are drier than the bottom parts; nutrients wash down and accumulate at the bottom; depressions may drain poorly. In addition to these natural sources of soil variation, humans create additional variation through how they manage the soil in production. Before selecting and using a site for genotype evaluation, the breeder should know about the previous use of the land (history of use). The parcel of land of interest may have been differentially managed (e.g., different plant species were grown, or different tillage practices were imposed). The breeder should use all available data (e.g., yield records, management records) and visual observations to identify general patterns of variation at the proposed site. Experimental design techniques and other tactics can be used to minimize the effects of soil heterogeneity in field trials.

As previously indicated, breeders usually have selected sites at which they conduct their yield trials (e.g., experimental stations of universities). Are locations then a fixed variable in ANOVA? Some argue that locations are random effects since the breeder has no control over the meteorological factors that occur at locations. Most breeders also consider years of testing as random effects. It is important to have at least two replications of each treatment (genotypes) in each trial for estimating error variance.

#### *Tactics for reducing experimental error*

In order to correctly and effectively evaluate the desired variation, the researcher should eliminate or, more real-

istically, minimize extraneous variation. Some errors come from natural soil variability whereas others are human in origin. The plant breeder may use certain field plot techniques to minimize experimental errors.

**Use of border rows** Different genotypes differ in various plant characteristics to varying extents (e.g., growth rate, size, height, nutrient and moisture uptake). When planted next to each other, interplot competition may cause the performance of one genotype to be influenced by another in the adjacent plot. In early generation or preliminary trials, which usually include large numbers of genotypes, breeders often use fewer rows (e.g., two rows) in planting a plot in order to save resources. In advanced yield trials, four-row plots are customarily used. Data are collected on the middle rows only, because they are protected from **border effects**. Some breeders minimize border effects by increasing interplot spacing or using a common genotype for planting the border of all plots in a test in which row plots are few (one or two rows). To reduce interplot competition, the materials may be grouped according to competitive abilities.

**Proper choice of plot size and shape** Several factors affect the optimum plot size to use in field evaluation of genotypes (breeding objectives, stage of breeding, resources available, equipment). Evaluation of an  $F_2$  segregating population is often based on individual plant performance (i.e., individual plants are essentially plots). Consequently, the plants are adequately spaced to allow the breeder enough room to examine each plant.

Some breeders use what are called **microplots**, consisting of planting hills or short rows of test plants. This tactic is used in the early stages of genotype evaluation as a quick and inexpensive way of eliminating inferior genotypes.

Row plots are commonly used by breeders for genotype evaluation. The size and shape depends on the plant species, the land available, and the method of harvesting. Generally, row plots are rectangular in shape.

**Adequate number of replications** Replication or repetition of treatment in a test is critical to statistical analysis, providing a means of estimating statistical error. The number of replications used usually varies between two and four; fewer replications may be used in early evaluation of genotypes while advanced yield tests usually have four replications. The number of replications depends on the accuracy desired in the analysis, the

resources available (land, seed, labor), and the statistical design used. Replications increase the precision of the experiment and help the plant breeder more effectively evaluate the genotypes to identify superior ones.

**Minimizing operator errors** Data collection and analysis provide opportunities for human errors to occur. Computer software such as MSTAT (a statistical package developed by Michigan State Crops and Soil Science Department) will allow the breeder to generate a customized data collection book. Where a machine or equipment is to be used, it should be properly serviced (cleaned, calibrated).

The plots should be planted uniformly and managed uniformly (i.e., uniform spacing, fertilizing, irrigating). Border rows should be planted uniformly. Each plot is not enclosed in border rows. The plants at the end of rows have a competitive advantage over those in the inner part of the rows. Mechanized harvesting usually starts at the first plants of the middle row and proceeds to the last plants. This may introduce yield inflation and may require adjustment of plot yields.

### Principles of experimental design

There are many experimental designs used to allocate treatments to experimental units. However, they are all based on three basic principles – replication, randomization, and local control.

- 1 **Replication.** The principle of replication is critical for estimating experimental error, as previously indicated. It is also used for reducing the magnitude of error in an experiment.
- 2 **Randomization.** This is the principle of assigning treatments to experimental units such that each unit has an equal opportunity of receiving each treatment. This action eliminates bias in the estimation of treatment effects and makes the experimental error independent of treatment effects – a requirement for a valid test of significance of effects. Systematic arrangement allocates treatments to experimental units according to a predetermined pattern.
- 3 **Local control.** Sometimes, researchers find it more efficient to impose restrictions on randomization to further minimize experimental error. This is appropriate when there is a gradient in an environmental factor (e.g., fertility, moisture). Fertility is different at the top of a slope than at the bottom as mentioned before. Rather than ignoring this obvious variation, a technique called **blocking** may be used to divide

the field into distinct areas, maximizing the variation between blocks and increasing the homogeneity within blocks. Statistical analysis is then used to extract interblock variation, thereby reducing the total error in the experiment.

## Field plot designs

Plant breeders use experimental designs to arrange genotypes in a trial to minimize experimental error. The designs vary according to the purpose of the evaluation, the nature of the genotypes (e.g., segregating or non-segregating), the number of genotypes, the stage of a breeding program (e.g., preliminary or advanced yield trials), and resources.

### Evaluating single plants

#### No design arrangement

Breeders using certain methods may select among segregating plants, starting in the  $F_2$  generation, on a single plant basis. Generally plants are spaced in a completely random arrangement.

#### Advantages

- 1 Inexpensive and easy to conduct.
- 2 Large number of genotypes can be evaluated at any one time.

#### Limitations

- 1 If a large land area is involved, the chance of soil heterogeneity effect increases. Inferior plants growing in fertile soil may outperform superior plants growing in less fertile spots in the field.
- 2 It is suitable for evaluating plants on the basis of traits with high heritability but less effective for evaluating traits with low heritability.

**Modifications** It is helpful to plant rows of standard cultivars in adjacent plots for comparison, to aid in efficient and effective selection of superior genotypes.

#### Grid design

First proposed by C. O. Gardener, the grid design entails subdividing the land into smaller blocks. The rationale is that smaller blocks are likely to be more homogeneous than larger blocks. Plants are selected based on comparison among plants within each block only.

### **Advantages**

- 1 Reduces the effects of soil heterogeneity.
- 2 No precise arrangements are needed, hence conventional plot planters can be used.
- 3 Suitable also for selecting on the basis of plant traits with low heritability.
- 4 Easier and more effective to compare plants within a small group than in a large group.
- 5 Selection intensity can be varied by selecting more than one plant per block.

### *Honeycomb (hexagonal) design*

Proposed by A. Fasoulas, a key feature of this design is a planting arrangement in which each plant is equidistant from others in a hexagonal pattern. Furthermore, the spacing is determined to remove interplant competition. Plants are selected only if they are superior to all in their hexagonal units. The selection intensity can be varied by widening the hexagon. When a plant in the immediate hexagon outyields its surrounding six plants, it is selected by 14.3% selection intensity. If it outyields the 18 plants within the second concentric hexagon, it is selected by 5.3% selection intensity, and so on. When using this method in breeding, the unit of selection at all stages in the breeding program are single plants, not plots.

### **Disadvantages**

- 1 Conventional equipment can not be used for planting.
- 2 It is more complex to conduct.

### **Evaluating multiple plants**

#### *Unreplicated tests*

When conducted, unreplicated tests often entail planting single rows of genotypes with check cultivars strategically located for easy comparison. A breeder may use this design to evaluate a large number of genotypes quickly, in order to eliminate inferior ones prior to more comprehensive field trials.

### **Advantages**

- 1 They save space.
- 2 They are less expensive to conduct.
- 3 Large number of genotypes can be quickly evaluated.

### **Disadvantages**

- 1 Such tests are susceptible to the effects of field heterogeneity. There is no other plot of the genotype in the

test for confirming performance. Poor soil may mask the genotypic potential of a superior genotype by causing it to perform poorly.

- 2 Experimental error estimation is problematic.

#### *Replicated tests*

Replication in field plot technique entails the representation of a particular entry (genotype) multiple times (usually 2–4) in a test. With multiples of each entry, the important and critical design consideration is how to arrange duplicates of genotypes in the field. The statistical concept of randomization requires treatment allocation to be by chance, such that each genotype has an equal chance of being allocated to each available plot. Even though randomization may not be imposed on certain occasions for practical reasons, plant breeders normally use randomization in advanced trials.

Different types of experimental design are used to conduct replicated trials. Designs impose varying degrees of restriction on randomization. A major consideration in plant breeding research is the number of entries to include in an evaluation. As stated elsewhere, plant breeding is a numbers game. The numbers are larger in the early part of the program. Three categories of experimental designs are used in plant breeding.

- 1 **Complete block designs.** These designs are suited to evaluating a small number of entries. Each block contains at least one complete set of entries (genotypes). That is, the number of replications and the number of blocks are the same.
- 2 **Incomplete block designs.** These designs are suited to evaluating a very large number of entries. Under such conditions, complete blocking is impractical because of the large numbers. Instead, each block contains only part of the complete set of entries being evaluated in the study. Hence, the number of replications and the number of blocks are not the same.
- 3 **Partially balanced designs.** These designs are generally complex to use. Some pairs of treatments occur in the same block an equal number of times and hence comparisons among treatments are not equally precise.

#### *Complete block designs: completely randomized design (CRD)*

This design assumes that the entire experimental area is homogeneous, hence there is no need for local control.

**Advantages**

- 1 It is the simplest of the designs to use and analyze.
- 2 It yields the highest number of degrees of freedom for error, making the level of error to be of the relatively lowest magnitude compared to other designs.
- 3 Missing data, when they occur, do not complicate the analysis.

**Disadvantages** It is not conducive for field studies for obvious reasons of soil heterogeneity.

A typical ANOVA table for a CRD for 15 entries and four replications is as follows:

Source	df
Treatment	14
Error	45
Total	59

*Complete block designs: randomized complete block design (RCBD)*

In this design, there are as many blocks as there are replicates. Each replicate of a genotype is represented in each block. This design is suitable for a test involving a small number of entries. It is the most widely used experimental design in plant breeding.

**Advantages**

- 1 It is flexible, being applicable to small as well as moderately large entries.
- 2 It is relatively easy to conduct. The field layout and statistical analysis are relatively straightforward.
- 3 Unbiased error can be effectively estimated.

**Disadvantages**

- 1 It is applicable to evaluations in a homogeneous environment.
- 2 It is not suitable for a large number of entries.

A typical ANOVA for an RCBD for 15 entries and four replications is as follows:

Source	df
Blocks	3
Treatment	14
Error	42
Total	59

*Complete block designs: Latin square*

In Latin square design, blocking is used to control environmental variation in two directions. It is best used

when field variation occurs in two directions, perpendicular to each other.

*Incomplete block designs: lattice design*

The lattice design is a preferred design for evaluating a large number of genotypes.

**Advantages**

- 1 It allows unbiased error to be estimated to determine the origin (genetic or environmental) of observed variation among genotypes.
- 2 It provides more effective comparison among genotypes.

**Disadvantages** Randomization and statistical analysis for design can be challenging (without the use of a computer).

**Materials, equipment, and machinery for field evaluation of genotypes**

Plant breeding, even on a small scale, entails the evaluation of large numbers of genotypes. Modern technology has enabled certain aspects of the field testing process to be mechanized. Some of the materials, equipment, and machinery commonly used in plant breeding field tests are discussed here.

**Materials**

- 1 **Field plan.** A notebook or field record book is prepared to show the layout of plots according to the experimental design. The layout also shows the treatment assignments to plots.
- 2 **Labels/stakes.** Computer software such as MSTAT will allow the plant breeder to select a design, allocate treatment to plots, print labels for seed envelopes, print a field layout, and print record books for the study, among other options. Stakes (wooden, plastic) are prepared to identify the plots.
- 3 **Seed envelopes.** The computer labels are pasted on paper envelopes. Some breeding programs use cloth bags, in which case labels are pasted on card labels and fastened to the sack by means of wires. Breeders may use permanent ink pens to prepare their labels.
- 4 **Seed treatment.** Treatment of the seed with fungicides may be needed. A bucket (or an appropriate container) may be needed for the process.
- 5 **Record books.** Data collection is facilitated by having record books printed according to the field plan and traits to be scored.

- 6 Statistical package for computers.** Data collected from the field are analyzed on the computer using appropriate software.

### Equipment

- 1 Seed counter.** Based on the plot length and spacing, seed packets are prepared containing the appropriate number of seeds for the appropriate number of rows.
- 2 Seed trays/boxes.** To facilitate planting, the seed envelopes are prearranged in order (according to the field layout) and set in trays. The trays are appropriately numbered.
- 3 Plot planters.** Mechanized planting is necessary for a large breeding program. Customized planters may be purchased for specific crops. Often, plot planters designed for various types of crops (e.g., small grains, bean) may be obtained and adjusted appropriately for the grain size and spacing desired. In the absence of tractor-based planting, other smaller motorized planters are available for various crops.
- 4 Harvesting.** Plot combines are available for harvesting, threshing, and bagging of seeds from small plots. The preparation of seed packets (labels, bags, etc.) is required for harvesting.
- 5 Computer/data loggers.** Some breeders may be able to computerize their field data collection to enter data directly into the computer. Where this is not possible, data from field record books are entered into the computer at a later date.

### Crop registration

After the formal release of the cultivar, it may be **registered**. In the USA, this voluntary activity is coordinated by the Crop Science Society of America (CSSA).

#### *Objectives of crop registration*

According to the CSSA, crop registration is designed to inform the scientific community of the attributes and availability of the new genetic material, and to provide readily accessible cultivar names or designations for a given crop. Further, crop registration helps to prevent duplication of cultivar names. Complete guidelines for crop registration may be obtained from the CSSA. Excerpts of these guidelines are discussed in this chapter.

#### *What can be registered?*

Over 50 crops and groups of crops may be registered. Other categories include grasses, legumes, and oilseeds.

Subcommittees have been established to review the registration manuscripts for various crops. Hybrids may not be registered. Eligible materials may be cultivars, parental lines, elite germplasm, genetic stocks, and mapping populations. The cultivar to be registered must have demonstrated its utility, and provide a new variant characteristic (e.g., disease or insect resistance, tolerance to stress). For breeding lines (parental lines), the breeder must include information on sources of cytoplasm and restoration information. Germplasm does not need to be commercially valuable in its present form, but must possess a demonstrable merit (e.g., unique trait, exotic background) that has potential for commercial utility when used in a breeding program. Genetic stocks are primarily used for basic genetic studies. They should be useful and unique. Key mapping populations should have high intrinsic value and utility (e.g., may be used to establish species representative or landmark molecular maps).

#### *Registration procedure*

First, the material to be registered should have been released by the breeder or the organization. A seed sample must be deposited in the National Seed Storage Laboratory (NSSL) prior to or at the time of submission of the manuscript. A registration packet is obtained, completed, and submitted to the appropriate subcommittee responsible for the crop. The packet includes a manuscript to be prepared by the breeder and a signed copy of the NSSL storage application form.

The manuscript should include the following information (see CSSA guidelines for details):

- 1 Name or identification assigned at time of release.
- 2 Scientific name (complete).
- 3 Experimental number or designation during development.
- 4 Names of agencies, organizations, or institutions involved in the development and evaluation; names of those officially releasing the plant materials.
- 5 Brief description of the material (including distinguishing features from like types), breeding procedures, pedigree, and comparative performance data (if applicable).
- 6 Probable regions of adaptation, generations of seed increase, and area of production for cultivars.
- 7 Agency or institution with responsibility for maintenance of the basic stock of these materials.
- 8 Any limitation on availability of the materials.

Mapping populations have additional specific requirements.

The manuscript is required to be prepared using the *ASA Style Manual* and following specific instructions regarding order of topics, spacing, font, and other instructions as customary for publishing in scientific journals.

## Variety protection

In addition to registration, a breeder may seek legal protection of the cultivar in one or several ways as discussed in detail in Chapter 15. A common protection, the Plant Variety Protection, or the Plant Breeders' Rights, is a *sui generi* (of its kind) legal protection, as discussed.

---

## References and suggested reading

- Annicchiarico, P. 2002. Genotype  $\times$  environment interaction. Challenges and opportunities for plant breeding and cultivar recommendation. FAO Plant Production and Protection Paper No. 174. Food and Agriculture Organization, Rome.
- Ceccarelli, S. 1994. Specific adaptation and breeding for marginal conditions. *Euphytica* 77:205–219.
- Cooper, M., and D.E. Byth. 1996. Understanding plant adaptation to achieve systematic applied crop improvement – A fundamental challenge. In: *Plant adaptation and crop improvement* (Cooper M., and G.L. Hammer, eds). CABI, Wallingford, UK.
- Eberhart, S.A., and W.A. Russell. 1966. Stability parameters for comparing varieties *Crop Sci.* 6:36–40.
- Finlay, W., and G.N. Wilkinson. 1963. The analysis of adaptation in a plant breeding programme. *Aust. J. Agric. Res.* 14:742–754.
- Koemel, J.E., Jr., A.C. Guenzi, B.F. Carver, M.E. Payton, G.H. Morgan, and E.L. Smith. 2004. Hybrid and pure line hard winter wheat yield and stability. *Crop Sci.* 44:107–113.
- Tigerstedt, P.M.A. 1994. Adaptation, variation, and selection in marginal areas. *Euphytica* 77:171–174.
- Tollenaar, M., and E.A. Lee. 2002. Yield potential, yield stability and stress tolerance in maize. *Field Crops Res.* 75:161–169.
- Yan, W. and L.A. Hunt. 2001. Interpretation of genotype  $\times$  environment interaction for winter wheat yield in Ontario. *Crop Sci.* 41:19–25.
- Yan, W., and N.A. Tinker. 2005. An integrated biplot analysis system for displaying, interpreting, and exploring genotype  $\times$  environment interaction. *Crop Sci.* 45:1004–1016.

---

## Outcomes assessment

### Part A

Please answer these questions true or false:

- 1 A breeder is required to register his or her cultivar upon release.
- 2 A non-crossover  $G \times E$  interaction indicates a lack of differential performance.
- 3 Hybrids may not be registered.

### Part B

Please answer the following questions:

- 1 Give the organization responsible for crop registration in the USA.
- 2 Give the purpose of a field trial in plant breeding.
- 3 Give the kinds of cross that can be registered.
- 4 Distinguish between the breeder's trial and the official trial.

**Part C**

Please write a brief essay on each of the following topics:

- 1 Discuss  $G \times E$  interactions in plant breeding.
- 2 Discuss the importance of stability analysis in plant breeding.
- 3 Discuss the crop registration process.
- 4 Discuss how plant breeders minimize experimental error in field trials.



## 24

# Seed certification and commercial seed multiplication

---

### Purpose and expected outcomes

*The ultimate goal of a plant breeder is to make available the new cultivar that has been developed to crop producers to be multiplied for use by consumers. This tail end of the plant breeding process constitutes a huge industry with its own customs and regulations. After the breeder has conducted yield trials and selected the genotype for release as a commercial cultivar, there are certain customary steps that are followed by the breeder and other appointed certifying agencies to release the cultivar into the public domain for access by consumers. Farmers can purchase seed that has been certified for use by the appropriate authorities. Consequently, plant breeders should be familiar with this process. After studying this chapter, the student should be able to:*

- 1 Discuss the concept of seed certification.
  - 2 Discuss the official agencies and their roles in the seed certification process.
  - 3 Discuss the commercial classes of seed.
  - 4 Discuss the certification process.
  - 5 Discuss the role of improved seed in agriculture.
- 

### Role of improved seed in agriculture

Seed is critical to the success of the agricultural enterprise. In traditional agriculture, seed is usually an integral part of the farmer's operation. The technology used in such production systems is essentially developed *in situ* – seed is obtained from the crop produced on the farm; additional fertility is obtained from livestock. In modern agriculture, however, seed development and production is a separate enterprise. Farmers are supplied seed by seed companies. Agrochemicals are widely used to provide the needed cultural environment for the high

yield potential of the improved cultivars to be fully expressed for optimal crop productivity.

### Yield gains

The major food crops of the world (wheat, corn, sorghum, rice, soybean) and industrial crops (e.g., cotton) are produced by seed. In these major crops, dramatic increases in yield have been observed over the past 70 years, due to a combination of the role of plant breeding, mechanization, and improved management practices (especially pest control and fertilizers).

In the USA, the most significant increases have been recorded in corn, in which average yield per acre rose from 30 bushels in 1930 to about 70 bushels in 1970 and then to 140 bushels by the mid 1990s (see Appendix 2 for conversion rates of units). The trends in other major crops are similar, albeit not as dramatic. Soybean yield rose fourfold between 1930 and 1998, while wheat expanded by a modest 2.5% increase during the period.

### **Seed market**

The seed market is very lucrative worldwide. The US leads the world in seed market size – about US\$15.7 billion in 1997, representing about a 20% share of the world market. In the same year, China and Japan used \$3.0 and \$2.5 billion of improved seed, respectively. In the US, total seed expenditure rose from \$500 million in 1960 to over \$6.7 billion in 1997. Soybean and wheat led the total amount of seed used in 1997 with 2.06 million and 3.08 million tons, respectively. Seed use is a factor of acreage planted, seeding rate per acre, cropping practice, and variation in agroecological factors. In 1997, seed use in the USA for the major field crops totaled 6.5 million tons.

The USA is a net exporter of seed, attaining an export growth from \$305 million in 1982 to \$698 million in 1996. Major importers of US seed are Mexico, Canada, Italy, Japan, France, the Netherlands, and Argentina, accounting for about 72% of the total US export in 1996. On the other hand, US seed import grew from \$87 million in 1985 to \$314 million in 1996.

### **Regulations in the seed industry**

Plant breeders are particularly at risk of having their creation or invention illegally used by competitors. This is because their inventions (cultivars) or the knowledge that led to them, are readily transported, imitated, or reproduced with little difficulty and at low cost. Once released for sale and planted, it is easy for competitors to have access to the seed. Consequently, there are laws (intellectual property rights) in place to protect seed developers (see Chapter 15).

Apart from regulations that protect the plant breeder, there are regulations that also protect the consumer. The two key avenues for consumer protection are varietal registration and seed certification, as discussed in this chapter. Also, there are regulations that protect the environment, especially when the invention derives from genetic engineering.

## **Role of the private sector in the seed industry**

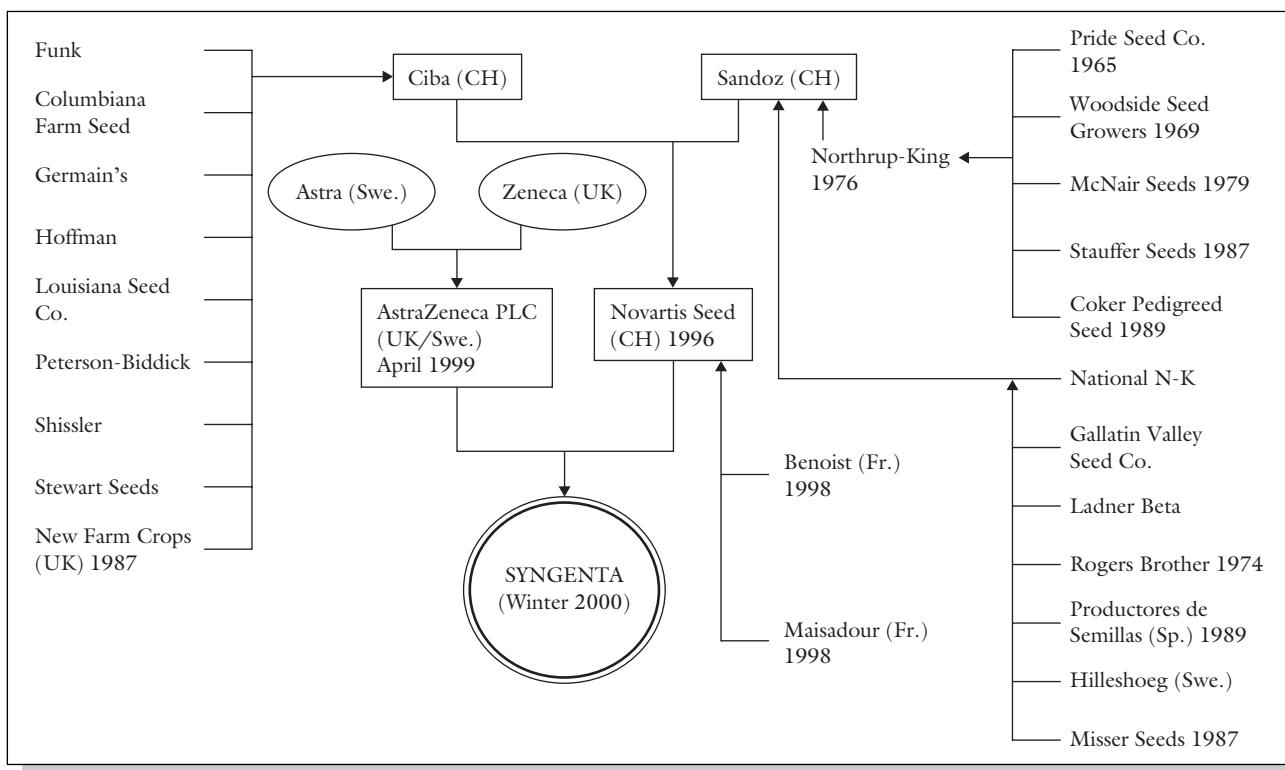
### **Early history**

The first seed company in North America was established by David Landreth in 1784. He published a seed (vegetables) catalog in 1799. In 1976, W. Atlee Burpee established vegetable and flower seed companies. US crop producers practiced saving seed from the previous crop and sharing seed with neighbors for planting in the early 1900s. Even though there were commercial seed producers, commercial seed was not well patronized by farmers until seed certification programs that provided quality assurance to farmers were introduced. These early entrepreneurs were small-scale family-owned enterprises, focusing on multiplying improved cultivars for the public breeding system.

The single most influential crop in the growth of the seed industry was corn. Open-pollinated varieties of corn dominated production until the end of the 19th century. The development of superior-yielding hybrid corn marked the turning point in the shift toward commercial seed by the 1930s. An estimated 150 companies were in operation in the early 1930s, some of them devoting resources to plant breeding in addition to replication of seed in the corn production industry. By 1995, an estimated 95% of corn acreage was planted to commercial hybrid seed. Hybrid corn was the primary business of the seed industry by 1944. The domination of the seed industry by the private sector began in the late 19th century. Research and development was intensified, resulting in the continuous development of higher performing hybrids.

### **Growth of the seed industry**

The Plant Variety Protection Act of 1970 and other amendments gave the impetus for growth in the seed industry. Heretofore, the private industry mostly had no proprietary rights over their inventions (except for the hybrid industry). This situation provided no incentive for commercial seed companies to develop new cultivars. The landscape of the seed industry has experienced repeated alterations over the past 30 years, as mergers and acquisitions eliminated many smaller companies from the scene, and large companies became dominant. A significant fact to note is that many of these acquisitions were made by pharmaceutical and petrochemical companies, many of which were multinational (e.g., Ciba-Geigy, Sandoz, Upjohn, Royal Dutch (Shell)).



**Figure 24.1** Evolution of the Syngenta AG seed company. CH, Switzerland; Fr., France; Sp., Spain; Swe., Sweden; UK, United Kingdom. (After Fernandez-Cornejo 2004.)

However, by the end of the 1980s, many of these chemical and industrial manufacturing companies exited the seed market. By 1989, the industry leaders were Pioneer-HiBred, Sandoz, Asgrow, and Limagrain.

In an attempt for the exiting companies to increase their market shares for high profits, mergers and acquisitions continued in the 1990s. From this reshuffle, Monsanto, Norvartis, and AgrEvo gained significant ground, structuring their businesses around the “life science” image (i.e., they developed chemicals, seeds, foods and food ingredients, and pharmaceuticals based on the application of the principles and practices of biotechnology and genetics). In terms of the US market shares, the leaders (total sales) in 1997 for the major seed crops were Pioneer HiBred (\$1,178 million), Monsanto (\$541 million), Novartis (\$262 million), Dow Agroscience/Mycogen (\$136 million), Golden Harvest (\$93 million), AgrEvo/Cargill (\$93 million), and Delta and Pine Land (\$79 million). The total market share for Pioneer was 33.6%.

The shuffling continues, the most recent being the acquisition of Novartis by Monsanto in 2005, and prior

to that Syngenta (formerly Novartis/Astra Zeneca), Aventis (formerly Hoechst and Rhone Poulenc – later acquired by Bayer), Dupont (incorporated Pioneer), and Monsanto/Pharmacia merged. A diagram showing the evolution of multinational companies engaged in the seed industry is presented in Figure 24.1, as put together by J. Fernandez-Cornejo of the US Department of Agriculture (USDA).

### General operation steps of the seed industry

The specific details vary among crop species. However, four basic elements, as identified by Fernandez-Cornejo, occur in the seed industry from development to marketing of the seed:

- 1 **Seed development.** This stage in the seed industry entails research and development of technologies and the application of science to develop new seed. This is the plant breeding stage where improved seed is

developed. It is very capital intensive and occurs in both the private and public sectors, as previously noted.

- 2 **Seed production.** Once a cultivar with potential has been developed, certain steps are followed to increase the seed to make it available to farmers. These steps are described next.
- 3 **Seed conditioning.** Seed conditioning is the process of readying the seed for market, whereby the certified seed is properly dried, cleaned, sorted, treated (where applicable), and packaged for sale. Tests for quality standards are conducted (see later in this chapter).
- 4 **Seed marketing and distribution.** Seeds are marketed at different levels, including direct marketing by the seed companies, or through licensed distributors. Local distribution may involve farmer dealers, farmers' associations, sales persons of the seed companies, private wholesalers, and retailers.

### Cultivar release process

This is the ultimate goal of a plant breeding program – to release a cultivar with higher yield and superior performance to existing cultivars. After analyzing the data from the field test, taking into account the  $G \times E$  interactions as well as the stability analysis, and considering local issues (e.g., market needs), the breeder will elect to declare one genotype a **cultivar** for release to the agricultural system for use by producers. As previously discussed, this cultivar may be released for specific adaptation or broad adaptation. One of the decisions during cultivar release is naming the new cultivar. Whereas there are no rules for naming cultivars, it is often done thoughtfully and strategically. Some breeders name cultivars in honor of people or special things of meaning to the breeder or the organization. Sometimes, the name may reflect a special attribute (e.g., physical appeal, color, size, nutritional quality, etc.) of the cultivar that producers can readily identify with, or it may be a town or locality in the production region. In the case of small grains in the USA, a cultivar name must be approved by the National Small Grains Authority.

The specific steps of the cultivar release process vary among countries and even among crops in the same country. In addition to natural seed organizations or committees, there may be regional, state, and crop-specific bodies that oversee the process. The breeder (or organization) is required to submit certain specific data to the appropriate review board for review and recommendation for release. Generally, the committees will evaluate the data to ascertain the genetic distinctness,

uniformity, stability of performance, and general agricultural merit.

### Multiplication of pedigree seed

Pedigree seed is seed of a named cultivar that is produced under supervision of a certifying agency for compliance with genetic purity and identity of the original source. After developing, naming, and registering (and securing other legal rights), the breeder makes the cultivar available for commercial multiplication for sale to the public. In introducing a new cultivar into the food production system, the obligation to maintain it in a marketable form so that it can be released at intervals, as needed, in its authentic form for commercial multiplication, is the responsibility of the breeder.

### Classes of seed in a certification system

This is a limited generation system for maintenance of pedigree seed. To maintain the original genetic identity and purity, multiplication of a newly released cultivar is limited to four generations or classes of seed, the main difference among them being the quantity of material generated. These classes are **breeder seed, foundation seed, registered seed, and certified seed**. These categories essentially represent multiplication classes in the ancestry of certified seed. The Organization for Economic Cooperation and Development (OECD) terminology for these classes are **pre-basic, basic, first generation certified, and second or successive-generation certified seed**, respectively.

#### Breeder seed

Breeder seed (or vegetative propagating material) is in the direct control of the plant breeder (or organization) responsible for developing the cultivar. The breeder assumes responsibility for preserving the genetic constitution of the cultivar, being the one uniquely familiar with the true genetic identity of the product. This class of seed is expected to have the highest level of genetic purity of any class of seed. The quantity held in reserve is variable but may range from a few kilograms to several sacks. It is used to establish the next class of seed (the foundation seed). The breeder seed plot (especially for self-pollinating species) comprises progeny rows of seed from individual plants (called **nucleus seed**) from the previous planting. Upon inspection, all deviant rows (possibly from natural outcrossing) are discarded. Seed from rows that are true to type are bulked to form the



## Industry highlights

### *Public release and registration of "Prolina" soybean*

Joe W. Burton

USDA Plant Science Building, 3127 Ligon Street, Raleigh, NC 27607, USA

Release and registration procedures for public soybean varieties will vary somewhat depending on the institution(s) that a soybean breeder is associated with. "Prolina" was developed in a USDA-Agricultural Research Service (ARS) program in cooperation with the North Carolina Agricultural Research Service (NCARS). Thus, approval from both institutions was required to release the "Prolina" soybean to the public for farm production. The procedures for this are outlined below.

- 1 Obtain enough data to demonstrate that the new soybean line is different from other available varieties** and can potentially contribute to soybean production in some geographic area (usually a state or growing region).
- 2 Prepare a case justifying release** that includes data summaries with appropriate test statistics comparing the new line with current varieties.
- 3 Present this case to the Breeders Release Board** of the College of Agriculture and Life Sciences at North Carolina State University. The Board evaluated the case for "Prolina" and approved its release.

#### Registration of 'Prolina' Soybean

"Prolina" soybean [*Glycine max* (L.) Merr.] (Reg. no. CV-393, PI 597389) was developed by the USDA-ARS, in cooperation with the North Carolina Agricultural Research Service. It was released in 1996 to provide a cultivar of Group VI maturity with increased seed protein concentration. Prolina is most adapted to production areas between 33° and 37° N latitude.

Prolina is the bulk of two  $F_8$ -derived lines selected from the first cycle of a recurrent selection population designated NRS4. The population originated from matings of 10 high-protein lines with the cultivars 'Bragg', 'Ransom', and 'Davis' (5,6,7,8). The high-protein parents (470–490 g kg<sup>-1</sup> protein) were  $F_3$  lines from Cycle 7 of Population 1A in a recurrent selection program for increasing seed protein (4). Parents of Population 1A were D55-4110 and N56-4071. 'Ogden' and 'CNS' were parents of D55-4110 (10,11). Maternal grandparents of N56-4071 were 'Volstate' and Ogden (10). The male parent of N56-4071 was a sister line of Lee.

In the initial population development, seven or eight matings of each parental combination resulted in 234  $F_1$  hybrids. These were selfed to produce 234  $S_1$  families. A restricted index was applied to this initial population of  $S_1$  families. The index was designed to increase average yield and maintain the average protein concentration at constant level (8). Modified pedigree selection was applied to  $S_1$  families chosen in the field index selection cycle. Prolina was initially tested as an  $F_6$  line in 1987 under the designation N87-984. Because of heterogeneity for plant height within the line,  $F_9$  lines were derived from N87-984 using single seed descent. These  $F_9$  lines were evaluated in multiple North Carolina locations in 1991. The two lines most desirable in terms of uniformity, protein concentration and seed yield were bulked in 1992. The N87-984 designation was maintained.

Prolina was tested in the Uniform Preliminary VI nursery in 1993 (9) and in the North Carolina Official Variety Trials in 1992, 1993, and 1994 (1,2,3). In the Uniform Preliminary tests (7 locations), Prolina matured 5 d earlier in full-season planting than the check cultivar Centennial. Average seed protein concentration of Prolina was significantly greater than that of Centennial (461 vs. 428 g kg<sup>-1</sup>,  $P < 0.01$ ), and seed oil concentrations of the two were 200 and 198 g kg<sup>-1</sup>, respectively. Average yield of Prolina was 13% less than Centennial. In North Carolina variety trials

(1992–1994), yields of Prolina were equal to those of Centennial. Prolina has yellow seeds with shiny luster, buff hilum, purple flowers, gray pubescence, tan pod walls, and determinate growth habit.

In 1994, breeder seed was provided to North Carolina Foundation Seed, Inc. Seed was distributed to other states by request and according to seed supply. The North Carolina Agricultural Research Service will be responsible for maintaining breeder seed. Small samples (500 seeds) of Prolina can be obtained from the corresponding author for at least five years.

J. W. BURTON,\* T. E. CARTER, JR., R. F. WILSON (12)

#### References and Notes

1. Bowman, D.T. 1992. Measured crop performance. Part IV. Soybean. Dep. of Crop Science, North Carolina State Univ., Raleigh.
2. Bowman, D.T. 1993. Measured crop performance. Part I. Soybean. Dep. of Crop Science, North Carolina State University, Raleigh.
3. Bowman, D.T. 1994. Measured crop performance. Part I. Soybean. Dep. of Crop Science, North Carolina State University, Raleigh.
4. Brim, C.A., and J.W. Burton. 1979. Recurrent selection in soybeans: II. Selection for increased percent protein in seeds. *Crop Sci.* 19:494–498.
5. Brim, C.A., and C. Elledge. 1973. Registration of 'Ransom' soybeans. *Crop Sci.* 13:130.
6. Cavinness, C.E., and H.J. Walters. 1966. Registration of 'Davis' soybeans. *Crop Sci.* 6:502.
7. Hinson, K., and E.E. Hartwig. 1964. 'Bragg' and 'Hardee' soybeans. *Crop Sci.* 4:664.
8. Holbrook, C.C., J.W. Burton, and T.E. Carter, Jr. 1989. Evaluation of recurrent restricted index selection for increasing yield while holding seed protein constant in soybean. *Crop Sci.* 29:324–329.
9. Kenty, M.M., and S.D. Mosley. 1994. Uniform Soybean Tests, Southern States. 1993. USDA-ARS, Stoneville, MS.
10. Morse, W.J., J.L. Carter, and L.F. Williams. 1949. Soybeans: Culture and varieties. USDA Farmer's Bull. 1520 (revised).
11. Weiss, M.G. 1953. Registration of soybean varieties: IV. *Agron. J.* 45:570–571.
12. USDA-ARS, 3127 Ligon St., Raleigh, NC 27607. Registration by CSSA. Accepted 31 July 1998. \*Corresponding author (joe\_burton@ncsu.edu).

Published in *Crop Sci.* 39:294–295 (1999).

**Figure 1** A copy of the registration article submitted to *Crop Science*.



**Figure 2** A copy of the registration certificate for “Prolina” soybean.

- 4 **Prepare a release notice** that will be distributed to other soybean breeders (mainly in the USA). The release notice for “Prolina” was prepared showing that it was a release of the USDA-ARS administrator who has oversight for the release of germplasm. That person reviewed and approved the notice and added their signature to the notice. (Occasionally, the administrator will require some revision of the notice prior to signing it.) The release notice was then sent to the Director of NCARS for his signature.
- 5 **Distribute the approved and signed notice** to all US soybean breeders. The release notice for “Prolina” was sent directly to all public soybean breeders. It was also sent to the secretary of the Commercial Soybean Breeders organization, who distributed it to the other publicity organizations of USDA-ARS and NCARS, who then prepared and distributed press releases.
- 6 **Write a registration article** for the journal of *Crop Science*. The registration manuscript for “Prolina” was prepared after its release following journal guidelines for registration articles. It was submitted to the editor who handles the review and acceptance of soybean germplasm registrations. It was peer reviewed and accepted contingent on some revisions. After revision, the journal accepted the article, gave “Prolina” a registration number, and published it in *Crop Science* (Figure 1).
- 7 **Submit seed samples** to the USDA-ARS soybean germplasm collection in Urbana, Illinois and to the National Seed Storage Laboratory at Fort Collins, Colorado. Seed samples of “Prolina” were sent to both these institutions, as requested.

The Crop Science Society of America then issued a registration certificate for “Prolina” (Figure 2).

breeder seed. The nucleus seed may be maintained as stock seed.

Producers and consumers do not have access to the breeder seed, partly because of its small quantity, but mainly because of laws governing the marketing of seed. The breeder’s seed is to satisfy the official certification process for purity, quality, health, and uniformity so that the producer has access to high quality seed.

#### *Foundation seed*

Foundation seed (or basic seed) is the immediate or first generation increase of the breeder seed. The genetic purity should be very close to the breeder seed. It is used to produce certified seed directly or through registered seed. It is produced under supervision (of the breeder or developer of the cultivar or their representative).

### *Registered seed*

Registered seed is produced directly from breeder seed or from foundation seed. It is the source of certified seed. It is often grown by farmers under contract with a seed company. The material is subject to inspection in the field as well as harvested seed. Farmers may commercially plant this seed.

### *Certified seed*

Certified seed may be produced from foundation, registered, or certified seed. It is grown in isolation under prescribed conditions for the crop such that it meets the genetic identity and purity of the cultivar required for approval by the certifying agency upon both field and seed inspection. This class of seed is that usually available for planting by producers.

### **Maintaining genetic identity of the breeder seed**

#### *Properties of the breeder seed*

The breeder must increase the original seed to have enough to store and supply to growers in the commercial seed multiplication chain. This seed (breeder seed) must be of high purity, quality, health, and uniformity.

- 1 Purity.** The seed must be free from contamination from other cultivars of the same species, as well as other foreign materials (e.g., weed seeds, pollen). The breeder seed is the most authentic and genetically pure seed. To obtain this high level of purity, the seed must be produced in a location distant from all sources of possible contamination and harvested and processed with care (using clean equipment).
- 2 Quality.** High quality seed is one that is fully mature and produced under an optimal environment so that all the traits of the cultivar are optimally expressed. It should be harvested at optimal moisture content.
- 3 Health.** The seed should be free from diseases and damage from insects. This is difficult to achieve, especially if the disease is seed borne.
- 4 Uniformity.** Cultivar uniformity depends on the kind of cultivar that was bred. Clonal cultivars, hybrids, and inbred lines can be uniform to a great extent; synthetics and other open-pollinated cultivars are less uniform.

#### *Causes of loss of genetic purity of seed*

Genetic deterioration of the breeder seed may arise from several sources.

**1 Mechanical admixture.** Failure to clean planting, harvesting, and processing machinery and equipment used in seed production may result in the physical transfer of seeds of one cultivar to another. When the contaminated cultivar is planted, more of the offending seed will be produced to further pollute the genetic purity of the cultivar of interest (unless the offending genotype has distinct morphological features that would enable the breeder to readily rogue them out before harvesting).

**2 Natural outcrossing.** Natural outcrossing by wind or insects may occur where the new cultivar is planted close to another cultivar. As previously indicated, the breeder seed should be planted in isolation to avoid or decrease the chance of natural outcrossing with other cultivars in the vicinity.

**3 Mutations.** Spontaneous mutations may arise during multiplication of seed.

**4 Growing in area of non-adaptation.** Growing the cultivar in an area to which it is not adapted could induce developmental variations.

#### *Prevention of loss of genetic purity*

There are a number of strategies to prevent the loss of genetic purity in breeder seed.

- 1 Plant in isolation.** Planting the breeder seed in isolation reduces the opportunities for outcrossing.
- 2 Rogue out off-types.** All atypical material should be rogued out of the field before harvesting.
- 3 Plant only in areas of crop adaptation.** This reduces the incidence of developmental changes.
- 4 Enforce quality control measures** (e.g., sanitation, cleaning of equipment).

### **Concept of seed certification**

**Seed certification** is a legal mechanism established to ensure that pedigreed seed produced by a plant breeder reaches the public in its highest quality, with its original genetic identity, and with the highest genetic purity. A certified seed is required to be identified by a label that attests to the quality of the product by guaranteeing its genetic constitution, level of germination, and results of seed analysis. Certification is conducted by certifying agencies according to prescribed guidelines for the crop and the Association of Official Seed Certifying Agencies (AOSCA) (most certifying agencies are members of this agency). It includes field inspection and seed inspection of crops grown under prescribed conditions. Following

a request by the AOSCA, the Federal Seed Act of 1970 was enacted to set the minimum standards for all certified seeds produced in the USA.

Usually, seed certification agencies in the USA are organized as state crop improvement associations that operate in collaboration with state agricultural extension stations. The operation often includes agricultural extension services and the state department of agriculture. The expenses of their operation are covered

partly by inspection and certification fees they charge growers.

The specific guidelines vary among crop species and regions of production. For example, the kinds and proportion of weed seed in commercial seed that is tolerable varies from one region to another. Certification services are available for field crops, turf grasses, vegetables, fruits, vegetatively propagated species, woody plants, and forbs.



## Industry highlights

### *Plant variety protection in Canada*

B. Riché, and D. J. Donnelly

Plant Science Department, McGill University, Ste. Anne de Bellevue, QC H9X 3V9, Canada

#### *International Union for the Protection of New Varieties of Plants*

The International Union for the Protection of New Varieties of Plants (UPOV) is an intergovernmental organization that provides for the recognition of plant breeders' rights on an international basis.<sup>1</sup> The acronym UPOV comes from the French name of the organization, Union internationale pour la protection des obtentions végétales. UPOV was established by the International Convention for the Protection of New Varieties of plants (the UPOV Convention) in 1961. There have been subsequent UPOV Conventions (1972, 1978, and 1991). UPOV member countries are bound by the UPOV convention that applied at the time they joined UPOV, or to subsequent conventions that they may have adopted. The different UPOV conventions differ in the minimum scope and duration of protection, farmers' privileges and breeders' exemptions. As of November 9, 2004, 58 countries were signatories to the UPOV Conventions.

The mission of UPOV is to "provide and promote an effective system of plant variety protection, with the aim of encouraging the development of new varieties of plants, for the benefit of society" (UPOV 2004).

The breeding of new plant varieties requires large investments in terms of technology, infrastructure, and human resources. UPOV member countries acknowledge the achievements of breeders of new plant varieties by granting them intellectual property rights called plant breeders' rights. Plant breeders' rights stimulate investment in plant breeding, variety improvement, and commercial propagation of plant material by individuals, private and public breeding organizations, and firms. UPOV also assists countries in the introduction of plant variety protection legislation. The UPOV conventions define the principles of plant variety protection that must be included in the national laws of the member countries. This promotes international harmonization and cooperation between the members.

According to the UPOV conventions, a variety is granted protection if it is new, distinct, uniform, and stable. In addition, the variety must be designated by an acceptable name (naming criteria are described in Article 20 of the 1991 Convention), and the breeder must pay the application and annual fees.

According to the 1991 UPOV Convention, the breeders' rights are granted for a period of not less than 20 years, or for not less than 25 years in the case of trees and vines. The breeders' rights include exclusive control of the protected variety and all aspects of its propagation. Without the authorization of the breeder, no one is allowed to propagate, sell, or market the plant in any way. However, there are two exceptions to the breeders' rights. Firstly, protected varieties may be used by other breeders for breeding and developing new plant varieties. This is called the breeder's exemption. Secondly, farmers may save and use their own seed of protected varieties for the purpose of re-sowing on the same farm (but not for the purpose of selling the seed) without infringing on the breeder's right. This is called the farmer's privilege. It is an optional exception, which means that each member country may decide to include it or not in their legislation.

<sup>1</sup>UPOV, 34, chemin des Colombettes, CH-1211 Genève 20, Switzerland; tel. (+41-22) 338 91 11; fax: (+41-22) 733 03 36; email: upov.mail@wipo.int; internet: <http://www.upov.int>.

### ***Canada's Plant Breeders' Rights Act***

Canada's Plant Breeders' Rights (PBR) Act was brought into force in 1990. Canada's PBR Act adheres to the terms of the 1978 UPOV Convention. Canada signed the 1991 UPOV Convention on March 9, 2002, to signify our intention to ratify. However, to adopt the amended convention, changes must be made to Canada's PBR Act. There has been a PBR Act Amendments consultation, which ended on March 8, 2005. The purpose of the consultation was to assess stakeholder views regarding the following proposed amendments:

- 1 The extension of plant breeders' rights to enclose conditioning, exporting, and importing of propagation material.
- 2 The extension of the minimum period of protection from 15 to 20 years,
- 3 To allow 1 year of sale prior to an application for protection.
- 4 To provide applicants with provisional protection while their application is pending (between the filing of the application and the grant of the right).
- 5 To explicitly state the farmers' exemptions in the PBR Act.

The Canadian Food Inspection Agency (CFIA) is presently reviewing comments and will seek the PBR Advisory Committee's views on the possibility of amending the PBR Act to be in conformity with the 1991 UPOV Convention. The PBR Advisory Committee includes representatives of plant breeders, farmers, growers, industry, and other stakeholders.

According to the 1991 UPOV Convention, a new member of the Union must offer protection to all plant genera and species within 5 years. Since 1998, Canada's PBR Act provides protection to all plant species, excluding algae, bacteria, and fungi.

The Ten Year Review of Canada's PBR Act, published in 2002, concluded that the PBR Act contributed to an increase in investment in plant breeding, research infrastructure, and technologies. In addition, it contributed to an improvement in the mechanism by which to obtain foreign varieties.

### ***Protection inside Canada***

An application for a plant breeder's right must be submitted to the PBR Office (PBRO), which is part of the CFIA. The PBRO examines applications to determine if applicants are eligible to receive a grant of rights. The candidate variety is submitted to comparative tests and trials to determine if it meets the requirements for distinctiveness, uniformity, and stability (DUS). In Canada, these trials are conducted by the breeder/applicant or by someone contracted by the applicant.

To begin the examination process, the applicant must submit a request for a PBR site examination for each candidate variety. An examiner from the PBRO visits the trial site to verify the comparative tests and trials and to see if the candidate variety meets the DUS requirements. Following the site examination, the applicant must submit a completed objective description form, a description of the comparative tests and trials, and comparative photographs that demonstrate the characteristics which distinguish the candidate variety from the reference variety(ies). Reference varieties are varieties of the same species that are the most similar morphologically to the candidate variety and that are grown in Canada at the time of filing the application.

A PBR examiner then prepares a description of the candidate variety and a summary of its distinguishing characteristics. This is published in the *Plant Varieties Journal* and submitted to a 6-month objection period. If, after this period, no objections have been raised by the public, the variety is eligible for a grant of rights.

After the right has been granted, the holder of the right has the responsibility to maintain the protected variety. He must also be able to supply the Commissioner of the PBRO with a sample of propagating material at any time during the term of protection. The Commissioner may inspect the facilities used to maintain the variety. If the holder of the right fails to maintain the variety, is unable to supply propagating material of the variety, or fails to pay the annual renewable fees, the Commissioner may revoke the PBR.

### ***Protection outside Canada***

If a plant breeder's right is granted in Canada, the variety protection is only valid in Canada. To extend variety protection to another member country, the plant breeder must apply separately to the appropriate authority of that country. For example, for protection in the United States, applications must be directed to the United States Patent and Trademark Office (USPTO) in the case of asexually propagated plants (except tubers) or to the Plant Variety Protection Office (PVPO) in the case of sexually reproduced and tuber-propagated plants. In Europe, applications can be directed to the Community Plant Variety Office (CPVO). Through this office, breeders can obtain protection inside all European Union countries with one application.

The addresses and websites of the plant variety protection offices in the UPOV member states can be found on the UPOV website (<http://www.upov.int>). A list of the taxa protected in the UPOV member states was published by UPOV on October 15, 2005, and is also available on the UPOV website.

### ***Naming protected plant varieties***

When applying for a PBR, the applicant must propose a denomination that conforms to the CFIA Variety Naming Guidelines and to the UPOV Recommendations on Variety Denominations. The variety denomination is approved by the PBRO only after its publication in the *Plant Varieties Journal* and an objection period of 6 months.

The plant breeder may also give a trademark name to the new plant variety. A trademark is a word, symbol, or design used to distinguish the wares or services of one person or organization from those of others in the marketplace. The trademark name is followed by the symbol ™.

A trademark may be registered with the Trademarks Office to protect it from misuse and imitation. A registered trademark is one that is entered on the federal government's Trademarks Register (the official listing of registered trademarks) and is followed by the symbol ®. Registration of trademarks is not required but is highly recommended. Registering your trademark with the Trademarks Office gives you the exclusive right to use the trademark across Canada for 15 years. The right is renewable every 15 years thereafter. An application for trademark registration can be filled in online on the Canadian Intellectual Property Office (CIPO) website. The application then goes through a rigorous examination process to ensure that it meets all requirements of the Trademarks Act.

### ***Plant variety protection or plant patents?***

In Canada, higher life forms (animals, plants, seeds, mushrooms) are not patentable. However, transformed plant cell lines and plant cell cultures are patentable. Genetically altered plants derived from these cell lines (such as Roundup Ready® soybean) are registered under the PBR in the regular way.

A method for producing a higher life form may be patentable if this method "requires significant technical intervention by man and is not essentially a natural biological process which occurs according to the laws of nature, for example, traditional plant cross-breeding" (CIPO 2005).

The federal agency responsible for granting patents in Canada is the Patent Office, which is part of the CIPO, an agency of Industry Canada.

In Europe, the European Patent Convention (Section 53) explicitly excludes from patentability "plant or animal varieties or essentially biological processes for the production of plants or animals". However, the European Patent Office interprets this narrowly, and patents have been granted on plants since 1999, provided they do not meet the strict UPOV criteria for plant varieties. For example, plant patents may be granted in Europe if the plants are the product of gene technology.

In the United States, patents are granted to asexually propagated plants (except tubers). Genetically modified plants are patentable. Plant variety protection is granted to sexually reproduced and tuber-propagated plants.

### ***Protected plant variety and plant patent databases***

The CFIA provides on its website a complete and regularly updated list of the plant varieties that are protected in Canada: <http://www.inspection.gc.ca/english/plaveg/pbrpov/pbrpove.shtml>.

For US plant patents, the USPTO provides full-text and full-page image databases of patents issued since 1976 and of published applications: <http://www.uspto.gov/patft/>. For US protected plant varieties, the PVPO posts on its website a regularly updated public version of the certificate status database: <http://www.ams.usda.gov/science/PVPO/CertificatesDB.htm>.

For protected plant varieties in the European Union, the CPVO provides a database of the applications and titles in force: <http://www.cpvo.eu.int/>. The European Patent Office provides a worldwide patent database: <http://ep.espacenet.com/>

### ***References and further reading***

- CIFA. 2002. Date modified: 2002-07-17. The Canadian Food Inspection Agency: Contributing to the quality of Canadian life. <http://www.inspection.gc.ca/english/plaveg/pbrpov/10yre.shtml>.
- CIFA. 2003. Date modified: 2003-10-07. Guidelines for conducting PBR comparative tests and trials. <http://www.inspection.gc.ca/english/plaveg/pbrpov/testsessaise.shtml>.
- CIFA. 2004a. Date modified: 2004-11-19. Proposed amendments to the Plant Breeders' Rights Act to bring existing legislation into conformity with the 1991 UPOV Convention. <http://www.inspection.gc.ca/english/plaveg/pbrpov/ammende.shtml>.
- CIFA. 2004b. Date modified: 2004-01-09. Ten year review of Canada's Plant Breeders' Rights Act. <http://www.inspection.gc.ca/english/plaveg/pbrpov/10yre.shtml>.
- CIFA. 2005. Date modified: 2005-01-13. Guide to Plant Breeders' Rights – a general overview of the PBR system. <http://www.inspection.gc.ca/english/plaveg/pbrpov/guidee.shtml>.
- CIPO. 2004. Date modified: 2004-06-14. A guide to trade-marks. [http://strategis.gc.ca/sc\\_mrksv/cipo/tm/tm\\_gd\\_main-e.html](http://strategis.gc.ca/sc_mrksv/cipo/tm/tm_gd_main-e.html).
- CIPO. 2005. Date modified: 2005-02-18. Manual of Patent Office practice. [http://strategis.ic.gc.ca/sc\\_mrksv/cipo/patents/mopop/chap12-e.html#12.04.01](http://strategis.ic.gc.ca/sc_mrksv/cipo/patents/mopop/chap12-e.html#12.04.01).
- European Patent Office. 1973. European Patent Convention of 5 October 1973. <http://www.european-patent-office.org/legal/epc/e/ma1.html#CVN>.
- UPOV. 2002. The UPOV system of plant variety protection. [http://www.upov.int/en/about/upov\\_system.htm](http://www.upov.int/en/about/upov_system.htm).
- UPOV. 2004. International Union for the Protection of New Varieties of Plants: What it is, what it does. UPOV Publication No. 437(E).

## Seed certification process

There are several key events in the seed certification process that includes paperwork and fieldwork. The specifics may vary among countries.

### *Application for certification*

The developer (breeder, agency, company) wishing to certify a new cultivar must first apply and receive approval from the appropriate authority in the state. In the USA, it is the responsibility of each state to establish the specific protocol for producing each class of seed, as well as the standards for genetic purity. These local standards, however, should not be inferior to those set by AOSCA.

To apply for certification, the applicant is required to supply information including the following: history or origin of the cultivar, documentation of evaluations conducted in comparison with other cultivars, a detailed description of how the classes of seed (breeder, foundation, registered, certified) will be maintained, the number of generations that can be grown from the breeder seed (or number of harvests from one generation in the case of perennial crops), source of seed for planting, and cropping history of the land.

### *Source of seed*

The grower must use a seed class such as foundation or registered seed of an approved cultivar for initiating the certification process. The seed requirement is that the starting seed must produce in the next generation a class of seed that can be verified by a crop certificate (e.g., foundation seed to produce registered seed or registered seed to produce certified seed).

### *Site selection (land)*

A key requirement is that the land should not be a source of contamination from volunteer plants or noxious weeds. To this end, the land must not have grown the crop in the recent (5 years) past (unless the previous crop was the same cultivar and of a properly certified class). Both primary and secondary noxious weeds (especially those whose seeds are hard to separate from the crop seeds) are intolerable. The site should be adequately isolated to exclude contamination. This is especially critical for open-pollinated species.

### *Management in the field*

Once planted, the field must be kept free of weeds. Off-types must be rogued out, preferably before they flower.

### *Field inspection*

While growing in the field, the crop is subject to inspection by authorized personnel from seed certifying agencies. The inspector looks for the presence of noxious weeds, seed-borne diseases, and any factors that might affect the purity of the cultivar.

### *Harvesting and processing*

The equipment for harvesting must be thoroughly cleaned to avoid mechanical mixtures from occurring. The harvested seed is cleaned, conditioned, and bagged. Representative samples are drawn from each lot by officials from the seed certifying agency. Once determined to meet or exceed the minimum standards set by the association (e.g., Crop Improvement Association), the seed is declared officially certified, and receives the official tag. The tag indicates the results of seed analysis (percent germination, inert matter, weed seed, etc.).

## Seed testing

The information mandated on a seed tag that accompanies a certified seed is obtained from laboratory evaluations, collectively called **seed testing**. The Federal Seed Act of 1939 mandated the use of seed tags or labeling. Seed testing provides the information to meet legal standards, determines seed quality, and establishes the rate of sowing for a given stand of seedlings. The USDA is the source of the procedures for seed testing in the US. The **Association of Official Seed Analysts** also publishes additional seed testing procedures for flowers, trees, and shrub species. At the international level, the **International Seed Testing Association** publishes international rules for seed testing.

A seed testing lab conducts tests in five primary categories – germination (viability), purity, vigor, seed health, and noxious weed seed contamination. The first step in seed testing is **seed sampling**. The rules of seed testing provide guidelines for the proper sample size to be submitted for seed testing pertaining to a particular species. Seed testing or **seed analysis** consists of the following tests:

- 1 **Seed germination test (viability)**. Seed viability is determined by conducting the standard germination test to determine the germination percentage (the percent of normal seedlings produced by the pure seed). The common methods of germination tests are the **rolled towel test** (seeds rolled in moist paper towel) or **Petri dish test** (seeds placed on absorbent

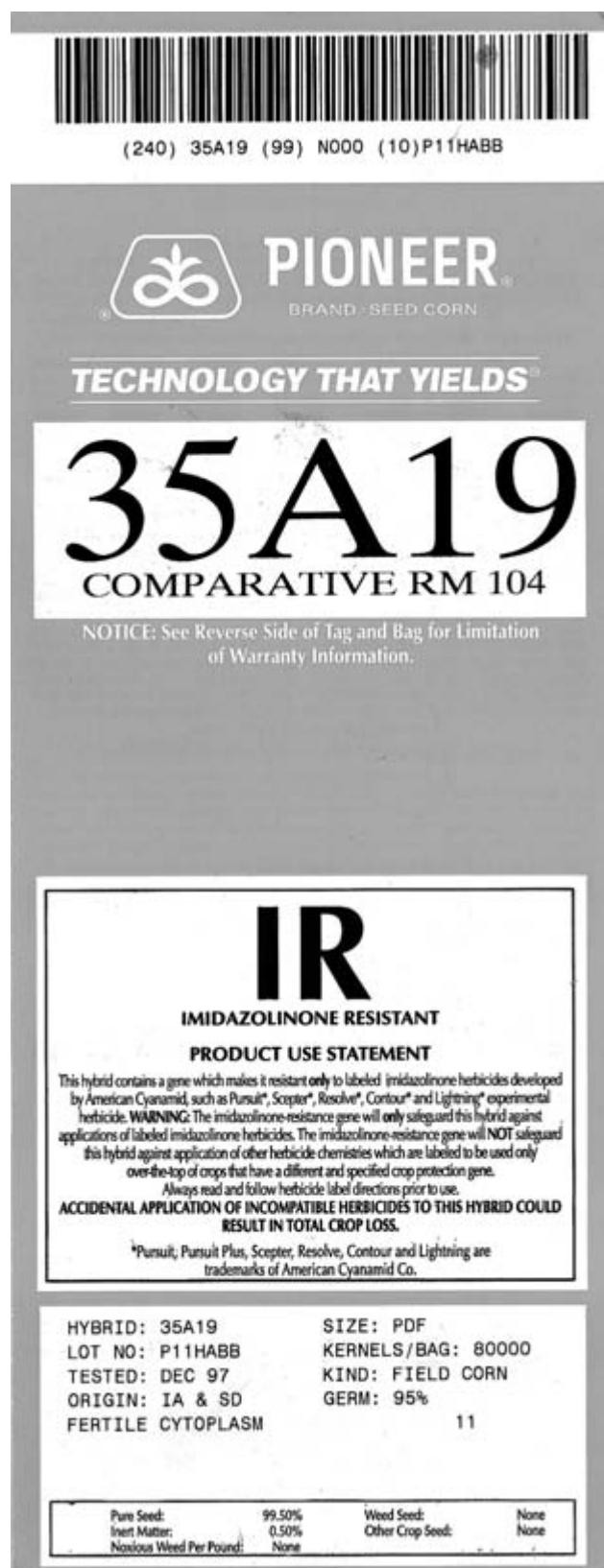
paper in dishes). The rolled towel is kept at 20°C for 16 hours, followed by 30°C for 8 hours. The germinated seedlings are counted. Seed viability may also be determined by a biochemical test, the **tetrazolium test**. Seeds are soaked in 2,3,5-triphenyltetrazolium chloride solution. Living tissue changes color to red while non-living tissue remains uncolored.

- 2 Purity test.** Seed purity is determined at two levels – genetic and physical. The sample seed should reflect the physical features of the cultivar and be distinguished from other seeds and weed seeds. A genetic purity test requires a chemical test that may be as relatively simple as isozyme analysis, or as sophisticated as DNA profiling or fingerprinting (see Chapter 14).
- 3 Vigor testing.** Seed vigor determines the capacity of seed to emerge rapidly and uniformly and develop into normal seedlings under a range of conditions. Common tests of vigor include **accelerated aging** (seeds are subjected to high temperature of between 40 and 45°C, and high humidity, before conducting a germination test), **cold test** (seeds in a rolled towel or containers are held at 10°C for 7 days before moving to a 25°C environment), and **electrical conductivity** of seed leakage.
- 4 Seed health.** The seed sample is examined for the presence of pathogens by visual inspection, seed incubation, or biochemical testing.
- 5 Noxious weed seed.** This is seed from a species that sooner or later becomes aggressive and difficult to control. The weed is officially prohibited from being reintroduced into a production region. A noxious seed in one state may not be classified as such in another state.

### Tagging commercial seed

Once tested, the seed is ready to be made available to consumers, but not until it is properly identified with a tag or label. A tag on the seed bag identifies each class of seed. Customarily, a **white tag** identifies breeder or foundation seed, while a **purple tag** is used to identify registered seed. A certified seed receives a **blue tag**. On some occasions, a **green tag** is used to identify a cultivar that the developer opts not to have certified but nonetheless has achieved the standards of certified seed. Such a cultivar may be simultaneously released by different companies under different brand names (called **branded cultivars**) for marketing purposes. An example of a tag and the information it displays is presented in Figure 24.2 and Table 24.1.

**Figure 24.2** A sample seed tag. (Courtesy of Pioneer Seed Company.)



**Table 24.1** Information on a seed tag.

Item	What it means
Kind and variety	States variety name; term “mixture” is used to refer to seed with more than one component
Lot identification	Identifies a specific amount of seed of uniform quality associated with a specific seed test
Pure seed	Indicates percent purity as it relates to the kind and variety of crop species indicated
Other crop seed	Percent by weight of other crop seed as contaminants
Weed seed	Percent by weight of weed seed present
Inert material	Percent by weight of foreign material such as chaff, stones, cracked seed, soil, etc.
Prohibited noxious weeds	Weed seeds prohibited in the variety to be sold (e.g., field bindweed, Canada thistle)
Restricted noxious weeds	Weed seeds that may be present up to an allowable limit (90 seeds per pound of pure seed), e.g., quackgrass, dodder, and hedge bindweed
Germination	Percent of seed that germinates to produce normal seedlings during a standard seed analysis
Date of test	Year and month in which seed test was conducted
Origin	Source of seed – where seed was produced
Name and address	Name of seed company or seller of the seed

### International role in seed certification

International involvement in seed certification is critical because of the international trade of seeds. In North America, the first official international effort was the International Crop Improvement Association (ICIA) founded in 1919. Its objectives were:

- 1 To assist members in promoting the production, identification, distribution, and use of certified seed and other propagating materials of superior crop varieties.
- 2 To establish minimum standards of seed production, storage, and handling.
- 3 To assist in standardization of certification requirements and procedures to the end that all certified seed will be as good as or better than an accepted, minimum quality.
- 4 To inform the public as to the value of certified seed and encourage its widespread use through approved educational media.
- 5 To develop cooperation with all individuals, groups, and organizations directly or indirectly interested in the improvement of crops.

The ICIA gave birth to the Association of Official Seed Certifying Agencies (AOSCA) to coordinate, standardize, and establish minimum standards for genetic purity and to identify minimum standards for seed quality for all classes of pedigreed seed.

At the international level, the Organization for Economic Cooperation and Development (OECD) was

established in 1966 to facilitate the international movement of certified seed. The OECD has developed five schemes for: (i) oil seed and forage seed; (ii) cereal seed; (iii) vegetable seed; (iv) sugar beet and fodder; and (v) corn. Each participating country is responsible for enforcing and supervising the certification process to conform to international standards and to issue appropriate official documentation to accompany commercial seed.

There are various national and regional seed associations around the world. In Canada, there is the Canadian Seed Trade Association, while the US has the American Seed Trade Association. Within the US, there are several regional groups such as the Western Seedmen's Association and the Southern Seedmen's Association.

### Production of conventional seed

Conventional seed production is often done in regions with drier climates. Drier growing conditions reduce the incidence of various diseases and hence result in higher quality seeds. Furthermore, if the growing season is long, growers can use proper irrigation management to maximize crop yield. In the USA, companies that produce vegetable seeds, flower seeds, and seeds of forage legumes and grass crops are concentrated in the western regions.

The key considerations in commercial seed production are:

- 1 Site (location) selection.** The location should be such that the seed can be produced in isolation to reduce contamination from natural outcrossing. Even though drier climates are desired for seed production, it is best to produce seed in regions of crop adaptation to reduce genetic shifts. The site should be well drained and of good fertility. Cropping history is critical, since a piece of land that previously carried a different cultivar of the crop may produce volunteer plants in the field that could be a source of admixture.
  - 2 Field preparation.** Apart from developing a good seed bed, the key in field preparation is to control weeds to avoid contamination of the harvested commercial seed with seed that reduces the quality of the product.
  - 3 Management.** This includes additional weed control, disease and pest control, fertilization, and irrigation as needed for optimal seed yield and healthy, disease-free products.
  - 4 Harvesting.** The seed should be harvested at the proper stage of maturity for the species and the right moisture content, without being in jeopardy of lodging and shattering, where applicable. It is best to harvest at a moisture content that requires no post-harvest drying. Sorghum and wheat can be safely stored at 13% or less seed moisture, whereas sunflower is best stored at 9.5–13% seed moisture. Where artificial drying is needed, seed may be dried at 40°C for 6–8 hours.
  - 5 Drying and storage.** It is best for seed to dry in the field to the desired moisture content. However, when needed, seed should be dried after harvesting to the level prescribed for the species.
  - 6 Conditioning.** Conditioning is done to remove inert material and weed seeds. Also, seed is sampled and analyzed (seed testing) to provide the data needed for producing a seed tag to accompany the product. Foundation fields require isolation to avoid contamination. Off-types should be rogued out before pollination, as previously indicated. Harvesting should be done carefully to avoid mixing of seed of different cultivars. Harvesting equipment should be thoroughly cleaned to remove any seed from the previous harvesting operation.
- Seed conditioning entails cleaning, sorting (by size), seed treatment with pesticides, and packaging (bagging). Seed is cleaned to remove inert material, weed seeds, broken seed, and other undesirable materials (plant debris). The grain is sorted into size classes where applicable. Fungicides may be used to dress seeds to protect them from soil-borne pathogens.

Clean, treated seed is packaged for the market. When needed, seed is stored in cold dry place (e.g., relative humidity of 50% and temperature of 10°C).

- 7 Packaging.** Clean seed is bagged and tagged for marketing.

Isolation is a key provision in the field production of crop seed, to prevent mechanical admixture of seed, especially when several cultivars of the same species or related species are being simultaneously produced. In the production of seed from self-pollinated species, cultivars only need to be separated by about 3–6 m of uncropped or mowed strips. On the other hand, spacing between different cultivars of cross-pollinated species is usually in excess of 200 m, reaching over 1,000 m in certain cases. Techniques used to reduce mixing from cross-pollination include the use of windbreaks between cultivars, and laying plots such that the prevailing wind effect is minimized. Because insects tend to first visit flowers at the edge of the field, crop fields should be square in shape for insect-pollinated species. Furthermore, the border plants may be discarded at harvesting.

## Production of hybrid seed

Hybrid seed production was discussed in detail in Chapter 18. Most commercial hybrids are single crosses ( $A \times B$ ). The success of commercial hybrid seed production is the availability of adequate foundation seed. Foundation seed is derived from crossing inbred lines. Cytoplasmic male-sterility (CMS) techniques may be incorporated to eliminate the need for emasculation. In corn, artificial emasculation by detasseling may be used. It is critical that adequate pollen be available for maximum seed set. Pollen shed varies from one growing environment to another. Corn producers often use a planting pattern consisting of a ratio such as 1 : 4 of parent row to seed parent rows, or a ratio of 1 : 2 : 1 : 4 of pollinator to seed producing rows. The female : male ratio for sorghum ranges from 3 : 1 to 6 : 1, while sunflower producers use ratios of 2 : 1 to 7 : 1 to optimize seed set. Producing hybrids of insect-pollinated species may require the aid of artificial colonies of bees for effective pollination. Seed set may also be optimized by synchronizing the flowering of the parents in a cross. It is important for the parent lines to be genetically pure to reduce the need for roguing, which can be expensive if the production field is large.

---

## References and suggested reading

- Boswell, V.R. 1961. The importance of seeds. In: Seeds: The yearbook of agriculture, 1961. USDA, Washington, DC.
- Duvick, D.N. 1977. Genetic rate of gain in hybrid maize yield during the past 40 years. *Maydica* 22:187–196.
- Fernandez-Cornejo, J. 2004. The seed industry in US agriculture. Agriculture Information Bulletin No. 786. USDA/ERS, Washington, DC.
- 

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Farmers plant registered seed.
- 2 Seed certification in the USA is conducted by the AOSCA.
- 3 The breeder seed has the highest genetic purity of any class of seed.

### Part B

Please answer the following questions:

- 1 What does the acronym AOSCA stand for?
- 2 Give the class of commercial seed that is released to farmers.
- 3 A white seed tag identifies a certified seed.
- 4 Give two specific pieces of information found on a seed tag.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the commercial classes of pedigree seed.
- 2 Discuss the commercial production of hybrid seed.
- 3 Discuss the international role in seed certification.
- 4 Discuss the role of multinational corporations in the seed industry.



# 25

# International plant breeding efforts

---

## Purpose and expected outcomes

*The purpose of this chapter is to discuss the plant breeding efforts at international institutes and their impact on world food supply. Most of their efforts are directed at developing countries. Modern plant breeding is significantly responsible for the tremendous success of the agriculture of developed economies. It takes tremendous amounts of resources – human and financial – to undertake modern plant breeding research for developing new and improved cultivars for producers. The research infrastructure of most developing countries and the political support available limit the effectiveness of local scientists in addressing crop improvement needs. Because of lack of economic markets in developing countries, the multinational corporations that dominate the commercial seed market in developed countries find it unattractive to invest in the improvement of crops that are of importance primarily to developing countries.*

*Consequently, plant breeding efforts in developing countries depend on philanthropic organizations and the international agricultural centers they aid, for a significant support of their local breeding programs.*

*After studying this chapter, the student should be able to:*

- 1 List all the International Agricultural Research Centers and indicate their mandate crops.
  - 2 Discuss the contributions of the International Agricultural Research Centers to world crop improvement.
  - 3 Discuss the role of the International Agricultural Research Centers in germplasm collection and maintenance.
  - 4 Discuss plant breeding efforts by national programs in developing countries.
- 

### Brief overview of plant breeding in developed countries

The purpose of this very brief overview is to provide a basis for contrasting plant breeding in developed countries with those in developing countries. Plant breeding in developed economies is conducted in both the public and private sectors. In the USA, the land grant university system ensures the training of professional plant breeders, whereas its researchers actively engage in plant breeding research, resulting in the development of new technologies and the development of new plant cultivars. Public sector agricultural research is well funded by the government at both state and national levels, and

primarily has a not-for-profit philosophy. Researchers use both conventional and modern technologies in their research.

Private sector research in plant breeding is significant in most developed economies. It is dominated by multinational corporations and is primarily for profit. Examples of such entities are Monsanto, Novartis, and Du Pont. These entities focus on high value crops (e.g., corn, wheat, rice, soybean) that are grown widely over the world. Unlike the public sector, patent rights protect the inventions of private corporations. Even though patents exist in the public sector as well, access to such protected materials is often much easier than access to those in the private sector. The issue of intellectual

property is a major one in plant breeding in developed countries, with consequences for crop improvement in developing countries.

### Brief overview of plant breeding in developing counties

Whereas agricultural research in developed countries is generally well organized and well funded, formal agricultural research in most developing countries is limited and underfunded. Typically, plant breeding research in these countries is conducted primarily in the public domain at national agricultural research stations and the local universities. Further, the emphasis of breeding is on improving a few of the major food and cash crops of local importance. Many farmers use landraces to grow many of the staple crops. Improved cultivars are obtained through plant introductions and the limited efforts of local research stations.

In terms of human capital, developing countries often depend on developed countries for the training of high level (graduate) researchers (e.g., plant breeders). Plant breeders in developing countries depend largely on conventional breeding technologies, since the exorbitant cost of some of the modern technologies (e.g., genetic engineering) prohibits their adoption. Cultivars are developed so that the farmer can maintain them, that is, seed can be saved for planting the next season's crop from the current season's harvest. Commercial seed companies are non-existent in most developing countries. It should be pointed out that the relatively more technologically advanced third world countries have very well funded plant breeding programs, which have produced outstanding results. Such countries include India, China, Brazil, and South Africa.

### Plant breeding efforts in sub-Saharan Africa

One of the regions of the world that frequently experiences food deficits and famines is sub-Saharan Africa. This region also has some of the most heterogeneous agroecological conditions, coupled with some of the most unstable political systems. Research indicates that some progress has been made in the development of research infrastructure, including the development of human capital and research capabilities. There were about 2,000 full-time equivalent researchers in 1961 and about 9,000 in 1991, with over 90% of them being

Africans. Unfortunately, research expenditure over the same period declined dramatically. In 1991, agricultural research spending averaged about 0.73% of gross domestic product with donor funding for agricultural research accounting for about 43% of total expenditures. It should be pointed out that there is marked variability in these statistics. For example, whereas Nigeria received only 6% of its agricultural research funding from donors, countries such as Senegal and Zambia received more than 60% of their funds from donor sources. An estimated 40% of the budget of the International Agricultural Research Centers (IARCs) is disbursed to efforts in Africa.

The most important crops produced in developing countries by acreage are maize, sorghum/millet, and root and tuber crops. The increase in production of major crops from 1971 to 1997 averaged more than 2.0%. Yield increase from crop improvement accounted for 70% of the increase in wheat production. Compared to Asia and Latin America, African research emphasis is on crop management against the complex agroecological conditions that prevail, rather than the improvement of maize. Input from the IARCs (through CIMMYT and IITA – see Table 25.1 for a list of the centers and their names in full) is very strong for maize. Cultivar releases include open-pollinated varieties (OPV) and hybrids, the latter being more common in East and South Africa, while OPVs and local cultivars dominate production in West and Central Africa. Between 1981 and 1990, cultivars of maize used in these regions included about 62% OPVs and 38% hybrids. The yield increase of hybrids over local varieties averages about 40%. Yield gain from OPVs is about 14–25% over local varieties.

Sorghum and millet are the second and third most important cereal crops, respectively, in Africa. They have the dubious title of "poor man's crops". Crop improvement efforts have been significantly supported by ICRISAT and the International Sorghum and Millet Collaborative Research Support Program (INTSORMIL-CRSP). Other success stories include hybrid sorghum in Sudan, semidwarf rice for irrigated production in West Africa, and disease-resistant potatoes in East and Central Africa.

On the whole, the National Agricultural Research Stations, with support from the IARCs, have made some progress in getting improved cultivars into agricultural production. However, a report in 1991 by the Special Program for African Agricultural Research (SPAAR) suggests that in order for agricultural research to remain a catalyst for modernizing African agriculture,

**Table 25.1** The 16 centers supported by the Consultative Group on International Agricultural Research (CGIAR).

International center	Founded	Location	Key mandate crop/activity
International Rice Research Institute ( <b>IRRI</b> ) ( <i>International Center for the Improvement of Maize and Wheat</i> )	1960 1966	Los Banos, Philippines El Batan, Mexico	Rice Wheat, maize
International Institute of Tropical Agriculture ( <b>IITA</b> ) ( <i>International Center for Tropical Agriculture</i> )	1967 1967	Ibadan, Nigeria Cali, Columbia	Maize, cassava, cowpea, soybean, yam Cassava, beans, tropical forages, rice
Centro International de la Papa ( <b>CIP</b> ) ( <i>International Potato Center</i> )	1970	Lima, Peru	Potato, sweet potato, Andean root and tuber crops
Africa Rice Center ( <b>ARC</b> ) (formerly <i>West African Rice Development Association (WARDA)</i> )	1970	Bouake, Côte d'Ivoire	Rice improvement for Africa
International Crops Research Institute for the Semi-Arid Tropics ( <b>ICRISAT</b> )	1972	Patancheru, India	Groundnut, pearl millet, sorghum, pigeonpea
International Plant Genetic Resources Institute ( <b>IPGRI</b> ) (formerly <i>International Board of Plant Genetic Resources (IBPGR)</i> )	1974	Rome, Italy	Genetic resources, promotes biodiversity
International Center for Agricultural Research in the Dry Areas ( <b>ICARDA</b> )	1975	Aleppo, Syria	Barley, lentils, faba bean, wheat
International Food Policy Institute ( <b>IFPRI</b> )	1975	Washington, DC, USA	Food policy issues
International Service for National Agricultural Research ( <b>ISNAR</b> )	1979	The Hague, Netherlands	Institutional development, policy, dissemination of information
International Water Management Institute ( <b>IWMI</b> ) (formerly <i>International Irrigation Management Institute (IMI)</i> )	1984	Colombo, Sri Lanka	Irrigation in developing countries
Center for International Forestry Research ( <b>CIFOR</b> )	1993	Bogor, Indonesia	Forestry issues in the tropics
International Livestock Research Institute ( <b>ILRI</b> ) (formerly <i>International Livestock Center for Africa (ILCA)</i> , <i>International Laboratory for Research on Animal Diseases (ILRAD)</i> )	1994	Nairobi, Kenya	Livestock production issues in Africa
World Fish Center ( <b>WFC</b> ) (formerly International <i>Center for Living Aquatic Resources Management (ICLARM)</i> )	2002	Penang, Malaysia	Fisheries and other aquatic resources
World Agroforestry Center ( <b>WAC</b> ) (formerly <i>International Center for Research in Agroforestry (ICRAF)</i> )	2002	Nairobi, Kenya	Improved agroforestry systems

important issues to be considered include: size of the National Agricultural Research Stations, commodity research programs, relative emphasis on testing versus breeding of cultivars, allocation of resources to different research activities in the various geographic regions, and low salaries and the consequent high turnover among local scientists.

### **International crop research centers**

The frequent food deficits in developing countries often prompt the international community to intervene in the local food and agricultural production systems of these parts of the world. International involvement in the agriculture of developing countries led to a concerted effort to boost international agricultural research, especially in the tropical regions of the world, where the need is most urgent.

The initial efforts by the Ford Foundation and Rockefeller Foundation led to the establishment of four international agricultural research centers (acronyms explained in Table 25.1):

- 1 CIAT in Columbia, focusing on general tropical agriculture.
- 2 CIMMYT in Mexico, focusing on tropical maize and wheat.
- 3 IITA in Nigeria, focusing on tropical agriculture.
- 4 IRRI in Philippines, focusing on rice.

One of the most dramatic impacts on tropical agriculture, dubbed the Green Revolution, was associated with two of these centers, CIMMYT and IRRI. As discussed elsewhere in this book, the Green Revolution was responsible for increased yields in wheat and rice through breeding of high-yielding and environmentally responsive cultivars of these major world food grains. This outstanding success prompted a discussion in the world community to extend the impact of the international agricultural research centers beyond Asia, which was the major beneficiary of the earlier efforts. Led by the World Bank and supported by the Food and Agricultural Organization (FAO) and United Nations Development Program (UNDP), the Consultative Group on International Agricultural Research (CGIAR) was formed in 1971. The nucleus of 18 member nations has since increased to 58. Similarly, CGIAR centers have increased from the founding four to the current 16 (Table 25.1). Each of these centers is autonomous, with its own charter, international board of trustees, and

staff. In 2001, the centers teamed up with an organization, Future Harvest, to build support for international research. The IARCs have since become known as **Future Harvest Centers**.

### **CGIAR centers and their mission**

The CGIAR centers are internationally recognized and respected for their leadership role in advancing agricultural research and crop productivity in developing countries. Unlike for-profit multinational companies, the CGIAR centers undertake the development of crops that may not be profitable to the private sector, but nonetheless are important to alleviating hunger in poor countries.

### **Structural organization and mission**

The current composition of member nations gives CGIAR a strong north-south identity, with 20 countries from the south and 21 from the north. Its original center-focused, scientific research has shifted to a strategy that is partnership oriented. CGIAR has established partnership committees with non-governmental organizations (NGOs) and the private sector. Funding for its major research activities is in excess of US\$300 million annually. The centers attract highly qualified researchers from around the globe, and include scientists from the region of its immediate influence.

The mission of the CGIAR has been modified over the years to reflect its priority focus on food security and poverty eradication. The current mission statement reads as follows:

To contribute to food security and poverty eradication in developing countries through research, partnership, capacity building, and policy support, promoting sustainable agricultural development based on the environmentally sound management of natural resources.

### **Location and mandate of the CGIAR centers**

The 16 CGIAR centers are located primarily in the tropical regions of the world. Four of these centers are located in Africa, with five in Asia. Each of these centers has mandate plants or animals on which it focuses its research efforts. These mandated areas are briefly summarized in Table 25.2. Some of the centers are located in regions with a diversity of mandate crops.

**Table 25.2** The focus crops for research by the Consultative Group on International Agricultural Research (CGIAR).

Cereals	Food legumes
Rice	Chickpea
Wheat	Cowpea
Maize (corn)	Beans
Barley	Lentil
Sorghum	Pigeonpea
Millet	Soybean
Roots, tubers, banana, plantain	Oil crops
Cassava	Coconut
Potato	Groundnut (peanut)
Sweet potato	
Yam	
Banana, plantain	

### Research emphasis

CGIAR research addresses issues involving plants, animals, soil, water, and policy, as they impact productivity and the management of these natural resources for the benefit of developing countries. The goals of the research programs may be summarized as follows:

- 1 Poverty reduction.** Contributing to food security and poverty eradication in developing countries are paramount goals of CGIAR research. Research is conducted into food crops that are staples of developing countries, to produce high-yielding and environmental stress-resistant cultivars of high nutritive value.
- 2 To improve the management of soil water and nutrients, and the practice of forestry and agroforestry.** The goal of this research emphasis is to minimize the use of agrochemicals and other practices that are environmentally intrusive.
- 3 The promotion of biodiversity.** To this end, CGIAR embarks on the collection and maintenance of plant genetic resources for its mandate plants.
- 4 The application of biotechnology.** The goal is to promote the judicious, safe, and ethical application of appropriate biotechnology for the benefit of developing countries.

### Mandate crops

CGIAR scientists conduct research on cereals, roots, tubers, banana, plantain, food legumes, and oil crops. Other research areas are forestry, agroforestry, fisheries,

livestock, and water management. A brief overview of research in selected crops is given here.

### Wheat

Research in wheat is concentrated at CIMMYT and ICARDA. About 9.4% (US\$27.5 million) of the 2001 commodity investment budget of CGIAR was devoted to wheat research. CIMMYT is considered the world center for the breeding of bread wheat, durum wheat, and triticale. Over 20 million hectares of the world's wheat are grown to cultivars developed at these CGIAR research centers. New wheat growing areas have been established in West Africa and North Africa. Cultivars have been developed with traits such as dwarf stature, disease resistance, efficient water and nutrient use, and tolerance to environmental stress.

### Maize

Research in maize is conducted at CIMMYT and IITA. In 2001, maize research represented about 8.3% (US\$24.4 million) of the CGIAR commodity investment budget. Major traits improved in maize are drought resistance, yield, protein quality, and resistance to maize streak and downy mildew.

### Rice

Rice improvement is conducted at IRRI, ARC, and CIAT. The 2001 commodity investment budget allocation to rice research amounted to US\$48.1 million, representing 16.4% of the budget. The research focus includes improving yield potential, developing hybrids for the tropics, and pest resistance.

### Barley

Barley research is conducted mainly at ICARDA. Over 100 barley cultivars have been developed for use in over 30 countries. Barley research allocation from the 2001 CGIAR commodity investment totaled US\$3.9 million, representing 1.3% of the budget.

### Sorghum

CGIAR spent about 1.8% (US\$5.3 million) of its commodity investment budget on sorghum research, which is conducted at ICRISAT, the world's center for grain sorghum research. The research objectives include developing early maturing cultivars, and disease resistance (e.g., to sorghum midge).

### Soybean

Soybean research is conducted at IITA. Breeding objectives include improved capacity to fix nitrogen without inoculation, high yields, and resistance to shattering. In 2001, CGIAR spent about 0.4% (representing US\$1.3 million) of its commodity investment budget on soybean improvement.

### Potato

Potato research is conducted at CIP, where breeding objectives include disease resistance (e.g., to *Phytophthora infestans*). The budget for the program is about 4.8% (representing US\$14 million) of the CGIAR commodity investment budget.

## Industry highlights

### Plant breeding research at ICRISAT

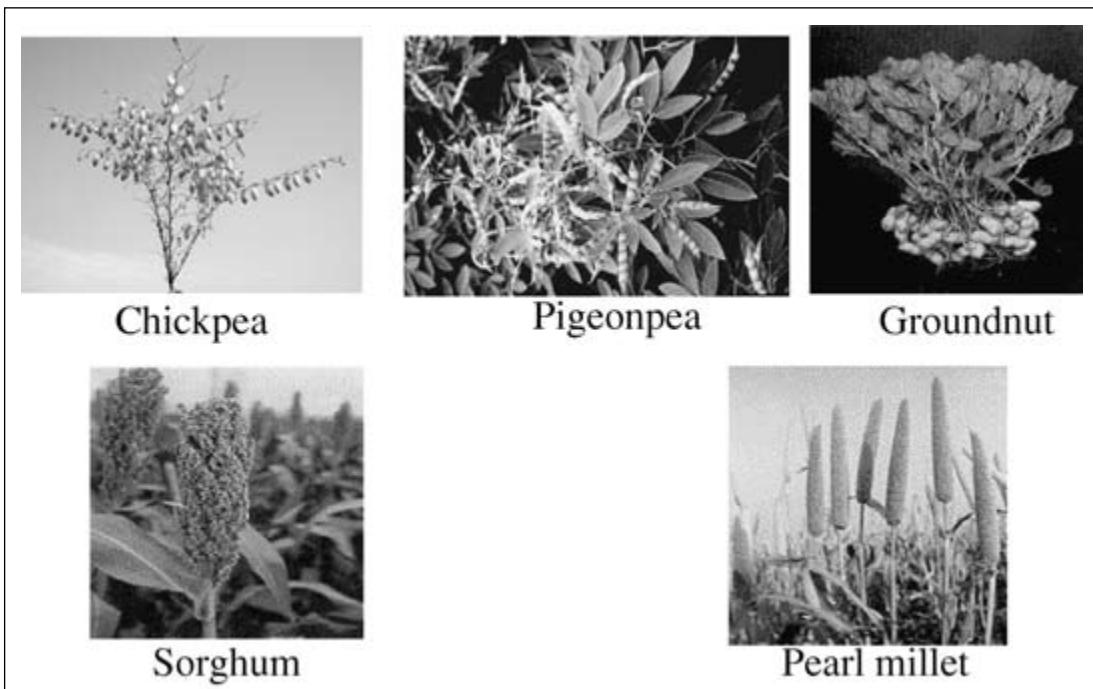
P. M. Gaur, K. B. Saxena, S. N. Nigam, B. V. S. Reddy, K. N. Rai, C. L. L. Gowda, and H. D. Upadhyaya

International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, AP, India

#### Introduction

The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) has a global mandate for the improvement of chickpea (*Cicer arietinum* L.), pigeonpea (*Cajanus cajan* (L.) Millsp.), groundnut (*Arachis hypogaea* L.), sorghum (*Sorghum bicolor* (L.) Moench), and pearl millet (*Pennisetum glaucum* (L.) R. Br.) (Fig. 1). These crops are grown in about 100 million hectares globally, predominantly under rainfed conditions by resource-poor farmers of the semiarid tropics.

ICRISAT has assembled over 104,000 accessions of these crops (17,258 chickpea, 13,632 pigeonpea, 15,419 groundnut, 36,774 sorghum, and 21,594 pearl millet) through donations by various genebanks and national programs and joint explorations. These valuable genetic resources preserved in ICRISAT's genebank at Patancheru, India have contributed significantly in



**Figure 1** The five mandate crops of ICRISAT.

strengthening breeding programs of ICRISAT and the National Agricultural Research Systems (NARS) globally. Close to 1.2 million samples of these crops have so far been distributed to the NARS and ICRISAT scientists.

The crop improvement activities are conducted at ICRISAT's locations in India and Africa, and jointly with many national program scientists globally, wherever the mandate crops are cultivated. In African regions, the development of varieties in all five crops continues to be the primary objective, while in Asia (specifically for India), the present emphasis is towards development of varieties in chickpea, groundnut, and pigeonpea, and hybrids in sorghum, pearl millet, and pigeonpea. Towards this goal, ICRISAT develops segregating materials, populations, advanced breeding lines, and hybrid parents and supplies these to scientists in NARS, non-government organizations (NGOs), and the private sector for evaluation and selection at their locations and utilization in their breeding programs. Based on performance in local, regional, or national trials, varieties/hybrids are released or notified by the various national programs according to their own protocols and procedures.

### **Breeding objectives**

The crop breeding priorities and strategies at ICRISAT have been dynamic, guided by a changing scenario of agriculture and development of new technologies, and are reviewed and revised periodically based on the feedback from NARS scientists, extension personnel, farmers, consumers, and industry. Improved yield potential (mostly for grain, but more recently also for fodder in sorghum, pearl millet, and pigeonpea) is the common and foremost important breeding objective in all crops. The other major objectives include genetic improvement of resistance/tolerance to diseases (*Fusarium* wilt (FW), *Ascochyta* blight, and *Botrytis* gray mold in chickpea; wilt and sterility mosaic in pigeonpea; rust, early and late leaf spots, and rosette and bud necrosis in groundnut; grain mold, anthracnose, and charcoal rot in sorghum; downy mildew in pearl millet), insect pests (*Helicoverpa* pod borer in chickpea; *Helicoverpa* and *Maruca* pod borers in pigeonpea; shoot fly, stem borer, midge, and head bug in sorghum), and abiotic stresses (drought and cold in chickpea; drought in groundnut; drought, salinity, and acidity in sorghum); adaptation (early maturity in chickpea, pigeonpea, and groundnut) and quality of grain (reduction in aflatoxin contamination in groundnut) and fodder (sorghum, pearl millet, pigeonpea).

### **Breeding methods and techniques**

#### **Conventional breeding methods**

**Chickpea** A combination of the pedigree and bulk methods is generally used for selection after hybridization in this highly self-pollinated legume. The early segregating generations ( $F_2$  and  $F_3$ ) are invariably grown in FW-sick nurseries and the surviving plants (resistant to *Fusarium* wilt) are harvested as bulk. The selection of single plants starts from  $F_4$  or later generations. Progeny evaluation is carried out in  $F_5$ – $F_7$  generations. High-yielding and nearly uniform progenies are bulked for replicated yield tests. The backcross method is used only occasionally to incorporate one or few traits from a germplasm line, sometimes a wild species, to a well-adapted variety. Rapid generation advancement in a greenhouse following a single-seed descent (SSD) method is generally used for development of recombinant inbred lines.

**Pigeonpea** Pigeonpea is a partially outcrossed crop (outcrossing up to 30%), and the breeding methods generally recommended for self-pollinated crops are used. Recurrent outcrossing and selfing within landraces has resulted in pigeonpea being heterozygous as well as heterogeneous for important agronomic traits. Such landraces contain a tremendous amount of genetic variability, which has been utilized very effectively to select/breed high-yielding pure-line varieties. Besides this, hybridization and pedigree selection is widely used. The limited natural outcrossing has been successfully exploited for increasing yield and stability through the development of commercial hybrids using genetic male sterility. Currently, cytoplasmic male sterility (CMS) is being used to develop and commercialize hybrids.

**Groundnut** Being a completely self-pollinating crop, the pedigree method is the most commonly used breeding method in groundnut. This allows breeders to practice selection for highly heritable traits such as plant type, pod and seed size and shape, and testa color in early segregating generations. Selection for quantitative traits such as yield and seed composition is made in later generations. SSD and recurrent selection have been used very sparingly. Only limited use of backcrossing has been made, particularly in situations where one of the parents is a primitive landrace or a compatible wild species.

**Sorghum** A trait-based pedigree breeding approach is being used in sorghum, in which the families are used as the selection units for resistance response, and individuals within the resistant families are used as selection units for grain yield. Also, a simple, mass selection-based, recurrent method is being used to improve male-sterile ( $ms_3$  and  $ms_7$  genes) populations to develop trait-based gene pools. Simultaneous testcrossing and backcrossing the selected maintainer plants, along with selection for resistance traits and grain yield, in the trait-based breeding programs is carried out for the purpose of improving male-sterile lines for a specific resistance trait and high grain yield through heterosis.

**Pearl millet** Being a highly cross-pollinated crop (> 85% outcrossing), pearl millet provides an opportunity for exploitation of heterosis. Various forms of recurrent selection have been used to develop open-pollinated varieties (OPVs). The availability of several alternative CMS systems and their restorers has enabled large-scale commercial exploitation of single-cross hybrids in

India. Pedigree breeding has been used in populations developed by recurrent selection, albeit on a limited scale, to develop hybrid parents. Various forms of pedigree breeding have been extensively used in populations derived from hybridization between lines to develop hybrid parents. Backcross breeding has been extensively used in developing partially converted dwarf versions of several composites. Of course, backcrossing remains the only option to develop male-sterile line (A-line) counterparts of maintainer lines (B-lines).

### Marker-assisted breeding

Marker-assisted selection (MAS) is being considered as a potential method to hasten and improve the precision and effectiveness of crop improvement. ICRISAT has established a high-throughput applied genomics laboratory and identified molecular markers for several important traits, such as the stay-green trait and resistance to shoot fly and *Striga* in sorghum, and downy mildew resistance and terminal drought tolerance in pearl millet. Research is underway to identify markers for root mass and resistance to *Ascochyta* blight, *Botrytis* gray mold, and *Helicoverpa* pod borer in chickpea; and FW resistance and fertility-restorer genes in pigeonpea. MAS has been successfully practiced for some traits. For instance, marker-assisted backcross breeding was used to incorporate resistance to downy mildew in the pearl millet single-cross hybrid HHB 67. Marker-assisted breeding for terminal drought tolerance in pearl millet is in progress.

### Transgenics

Transgenics have been developed in pigeonpea and chickpea with resistance to *Helicoverpa* pod borer by using the *Bt Cry1A(b)* gene derived from the bacterium *Bacillus thuringiensis* and the soybean trypsin inhibitor (*SbTI*) gene. Molecular characterization and insect bioassays are in progress. Efforts are also being made to develop transgenics in chickpea for tolerance to abiotic stresses such as drought and low temperatures. Transgenics have been developed in groundnut for several genes such as those encoding for viral coat protein of Indian peanut clump virus (IPCV) and groundnut rosette assistior virus (GRAV), replicase of IPCV, *Bt Cry1A(b)*, and chitinase from rice. In cereals, transgenics have been developed for resistance to stem borer in sorghum and are currently under greenhouse testing.

### Major accomplishments

#### Chickpea

- 1 Short duration varieties (85–100 days at Patancheru, 17.4°N) have been developed that can escape terminal drought and provide wider adaptability to the crop, e.g., ICCV 2 and KAK 2 in *kabuli* type and ICC 37 and JG 11 in *desi* type.
- 2 Super-early *desi* chickpea lines, ICCV 96029 and ICCV 96030, have been developed, which mature in 75–80 days at Patancheru. These lines are being extensively used in crossing programs as source of earliness by NARS in many countries.
- 3 High root biomass has been identified as an important trait for drought avoidance in terminal drought conditions. Lines with a greater degree of drought tolerance (e.g., ICCV 98901–98907) were developed by combining the large root traits of ICC 4958 with the few pinnules trait of ICCV 5680.
- 4 Most chickpea cultivars are susceptible to chilling temperatures at flowering. A number of cold-tolerant lines (e.g., ICCVs 88502, 88503, 88506, 88510, 88516) have been developed that are able to set pod at a low temperature.
- 5 Several varieties with high and stable resistance to FW, the most important root disease of chickpea, have been developed (e.g., ICCV 10, ICCC 37, JG 11).
- 6 Breeding lines with moderate to high levels of resistance to the important foliar diseases *Ascochyta* blight (e.g., ICCV 04512, ICCV 04514, ICCV 04516) and *Botrytis* gray mold (e.g., ICCL 87322, ICCV 88510), have been developed.
- 7 Sources of high-level resistance are not available for pod borer (*Helicoverpa armigera* Hubner), which is the most important pest of chickpea worldwide. Several breeding lines/cultivars have been developed with some level of resistance, e.g., ICCV 7, ICCV 10, ICCL 86102, and ICCL 86103. Further efforts are being made to combine different mechanisms of resistance identified in the cultivated and wild germplasm.

#### Pigeonpea

- 1 Extra-early and early maturing (90–120 days at Patancheru, 17.4°N) photoinsensitive varieties/lines have been developed that made cultivation of pigeonpea possible in a range of environments. Extra-early lines (e.g., ICPL 88039) allow farmers to take two crops (pigeonpea and wheat) in a year.
- 2 FW and sterility mosaic are major pigeonpea diseases. A number of varieties with high resistance to FW and sterility mosaic have been developed and some of these have combined resistance to both diseases (e.g., ICPL 87119).
- 3 *Helicoverpa* and *Maruca* pod borers are the major insect pests. Sources of moderate resistance have been identified and a moderately resistant variety ("Abhaya") has been released.
- 4 The high protein trait was successfully transferred from wild species *Cajanus scarabaeoides*, *C. sericeus*, and *C. albicans* to the cultivated species without sacrificing grain yield or seed size (e.g., HPL 40).
- 5 Commercial hybrids were initially developed using genetic male-sterility systems (e.g., ICPh 8). These provided, on average, 30–35% more yield and greater stability in yield than the pure-line cultivars. Recently, three stable CMS systems have been developed and corresponding fertility restores have been identified to overcome the problems of hybrid seed production associated with genetic male-sterility systems. Efforts are being made to develop commercial hybrids using CMS systems.

### Groundnut

- 1 Several drought-tolerant and high-yielding varieties have been developed, which perform well under rainfed conditions, e.g., ICGS 5, ICGS 44, ICGS 76, and ICG (FDRS) 10 for mid-season drought, and ICGS 11 and ICGS 37 for end-of-season drought in India and ICGV 86021 for end-of-season drought in Indonesia.
- 2 Varieties with high levels of resistance to rust and moderate levels of resistance to late leaf spot have been developed and released by national programs to farmers in Asia (ICG (FDRS) 10 and ICGV 86590 in India) and Africa (ICG (FDRS) 4 in Mali). Resistance to early leaf spot has been introgressed from wild *Arachis* species.
- 3 Varieties with field resistance to peanut bud necrosis disease, a widespread disease in Asia, have been developed (e.g., ICGS 11, ICGS 44, and ICGS 37). These varieties are generally resistant/tolerant to thrips, the vector of the disease. Some genotypes also show tolerance to the virus (ICGV 86031 and ICGV 86029).
- 4 Groundnut rosette disease (GRD) is endemic to Africa and its nearby islands. A short-duration GRD-resistant variety has been developed (ICGV-IS 96894) and released as "Samnut 23" in Nigeria.
- 5 The contamination of groundnut by aflatoxins is a serious problem in most groundnut-producing countries. Genetic variation has been identified for preharvest seed infection, *in vitro* seed colonization, and aflatoxin production. These resistances have been transferred to superior agronomic backgrounds, e.g., ICGV 88145 and ICGV 89104.
- 6 Short-duration varieties are required where the growing season is short, crops suffer end-of-season drought, early frost occurs, and in multiple cropping system. Using the thermal time concept, several short-duration and high-yielding cultivars have been developed (e.g., ICGV 86143 as "BSR 1" in India; ICGV 86015 as "Jayanti" in Nepal; BARD 92 in Pakistan; HL 25 in Vietnam; ICGV 93382 as "Sinpadetha 7" in Myanmar; ICGV 86072 as "BARI Groundnut 5" in Bangladesh; and ICGV 93437 as "Nyanda" in Zimbabwe). It has also been possible to combine early maturity with high-yield potential and tolerance to rust, late leaf spots, and low temperature in ICGV 92267.
- 7 Groundnut genotypes belonging to Spanish types have non-dormant seed, and rains prior to harvest in such genotypes can cause seeds to sprout in the ground, resulting in loss of yield and poor quality of produce. A fresh seed dormancy trait has been successfully introduced in Spanish types from Virginia types, and several short-duration Spanish cultivars with a fresh seed dormancy of 2–3 weeks have been developed, e.g., ICGVs 86155, 86156, 86158, 87378, 87921, and 93470.

### Sorghum

- 1 Several high-yielding varieties have been developed and released in several countries in Africa and Asia for rainfed, drought-prone areas. Some varieties are popular in many countries (e.g., ICSV 112 in Zimbabwe, Kenya, Swaziland, Malawi, and Mozambique; ICSV 111 in Cameroon, Chad, and Nigeria). Some varieties have been bred for dual purposes (grain and stover), e.g., ICSV 112 and ICSV745. A variety NTJ 2 is highly popular for its "roti" making quality.
- 2 Several ICRISAT-bred improved hybrid parents have been extensively used by both public and private sector research organizations to develop and market hybrids in Asia. More than 30 hybrids, based on ICRISAT-bred parents, have been released in India and China. Notable among them are JKSH 22 in India, and Lio Za 4, Longsi 1, Jinza 12, and Gilaza 80 in China.
- 3 *Striga*, an obnoxious obligate parasitic weed is one of the most important biotic yield constraints in Africa, although less important in Asia. Several *Striga*-resistant varieties (e.g., "Framida" in Burkina Faso and Ghana, "SAR 1" in India) and seed parents (e.g., ICSAs 579, 583, 584, 588, 592) have been developed.
- 4 Grain mold is an important disease of sorghum in Asia and Africa. Many grain mold-resistant varieties have been developed. Among them, PVK 801, besides being grain mold resistant is a dual-purpose variety with a good quality stover.
- 5 Shoot fly, stem borer, and midge are the major insect pests of sorghum. Midge-resistant white grain varieties in tan color background (e.g., ICSV 745, PM 13654) have been released in Australia. Several grain mold-resistant (e.g., ICSAs 300, 369, 400, 403, 404) and shoot fly-tolerant (e.g., ICSA 419 and ICSA 435 for rainy season, ICSA 445 and ICSA 452 for post rainy season) CMS-based seed parents have been developed.

### Pearl millet

- 1 About 60 OPVs – 40 in Asia and 20 in African regions – have been developed primarily for grain yield and downy mildew resistance. The most popular OPVs include WC-C75 and ICTP 8203.
- 2 Several hybrids and hybrid parents have been developed with resistance to downy mildew, ergot, and smut. ICMH 451, the first ICRISAT-bred hybrid developed in 1986, covered an area of over 1.0 million ha by the mid-1990s. Later ICRISAT developed and disseminated a wide range of breeding materials and over 90 male-sterile lines for use by NARS and the private sector in hybrid development. Of the 70 pearl millet hybrids released in India, about 60 are based on ICRISAT-bred A-lines or on A-lines developed by the public and private sectors from ICRISAT-bred germplasm.
- 3 Alternative and more stable CMS sources were identified and characterized. The A<sub>1</sub> CMS system was not associated with susceptibility to downy mildew, ergot, and smut. Using the stable A<sub>4</sub> CMS system, it was also shown that it is possible to quickly develop male-sterile populations for use in breeding interpopulation hybrids.
- 4 Topcross, three-way, and interpopulation hybrids forms were identified to have numerous advantages over single-cross hybrids, in terms of seed production economy and reduced vulnerability of downy mildew, ergot, and smut diseases. Topcross hybrids were suggested to be the most efficient route to combine the high-yield potential of improved seed parents and the adaptation of landrace-derived populations.

## Selected accomplishments

The impact of the CGIARs on the agriculture of developing countries includes the following:

- 1 Food production and human nutrition.** Researchers have developed over 300 cultivars of wheat and rice and more than 200 cultivars of maize for farmers. Some of these cultivars, such as quality protein maize (QPM), have augmented the nutritional profiles of major food crops, enhancing the health of consumers.
- 2 Natural resources conservation.** CGIAR-developed technologies have been adopted by farmers resulting in reduced environmental degradation and conservation of water, soil, and biodiversity.
- 3 Reduction in pesticide use.** This is a result of the development of pest-resistant cultivars and improved farming practices for use in developing countries.
- 4 Germplasm conservation.** CGIAR maintains over 6,000 accessions of germplasm, representing over 3,000 species (including crops, forage, and pasture species).
- 5 Capacity building.** CGIAR has trained over 75,000 scientists and technical staff from developing countries.

## Agricultural biotechnology efforts in developing countries

Whenever the subject comes up, the role biotechnology can play in the humid tropics is often identified to be alleviating hunger. Then, there is also the ongoing debate about whether or not developing economies and donor nations and agencies should exploit biotechnology in addressing the food security of developing nations.

### Overview of world food issues

Because of the expected population expansion and increasing land erosion, food security in developing countries is a major concern to the international community. Whereas population growth is leveling off in developed countries, most of the estimated 5 billion additional people on earth by 2030 will inhabit the poor regions of the humid tropics. It should be pointed out immediately that it is an oversimplification to equate hunger alleviation with food security. Associated issues such as effective and efficient distribution networks, effective management of production resources (land, water), and government pricing policy critically impact the success of any food security undertaking. Further,

food security in these economically disadvantaged areas is intertwined inextricably with disease and environmental degradation. Poor soils and poor production management result in low crop yield, malnutrition, and a variety of health issues. Some observe that the medical problems of Africa are inseparable from the lack of food. Needless to say that tackling third world food security is a challenging proposition that requires careful planning and an integrated approach.

Promotion of agricultural biotechnology in developing economies should be accompanied by a promotion of improved agricultural practices. This way, the ecological limits of population growth can be expanded by utilizing existing farmland more productively and also by improving crop harvests. It has been pointed out by some experts that the current agricultural biotechnologies do not increase the productivity *per se* of plants. Instead, they lower pre- and postharvest losses by up to 25%. In terms of strategy, it is suggested that in view of the problems with food distribution and local environments, agriculture in the humid tropics must be indigenous and very productive. Consequently, the infusion of foreign technologies must proceed cautiously. Further, the technology of gene transfer must be developed *in situ*, at least in some of the tropical developing countries, to ensure that it responds to local conditions. This strategy will also ensure that the technology is more readily acceptable to the local government, the local scientific community, plant breeders, and also the local population.

Biotechnology is very capital and knowledge intensive. Such commitments are woefully inadequate in many developing economies. However, because the private sector plays a dominant role in setting biotechnology research and development targets, and because economic returns on investment is critical to investors, crops that benefit developing countries receive little attention because they are of little commercial interest. Consequently, the talk of improving tropical crops falls to other agencies (e.g., the United Nations) where there is little pressure to make profit. The need to infuse biotechnology into the agriculture of developing countries is further underscored by the fact that most of these countries have agricultural-based economies. Improving agriculture is hence a major avenue for improving such economies. The productivity potential of the major food crops of the developing world is far from being attained. Doubling the current level of productivity of staple food crops will make a significant impact on the food security of these nations.

### **Barriers to commercializing agricultural biotechnology in developing countries**

Developing countries are as diverse as they are numerous. The common differences among them are based on political, socioeconomic, and geographic factors. The existing technology capacity for biotechnology in these nations ranges from nil to adequate. The levels of resource (human, capital) endowment also vary widely, and so do domestic conditions regarding politics, government policies, scientific knowledge base, and macroeconomics. It is inaccurate, therefore, to lump all developing nations into one category. Rather, two general groups may be identified: (i) countries in economic transition; and (ii) countries at the bottom, with no or very limited existing infrastructural capacities to exploit biotechnology. Most of the countries in the bottom tier are located in Africa. Even so, some countries in Africa have begun to put in place mechanisms for commercially exploiting biotechnology. These include Kenya, Zimbabwe, Nigeria, and South Africa. In the Caribbean region, Cuba has implemented significant biotechnology programs. There are also the newly industrialized nations in Asia (e.g., China, India) and Latin America (e.g., Brazil, Mexico).

There are several major barriers to commercializing biotechnology in developing countries:

- 1 Lack of appropriate technology.
- 2 Limited infrastructure for exploiting biotechnology.
- 3 Intellectual property rights.
- 4 Biosafety issues.
- 5 Lack of market mechanisms.
- 6 The biotech debate being waged in the potential donor countries.
- 7 Local and regional politics.
- 8 Poverty and disparities.

It might appear that a lack of appropriate technology would be a major barrier to attempts at applying biotechnology to benefit the needy in the developing world. This is so because most of the research and product development occurs in developed countries and is targeted to solving problems in their regions. However, existing technologies can be adapted for use in developing countries, while some new and unique technologies may have to be developed *in situ* in these nations to be effective. The issue of intellectual property rights is also a potential key barrier. Companies in developed countries own most of the patents for the technologies that would be deployed in poor regions. The commercial companies would have to be adequately compensated, in most cases, to allow access to their inventions.

These two factors notwithstanding, some experts believe that the primary barrier to successful exploita-

tion of biotechnology in developing countries is the lack of market mechanisms that normally constitute the driving force behind the research and development process. In terms of agriculture, one of the clearly accessible markets is the seed market, especially those for cash crops. Major seed companies in the USA (e.g., Monsanto) and Europe (e.g., Sandoz) have an interest in accessing this market. If profitable markets exist for biotechnology, companies in industrialized nations with resources will be enticed to invest in the third world oriented projects. However, if the objective of biotechnology exploitation in developing countries is to benefit the poor and needy, then other avenues beside business ventures need to be sought. Developing countries also need to implement biosafety guidelines in order to conform to international regulations for conducting biotechnology research. As the biotechnology debate goes on in developed countries, attempts by developing countries to advance their biotechnology efforts becomes needlessly entangled in the debate and adversely affected. Some opponents of biotechnology tend to think that multinational corporations are only profit-oriented, and look for opportunities to exploit developing countries.

Apart from barriers that may originate outside the developing world, local and regional politics in developing countries pose a significant barrier to the adoption of biotechnology. Local governments are responsible for developing or implementing biosafety regulations, honoring intellectual property rights, supporting local research and development efforts, accepting biotechnology as a viable tool for helping local agriculture, and putting in place the environment for overseas partnerships to be successful. The issue of poverty is important in the adoption of any technology. Most of the agricultural production in developing countries is undertaken by the rural poor. The concern always is how they can afford new technologies. The other critical concern is about the distribution of benefits or the impacts of technology. A criticism of the Green Revolution is that it marginalized the poorer producers, while bringing most of the economic benefits to the already richer producers.

### **Role of international initiatives in agricultural biotechnology**

With proper caution and good planning, biotechnology can be successfully implemented in developing countries to improve agricultural production. It is important that any effort be approached from the angle of partnerships and collaboration. Overseas partnerships should include the public and private sectors, as well as inter-

national entities. More importantly, every partnership should involve the developing countries directly and be implemented in the social context. Including the developing countries makes the technology more readily acceptable and facilitates its adoption. It would also help if the developing countries feel they are not being taken advantage of, or being forced to accept what they do not want.

Because of the prohibitive costs of participation in the exploitation of biotechnology by many developing countries, a variety of international initiatives exist for supporting countries to plan or implement research and extension programs in biotechnology. Currently, most of these efforts are directed towards food-based biotechnology, and involve bilateral initiatives by governments from the developed world, private foundations, tripartite arrangements, and efforts by the United Nations system.

## Summary

- 1 International plant breeding is conducted primarily by the IARCs, which are mainly located in the tropical regions.
- 2 There are 16 such research centers administered by CGIAR.
- 3 The funding for CGIAR comes from member nations and philanthropic foundations such as the Ford and Rockefeller Foundations.
- 4 The research at the CGIAR centers focus on crops that feed the world.
- 5 Their research emphasis is on poverty reduction, promotion of biodiversity, application of biotechnology, and management of natural resources.
- 6 The CGIAR centers are repositories for germplasm for mandate crops.

---

## Reference and suggested reading

Acquaah, G., G. Ude, K. Mataud, and N.J. Tonukari. 2006. Agricultural biotechnology and developing countries: prospects, challenges and impact. In: Floriculture, orna-

mental and plant biotechnology, Vol. IV (Teixeira da Silva, J.A., ed.), pp. 314–327. Global Science Books, Ikenobe, Japan.

---

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 CIAT is located in Mexico.
- 2 IRRI has a mandate for corn research.
- 3 There are 26 CGIAR centers.
- 4 The World Bank is a funding member of CGIAR.

### Part B

Please answer the following questions:

- 1 Give the full name of the following acronyms:  
CIMMYT .....  
IITA .....  
CIAT .....
- 2 Give a specific research emphasis of CGIAR.
- 3 Give three of the mandate crops of CGIAR.
- 4 The ARC is located in .....

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss funding support for CGIAR.
- 2 Discuss the major selected accomplishments of CGIAR.
- 3 Discuss the structural organization and mission of CGIAR.

# 26

## Emerging concepts in plant breeding



### Purpose and expected outcomes

*The methods and techniques of plant breeding change over time as science and technology advance. Whereas new technologies have and continue to be incorporated into plant breeding to facilitate the selection process (leading to the concept of molecular breeding), the classic methods of breeding appear to have remained relatively unchanged in recent times. No new major breeding methods have been developed in recent times, even though many biotechnological techniques have been introduced. Nonetheless, a few approaches are receiving some attention, even though they are not yet mainstream breeding concepts. These are the concepts of decentralized participatory plant breeding and organic breeding. These concepts are being formally introduced in this textbook to enable students to become engaged in the discussion of some of these topics while in training.*

*After completing this chapter, the student should be able to:*

- 1 Explain the concept of decentralized participatory plant breeding.
- 2 Discuss the scientific basis of decentralized participatory plant breeding.
- 3 Discuss the advantages and disadvantages of decentralized participatory plant breeding.
- 4 Discuss the concept of organic plant breeding.
- 5 Discuss the concept of naturalness.
- 6 Discuss the acceptable techniques for organic breeding.

### Concept of centralized plant breeding

The traditional approach to plant breeding is for researchers to independently initiate and conduct a plant breeding program at a specific research station or institute, without input from farmers (**centralized plant breeding**). Then, they evaluate advanced generations of genotypes over selected locations, after which they release a genotype as a new cultivar. At the very best, farmers are occasionally invited to observe finished products during field days. Even on such occasions, their input is limited to choosing among different finished products. Some breeders believe that engaging farmers at some point in the actual breeding process is advantageous. There is a disconnection between the site

of selection (site of breeding) and the target environment (where the product will be used). Consequently, as S. Cicarelli pointed out, the selection efficiency decreases as the selection environment becomes increasingly different from the target environment. In developed countries, crop producers often have facilities to duplicate the favorable selection environments that occur at the research stations. Unfortunately, in developing countries, breeders address the problems of poor farmers, who operate in unfavorable conditions, from their national research stations (and from the International Agricultural Research Centers, IARCs). As previously discussed (see Chapter 23), plant breeders interpret  $G \times E$  interactions in their cultivar release decision-making process. Some researchers such as

Cicarelli contest that most farmers are interested in avoiding or reducing temporal variability, while most breeders focus on avoiding geographic variability. To achieve temporal variability, breeders should develop heterogeneous cultivars rather than uniform ones. This will give yield stability over time.

### **Concept of decentralized participatory plant breeding**

**Decentralized participatory plant breeding** is the concept of actively involving farmers in the plant breeding process, rather than simply delivering a prepackaged seed product to farmers. Farmers are involved in the selection process of the early segregating populations so that the final products are adapted to the target environments in which they will be used for production. This approach to breeding should be distinguished from a multilocation performance evaluation of breeding lines that is part of most conventional cultivar development programs. Also, participation of the farmer should be emphasized in order to make decentralized breeding accomplish its objective of taking advantage of the farmers' knowledge of the crop and the production environment.

The concept of participatory plant breeding is not novel, since farmers, from the time of the invention of agriculture, have selected among existing variability to advance genotypes with useful characteristics. In fact, farmers in developing countries continue this practice by selecting and saving seed from the most appealing plants in the current season's crop for planting the next season's crop. However, the modern application of decentralized participatory plant breeding is attributed to N. W. Simmonds who in 1984 used the term "**decentralized selection**" to refer to the selection process with emphasis on selecting for specific adaptation to specific environments, rather than evaluating the mean performance of genotypes across different environments.

The scientific basis of this approach to breeding rests on  $G \times E$  and its interpretation. In Chapter 23, the concept of  $G \times E$  and its importance in plant breeding was discussed in detail. It was indicated that the type of  $G \times E$  interaction determines the type of cultivar a plant breeder releases for use by farmers. Where evaluations reveal a crossover  $G \times E$  interaction (i.e., the rank in genotypes changes in different environments), the breeder is confronted with a more complex decision regarding the kind of cultivar to release. Some argue that the traditional action by most plant breeders is to

release genotypes with the highest average performance (yield), discarding the best or worst performers at the extremes of the scale. This habit has been described as "negative interpretation of  $G \times E$  interactions", and is motivated by a desire to release a cultivar with wide adaptation for seed production. This approach to plant breeding suits the crop production practices of developed countries where production conditions can be readily manipulated or supplemented to become conducive to optimal plant performance. However, in developing economies, most farmers produce crops under marginal conditions. Consequently, it is important to consider genotypes that perform best under favorable conditions as well as those that perform well under less favorable conditions. In other words, selection should be for specific adaptation, both favorable and less favorable. Such a selection approach is described by some as "positive interpretation of  $G \times E$  interactions".

### **Conventional plant breeding versus decentralized participatory plant breeding**

There are certain key features, with advantages and disadvantages, of conventional and decentralized participatory plant breeding approaches.

#### **Key features of conventional breeding**

The key features of conventional plant breeding may be summarized as follows:

- 1 Breeders formulate breeding objectives and initiate cultivar development at their research facility.
- 2 Promising genotypes in advanced stages of cultivar development are evaluated at selected sites by breeders.
- 3 Superior genotypes that are uniform and have wide adaptation are released through formal channels.
- 4 Farmers may visit on-farm trials on field days at research stations, but are not actively involved in the breeding process.
- 5 Breeders continue to develop superior cultivars to replace older cultivars.

#### **Advantages**

- 1 The process is generally simplified, not having the added management challenges of supervising farmers.
- 2 Usually, only one genotype is released as a cultivar.

### *Disadvantages*

- 1 Genotypes that may perform well under marginal conditions are discarded during the early part of the selection process.
- 2 The cultivar released is adapted to wider regions, rather than a specific region.
- 3 Adoption of the new cultivar by farmers is not guaranteed.
- 4 New cultivars are needed every so often.

### **Key features of decentralized participatory breeding**

The key features of decentralized participatory plant breeding include the following:

- 1 Breeders formulate breeding objectives and initiate the breeding process at the research station.

- 2 Early segregating populations are evaluated in the target environments in which the products would be used for crop production.
- 3 Farmers are involved in the breeding process, contributing their intimate knowledge of the local environment and the crop.
- 4 Cultivars are released based upon specific adaptation to the growing environments.
- 5 Focus is on adaptation over time (i.e., stability).

### *Advantages*

- 1 Cultivars released are readily adopted by farmers and are well adapted to the production environment.
- 2 Variability that may have been discarded in the early stages of conventional breeding may be adapted to specific farmers' fields.



## **Industry highlights**

### *An example of participatory plant breeding: barley at ICARDA*

S. Ceccarelli and S. Grando

International Center for Agricultural Research in the Dry Areas (ICARDA) PO Box 5466, Aleppo, Syria

Decentralized selection, defined as selection in the target environment, has been used by the International Center for Agricultural Research in the Dry Areas (ICARDA) in its barley breeding program to avoid the risk of discarding useful lines because of their relatively poor performance at the research stations. Decentralized selection is a powerful methodology to fit crops to the physical environment. However, crop breeding based on decentralized selection can still miss its objectives if its products do not fit the farmers' specific needs and uses.

Participation of farmers in the initial stages of breeding, when the genetic variability created by the breeders is untapped, will fully exploit the potential gains from breeding for specific adaptation through decentralized selection by adding farmers' perceptions of their own needs and farmers' knowledge of the crop. Therefore, farmers' participation has been the ultimate conceptual consequence of a positive interpretation of genotype  $\times$  environment interactions, i.e., of breeding for specific adaptation.

At ICARDA, the gradual change from centralized non-participatory to decentralized participatory barley breeding was implemented in Syria between 1997 and 2003 in three steps. The model and concepts developed during these developments were gradually applied in Tunisia, Morocco, Eritrea, Yemen, Jordan, and Egypt.

#### ***Step 1: selection phase***

The first was an exploratory step with the main objectives of building human relationships, understanding farmers' preferences, measuring farmers' selection efficiency, developing scoring methodology, and enhancing farmers' skills. The exploratory work included the selection of farmers and sites, and the establishment of one common experiment for all participants. The experiment, described by Ceccarelli et al. (2000, 2003), included 208 plots and was grown in two research stations and nine villages. All possible combinations of selection were conducted, namely centralized non-participatory (breeders on station), centralized participatory (farmers on station), decentralized non-participatory (breeders on farm), and decentralized participatory (farmers on farm). The results indicated that: (i) farmers can handle large number of entries, can take a number of observations during the cropping season, and develop their own scoring methods; (ii) farmers select for specific adaptation; (iii) for some broad attributes, selection is mostly driven by environment; (iv) there is more diversity among farmers' selections in their own fields than among farmers' selections on research stations, and among breeders' selections, irrespective of where the selection was conducted; (v) the selection criteria used by the farmers are nearly the same as those used by the breeders; and (vi) in their own fields, farmers are slightly more efficient than the breeders in identifying the highest yielding entries, although the breeders are more efficient than the farmers in selecting in the research station located in a high rainfall area, but less efficient than the farmers in the research station

located in a low rainfall area. Therefore, the first step indicated that there is much to gain, and nothing to lose, in implementing a decentralized participatory plant breeding (PPB) program.

### **Step 2: methodology**

The second step was mostly about methodologies and consisted of the implementation of the breeding plan, the choice and testing of experimental designs and statistical analysis, the refinement of farmers' selection methodology, and eventually in initiating village-based seed production activities.

From a breeding point of view, the major features of the second phase were: (i) a different role of the two research stations, one of which was not used, while the second, located in an area with more reliable rainfall, was used for seed multiplication; (ii) an increase in the number of farmers involved in the project; and (iii) the initiation of village-based seed production. The details of the second phase, such as number of lines, plot size, type of germplasm, selection criteria, and seed production issues, were discussed in meetings with farmers in each village. The host farmers and a number of neighbors attended these meetings, which were organized by the host farmers. In the case of the type of germplasm, the farmers generally expressed preferences for the seed color (black or white) and the row type. In one village, farmers wanted to test the breeding lines in two different rotations (barley–barley and vetch–barley), and in another village in deep and shallow soil.

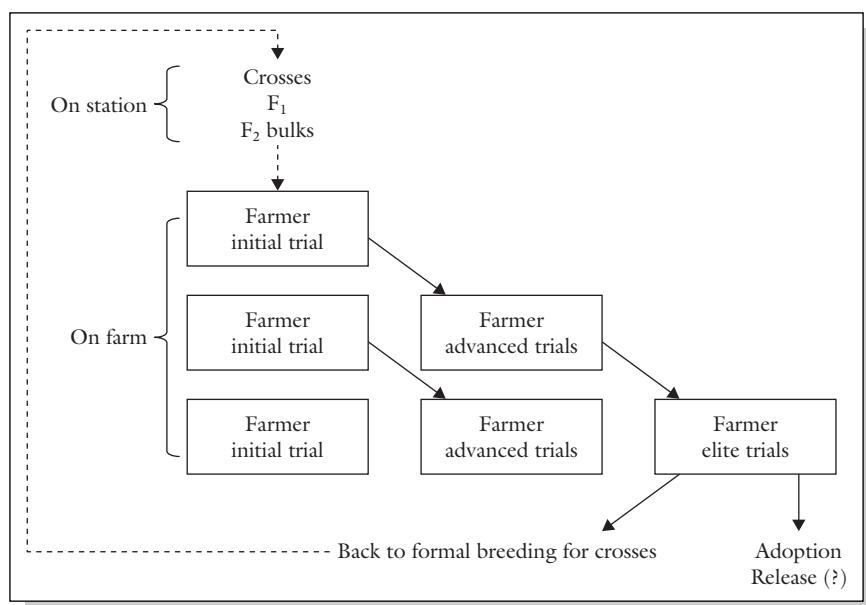
The model of plant breeding we use in Syria and in a number of other countries is a bulk pedigree system, in which the crosses are done on the station, where we also grow the  $F_1$  and the  $F_2$ , while in the farmers' fields we yield test the bulks over a period of 3 years (Figure 1). The testing starts from the  $F_3$  bulks in trials called farmer initial trials (FITs), which are unreplicated trials with 165 entries, five common checks, and 30 systematic check plots (with one or two check cultivars). This allows the evaluation of 165 new breeding materials every year.

In parallel, we conduct on-station pure-line selection by collecting heads within the  $F_3$  bulks selected by the farmers. The  $F_4$  head-rows are promoted to the  $F_5$  screening nursery only if farmers select the corresponding  $F_4$  bulks. The process is repeated in the  $F_5$  and the resulting families, after one generation of increase, return as  $F_7$  in the yield testing phase. Therefore, when the model is fully implemented, the breeding material that is yield tested includes new bulks as well as pure lines extracted from the best bulks of the previous cycle.

The breeding materials selected from the FITs are yield tested for a second year in the farmer advanced trials (FATs), which are grown by between four and eight farmers in each village. Within a village the FATs contain the same entries, but the type and the number of entries and checks varies in each village. The number of FATs in each village depends on how many farmers are willing to grow this type of trial. Each farmer decides the rotation, the soil type, and the amount and time of application of fertilizer. Therefore, the FATs are planted in different conditions and management systems. During selection, farmers exchange information about the agronomic management of the trials, and rely on this information before deciding which lines to select. Therefore, one of the advantages of the program is that the lines start to be characterized for their responses to environmental or agronomic factors at an early stage of the selection process.

The entries selected from the FATs are yield tested for a third year in the farmer elite trials (FETs), which are grown by between four and eight farmers in each village. These entries are also used on station as parents in the crossing program. The three types of trials are planted by scientists using plot drills and are entirely managed by the farmers.

During selection some farmers are assisted by a researcher (Figure 2). Some farmers select at various stages but the majority make their selection when the crop is close to full maturity. Using a scoring method from 0 = discarded to 4 = most desirable, farmers express their opinion on each individual entry.



**Figure 1** The scheme of decentralized participatory barley breeding implemented in Syria.



**Figure 2** Farmers in Eritrea making visual selections assisted by a researcher.

Extension Service. This was to be done in such a way that at the end of the process each province would implement all the various PPB activities within its boundaries, with overall coordination shared between ICARDA and the Ministry of Agriculture. Therefore, one component of the initial step in scaling-up was a training program for the researchers and extension staff on all aspects of PPB.

As a result, the PPB program was extended from five to seven provinces and from 11 to 25 villages (Figure 3) with between 15 and 30 farmers per village. Such a large network of farmers will facilitate the access of non-participating farmers to the products of PPB, and to their large-scale adoption. For this to be possible, village-based seed production will play a key role.

One of the examples of the success that the PPB project is having is offered by the variety "Zanbaka", which about 10 years ago went through the conventional system and was rejected from being released. When it entered the PPB program it began to be slowly adopted, until the drought in 2000 forced the farmers to use all the available seed to feed their sheep. ICARDA then distributed 5 t of seed, which was planted on about 50 ha. Within 2 years the variety has reached 3,500 ha in an area receiving 150–250 mm rainfall and where conventional breeding never had any impact. Similar initial successes have been observed in Egypt, where new barley varieties have out-yielded the local varieties by between 30% and over 100% in four villages, and in Yemen, where two varieties of barley and two of lentil have been adopted by farmers.

### References

- Ceccarelli, S., S. Grando, M. Singh, et al. 2003. A methodological study on participatory barley breeding. II. Response to selection. *Euphytica* 133:185–200.

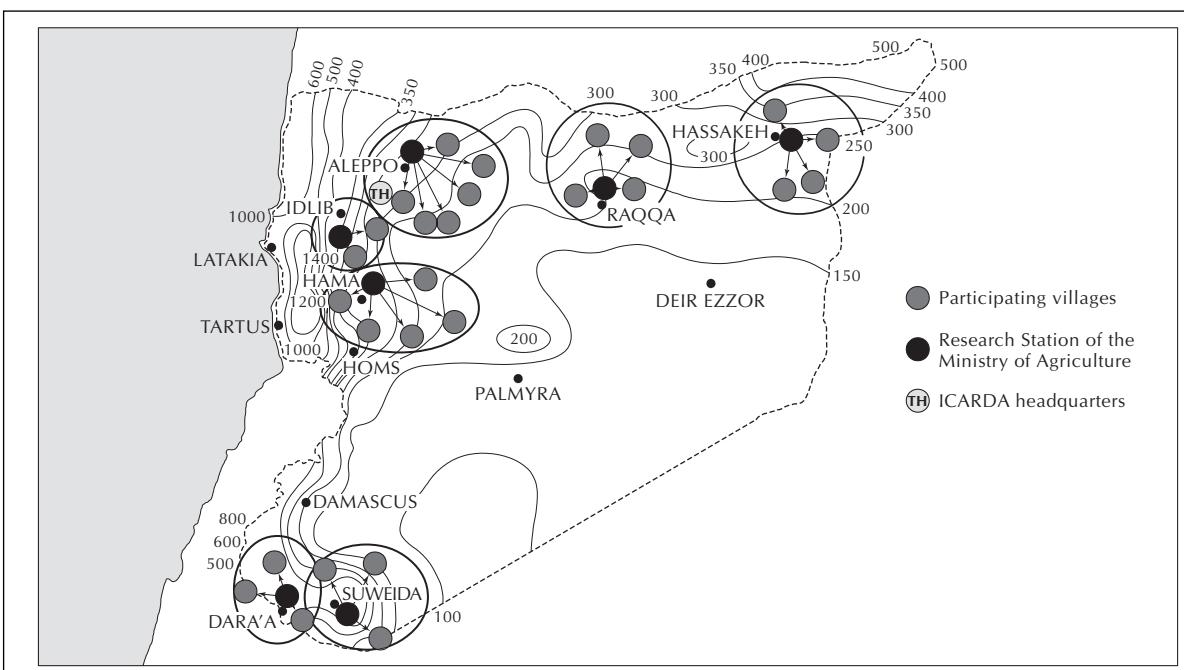
In each trial, the scientists record plant height, spike length, grain yield, total biomass and straw yield, harvest index, and 1,000 kernel weight. On-station scientists record the number of days to heading and days to maturity. The data are subjected to different types of analysis, such as spatial analysis of unreplicated or replicated trials (Singh et al. 2003). The environmental standardized best lineal unbiased predictors (BLUPs) obtained from the spatial analysis are used to analyze genotype  $\times$  environment interactions with the GGE Biplot software (Yan et al. 2000).

One farmer's concern was the seed multiplication of the selected lines. Farmers requested a full control of this operation to avoid mechanical mixture. To address this concern, we established, in four of the eight villages, small seed units consisting of a seed cleaner that also treats the seed with fungicides against seed-borne diseases. The unit has a limited capacity (about 400 kg/h) but allows farmers full control of the seed quality of their selections in the various stages of the breeding program. This is the first step towards the creation of village-based seed production activities.

### Step 3: project extension

We soon recognized that the work described above, will not be able to reach a large number of villages and farmers and hence have an impact at national level. Therefore, in the third phase the emphasis was on institutionalization and scaling-up.

The first step in this direction was the organization of a workshop, with the participation of the farmers, researchers (including heads of research stations of agricultural offices, and research policy-makers), the Seed Organization, the Extension Service, and the Minister of Agriculture. The discussions covered the relationships between PPB, seed production, and variety release. The mechanism agreed upon for scaling-up PPB was a gradual transfer of responsibilities from ICARDA scientists to GCSAR (General Commission for Scientific Agricultural Research) scientists and the staff of the



**Figure 3** The participatory barley breeding scheme in Syria has been extended to seven provinces and 25 villages.

- Ceccarelli, S., S. Grando, R. Tutwiler, et al. 2000. A methodological study on participatory barley breeding. I. Selection phase. *Euphytica* 111:91–104.  
 Singh, M., R.S. Malhotra, S. Ceccarelli, A. Sarker, S. Grando, and W. Erskine. 2003. Spatial variability models to improve dryland field trials. *Exper. Agric.* 39:1–10.  
 Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavnics. 2000. Cultivar evaluation and mega-environment investigation based on the GGE Biplot. *Crop Sci.* 40:597–605.

#### Further reading

- Ceccarelli, S., E. Bailey, S. Grando, and R. Tutwiler. 1997. Decentralized, participatory plant breeding: A link between formal plant breeding and small farmers. In: *New frontiers in participatory research and gender analysis*. Proceedings of an International Seminar on Participatory Research and Gender Analysis for Technology Development, Cali, Colombia, pp. 65–74.  
 Ceccarelli, S., and S. Grando. 2002. Plant breeding with farmers requires testing the assumptions of conventional plant breeding: Lessons from the ICARDA barley program. In: *Farmers, scientists and plant breeding: Integrating knowledge and practice* (Cleveland, D.A., and D. Soleri, eds), pp. 297–332. CABI Publishing, Wallingford, UK.  
 Ceccarelli, S., S. Grando, A. Amri, et al. 2001. Decentralized and participatory plant breeding for marginal environments. In: *Broadening the genetic base of crop production*. (Cooper, H.D., C. Spillane, and T. Hodgink, eds), pp. 115–136. CABI, New York/FAO, Rome/IPRI, Rome.  
 Ceccarelli, S., S. Grando, E. Bailey, et al. 2001. Farmer participation in barley breeding in Syria, Morocco and Tunisia. *Euphytica* 122:521–536.  
 Ceccarelli, S., A. Grando, and F. Capettini. 2001. La participación de los agricultores en el mejoramiento de cebada en el ICAR-DAS. In: *Memorias de la Conferencia Internacional sobre: Futuras Estrategias para Implementar Mejoramiento Participativo en los Cultivos de las Zonas Altas en la Región Andina*, 23–27 de Septiembre del 2001, Quito, Ecuador (D. Daniel, ed.), pp. 25–54.  
 Soleri, D., D.A. Cleveland, S.E. Smith, et al. 2002. Understanding farmers' knowledge as the basis for collaboration with plant breeders: Methodological development and examples from ongoing research in Mexico, Syria, Cuba and Nepal. In: *Farmers, scientists and plant breeding: Integrating knowledge and practice* (Cleveland, D.A., and D. Soleri, eds), pp. 19–60. CABI Publishing, Wallingford, UK.

- 3 The approach favors producers in developing countries where the practice of low input agriculture is prevalent, and farmers use highly adapted landraces.
- 4 Farmers are able to select for traits they actually need that may not be appealing to plant breeders.

#### *Disadvantages*

- 1 Including farmers in the decision-making process may be time-consuming.
- 2 Several cultivars are released simultaneously.
- 3 Replacing cultivars with new ones is not a straightforward operation.

### **New approach to international-national collaborative breeding**

The IARCs devote resources to assisting national agricultural programs to increase agricultural production by developing improved cultivars for farmers. Generally, the international-national collaboration has been, as some have observed, “top-down”, whereby the international programs develop new genotypes that are then evaluated by national programs. From testing these finished or near-finished products, the local breeders recommend superior ones for release as cultivars for local production. This trend of plant “introduction” has the tendency to displace locally adapted cultivars.

In view of the preceding discussion, it is being proposed that early generation breeding material be evaluated by local breeders in developing countries as illustrated by Cecarelli in the box in this chapter. This approach will increase the chance of exploiting positive  $G \times E$  interactions. It will also make breeding programs in developing countries more self-reliant. The needs of small producers will be more effectively met, with the release of cultivars that suit their specific production package.

### **Concept of organic plant breeding**

The concept of organic breeding is relatively new and less developed than the concept of participatory plant breeding. **Organic agriculture** is simply agricultural production without the use of synthetic inputs (e.g., pesticides, fertilizers, herbicides). To achieve this, organic producers apply agroecological principles that promote the self-regulating capacity of the agroecological system (i.e., self-regulation of the soil, plants, and animals).

#### **Issue of “naturalness”**

Organic production is also described as the “natural” way of production (alluding to the absence of synthetic inputs). Some researchers insist that naturalness of organic agriculture should not be limited to the absence of synthetic inputs, or adherence to ecological principles, but also to an acknowledgement of **integrity** and wholeness in the production system. The concept of integrity implies a belonging to a specific natural entity (just like the concept of species that can freely interbreed, but are genetically separated from others belonging to another natural group).

To be organic, this naturalness must be accorded proper consideration in how plants are propagated, cultivated, or genetically manipulated to be in conformity with the ideals of true organic agriculture. Further, plants are ascribed an **intrinsic value** that indicates that they are ethically relevant (in accordance with the attitude society has towards nature).

#### **Need for organic plant breeding**

Organic plant breeding is a relatively new concept of crop improvement. Currently, organic crop producers depend on seed and other propagules that are developed by conventional breeding procedures. If the integrity of an organic production system is to be maintained, the planting material used to initiate production should have organic origin.

There are several reasons why organic breeding is needed to service the organic production industry:

- 1 Genetically modified (GM) crop varieties are not allowed in organic crop production. However, the trends in the production of some crops are to use GM seed or other materials. Efforts should be made to develop non-GM varieties for organic production.
- 2 Crop cultivars suited to organic production are different from those suited to conventional production. Successful cultivars should be adapted to specific soil and fertility conditions, be disease- and insect pest-resistant, and be competitive against weeds. Organic producers rely on the natural fertility of the soil to a large extent, hence the need for cultivars that optimally interact with existing conditions. Crop cultivars with architecture and structure that reduce disease incidence are desirable.
- 3 There is a need to preserve the integrity of plants. Conventional plant breeding methods sometimes violate natural barriers (genetic engineering, wide crosses) and consequently the integrity of plants.

## Principles of organic plant breeding

The integrity approach to organic agriculture does not imply that breeders cannot manipulate plants, but rather that the tools and techniques used in breeding, propagation, and cultivation of plants should not violate this integrity. Integrity of plants pertains to characters such as their nature, wholeness, species-specific characteristics, and their being in balance with species-specific environments.

Four levels of plant integrity have been proposed:

- 1 Integrity of life.** This is defined as the state of wholeness or completeness of a living organism that allows it to perform all of its functions in a more or less autonomous fashion. Consequently, crop cultural practices that introduce synthetic chemicals may interfere with this self-regulating capacity of the plant, and hence be incompatible with organic farming.
- 2 Plant-specific integrity.** This is the state of wholeness or completeness of a plant that allows it to perform all of its plant-specific functions. Plants and animals differ in specific ways at the cellular, whole organism, and functional levels. Growing plants in artificial environments (tissue culture, hydroponics) infringes on the plant's ability to perform its natural functions (natural interaction with the soil). Using techniques that reduce the natural reproductive capacity of plants is unacceptable practice in organic breeding. For example, using cytoplasmic male sterility (CMS) without fertility-restorer genes will cause the progeny from CMS hybrids to be sterile.

**3 Genotypic integrity.** This is defined as the state of wholesomeness or completeness of the species-specific genome. Plant breeding depends on variability for success. Genotypic integrity is not violated as long as the variation is natural in origin. However, genetic engineering technology, which allows the transfer of genes across natural barriers, breaches this integrity principle.

**4 Phenotypic integrity.** This is defined as the state of wholesomeness or completeness of an individual plant, including its health. This principle is violated when plants are developed (or cultivated) in a manner that makes them unable to maintain themselves or complete their life cycle in an organic production system without chemical protection. Chemical mutagenesis as a means of breeding violates this principle simply because chemicals are used in the process.

## Acceptable organic plant breeding techniques

In terms of creating variability, techniques that do not violate the integrity of plants include crossing cultivars, hybrid development with fertile  $F_1$ , testcrosses, backcrosses, and bridge crossing. However, techniques at the cell level (e.g., embryo cultures, somatic variation, ovary culture) and the DNA level (e.g., genetic engineering, protoplast fusion) are not acceptable.

In terms of methods of selection, mass selection, pedigree selection, and even DNA diagnostics and marker-assisted selection are considered compatible with plant integrity. The diagnostic tools are acceptable because they do not cause genetic modification of plants.

## References and suggested reading

- Ceccarelli, S. 1994. Specific adaptation and breeding for marginal conditions. *Euphytica* 77:205–219.
- Ceccarelli, S., and S. Grando. 1991. Environment of selection and type of germplasm in barley breeding for stress conditions. *Euphytica* 57:207–219.
- Heaf, D., and J. Wirt. 2001. Intrinsic value and integrity of plants in the context of genetic engineering. In: Proceedings of the Ifgene Workshop, 9–11 May 2001 (Heaf, D., and J. Wirt, eds), pp. 32–35. Natural Sciences Section, Dornach, Switzerland.
- Lammerts van Bueren, E.T., M. Tiemens-Hulscher, J. Jongerden, J.D. van Mansveld, A.P.M. den Nijs, and G.T.P. Ruivenkamp. 1999. Sustainable organic plant breeding. Louis Bolk Institute, Driebergen, the Netherlands.
- Lammerts van Bueren, E.T., P.C. Struik, M. Tiemens-Hulscher, and E. Jacobsen. 2003. Concepts of intrinsic value and integrity of plants in organic plant breeding and propagation. *Crop Sci.* 43:1922–1929.
- Simmonds, N.W. 1984. Decentralized selection. *Sugar cane* 6:8–10.
- Simmonds, N.W. 1991. Selecting for local adaptation in a plant breeding program. *Theor. Appl. Genet.* 82:363–367.
- Verhoog, H. 1992. The concept of intrinsic value and transgenic animals. *J. Agric. Environ. Ethics* 5(2):147–160.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Backcrossing is an acceptable breeding method in organic breeding.
- 2 Decentralized plant breeding involves farmers, particularly in the final field evaluation stages.
- 3 GM crops are not acceptable in organic crop production.

### Part B

Please answer the following questions:

- 1 What is organic breeding?
- 2 What is decentralized participatory breeding?
- 3 Give two specific advantages of decentralized plant breeding.
- 4 Why are DNA diagnostic tools acceptable in organic breeding?

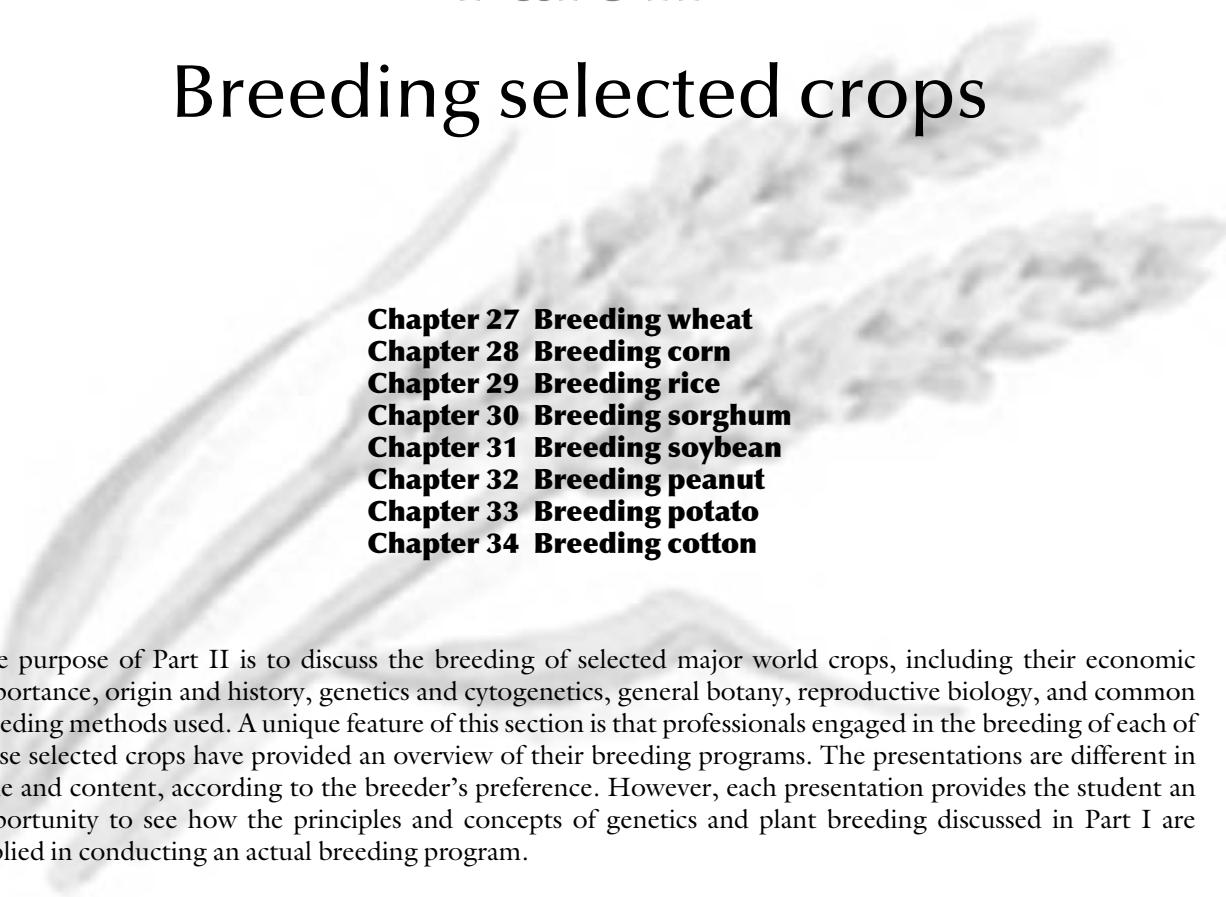
### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the principles of organic plant breeding.
- 2 Discuss the techniques acceptable for use in organic plant breeding.
- 3 Discuss the scientific basis of decentralized plant breeding.

# **Part II**

## **Breeding selected crops**

- 
- Chapter 27 Breeding wheat**
  - Chapter 28 Breeding corn**
  - Chapter 29 Breeding rice**
  - Chapter 30 Breeding sorghum**
  - Chapter 31 Breeding soybean**
  - Chapter 32 Breeding peanut**
  - Chapter 33 Breeding potato**
  - Chapter 34 Breeding cotton**

The purpose of Part II is to discuss the breeding of selected major world crops, including their economic importance, origin and history, genetics and cytogenetics, general botany, reproductive biology, and common breeding methods used. A unique feature of this section is that professionals engaged in the breeding of each of these selected crops have provided an overview of their breeding programs. The presentations are different in style and content, according to the breeder's preference. However, each presentation provides the student an opportunity to see how the principles and concepts of genetics and plant breeding discussed in Part I are applied in conducting an actual breeding program.

---

### **Purpose and expected outcomes**

*The purpose of each of the following chapters is to discuss the breeding of a particular crop. After studying this chapter, the student should be able to:*

- 1 Discuss the economic importance of the crop.
  - 2 Discuss the origin and adaptation of the crop.
  - 3 Discuss the genetics and cytogenetics of the crop that impact its breeding.
  - 4 Describe the germplasm resources for breeding the crop.
  - 5 Discuss the general botany and reproductive biology of the crop.
  - 6 Present the common breeding methods used in breeding the crop.
-

# Breeding wheat



## Taxonomy

Kingdom	Plantae
Subkingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Liliopsida
Subclass	Commelinidae
Order	Cyperales
Family	Poaceae
Genus	<i>Triticum</i> L.
Species	<i>Triticum aestivum</i> L.

## Economic importance

Wheat is the most important cereal grain crop in the world. It is the principal cereal grain crop used for food consumption in the USA and most parts of the world. In the US, it usually ranks fourth after corn, hay, and soybean, in that order of importance. Wheat is grown commercially in nearly every state in the US, with a concentration of production in the Great Plains, an area spanning the states from Texas to Montana. US Department of Agriculture (USDA) production trends indicate that in 1866 wheat was harvested from an area of 15.4 million acres, yielding an average of 11 bushels per acre (see Appendix 2 for conversion rates of units). By 1950, production occurred on 61.6 million acres, with an average yield of 16.5 bushels/acre. In 1990, the acreage was 69.2 million acres, with a yield average of 39.5 bushels/acre. The Central and Southern Plains (Texas, Oklahoma, and Kansas, among others) produces more than the Northern Plains (e.g., Montana, North Dakota), with the two regions accounting for two-thirds of US wheat production and about 80% of wheat acreage. Kansas leads all states in wheat production. Hard white wheat is also grown in the Central and Northern Plains.

World production of wheat in 2001 was 583.9 million metric tons, occurring on 219.5 million acres. World wheat consumption in that period was 590.6 million metric tons. Developing countries (excluding those in Eastern Europe or the former Soviet Union) account for nearly 50% of the world's wheat production, the leading producers being China, India, Turkey, Pakistan, and Argentina. The success of wheat production in these countries is credited to the impact of the Green Revolution that occurred in the 1960s and 1970s. In 2000, China produced 111.9 million metric tons while India produced 26.5 metric tons. Latin America and Asia (excluding China and India) each produce about 20 million metric tons a year. Wheat is produced in Europe, including the United Kingdom, Denmark, the Netherlands, Belgium, Switzerland, and West Germany.

## Origin and history

The origin of wheat is believed to be southwestern Asia. A cross between wild emmer wheat (*Triticum dicoccoides*) and *Aegilops squarrosa*, a grass, produced a spelt-like plant. This suggests that the common or bread wheat (*T. aestivum*) is descended from a cross between spelt and the progenitor of Persian wheat (*T. persicum*). Persian wheat occurs in the wild in the Russian Caucasus. The Persian wheat probably is descended from the wheat of the Neolithic Swiss lake dwellers, which in turn might have originated from a cross between einkorn and a grass, *Agropyron triticum*. Archeological findings indicate that emmer wheat was cultivated before 7000 BC. Similarly, wheat was cultivated in Europe in prehistoric times. In the US, wheat was first cultivated along the Atlantic coast in the early 17th century, moving westwards as the country was settled.

## Adaptation

Wheat is best adapted to cool temperate climates where rainfall is not excessive (40–60 cm per annum). Based on season of production, there are two types of wheat – winter wheat and spring wheat.

### *Winter wheat*

Winter wheat is sown in the fall so that it can have some growth before the onset of cold weather in winter. Growth ceases and the plants remain dormant through winter, resuming growth in spring for harvesting in summer. About two-thirds of US wheat is winter wheat. Winter wheat can survive cold temperatures as low as –40°C if protected by snow.

### *Spring wheat*

Spring wheat is planted in early spring and harvested in July–August. Spring wheat is less tolerant of low temperatures and is damaged by even a light frost of –2 to –1°C.

Wheat is a long-day plant. Short days of high temperatures stimulate tillering and leaf formation but delay flowering of wheat plants. Early maturing cultivars are available for production under any photoperiod conditions. However, the quality (nutritional uses such as baking) of wheat is influenced by the production environment. For example, growing hard wheats in soft wheat regions results in grains that are starchy or “yellow berry” (soft and starchy).

## History of breeding in the USA

Wheat is one of a few food crops (the others being corn and rice) that have been associated with the Nobel Peace Prize. The 1970 Nobel Peace Price awarded to Norman Borlaug, the father of the Green Revolution, recognized his contribution to agricultural productivity through the introduction of superior genotypes of wheat. These superior varieties were high yielding, shorter (semidwarf wheat), more lodging resistant, and responsive to high levels of fertilizer. A significant contributor to this effort was Orville Vogel, a USDA wheat breeder stationed at Washington State University. Under his leadership, the first successful commercial semidwarf wheat variety in the Western Hemisphere was released to farmers in 1961. This variety, “Gaines”, was a soft white wheat and yielded in excess of 100 bushels/acre under both dry-

land and irrigated production. In spite of its agronomic qualities, “Gaines” had milling quality problems. In response to the demands of the milling industry, a new selection with more desirable milling qualities, called “Nugaines”, was released in 1965. The role of wheat in the Green Revolution was discussed in Chapter 1.

## Commercial wheat classes

Wheat breeders specialize in one of the special market classes of wheat. There is a genetic basis for this classification. Wheat may be classified into seven groups based on time of year they are planted, and kernel characteristics (hardness, color, shape). However, for commercial production, the varieties may be narrowed down to six basic classes: hard red winter, soft red winter, hard red spring, hard white, soft white, and durum wheat. The hard red wheat accounts for about 40% of total US wheat production and is the dominant class in US wheat export.

### *Hard red winter wheat*

This is grown mainly in the Great Plains (Kansas, Oklahoma, Nebraska, Texas, Colorado). It is also grown in the former Soviet Union, Argentina, and Danube valley of Europe. It is used for bread flour.

### *Hard red spring wheat*

This class of wheat is grown in regions with severe winters in the north central states (North Dakota, Montana, South Dakota, Minnesota). It is also produced in Canada, Russia, and Poland. It is the standard wheat for bread flour.

### *Soft red winter wheat*

This class of wheat is grown predominantly in the eastern USA (Ohio, Missouri, Indiana, Illinois, Pennsylvania). It is also grown in Western Europe. Soft red winter wheats are used mainly for pastry, cake, biscuit, and household flour. For bread-making, it needs to be blended with hard red wheat flour.

### *White wheat*

White wheat (hard or soft) is produced in the four western states and in the northeastern states (Washington, Oregon, Michigan, California, New York). Some of this is club wheat. It is also produced in Northern, Eastern,

and Southern Europe, Australia, South Africa, South America, and Asia.

#### Durum wheat

Durum wheat is grown mainly in North Dakota, Minnesota, and South Dakota. Other smaller production states are California, Arizona, Oregon, and Texas. Elsewhere, it is grown in north Africa, Southern Europe, and the former Soviet Union. Durum wheat is used in making semolina, which is used for producing products such as macaroni and spaghetti.

### Germplasm resources

Plant breeders have access to over 400,000 accessions in natural and international germplasm banks. These banks include the USDA National Seed Storage Lab at Fort Collins, Colorado, the Centro Internationale de Mejoramiento de Maiz y Trigo (CIMMYT) in Mexico, and the N. I. Vavilov All-Union Institute of Plant Industry, St Petersburg, Russia. Over 40,000 accessions are held at Aberdeen, Idaho, as a working collection and parts of the United States National Small Grains Collection.

### Cytogenetics

The species of *Triticum* are grouped into three ploidy classes: diploid ( $2n = 2x = 14$ ), tetraploid ( $2n = 2x = 28$ ), and hexaploid ( $2n = 6x = 42$ ). The cytoplasmic male-sterility (CMS) gene used in modern wheat breeding is derived from *T. timopheevii*, a wild tetraploid variety. Three genomes (*A*, *B*, *D*) comprise the polyploid series of wheat. The *A* genome comes from *T. monococcum*, while the *D* comes from *Aegilops squarrosa* (or *T. tauschii*). The origin of the *B* genome is debatable. The genomic formula of the ploidy classes are *AA* or *BB* for diploids and *AABB* for tetraploid or emmer wheat. Common wheat (*T. aestivum*) is an allohexaploid of genomic formula *AABBDD*. In hexaploid wheat, the 21 chromosomes are divided into seven homeologous groups (partially homologous chromosomes) identified with numbers from 1 to 7. The three chromosomes within the *ABD* homeologous group usually share some loci in common for a specific trait. An example of this is that there are two genes for rust resistance on chromosome *2A*, three genes on *2B*, and three genes on *2D*.

Tetraploid and hexaploid wheat reproduce naturally as diploids ( $2n = 28$  or  $2n = 42$ ). This reproductive

mechanism is made possible by the presence of a gene on chromosome 5*B*, *Pb1*, which enables diploid pairing to occur. The *Pb1* gene causes truly homologous paring within the same genome. When absent, paring between one chromosome and a homeologous chromosome from another genome is possible.

The homeology that exists in its three component genomes allows the species to tolerate a range of aneuploidy. *T. aestivum* exhibits vigor and morphology similar to disomic wheat. Among other applications, aneuploidy has been used to locate genes that confer agronomically important traits (e.g., the *mlo* locus for resistance to powdery mildew). Classic wheat genetics was advanced through the work of E. R. Sears of the University of Missouri. He developed a compatible set of the possible 21 monosomics ( $2n - 1$ ) of wheat, and sets of related aneuploid forms in the hexaploid wheat cultivar, "Chinese Spring".

Introgression of alien genes is problematic because of the lack of crossability between hexaploid and diploid species, as well as the numerous problems that manifest at various stages in the ontogeny of the hybrid. Crossability genes (*kr1kr1*, *kr2kr2*) located on chromosomes 5*B*, 5*A*, and 5*D*, respectively) have been identified in the "Chinese Spring" wheat, which facilitates a wheat  $\times$  rye cross. Some breeders also use genetic bridges and chromosome number doubling to overcome problems with ploidy differences. In particular, alien autotetraploids of *Agropyron cristatum* and *Psathyrostachys juncea* have been used to overcome hexaploid  $\times$  diploid alien species crossability barriers. Generally, in practice, the parent with the higher ploidy is used as the female in crosses. However, successes with the reserve have also been recorded. Widening the genetic base of *T. aestivum* through intergeneric crosses often involves complex wheat and alien chromosome combinations. Research has shown that alien genes must be epistatic to those of wheat or interact with them to produce the desired effect. Modifications of the expression of disease- and pest-resistance genes usually occur when they are introduced into a new genetic background. Nonetheless, successes with spontaneous translocations have been reported in triticale  $\times$  wheat crosses. One of the notable induced translocations was conducted by Sears and involved chromosome 6*B* and an *Aegilops umbellata* chromosome, resulting in leaf rust resistance in the release cultivar, "Transfer".

Fertile wheat  $\times$  alien amphiploids can result from chromosome doubling, the most successful so far being triticale (wheat  $\times$  rye). Other wheat  $\times$  alien amphiploids

are less successful, being of poor fertility and often exhibiting undesirable alien traits. The technique of alien chromosome additions has been used in an attempt to reduce the undesirable effects introduced by the wild species.

## Genetics

As previously noted, dwarfing genes occur in wheat and have been used in breeding to develop cultivars with short stature (semidwarf wheat) (see Chapter 1). Early work in Japan produced dwarfing genes. Designated *Rht*, over 20 dwarfing genes have been identified, the most commonly used in wheat breeding including *Rht<sub>1</sub>*, *Rht<sub>2</sub>*, and *Rht<sub>8</sub>*. The first two, called the **Norin 10 dwarfing genes**, also belong to a group of dwarfing genes called gibberellic acid (GA)-insensitive dwarfing genes. Cultivars with these genes fail to respond to the application of GA. *Rht<sub>3</sub>* and *Rht<sub>10</sub>* genes confer extreme dwarfism on plants, the latter having a greater effect. Practical application to commercial breeding is yet to materialize. *Rht<sub>4</sub>* and *Rht<sub>8</sub>* plus others are called the GA-sensitive dwarfing genes. Monosomic analysis was used to locate the *Rht<sub>1</sub>* gene and *Rht<sub>2</sub>* gene on chromosomes 4A and 4D, respectively. Chromosome substitution can be used to transfer these genes in breeding programs. The dwarfing genes increase grain yield by increasing tillering and the number of seeds per plant.

Other genes of interest in wheat breeding include awnedness, pubescence, grain color, and glume color. The awnedness trait is inhibited by three dominant alleles at three independent loci. *Hd* conditions hooded awn, while *B<sub>1</sub>* and *B<sub>2</sub>* condition awnless or tipped awned phenotypes. A genotype of *hdb<sub>1</sub>b<sub>2</sub>* produces a bearded or fully awned phenotype. Pubescence in the glume and other parts of the plant is conditioned by a variety of dominant alleles, e.g., *Hg* producing hairy glume, while *Hp* conditions hairy peduncle. Red grain color is conditioned by three independent dominant alleles acting in additive fashion (*R<sub>1</sub>R<sub>2</sub>R<sub>3</sub>*), while white grain occurs when the genotype is *r<sub>1</sub>r<sub>2</sub>r<sub>3</sub>*. Consequently, when all three alleles occur in one genotype, the seed color is very dark red.

Anthocyanin pigmentation occurs in various parts of the plant. For example, red auricles are conditioned by a single dominant allele, *R<sub>a</sub>*. The red color of glumes is controlled by two dominant alleles, *R<sub>G1</sub>* and *R<sub>G2</sub>*, while photoinsensitivity is controlled by alleles at three independent loci, designated *ppd<sub>1</sub>*, *ppd<sub>2</sub>*, and *ppd<sub>3</sub>*.

## General botany

Wheat (*Triticum* spp.) is an annual plant. It has a spikelet inflorescence. A floret is composed of a lemma, palea, and a caryopsis or grain that has a deep furrow and a hairy tip or brush. The floret may be **awned** or **awnless**. Awned varieties are common in regions of low rainfall and warm temperatures. The presence of awns also tends to influence transpiration rate, accelerating the drying of ripe grain. Consequently, the tips of awnless spikes tend to be blasted in hot dry weather. The grain may also be amber, red, purple, or creamy white in color.

Under normal high density production conditions, a wheat plant may produce 2–3 tillers. However, when amply spaced on fertile soils, a plant may produce 30–100 tillers. The spike (head) of a plant may contain 14–17 spikelets, each spike containing about 25–30 grains. Large spikes may contain between 50 and 75 grains. The grain size varies within the spikelet, the largest being the second grain from the bottom and decreasing in size progressively towards the tip of the spike.

Wheat is predominantly self-pollinated. Anthers assume a pendant position soon after the flower opens. Blooming occurs at temperatures between 13 and 25°C starting with the spikelet around the middle of the spike and proceeding upwards and downwards. The wheat kernel or berry is a caryopsis about 3–10 mm long and 3–5 mm wide. It has a multilayered pericarp that is removed along with the testa, nucellus, and aleurone layers during milling. The endosperm makes up about 85% of a well-developed kernel. Below the aleurone layer occurs a complex protein called **gluten** that has cohesive properties. It is responsible for the ability of wheat flour to hold together, stretch, and retain gas as fermented dough rises. This property is available to the flour of only one other species, rye flour.

Wheat is classified based on three primary characteristics – agronomic, kernel color, and endosperm quality. There are two seed coat colors – **red** or **white**. Red is conditioned by three dominant genes, while the true whites comprise recessive alleles of all three genes. Most wheat varieties in the USA are red. Kernel hardness is classified into two – **hard** or **soft**. Upon milling, hard wheat yields coarse flour. White wheats, lacking this starch–protein complex, produce a higher yield of fine flour upon milling. Hard wheat is used for bread-making because its gluten protein is cohesive and elastic.

## Reproductive biology

### Floral biology

Wheat has a determinate, composite spike inflorescence. Each spike bears 10–30 spikelets, which are borne singly at nodes on alternate sides of a zig-zag rachis. A spike may be awnless, awnleted, or awned. A spikelet consists of 1–5 flowers (or florets) attached alternatively to opposite sides of the rachilla (central axis). Except in some club wheats, only two or three kernels mature, because one or more of the upper florets are usually sterile. A spikelet is subtended by a pair of empty bracts and glumes.

A floret consists of a lemma and palea, which enclose these stamens and a pistil, plus two lodicules that regulate the opening of the flowers and anthers. Wheat flowers bloom under temperatures of 13–25°C. The flowering is usually diurnal, the highest peak occurring in the morning, and a lower peak in the afternoon. Blooming begins in the spikelets located above the middle of the spike and proceeds both upward and downward. It takes about 2–3 days for a wheat spike to complete blooming, after the appearance of the first anthers. The flowering period may last from 14 to 21 days.

### Pollination

Wheat is predominantly self-pollinated with about 1–4% natural cross-pollination. Pollen shed usually starts inside the floret, but about 80% of anther dehiscence occurs outside the floret. The primary and secondary florets produce larger and more viable pollen grains than other florets. Wheat pollen remains viable for up to about 30 minutes after shedding. Once pollinated, the pollen tube growth starts within 15–60 minutes. Even though the stigma remains receptive for up to 13 days, it is most receptive within 3 days of anthesis. Xenia may occur when plants with the blue aleurone trait are used as males in a cross.

### Common breeding methods

A sample of some of the steps used at CIMMYT are summarized below as an example.

**F<sub>1</sub>** Make simple crosses. Evaluate on the basis of disease resistance, agronomic traits, and hybrid vigor. Bulk and harvest seed for F<sub>2</sub>.

**F<sub>2</sub>** Space plant 2,000–3,000 F<sub>2</sub> under optimal conditions (high fertility, moisture). Select plants based on disease resistance, lodging, tillering, maturity, etc.

**F<sub>3</sub>** Grow progeny rows in 2 m long three-row plots at dense spacing. Select desirable plots and then select and bulk the best heads in each plot. Selection environment is variable (irrigated, rainfed, acid soil, etc.)

**F<sub>4</sub>** Grow selected plants in dense-planted rows and treat as in F<sub>3</sub>.

**F<sub>5</sub>** Space plant 100 plants per plot of selected F<sub>4</sub> families under favorable conditions. Evaluate on the basis of disease resistance, desirable agronomic traits, and spike fertility.

**F<sub>6</sub>** Grow selected plants individually as F<sub>6</sub> plots of three rows, 2 m long. Select and bulk agriculturally superior lines for yield testing under various conditions (irrigated, rainfed, hot climate, acid soil, etc.).

Various approaches are adopted in wheat improvement. Some cultivars are developed through the introduction of genotypes and adapting them to new production environments. Evaluation of germplasm is also a way of identifying genotypes for use as parents in future breeding.

Modern wheat breeding depends primarily on hybridization to create variability for selection. Being a self-pollinating species, pure-line selection is often used in wheat improvement. As needed, backcrossing may be used to introgress desirable genes into existing commercial cultivars. Many traits in wheat are influenced by several genes rather than one or two, because wheat is a polyploid species. Consequently, it is uncommon for breeders to observe unexpected phenotypes in the F<sub>1</sub>. Hybridization may also bring together the three independent complementary genetic systems that condition lethals, partial lethals, or reduced productivity in the F<sub>1</sub>. Specific undesirable traits resulting from hybridization include hybrid chlorosis and grass-clump dwarfness. Hybrid necrosis is conditioned by a complementary two-gene, multiallelic system, whereas hybrid chlorosis is controlled by a complementary two-gene system. Genotypes with these undesirable genes are known. Breeders can reduce the incidence of these F<sub>1</sub> defects by carefully selecting parents for hybridization. The incidence of grass-clump dwarfness may be reduced by growing plants under high temperature and also using GA.

Hybrid wheat for commercial production has not been a practical breeding approach because of lack of sufficient heterosis upon crossing and low seed set. Other practical problems include the complexity of fertility restoration in wheat hybridization. Several major and minor genes are involved.



## Industry highlights

### *Bringing genomics to the wheat fields*

K. A. Garland-Campbell,<sup>1</sup> J. Dubcovsky,<sup>2</sup> J. A. Anderson,<sup>3</sup> P. S. Baenziger,<sup>4</sup> G. Brown-Guedira,<sup>5</sup> X. Chen,<sup>1</sup> E. Elias,<sup>6</sup> A. Fritz,<sup>7</sup> B. S. Gill,<sup>8</sup> K. S. Gill,<sup>9</sup> S. Haley,<sup>10</sup> K. K. Kidwell,<sup>9</sup> S. F. Kianian,<sup>6</sup> N. Lapitan,<sup>10</sup> H. Ohm,<sup>11</sup> D. Santra,<sup>9</sup> M. Sorrells,<sup>12</sup> M. Soria,<sup>2</sup> E. Souza,<sup>13</sup> and L. Talbert<sup>14</sup>

<sup>1</sup>USDA-ARS Wheat Genetics, Quality, Physiology, and Disease Research Unit, Washington State University, Pullman, WA 99164, USA; <sup>2</sup>Department of Agronomy and Range Science, University of California at Davis, Davis, CA 95616, USA;

<sup>3</sup>Department of Agronomy and Plant Genetics, University of Minnesota, Twin Cities, St Paul, MN 55108, USA; <sup>4</sup>Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583, USA; <sup>5</sup>USDA-ARS Plant Science Research Unit, North Carolina State University, Raleigh, NC 27606, USA; <sup>6</sup>Department of Plant Sciences, North Dakota State University, Fargo, ND 58105, USA; <sup>7</sup>Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA;

<sup>8</sup>Wheat Genetics Resource Center, Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA;

<sup>9</sup>Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164, USA; <sup>10</sup>Department of Soil and

Crop Sciences, Colorado State University, Fort Collins, CO 80526, USA; <sup>11</sup>Department of Agronomy, Purdue University,

West Lafayette, IN 47907, USA; <sup>12</sup>Department of Plant Breeding, Cornell University, Ithaca, NY 14853, USA; <sup>13</sup>Aberdeen

Research and Extension Center, University of Idaho, Aberdeen, ID 83210, USA; <sup>14</sup>Department of Plant Sciences and Plant

Pathology, Montana State University, Bozeman, MT 59717, USA

#### **Introduction**

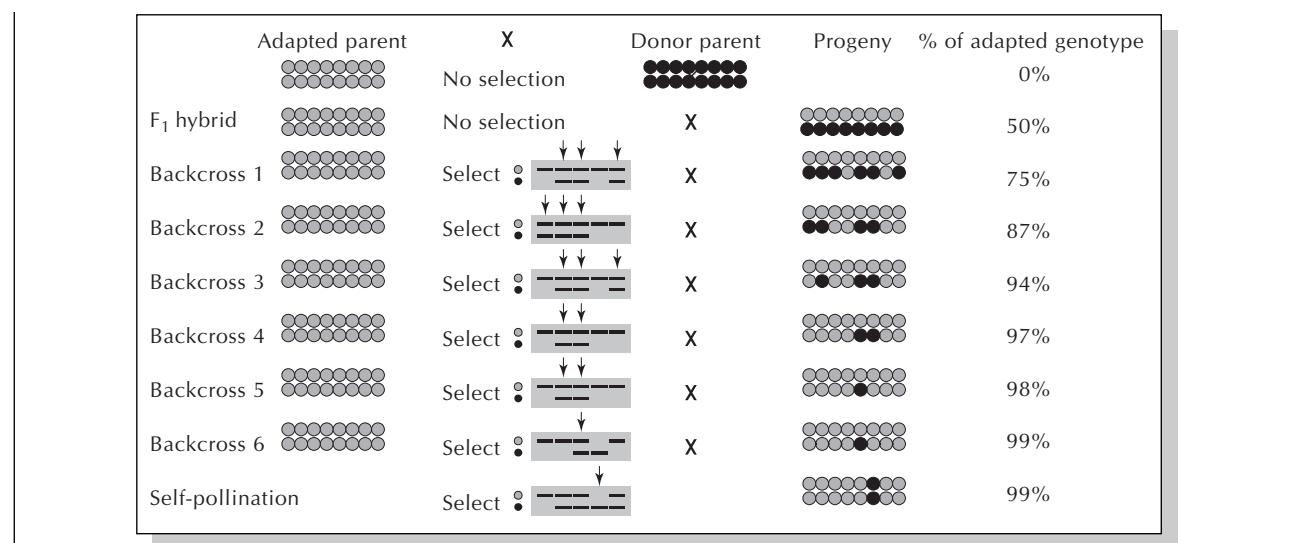
Wheat (*Triticum aestivum* (L.) em Thell) is well adapted to diverse climatic conditions around the world, and is grown in all regions of the United States. Wheat genetic improvement was the foundation of the Green Revolution and recent progress in the area of wheat genomics has been referred to as the beginning of a new Green Revolution. While wheat improvement through plant breeding still requires a great deal of phenotypic or trait assessment, new knowledge about the wheat genome enables plant breeders to target their breeding efforts with more precision than ever before.

*Triticum aestivum* (bread wheat) is hexaploid while *T. durum* L. (durum or macaroni wheat) is tetraploid. Hexaploid wheat has three genomes made up of seven chromosomes each labeled 1–7. The genomes are named *A*, *B*, and *D* and closely related homologous chromosomes are named with the genome name, e.g., 1*A*, 1*B*, and 1*D*. Durum wheat has the *A* and *B* genomes. Genes are frequently found in multiples, as gene orthologues located on each genome, that are closely related in their structure and function. Orthologous sets of wheat genes are named with the gene name, the genome name, and the orthologous set designation. Long and short arms of chromosomes are designated with *L* and *S*, respectively. For example, the genes for reduced height that were fundamental to the performance of wheat cultivars developed during the Green Revolution are named *Rht-B1b* and *Rht1-D1b* and are located on wheat chromosomes 4*B* and 4*D*.

Two growth habits are agriculturally important in wheat production, fall-seeded winter wheat and spring-seeded spring wheat. In warm climates, spring wheat is also seeded in the fall. Winter wheat requires a period of cold temperatures, or vernalization, to induce flowering. Wheat is classified in world markets according to its growth habit, grain texture (soft or hard), grain color (red or white) and gluten properties (strong or weak). These characteristics result in specific end-use properties of each class of wheat, and have led to targeted breeding efforts (e.g., bread versus confectionary uses).

Marker-assisted selection (MAS) is a method of rapidly incorporating valuable traits into new cultivars. Molecular markers, or DNA tags, that have been shown to be linked to traits of interest are particularly useful for incorporating genes that are highly affected by the environment, genes for resistance to diseases and pests, and to accumulate multiple genes for resistance to specific diseases and pests within the same cultivar – a process called gene pyramiding. One of the first wheat cultivars to be developed using MAS was the soft winter wheat cultivar “Madsen”, released in 1986 by the USDA-Agricultural Research Service (ARS) and Washington State University. “Madsen” was developed using the isozyme marker from the endopeptidase protein, EpD1b, to incorporate a gene for resistance to eyespot (*Tapesia yallundae*) (Allan et al. 1989). Since 1990, detailed molecular maps of wheat have been constructed that include more than 3,000 molecular markers and several important traits have been associated with DNA markers. Additional markers can be developed from the 8,000 expressed sequence tags (ESTs) that have been mapped in wheat. (Maps and references are available on-line (USDA-ARS 2005).)

If selected genes or chromosome segments are incorporated from donor parents into adapted wheat breeding lines for six backcross (BC) generations using markers to select for the targeted gene from the BC<sub>2</sub> generation, more than 99% recovery of the adapted recurrent parent is expected (Figure 1) (Hospital et al. 1992). Seven BC plants per generation are adequate to have a probability higher than 0.99 of recovering a single BC plant with the desired genotype. Linkage drag, which results when undesirable chromosome regions from the donor parent are carried along during backcrossing, can be reduced by advancing more than 10 plants per cross and selecting for those with higher percentages of the adapted recurrent parent genome.



**Figure 1** Six backcrosses with marker-assisted backcrossing and selection for heterozygote plus self-pollination with selection for homozygote results in progeny with less than 1% donor parent plus the desired gene.

### Current use of MAS in US wheat breeding

Several traits have been incorporated into advanced wheat breeding lines using MAS. We will focus on two examples that improve end-use quality and disease resistance. Additional details are available on-line (Dubcovsky & Soria 2005).

#### End-use quality improvements

Grain protein content is one of the major factors affecting bread-making and pasta quality. In spite of the importance of this character, progress in breeding for high grain protein content has been slow and difficult for two reasons: (i) most variation in protein content is due to environmental rather than genetic effects; and (ii) there is a strong negative relationship between grain protein content and grain yield, so that cultivars selected for high grain protein content tend to have low grain yields. A promising source of high grain protein content was detected in a survey of the wild tetraploid wheat relative, *T. dicoccoides*. It was crossed into the durum cultivar "Langdon" and the responsible gene was mapped on the short arm of chromosome 6B (6BS) (Joppa et al. 1997). This segment accounted for 66% of the variation in grain protein content observed in a cross between durum wheat and *T. dicoccoides*. The same chromosome segment was transferred to hexaploid wheat by R. Frohberg. The chromosome segment carrying the *HPGC* gene for high seed protein content from *T. dicoccoides* can be efficiently manipulated in tetraploid and hexaploid wheat with microsatellite markers (also known as simple sequence repeats or SSRs). The more useful markers include *Xgwm193* and a cleaved amplified polymorphic marker for the NOR locus (Khan et al. 2000). Results from field trials have been mixed, but generally indicate an increase in protein dependent upon genetic background and environment (K. Kidwell, personal communication).

Wheat grain end-use properties are affected markedly by endosperm texture. Hard wheat requires more grinding energy to reduce endosperm into flour and during this milling process a considerable number of starch granules become physically damaged. Soft wheats, by contrast, produce flours with smaller particles and lower levels of damaged starch. Damaged starch is of value in yeast-leavened products because in addition to absorbing water, it acts as a substrate for  $\alpha$ -amylase and creates a favorable environment for yeast growth. In contrast, chemically leavened soft wheat products have better texture if they are made from flour with small particle size, and low water retention capacity. Therefore, hard wheat lines have been selected for high damaged starch and higher hardness values and soft wheat flour for the opposite parameters. Differences in endosperm texture are associated with the complementary action of proteins puroindoline A (*pinA*) and puroindoline B (*pinB*), which are coded by genes located on the distal part of chromosome arm 5DS. Most hard wheats possess a glycine to serine mutation in puroindoline B (allele *pinB-D1b*) or they are devoid of puroindoline A (allele *pinA-D1b*) (Giroux et al. 1998). Cultivars carrying these different mutations differ in their grain hardness and in their milling and baking characteristics. Hard red spring wheats with the *pinB-D1b* allele have improved flour yield, milling quality, and loaf volume relative to sister lines with the *pinA-D1b* allele. Allelic variants at the *Ha* locus are available to modulate grain texture in soft wheats. The replacement of the distal part of 5AS by the distal 5A<sup>mS</sup> segment from *T. monococcum* (active *pinB-A<sup>m1a</sup>* and *pinA-A<sup>m1a</sup>*) results in a reduction of the hardness of soft cultivars (G. Tranquilli and J. Dubcovsky, unpublished). These different alleles have been manipulated to confer particular levels of hardness.

### Host plant resistance to disease and insect pests

Genetic resistance is the primary method of choice for controlling diseases in wheat, and has been proven repeatedly as an effective and environmentally sound method to control serious yield-reducing pathogens. The use of disease-resistant cultivars reduces the use of pesticides and thus contributes to a reduction in environmental contamination (Anderson 2000). Losses to all pests and diseases currently average 10–20% annually, and therefore potential savings to growers, not counting the elimination of costs of applying pesticides, are in the hundreds of millions of dollars (10% of the US wheat production has been worth about US\$500 million annually since 2000).

Genetic resistance has frequently resulted in selection pressure on the pathogen population, which then mutates to overcome the resistance. Two strategies, either a pyramid or a multiline approach, have been used to increase the durability of resistance genes. Gene pyramiding is the combining of two or more resistance genes. Resistance is more durable because additional mutations in the pathogen population are needed to overcome the resistance in the host plant. Multiline cultivars are made up of a series of closely related genotypes, each carrying different sources of resistance to a pest. The genetic diversity of the multiline results in balancing selection against the pest that enhances the durability of the resistance genes deployed in the multiline cultivar. Both strategies have been difficult to accomplish without MAS because once an effective resistance gene is present in a breeding line, it is difficult to screen for the incorporation of additional resistance genes using the plant phenotype. Few multilines have been released, in practice, because of the difficulty in introgressing so many different sources of resistance into visually and agronomically similar genotypes. MAS can be used effectively to combine different resistance genes into elite lines while maintaining pre-existing, effective resistance genes and to introgress several resistance sources into a recurrent parent.

On a global basis, the three rusts – leaf rust (*Puccinia recondita*), stem rust (*P. graminis*), and yellow or stripe rust (*P. striiformis*) – are among the most damaging diseases of wheat and other small grain crops. Besides reducing yield, the rust diseases also seriously affect the milling and baking qualities of wheat flour. Multimillion dollar yield losses have been attributed to leaf and stripe rust every year since 2000. Stripe rust alone caused a US\$360 million loss to the wheat crop in 2004 (updates are available online (Long 2005)). New races of leaf and stripe rust have been identified in the USA since 2000 that are virulent in many of the previously resistant cultivars. Fortunately, new resistance genes have been identified in wheat cultivars, and wild relatives of wheat have been introgressed into hexaploid wheat. Using MAS, six new leaf rust resistance genes and five stripe rust resistance genes have been introgressed into 58 lines and cultivars from eight different market classes. Leaf rust resistance genes include *Lr21*, *Lr39*, and *Lr40* from *Triticum tauschii*, *Lr37* from *T. ventricosum*, *Lr47* from *T. speltoides*, and *LrArm* from *T. timopheevi* subsp. *armeniacum*. Four major stripe rust resistance genes *Yr5*, *Yr8*, *Yr15*, and *Yr17* have been combined with the high temperature adult plant resistance gene identified in "Stephens". This resistance has been durable for over 25 years in the Pacific northwest. A quantitative trait locus (QTL) for this resistance has recently been identified, probably located on chromosome 6BS (K Campbell, unpublished data). A new stripe rust resistance gene *Yr36* has also been identified on chromosome 6BS, closely linked to the *HPGC* gene. Both genes can be simultaneously selected using PCR (polymerase chain reaction) markers *Xuhw89* and *Xgwm193* (J. Dubcovsky, personal communication). PCR-specific markers are available for *Lr2*, *Lr47*, and the linked group of resistance genes *Lr37-Yr17-Sr38* (Helguera et al. 2003). Microsatellite marker *Xgwm210* is linked to *Lr39* and *Xgwm382* to *LrArm*. Microsatellite markers *Xgwm18* and *Xgwm264* flank stripe rust resistance gene *Yr15* (Chen et al. 2003).

*Fusarium* head blight (FHB) is an important disease in common and durum wheat producing areas of the United States and Canada. An epidemic of FHB from 1993 to 1997 resulted in devastating economic losses to the wheat industry of the region, with 1993 estimates alone surpassing \$1 billion. FHB causes both severe yield reduction and decreased grain quality. In addition, infected grain may contain harmful levels of mycotoxins that prevent its use for human consumption or feed. Control of FHB has been difficult due to the ubiquitous nature and wide host range of the pathogen, and dependence of the disease upon unpredictable climatic conditions. In some parts of the USA, fungicides have been used to reduce losses, but this practice adds to grower costs, poses significant environmental risks, and is not always effective. Available resistance to FHB in wheat is quantitatively expressed, with a continuous distribution among progeny. Two major QTLs have been identified on chromosome 3BS from "Sumai 3" and on chromosome 3A from *T. dicoccoides*. Microsatellite markers, *Xgwm533* and *Xgwm493*, bracketing both QTL regions have been used to introgress these genes into 28 durum and common wheat cultivars. The selected QTL region from the "Sumai 3" chromosome arm 3BS explains up to 40% of the phenotypic variation in FHB resistance in one cross. Selection for the QTL region from chromosome arm 3AS in durum wheat has been done using SSR markers *Xgwm2* and *Xgwm674*. The QTL region on 3AL explains more than 37% of the phenotypic variation in a durum cross. Both FHB QTL regions are robust and are expressed in well-adapted genetic backgrounds (Liu & Anderson 2003).

### Practical use of MAS in forward breeding programs

MAS selection programs must be integrated at all times into existing breeding programs. Recurrent parents are selected from high-yielding, elite germplasm. Intermediate products of MAS are returned to the breeding programs for evaluation and crossing purposes. After each generation of backcrossing, selected heterozygous plants are self-pollinated and the  $BC_{1-3}F_2$  seeds planted as additional segregating populations. A strict backcrossing strategy is not expected to increase yield, except for the reduction of yield losses due to pathogens. Therefore, this backcrossing strategy should be used only as a complement of active "forward breeding" programs.

Completion of the trait introgression to the BC<sub>6</sub> generation not only provides superior cultivars for immediate use by growers, but also provides a unique set of near-isogenic experimental materials to rigorously address scientific questions in wheat breeding and wheat genetics. One such question regards the potential costs of pest-resistance genes in the absence of pest infestation. Several experiments have shown that some disease-resistance genes in plants come at a cost in fitness, or yield potential. In general, however, measurement of cost has been difficult due to the lack of defined genetic populations. The development of isogenic lines carrying different genes affecting quality in the same genetic background will facilitate studies of epistatic interactions. A cross between two isogenic lines in the same recurrent parent generates a mapping population segregating only for the targeted genes in an isogenic background. Current public awareness of biotechnology has been shaped by the GMO debate. This is unfortunate because genetically modified organisms are only one aspect of biotechnology. MAS is a valuable biotechnological tool for selection and reassembly of genes that already exist in *Triticeae*.

### References

- Allan, R.E., C.J. Petersen Jr., G.L. Rubenthaler, R.F. Line, and D.E. Roberts. 1989. Registration of "Madsen" wheat. *Crop Sci.* 29:1575–1576.
- Anderson, J.A. 2000. Marker-assisted selection of disease resistance genes in wheat. In: *Marker-assisted selection of disease resistance genes in wheat* (Kohli, M.M., and M. Francis, eds), pp. 71–84. La Estanzuela, Uruguay.
- Chen, X., M.A. Soria, G. Yan, J. Sun, and J. Dubcovsky. 2003. Development of user-friendly PCR markers for wheat stripe rust resistance gene Yr5. *Crop Sci.* 43:2058–2064.
- Dubcovsky, J., and M.A. Soria. 2005. MASWheat: Bringing genomics to the wheatfields. Available at <http://maswheat.ucdavis.edu/Index.htm>. Verified 8-Jul-05.
- Giroux, M.J., and C.F. Morris. 1998. Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. *Proc. Natl. Acad. Sci. USA* 95:6262–6266.
- Helguera, M., I.A. Khan, J. Kolmer, D. Lijavetzky, L. Zhong-q, and J. Dubcovsky. 2003. PCR assays for the *Lr37-Yr17-Sr38* cluster of rust resistance genes and their use to develop isogenic hard red spring wheat lines. *Crop Sci.* 43:1839–1847.
- Hospital, F., C. Chevalet, and P. Mulsant. 1992. Using markers in gene introgression breeding programs. *Genetics* 132:1199–1212.
- Joppa, L.R., C. Du, G.E. Hart, and G.A. Hareland. 1997. Mapping a QTL for grain protein in tetraploid wheat (*Triticum turgidum* L.) using a population of recombinant inbred chromosome lines. *Crop Sci.* 37:1586–1589.
- Khan, I.A., J.D. Prochnier, D.G. Humphreys, et al. 2000. Development of PCR based markers for a high grain protein content gene from *Triticum turgidum* ssp. *dicoccoides* transferred to bread wheat. *Crop Sci.* 40:518–524.
- Liu, S., and J.A. Anderson. 2003. Targeted molecular mapping of a major wheat QTL for *Fusarium* head blight resistant using wheat ESTs and synteny with rice. *Genome* 46:817–823.
- Long, D. 2005. USDA-ARS Cereal Disease Laboratory: Small grain losses due to rust in the US. Available at <http://www.cdl.umn.edu/loss/loss.html>. Verified 8-Jul-05.
- USDA-ARS. 2005. Graingenes: A database for *Triticeae* and *Avena*. Available at <http://wheat.pw.usda.gov/GG2/index.shtml>. Verified 8-Jul-05.

## Establishing a breeding nursery

### Field nursery

#### Layout

A wheat breeding nursery may contain various materials such as parental genotypes with desired traits, elite lines, special genetic stock, and sterility sources. It is advantageous to locate the crossing block close to the F<sub>1</sub> hybrid nursery to facilitate backcrossing, topcrossing, and hybrid parent comparisons. Wide between-row and within-row

spacing facilitates crossing and promotes good plant growth and development.

#### Planting

Researchers often raise their seedlings at a plant spacing of 30–60 cm between rows and 15 cm within rows for good tillering. Planting dates should be selected such that flowering is synchronized for crossing. The growth environment may also be modified to synchronize flowering. Some breeders of, in particular, winter wheat, clip the plants (not below the spike primordial to avoid killing it).

## Greenhouse nursery

Breeders may use a greenhouse to help synchronize flowering. Greenhouses enable plant breeders to manipulate temperature, photoperiod, light, and other plant growth environmental conditions to synchronize flowering. Increasing the photoperiod tends to reduce days to heading. A greenhouse nursery also enables the plant breeder to move pots around to facilitate crossing.

## Artificial pollination

### Materials and equipment

Emasculation is needed in wheat crossing. The equipment and materials used include scissors, curved and straight-type forceps, soda straws, squeeze-type bottles, paper clips, tags, and glassine bags ( $5 \times 17$  cm).

### Emasculation

Preparation of the female for hybridization by emasculation is done 1–3 days before normal anthesis. The anthers at this stage should be light green but not yellow or cream colored. The spikes to be emasculated should be carefully selected. Further, the stigmas should be clearly visible extending to about one-quarter the length of the floret. One to three of the upper and basal spikelets are usually non-functional and are removed with the scissors. In awned varieties, the awns are clipped. Only the primary and secondary florets are retained emasculated. To emasculate, gently but firmly grasp the spike between the thumb and index finger of one hand. Then, lightly press the thumb below the tertiary floret and pull the center florets downward and outward with the forceps. The three anthers in each floret are removed by carefully inserting a pair of forceps between the lemma and palea and spreading them. Care should be exercised to avoid damaging the stigma. Cover the emasculated spike with a bag.

Florets may also be emasculated by cutting off about one-third of the florets with the scissors to expose the anthers, which are then removed with the forceps. This procedure predisposes the florets to rapid desiccation and is favored under cool and humid environments. Another procedure entails cutting the primary and secondary florets at about 5–7 days before anthesis. The remaining parts of the anthers eventually degenerate without forming viable pollen. This procedure is quick but requires greater skill and experience for high success.

## Pollination

Emasculated flowers should be pollinated within 2–4 days for best results. Also, a mature and receptive stigma is one with fully developed feathery features, whereas recently extruded anthers provide good pollen. Forceps are used to remove anthers from the florets. The glassine bag is removed from the emasculated spike so the pollen can be gently brushed on one or several stigmas. A soft brush may also be used to transfer pollen to stigmas. CIMMYT in Mexico developed a pollination technique that is suitable when pollen is present in copious amounts. First, the upper third of each primary and secondary floret of the male spike is removed to allow the mature anthers to freely extrude to release the pollen. The top of the bag covering the female spike is cut to allow the male spike to be inverted into the bag, parallel to the female spike. The male spike is then vigorously rotated by twirling the peduncle between the thumb and forefinger. The bag is resealed, with the male spike in place, for an additional day for enhanced pollination.

In another method, the approach method, the female spike is positioned lower than the male spike and the two covered with a glassine bag. Such maneuvering is easier in the greenhouse. It can be used in the field if the female and male plants are planted in close proximity or the male spike is detached. To prolong the supply of pollen, culms of the detached spikes can be placed in water to keep them alive.

## Natural pollination

As previously indicated, about 1–4% outcrossing occurs in wheat, depending on the variety. To facilitate large-scale hybridization and eliminate the tedium of emasculation, repeated backcrossing is used to convert breeding lines into female CMS or A-lines. This scheme enables breeders to undertake composite cross-breeding, production of synthetics, and the hybridizations needed for recurrent selection. However, the stability of the male fertility-restoration gene is difficult to maintain, being conditioned by several genes and affected by the genetic background in which the genes occurs.

## Seed development

A pollinated spike is properly targeted for identification. Successful pollination and fertilization can be verified after 3–5 days of pollination. Kernel development is

detectable during the first week following pollination. At normal maturity, the tagged spikes are harvested and threshed.

### Common breeding objectives

The success of modern wheat cultivars is largely due to high-yield potential, wide agroecological adaptation, and high responses to agronomic inputs (fertilizers, irrigation). Yield components of wheat are spikes  $\times$  number of grains/spike  $\times$  weight of grain (or number of grains/unit area  $\times$  weight of grain in that unit). Breeders should determine what balance of these components to include in a cultivar for an agroecological niche. The right balance is determined taking into account the photoperiod, heat units, and the moisture and fertility status of the target area. For example, in one study, a top-yielding wheat cultivar, "Seri 82", produced 3,778 kg/ha in 98 days (maturity) at 22°12'N but 8,544 kg/ha in 140 days at 30°53'N. It is not possible to simultaneously select for all the yield components because of the presence of negative inter-correlations among them.

**1 Grain yield.** Breeding for high yield potential in wheat can be accomplished by hybridizing high-yielding genotypes and selecting transgressive segregants from the progeny with desired traits. However, it is the discovery and use of dwarfing genes that dramatically increased yield potential in wheat. Short-statured cultivars have high tillering capacity and also increased grain yield per spike. Manipulation of harvest index by incorporating semidwarf genes has resulted in high lodging resistance, high biomass, and high harvest index and consequently a high rate of partitioning of assimilates into the grain for higher grain yield.

**2 Yield stability.** Some breeders have used the concept of shuttle breeding (selecting  $F_2$  segregating populations in one location and the  $F_3$  in another, etc.) to develop cultivars with wide adaptation and high yield potential (e.g., cultivars such as "Siete Cerros" and "Pavon 76" developed at CIMMYT). The capacity to sustain high yield potential over a broad range of growing environments is desirable in a cultivar. Breeders conduct  $G \times E$  evaluations of genotypes in yield trials to identify those with yield stability.

**3 Agromorphological traits.** As previously discussed, short stature and lodging resistance are important breeding objectives in wheat breeding. Semidwarf cultivars are lodging resistant. Selection for short

stature often impacts other plant characteristics. For example, semidwarfs tend to be photoperiod insensitive, and have reduced seed size and protein content.

#### 4 Adaptation:

- (a) **Winter hardiness.** Winter-hardy wheat cultivars are needed in places where plants are likely to be exposed to unseasonable low temperatures. Regions with high rainfall tend to hold moisture in the soil for a longer time. Under wind chill temperatures and alternate freezing and thawing, wheat plants grown in such soils are prone to heaving. The red soft winter wheat types are more resistant to heaving injury than the hard red wheats. Selecting for winter hardiness should be done under natural conditions.
- (b) **Drought resistance.** Germplasm of the Crimean origin has drought resistance and narrow leaves. Similarly, durum wheats have drought resistance and are adapted to the drier production regions of North Africa and the Middle East. At CIMMYT, breeders have used a shuttle breeding approach in drought breeding. Generations  $F_2$ ,  $F_5$ , and  $F_6$  are tested under optimal conditions, while  $F_3$  and  $F_4$  are evaluated at reduced fertility and moisture conditions. The assumption is that input efficiency and input responsiveness can be incorporated into one genotype. Traits of interest under optimal conditions are disease resistance, good tillering capacity, head development, leaf retention, and grain plumpness. Under low input, breeders select for delayed leaf senescence, tiller viability, grain plumpness, reduced spike sterility, relative high yield, and relative higher pest resistance.
- (c) **Aluminum (Al) tolerance.** Tolerance to Al is needed in cultivars grown for production regions where the soils are acidic. Breeders may select for Al tolerance under artificial conditions in the lab. In fact, breeding efforts in Brazil produced genotypes with high Al tolerance. However, yield of those genotypes were poor. Breeders at CIMMYT have improved upon the yield potential of the Al-tolerant genotypes.
- 5 Disease resistance.** The strategy of using resistance genes singly in wheat breeding is not effective and hence not widely practiced any longer. Rather, a combination of multiple hypersensitive resistance genes is preferred. Rust-resistance genes ( $Lr_9$ ,  $Lr_{19}$ ,  $Lr_{24}$ ) and stem-resistance genes ( $Sr_{24}$ ,  $Sr_{26}$ ,  $Sr_{31}$ ) once deployed singly are now used in combination to promote stability to stem rust in North America and Australia. However, this strategy is effective only when breeding is centrally coordinated in the country and

when the production area is isolated from other areas where this system is not operational. Some failures of system have occurred in leaf rust fungus in Mexico.

- (a) **Rusts.** Diseases of importance in wheat production include stem rust (caused by *Puccinia graminis* Pers.), leaf rust (*P. recondite* Rob.), and stripe rust (*P. striformis* West). Stem rust is particularly devastating to wheat production. Researchers have identified numerous physiological races of rust species. Over 30 resistance genes have been discovered for *P. graminis* and *P. recondite*, and more than 16 genes for *P. striformis*.
- (b) **Smuts.** Wheat is attacked by a variety of smut diseases, the major ones being loose smut (caused by *Ustilago tritici*), common bunt (*Tilletia tritici*), and dwarf bunt (*T. controversa*). Yield can be drastically reduced by smut. Loose smut can destroy the entire spike, whereas other smuts cause the grain to be replaced with smut spores. Genes of resistance have been identified for loose and common smut diseases.
- (c) **Powdery mildew.** Genes encoding resistance to powdery mildew (caused by *Erysiphe graminis*) have been identified.

- (d) Other diseases of interest in wheat breeding include viral infections, blights, and rots.

#### 6 Insect resistance:

- (a) **Hessian fly.** The Hessian fly (*Mayetiola destructor*) is known to cause a wide range of effects to wheat plants including dwarfing of plants, reduced tillering, straw breakage, and increased susceptibility to winter injury. Some genotypes have tolerance to the pest, being able to compensate for reduce tillering by producing additional tillers. Several biotypes of the fly have been identified. Resistance genes have been identified. Sources of resistance include genotypes such as "Ribeiro", "Marquillo", and "Kawvale".
- (b) **Greenbug.** About eight biotypes of the greenbug (*Schizaphis graminum*) have been identified. Genes for many of these biotypes have been identified.

- 7 **End-use quality.** Wheat is used in bakery products and semolina. Various breeding programs focus on various market types. The hard wheats are used for bread, whereas the soft wheats are used for confectionary products and cookies. Durum wheats are used for pasta products.

## References and suggested reading

- Acquaah, G. 2005. Principles of crop production: Theory, techniques, and technology. Prentice Hall, Upper Saddle River, NJ.
- Hall, M.D., and D.A. van Sanford. 2003. Diallel analysis of *Fusarium* head blight resistance in soft red winter wheat. *Crop Sci.* 43:1663–1670.
- Rajaram, S., A. Mujeeb-Kazi, and R.P. Singh. 1997. Wheat. *In: Genetics, cytogenetics, and breeding of crop plants. Vol. 2. Cereal and commercial crops* (Bhal, P.N., P.M. Salimath, and A.K. Mandal, eds). Science Publishers, Inc., Enfield, NH.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Wheat is predominantly a self-pollinated species.
- 2 Wheat has spikelet inflorescences.
- 3 Cultivated wheat is a tetraploid.
- 4 Gluten is a complex starch found in wheat.

### Part B

Please answer the following questions:

- 1 What is vernalization?
- 2 Give the two groups of wheat on the basis of adaptation.

- 3 What specific kind of polyploidy is cultivated wheat?
- 4 Give three of the top major wheat producing states in the US.
- 5 Give the scientific name of cultivated wheat.

**Part C**

Please write a brief essay on each of the following topics:

- 1 Discuss three major breeding objectives in wheat breeding.
- 2 Discuss important diseases of wheat.
- 3 Discuss the market classes of wheat.
- 4 Briefly discuss the emasculation of wheat flowers for crossing.



# 28

## Breeding corn

### Taxonomy

Kingdom	Plantae
Subkingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Liliopsida
Subclass	Commelinidae
Order	Cyperales
Family	Poaceae
Genus	<i>Zea</i> L.
Species	<i>Zea mays</i> L.

### Economic importance

Corn or maize is the single most important crop in the USA. It is grown on more than 20% of the cropland. Most of the production occurs in the region of the US called the Corn Belt where six states, Iowa, Illinois, Nebraska, Minnesota, Indiana, and Ohio, account for about 80% of national production. Corn is grown in every state except Alaska. Iowa leads the nation with 22% of the total production. Corn is the fourth most important crop in the world, behind wheat, rice, and potato in total production. More than 327 million acres of corn are planted each year, worldwide. World yields average about 42 bushels per acre (see Appendix 2 for conversion rates of units).

There are six “corn belts” in the world – the US Corn Belt, Danube basin (southwest Germany), Po valley (north Italy), plains of north China, northeast Argentina, and southeast Brazil. On the world scene, the USA, China, Brazil, Mexico, France, and Argentina together account for 75% of the world’s corn production, the US accounting for about 40% of this total. Other producers include Romania and South Africa. In 1866, the US Department of Agriculture (USDA)

reported 30 million acres of corn were harvested at an average yield of about 24 bushels/acre. In the early 1990s, about 100 million acres were grown, with an average harvested yield of 28 bushels/acre. However, in 2000, USDA reported 67 million acres of harvested corn with an average yield of 118 bushels/acre. The dramatic increase in yield over the period is attributed to the adoption of hybrid seed and the use of fertilizers. Corn has the highest value of production of any crop in the USA, averaging 8 billion bushels worth US\$20 billion per year.

### Origin and history

Corn is arguably the most completely domesticated of field crops. Modern corn is incapable of existing as a wild plant. No wild form of corn is known. Its origin is probably Mexico, or Central America. It was produced as early as 6000 BC in Tehuacan, Mexico by the Mayan and Aztec Indian civilizations. It was taken north by the native Americans. Corn was dispersed to the Old World in the 16th and 17th centuries. The explorers introduced Indian corn to Europe and Africa. Modern varieties have larger cobs and a greater number and weight of kernels per ear as compared to the original Indian corn. The modern cultivated plant is believed to have been obtained through the process of mutation, coupled with natural selection, and mass selection by the American Indians. It is proposed that corn’s progenitor may be a domesticated version of teosinte, a wild grass that occurs in Mexico and Guatemala.

### Adaptation

Corn has a wide geographic adaptation. It is grown from as far north as 58°N latitude to 35–45°N latitude. It is

grown at below sea level to 4,000 m. Corn is adapted to warm temperatures.

### History of breeding in the USA

Corn is believed to have been domesticated more than 7,000 years ago. Deliberate attempts at improving corn began with the development of open-pollinated varieties that are still common in developing countries. Even though the first corn hybrid in the United States was developed in the early 1920s, most producers continued to use open-pollinated cultivars until the 1940s. Producers in the short season northern regions preferred flint (northern flint) corn varieties. However, early production acreage was dominated by the Corn Belt dent varieties. However, growers in the mid-maturity production areas preferred the southern dent open-pollinated corn races for their superior high yield potential. Natural crossing between the northern flint and southern dent varieties produced “hybrids” (intervarietal hybrids) with superior performance. Farmers were able to select and develop numerous open-pollinated varieties, one of the most successful being “Minnesota 13”, which was later widely used in early hybrid variety development.

The pioneering work of G. H. Shull of the Carnegie Institution of Washington started the move toward the development and spread of hybrid corn. His landmark publication “A pure-line method in corn breeding” laid the foundation for corn hybrid breeding. Later, E. M. East, while providing collaborative evidence for Shull’s work, at the same time discouraged the commercial development and use of hybrids. This was because the first hybrids were inferior and, further, the single-cross hybrid produced seed on an inbred line, making it an expensive undertaking. The first commercial hybrid, the “Copper Cross”, was a single-cross variety. This first commercial variety was produced in Iowa under contract by H. A. Wallace in 1923. Wallace founded the Hi-Bred Corn Company, which later became the Pioneer Hi-Bred Company. Commercial hybrid production became feasible when in 1918 D. F. Jones proposed the use of the double-cross hybrid, which was a product of two single crosses. This meant that a hybrid seed was produced on a relatively high-yielding single-cross hybrid female plant. Double-cross hybrids dominated corn production until the 1960s. The dominance of hybrid varieties was due to their superior characteristics, especially greater uniformity, higher yield, tolerance to biotic and abiotic stresses, and amenability to mechanization.

Plant breeders later (after the 1960s) developed superior inbred lines with high yield potential and stability. Single-cross hybrids became more productive and more uniform, eventually replacing the double crosses. Initially, corn breeding was primarily conducted in the public arena. However, private companies began developing superior lines in the 1960s for hybrid breeding, eventually dominating the commercial seed corn market. Corn yield has characteristically increased from about 64 bushels/acre with double crosses in 1959 to about 129 bushels/acre using single crosses.

It has become clear to breeders that most of the corn hybrids in North America originate from a few inbred lines developed in the public arena, severely narrowing the genetic base of corn and making them vulnerable. There is an effort by breeders in both private and public sectors to enhance the genetic base of corn through the germplasm enhancement of maize. In the latter parts of the 20th century, the use of biotechnology became a major feature of maize breeding, resulting in transgenic products, such as *Bt* corn. Nutritional augmentation breeding has also produced products such as high-lysine corn and quality protein maize (QPM).

### Types of corn

Corn belongs to the family Poaceae, other members of this family being *Zea*, *Euchlaena*, *Tripsacum*, *Coix*, *Chionachne*, *Polytoca*, *Scherachne*, and *Triobachne*. Of these, the closest relatives to corn are *Euchlaena* (teosinte) and *Tripsacum*. Intergeneric crosses between *Zea* and *Euchlaena*, and *Tripsacum*, *Saccharum*, and *Coix* have been successfully attempted, the products usually being sterile. *Zea × Euchlaena* (corn × teosinte) crosses are, however, often successful with fertile offspring.

Corn may be grouped into seven types on the basis of endosperm and glume characteristics – dent, flint, flour, pop, sweet, waxy, and pod corns. Of these, five are commercially produced (dent, flint, flour, sweet, and waxy corns).

**1 Dent corn (*Z. mays indentata*)**. Dent corn is the most widely cultivated type in the US. It is characterized by a depression (dent) in the crown caused by the rapid drying and shrinkage of the soft starch at the crown. Of the multiple colors available, the yellow or white kernels dominate commercial production.

**2 Flint corn (*Z. mays indurate*)**. Flint corn is predominantly comprised of corneous or hard starch that

encloses the soft starch at the center. The kernels are smooth, hard, and usually rounded at the top. This type of corn is grown widely in Europe, Asia, Central America, and South America. It is less widely grown in the US.

- 3 Flour corn (*Z. mays amylacea*).** As the name implies, flour corn consists almost entirely of soft starch, making the kernels soft. It has the shape of dent corn but shrinks uniformly upon drying. It is grown in the drier sections of the USA, mainly by American Indians, and also in the Andean region of Central and South America. Different kernel colors exist, the most common being white, blue, and variegated.
- 4 Popcorn (*Z. mays everta*).** Popcorn is an extreme form of flint corn. It has a very hard corneous endosperm with only a small portion of soft starch. The kernels are characteristically small and may either be pointed or have a rounded tip. Different colors exist, most corneous varieties being yellow or white. The kernel pops upon heating as a result of the unique quality of the endosperm that makes it resist the steam pressure generated, until it reaches explosive proportions.
- 5 Sweet corn (*Z. mays saccharata*).** This corn is characterized by a translucent and wrinkled appearance upon drying, and a sweet taste when immature. Standard sweet corn is a mutant of the dent corn with a mutation at the *sugary* (*sy*) locus. This mutation causes the endosperm to accumulate about two times more sugar than field corn. New mutants have been developed – *sugary enhanced* (*se*) and *shrunken-2* (*sh2*) or supersweet corn. Some sweet corn varieties are unable to convert sugar to starch. Sweet corn is grown as a winter crop in the southern US, especially Florida.
- 6 Waxy corn.** Waxy corn has a uniformly dull appearance. Instead of amylose, the starch of waxy corn consists of amylopectin, the result of *waxy* (*wx*) mutation. Ordinary corn consists of about 78% amylopectin (a high molecular weight branched chain) and 22% amylose (a low molecular weight straight chain).
- 7 Pod corn (*Z. mays tunicata*).** Pod corn has primitive features, each kernel being enclosed in a pod or husk, before the entire ear is enclosed in husks like other corns. Pod corn versions of the other types of corn (e.g., flint pod corn, dent pod corn) exist.

Similarly, corn that is indigenous to the USA (excluding sweet corn and popcorn) may be classified in up to nine or 10 races. Of these, the most important are the Corn Belt dents, the southern dents, and the northern dents.

## Germplasm resources

Over 13,000 accessions of corn are kept in storage at the Centro Internationale de Mejoramiento de Maiz y Trigo (CIMMYT) in Mexico, with duplicates of these accession held in the US (at the US National Seed Storage Laboratory, Fort Collins, Colorado), Columbia, and Peru. Many heterotic populations have been identified in maize, the most well known and exploited including “Reid” × “Lancaster” and “Lancaster” × “Stiff Stalk”. Others are the European flint × dents, “Tuxpeno” × “ETO”, and “Pantap × Suwan I”.

## Cytogenetics

Corn (*Zea mays* L.) is a diploid ( $2n = 20$ ) and a monocot of the family Poaceae (Gramineae), or grass family. The genus has four species: *Z. mays* (cultivated corn and teosinte), *Z. diploperennis* Iltis et al. (diploperennial teosinte), *Z. luxurians*, and *Z. perennis* (perennial teosinte). Of these four species, only *Z. mays* is widely grown commercially in the US. The closest generic relative of *Zea* is *Tripsacum*, which has seven species, three of which are known to occur in the US. Teosinte occurs in the wild in Mexico and Guatemala. Cultivated corn has 10 pairs of chromosomes ( $n = 10$ ). However, plants with 1–8 sets of chromosomes have been developed for various purposes.

In addition to the autosomes or normal or standard chromosomes (A-chromosomes), maize has supernumerary elements such as the B-chromosomes (also called supernumerary chromosomes). The role of the B-chromosomes in the cell varies from being practically a nuisance to having some definite function, depending on the organism. However, when the B-chromosomes number 10–15 or more, certain abnormalities may occur (e.g., reduced fertility, decreased vigor, aborted pollen, defective kernels). Seed is rarely produced with the presence of a B-chromosome number in excess of 25. Maize B-chromosomes are among those most widely studied in plants. They are suspected to influence the frequency of crossing over, among other roles. Thousands of translocation events have been described in maize. They are used for locating genes on chromosomes.

Monoploids (haploids) may arise spontaneously by pathogenesis (the unfertilized egg develops into a plant). Occasionally, paternal haploids develop by androgenesis. Haploid generation via these systems has low frequency of occurrence. The average frequency of this event in corn is estimated at one per 1,000 kernels. These lines

may be used to develop homozygous diploid inbred lines for hybrid production. They can also be used to convert inbred lines with male fertility to male-sterile cytoplasm. Tetraploid corn was shown to have gigas features (e.g., regarding leaves, tassels, ears, kernel size) but with reduced fertility. Tetraploid yellow corn also produces about 40% higher carotenoid pigment content than the diploid parent. Barbara McClintock conducted extensive cytological work on maize. She developed a complete primary trisomic series, using triploids, of which only trisomy 5, 7, and 8 can be distinguished and characterized phenotypically. Primary trisomics have been used to assign genes to specific chromosomes.

Intergeneric crosses between cultivated corn and related genera, teosinte (*Euchlaena* spp.) and gamma-grass (*Tripsacum* spp.), have been accomplished. The success is more common with annual teosinte (*Z. mexicana*) in which the annual strains ("Chalco", "Durango", "Florida" varieties) readily cross with cultivated corn. *Tripsacum* may be diploid ( $2n = 36$ ) or tetraploid ( $2n = 72$ ). It is a potential source of resistance to many diseases and insect pests of corn. The  $F_1$  is backcrossed to maize to remove all the *Tripsacum* chromosomes, leaving a plant that exhibits a pure maize phenotype. However, yield is reduced and so is agronomic suitability, making this introgression, overall, less attractive at the present time.

In addition to the 10 A-chromosomes, corn may have additional chromosomes, called B-chromosomes or supernumerary chromosomes, that are genetically inert (have no known impact on normal growth).

## Genetics

Corn is one of the plants that has been genetically widely studied. Hundreds of mutations have been identified in corn that impact traits such as plant height, endosperm characteristics, plant colors, insect resistance, disease resistance, stalk strength, and many other traits. Some of the significant genetic effects are discussed here.

### Xenia

**Xenia** is the immediate effect of pollen on the developing kernel. It may be observed when two varieties differing in a single visible endosperm trait are crossed. Xenia occurs when the trait difference is conditioned by a dominant gene present in the pollen. However, when dominance is incomplete, xenia will occur when either variety is the pollen parent. Xenia is important because

endosperm characteristics distinguish some of the major corn groups. For example, starchy endosperm is dominant over sugary (sweet) and waxy endosperm. A cross of starchy  $\times$  sugary exhibits xenia. Similarly, a cross of shrunken  $\times$  non-shrunken endosperm, waxy  $\times$  non-waxy endosperm, purple  $\times$  colorless aleurone, and yellow  $\times$  white (colorless) endosperm, all exhibit xenia.

Whereas xenia may result from simple dominance gene action, the effect is different in some instances. In the cross of flinty  $\times$  floury endosperm, the  $F_1$  is flinty ( $FFf$ ). However, the reciprocal cross of floury  $\times$  flinty endosperm produces an  $F_1$  with floury endosperm ( $ffF$ ), indicating the ineffectiveness of the dominant allele ( $F$ ) to overcome the double recessive ( $rr$ ) floury genes. Similarly, xenia in aleurone color depends on the combined action of five dominant genes (designated  $A1, A2, C, R$ , and  $Pr$ ).

### *Chlorophyll varieties*

Numerous leaf color abnormalities that affect the corn plant in both the seedling and mature stages have been identified. Chlorophyll-deficient mutations cause a variety of leaf colors such as albino, virescent, and luteus (yellow). Mature plants exhibiting golden, green-striped, and other leaf patterns are known.

### *Transposable elements*

Genomes are relatively static. However, they evolve, albeit slowly, by either acquiring new sequences or rearranging existing sequences. Genomes acquire new sequences either by mutation of existing sequences or through introduction (e.g., by vectors, hybridization). Rearrangements occur by certain processes, chiefly genetic recombination and transposable genetic elements (see Chapter 12).

The ancient allotetraploid origins and the presence of large numbers of transposable elements makes the maize genome complex. However, it is these very features that make the maize plant suitable for functional genomics studies. The number and variety of transposable elements facilitate insertional mutagenesis projects. Further, its allotetraploid-based gene redundancy allows scientists to characterize mutants that may be lethal in a diploid species.

### *Cytoplasmic male sterility (CMS)*

Male-sterility genes are among the most important mutations in corn from the standpoint of breeding.

Designated  $ms_1$ ,  $ms_2$ , . . .  $ms_n$ , over 20 of these genes are known. Sterility controlled by nuclear genes is also known in corn, even though it is not of practical application in hybrid corn production. In corn, two main sources of male sterility, controlled by the cytoplasm known as the T (Texas) and S systems, are known. Though deemed superior to the S, the T system, once most widely used, fell out of favor when overuse predisposed maize production in the US to maize leaf blight in 1920, leading to devastation of the industry. CMS-containing stock is used as seed parents to eliminate the cost of detasseling in commercial hybrid seed production. In addition to sterility genes, fertility-restoration ( $RF_n$ ) genes are used in corn. Some seed corn companies use mechanical detasselling for hybrid seed production.

#### Plant height

Studies have been conducted on the *brachytic 2* (*br2*) gene. This gene drastically reduces plant height. However, its practical exploitation in breeding maize has been limited because of the undesirable effects associated with its use (e.g., reduced broad leaves, delayed maturity, thick stem).

Restriction fragment length polymorphism (RFLP) probes capable of detecting more than 500 polymorphic loci have been developed in maize. A maize RFLP map has been generated.

### General botany

Corn is a monoecious annual and one of the largest of the cereals, capable of reaching 4.5 m in height. The male flowers (staminate) occur in the terminal panicle or tassel at the top of the stalk, while the female inflorescence (pistillate) is borne in the axils of leaves as clusters, called a cob, at a joint of the stalk. Long silks (long styles) hang from the husk of each cob. These pollen tubes are the longest known in the plant kingdom. As pollen receptors, each silk must be individually pollinated in order to produce a fruit or kernel. A fertilized cob (also called an ear) may contain eight or more rows of kernels. Furthermore, a stalk may bear 1–3 cobs.

Corn has a variety of morphological features. Some early maturing types maturing in 50 days may attain a height of 0.6 m and produce 8–9 leaves, while tall late maturing types (330 days) may attain a culm or stalk height of 6 m and bear 42–44 leaves. The hybrid corn varieties grow in the northern USA attain a height of

0.9–2.5 m, bear 9–18 leaves, and mature within 90–120 days. The central Corn Belt hybrids varieties range between 2.5 and 3 m in height, bear 18–21 leaves, and mature in 130–150 days. The varieties used on the Gulf coast and South Atlantic regions are much taller (3–3.6 m), produce more leaves (22–27), and tiller profusely, maturing late (170–190 days).

Corn has both seminal and adventitious roots. The seminal roots may number 3–5 and grow downward at the time of seed germination. The crown or coronal roots arise from the nodes of the stem, about 25–50 mm below the soil surface, and may number between 15 and 20 times as many as the seminal roots. The aerial roots (buttress, prop, or brace roots) arise at nodes on the stem above ground.

The number of rows of grain is variable among varieties, ranging between eight and 28. Each row may contain between 20 and 70 kernels. Most of the corn varieties grown in the US contain 14, 10, 12, or 18 rows of kernels, and average about 500 kernels per ear.

### Reproductive biology

#### Floral biology

Each spikelet consists of two **staminate flowers**. Each flower has three anthers that are pushed out of the spikelet as the filament elongates at anthesis. The exertion of the anthers is followed by opening of these structures to shed the pollen. Complete pollen shed may occur in just a few minutes, or over a longer period. A tassel may shed all of its pollen in 1 day or even over a period of about 1 week. The shed pollen pattern is dependent on the genotype and environmental factors such as temperature, humidity, and air movement. Corn produces pollen grains profusely. A normal plant tassel produces an estimated 25,000 pollen grains per kernel on each ear of corn.

#### Pollination

##### Pollen

Corn pollen is primarily dispersed by wind. Consequently, corn is about 95% cross-pollinated, most of the effective pollination of an ear originating from sources in the immediate vicinity of the ear. Pollen dispersal is favored by warm temperature and low humidity. Under such conditions, pollen dispersal may start at about 3 hours after sunrise in the US Corn Belt and last

for 1–3 hours. High temperatures of about 35°C may kill the pollen grain. Further, once dispersed into the atmosphere, the pollen grain may lose viability due to rapid desiccation within a few minutes.

#### *Pistillate flower*

The **pistillate flowers** are borne on a cob that arises from a husk formed at the sixth or seventh node on the stem below the tassel. The female spikelet occurs in pairs of one fertile ovary and one sterile ovary. This is the reason for the even number of rows of kernels on a corn ear. Occasionally, both ovaries in a pair become fertilized, producing kernels, resulting in crowding and irregular rows of kernels on the ear. A silk grows from the tip of each ovary until it emerges at the tip of the ear husk. Silks at the basal part of the ear usually emerge first. A silk is structurally a stigma and style and hence is receptive along its entire length. Temperature, soil moisture, and soil fertility, affect the rate of silk emergence. Adverse weather such as severe drought may delay or cause complete cessation of silk emergence.

#### *Receptivity of the stigma*

Corn is generally a protandrous plant (the male spikelets usually mature before the female spikelets). For the same plant, pollen shed usually precedes silk emergence by about 1–3 days. Silks are receptive soon after emergence and remain receptive for up to about 10 days. For optimum results, the emerged silk should be pollinated within 3–5 days after first silk emergence. Fertilization usually occurs within 12–24 hours of pollination. High temperatures or low humidities adversely affect stigma receptivity.

#### **Genetic consequences of reproductive biology**

- 1 Heterozygosity.** Being predominantly open-pollinated, natural populations of corn are highly heterozygous and genetically variable. In theory, each kernel on the ear could be produced from the fertilization of an ovule by a different pollen parent. A field of naturally pollinated corn in effect comprises a population of hybrids.
- 2 Xenia.** As previously discussed, xenia occurs in corn.

#### **Common breeding methods**

Three general approaches are used over the world for corn improvement for cultivar development –

germplasm introduction, population improvement, and hybrid breeding. These approaches are used to develop: (i) open-pollinated cultivars; (ii) population improvement cultivars; and (iii) hybrid cultivars. Both intra-population and interpopulation improvements are used in maize breeding.

Mass selection has been used by plant breeders to modify corn characteristics including height, maturity, ear characteristics, and grain yield. The method was used to develop cultivars for new production areas in the Corn Belt. Modified mass selection was used to achieve a 2.9% per cycle average yield gain. Mass selection in corn breeding in the US is now practically non-existent. However, the method continues to be used in developing countries by farmers and plant breeders to produce open-pollinated cultivars.

Other intrapopulation breeding methods used in corn breeding include ear-to-row (half-sib family), modified ear-to-row, half-sib family, full-sib family, modified full-sib family, and selfed family, with varying degrees of success. Genetic gains averaging 3.8% per cycle were recorded using modified ear-to-row, while a selfed family selection average of 6.4% per cycle for grain yield was recorded. Various interpopulation methods (recurrent selection, both half and full sib) have been used.

#### **Development of inbred lines in hybrid corn production**

Sources of inbreds for corn hybrid breeding have changed over the years. Before 1930, corn inbreds were isolated from open-pollinated varieties. Later, breeders used inbreds derived from single crosses, modified single crosses, and three-way crosses. Sometimes, backcrossing is used to enhance any inbred line in a specific way by introducing a gene to correct a deficiency. However, these inbreds were less effective for improving quantitative traits. To overcome this weakness, breeders resorted to developing inbred lines from recurrent selection populations that have been improved for specific quantitative traits (e.g., grain yield, stalk quality, disease resistance). Superior inbred lines are also produced by crossing other inbred lines with superior complementary traits and then selecting from the progeny.

Inbreds are developed by artificially controlled pollination. The  $F_1$  plants of a cross are hand-pollinated, followed by pedigree selection through 5–7 generations, or until a desirable level of uniformity in appearance and performance is achieved. The inbred lines are evaluated for combining ability. Single-cross hybrid corn is most commonly used in commercial production.

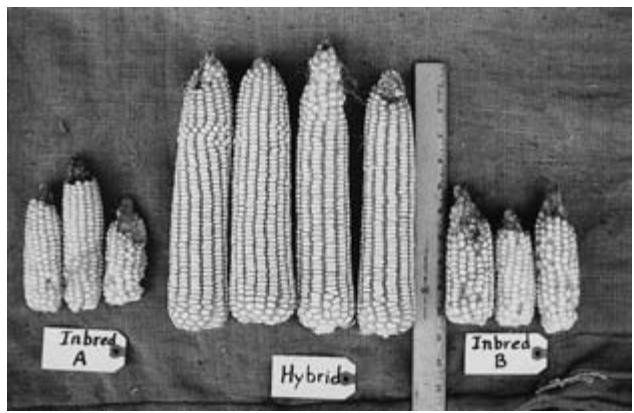


F. J. Betrán

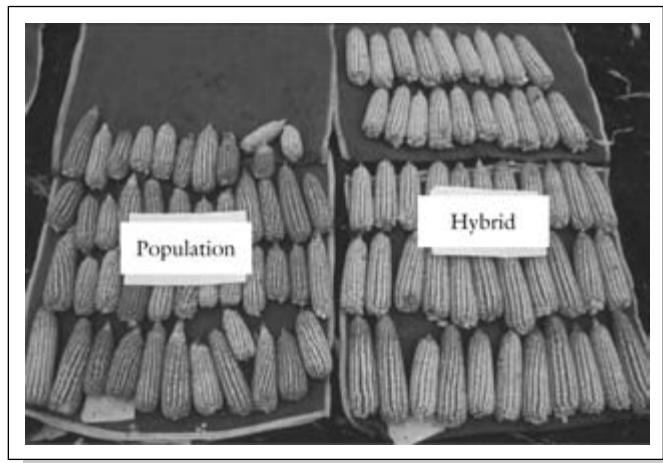
Texas A&amp;M University, College Station, TX 77843, USA

## Industry highlights

### *Hybrid breeding in maize*



**Figure 1** Heterosis in maize: A × B hybrid ears in the middle and corresponding parental inbreds A and B in both sides.



**Figure 2** Relative grain yields of an open-pollinated population and a single-cross hybrid developed from selected inbreds from the same population.

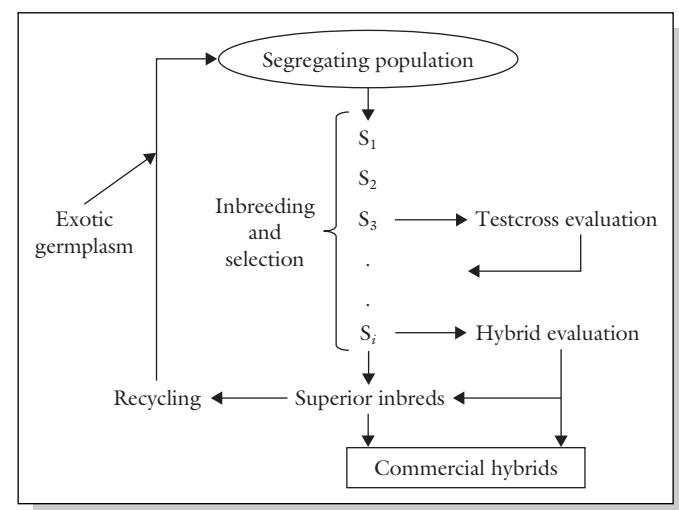
are therefore cyclical creating second-, third-, fourth-, etc. generation recycled improved inbreds (Figure 3). The incorporation and introgression of exotic germplasm brings new desirable alleles and genetic diversity to this recycling of elite lines (Goodman et al. 2000). A backcross or multiple backcross to the best parental inbred is used commonly to increase the probability of maintaining favorable combinations of alleles (Troyer 2001). Maize breeders use multiple trait, multistage, and multiple environment selection methods (Betrán et al. 2003). Multiple environment and multiyear inbred general combining ability values with

#### **Introduction**

Maize is a cross-pollinated species that shows high heterosis (i.e., superior performance of crosses relative to their parents) for grain yield (Figure 1). This high expression of heterosis is exploited in maize hybrids and constitutes the foundation of the maize seed industry. Maize hybrids were first developed in the United States in the mid-1930s and by the early 1960s practically all the maize area in the US was planted to hybrids (Duvick & Cassman 1999). Improved productivity and selection gain with the use of hybrids has stimulated increased investment in hybrid development, resulting in impressive genetic progress (Figure 2). Shull (1909) outlined the pure-line method in maize breeding suggesting the use of self-fertilization to develop homozygous lines that would be of use in hybrid production. This combination of inbreeding and hybridization constitutes the basis of maize improvement. The general process to develop maize hybrids starts with the creation of a source segregating breeding population that it is used to develop inbred lines through inbreeding and selection (Figure 3). Selected inbreds are then evaluated in hybrid combinations across locations to select superior hybrids and to estimate their combining abilities. The following is a brief description of the main components of this process from the starting breeding population to the commercial hybrids.

#### **Source breeding populations**

Different type of segregating populations can be used as the source in line development: open-pollinated cultivars (OPC), synthetic cultivars, single crosses, backcrosses, double crosses, related line crosses, and exotic germplasm. Overall, major emphasis goes to the use of breeding populations created by hybridization of complementary inbreds and the selection of progenies possessing the desirable traits from both parents (Hallauer 1990). Selection within  $F_2$  and backcross populations using pedigree breeding is the most important breeding method to develop maize inbreds. Breeding programs that emphasize pedigree selection within populations developed from elite inbred lines



**Figure 3** General scheme for the development of maize inbreds and hybrids.

depression also show high heterosis (e.g., grain yield). Vigor, plant size, grain yield components, and grain yield are reduced while time to flowering and incidence of barrenness increase with inbreeding (Hallauer & Miranda 1988). The improvement of inbred lines over the last decades has reduced the inbreeding depression considerably in temperate maize. Therefore, source germplasm oriented to develop hybrids has been selected for low inbreeding depression (particularly in the female parent because high yielding lines reduce the cost of the hybrid seed) in addition to high combining ability.

#### Breeding methods to develop inbreds

Development of inbred parents can follow different breeding methods such as pedigree breeding, backcrossing, bulking, single-seed descent, double haploids, etc.

Pedigree breeding is the most widely used breeding system to develop maize inbreds. Typically, specific crosses are made between inbred lines, and then self-pollination is applied to the  $F_1$  and subsequent generations to develop inbred lines that are superior to either parent (transgressive segregants) through genetic segregation and recombination. Selection is applied among progeny rows and among plants within  $S_1$  families. It is common to have replicated nurseries for the  $S_1$  families exposed to different disease, insect, or abiotic stresses. This process of selfing and selection is repeated in successive generations ( $S_2, S_3, S_4, S_5, \dots, S_n$ ) until homozygous elite inbreds are developed. Effective phenotypic selection and greater selection intensity can be applied in initial inbreeding stages for traits with high heritability such as pest resistance, maturity, morphological traits, etc.

The backcross breeding method is used widely in maize breeding to transfer one or a few traits/genes from the donor parent to the recurrent and most desirable parent. With the advent of genetically modified organisms, major emphasis is devoted to accelerate backcrosses to transfer the transgenes to elite inbreds. The use of DNA molecular markers has facilitated both the speed and accurate recovery of the recurrent parent, and the reduction of linkage drag.

The bulk method, where the seeds for each selfing generation are harvested in bulk, and single-seed descent, where one or a few seeds from each genotype are advanced each generation until approximate fixation is reached, are also used because of their simplicity and low space requirements. Double haploids derived from maternal (e.g., *stock6*) or paternal (e.g., *indeterminate gametophyte, ig*) gametes have been used to derive homozygous inbred lines instantaneously (Birchler 1994). However, this method of developing inbred lines has not been used extensively because of the absence of any possibility for phenotypic evaluation and selection, which generates an unselected, large sample of inbreds that needs to be evaluated for combining ability.

#### Breeding methodologies

Improvements of technologies such as off-season nurseries, managed environments for screening against biotic and abiotic stresses, adoption of experimental equipment (combines, planters, computers, etc.), and applications of molecular tools and biological research have increased the accuracy and efficiency of inbred development. Experimental screening techniques have been developed to increase heritability, such as in artificial insect infestation, disease inoculation, and environments managed for higher plant densities or specific abiotic stress factors.

validated performance in hybrids are considered carefully for choosing parents to start breeding populations. The hybrid testing program provides information about the best inbreds to initiate breeding projects.

#### Development of parental inbreds

##### Inbreeding and inbreeding depression

Hybrid development requires the development of parental inbred lines (Figure 3). The inbred parents used to produce the hybrids are developed through a process of inbreeding and selection. The consequence of inbreeding is the increase in homozygosity that leads to homogeneous expression of traits and to inbreeding depression (i.e., loss of vigor and productivity). Self-pollination is the most common and fastest system of inbreeding. As inbreeding reduces the genetic variation within families and increases the genetic variation among families, the efficiency of selection among lines increases while it decreases within lines. The level of inbreeding depression depends on the trait. Traits that show high inbreeding

depression also show high heterosis (e.g., grain yield). Vigor, plant size, grain yield components, and grain yield are reduced while time to flowering and incidence of barrenness increase with inbreeding (Hallauer & Miranda 1988). The improvement of inbred lines over the last decades has reduced the inbreeding depression considerably in temperate maize. Therefore, source germplasm oriented to develop hybrids has been selected for low inbreeding depression (particularly in the female parent because high yielding lines reduce the cost of the hybrid seed) in addition to high combining ability.

### *Development of hybrids*

#### Heterotic groups

Heterosis is related to the level of heterozygosity. If two inbreds are crossed, heterosis is a function of the dominance in those loci with different alleles in the inbreds. Therefore, the identification and development of heterotic groups of elite inbreds having different alleles at loci regulating productivity can contribute to hybrid performance. A heterotic group is germplasm that when crossed to germplasm from another heterotic group, tends to exhibit a higher degree of heterosis than when crossed to a member of its own group (Lee 1995). Heterotic patterns, which are composed by two reciprocal heterotic groups, were empirically established through testing and choice of lines to start breeding populations. Performance of lines in hybrids has been the main criteria of classification into heterotic groups: A × B hybrids were superior to A × A or B × B hybrids. Heterotic patterns for temperate maize are well established (e.g., "Reid's Yellow Dent" × "Lancaster Sure Crop" in the USA, and European flints × US dent in Europe). After establishment, heterotic patterns have been enhanced and optimized through selection and recombination. Multiple heterotic patterns have been developed as a result of intensive elite line recycling and specific emphasis across breeding programs. Testcross performance with representative testers has been used to group large number of inbreds to known heterotic groups. Recently, DNA molecular markers have been effective for assigning inbreds to heterotic groups (Melchinger 1999). The enhancement of heterotic response is improved by subsequent cycles of inbred line development. Increasing degrees of heterosis are observed after several cycles of hybrid selection due to increasing divergence of allele frequencies and selection of complementary alleles in the heterotic groups. In line recycling and in the development of source breeding populations, crosses among elite lines from the same heterotic group are preferred. Heterotic response is heritable and inbreds have heterotic reactions similar to their parents (Troyer 2001).

#### Correlation between inbred and hybrids, and hybrid prediction

The number of potential single crosses to evaluate increases substantially with the number of parental inbreds. The possibility of using inbred line information, as indicative of hybrid performance, is desirable to reduce the number of hybrid evaluations. The correlation between parental inbreds and hybrids depends on the trait. In general, the correlation is relatively high for some additively inherited traits (e.g., plant morphology, ear traits, maturity, quality characters) but is relatively low for grain yield. The correlation for grain yield has been consistently positive and significant but not high enough to predict hybrid performance. The correlations between parental genetic diversity estimated with molecular markers, pedigree, or phenotypic traits and hybrid performance also have been too low to have predictive value (Melchinger 1999). Methods based on linear mixed models have been adapted to maize to predict performance of inbreds in untested environments or hybrid combinations (Bernardo 1999). Although these recent approaches facilitate hybrid selection, hybrid testing is required ultimately to identify the inbreds with the best breeding values.

#### Hybrid testing and screening

Hybrid testing in several environments representative of the target area is executed in several testing stages. A good example of testing stages within a commercial breeding program is outlined by Smith et al. (1999):

- Stage 1** Testcross performance of experimental lines in a few locations (e.g., five).
- Stage 2** Hybrid evaluation of selected lines in more hybrid combinations and locations (e.g., 20).
- Stage 3** Hybrid evaluation in c. 50 locations on research plots in several hybrid combinations.
- Stage 4** Evaluation of best precommercial hybrids in c. 75 research plot locations and c. 200–500 on-farm locations.
- Stage 5** Hybrid performance verification in c. 75 research plot locations and 300–1,500 on-farm strip plot tests.

Efforts are allocated in preliminary tests to evaluate as many hybrids as possible in a few locations with intensive selection, leaving relatively few hybrids to proceed to the more advance precommercial stages. As the numbers of lines to be tested at various stages of inbreeding increase over time, their evaluation in all possible hybrid combinations is not feasible. Therefore, testcrossing with appropriate testers has been adopted extensively to evaluate the relative combining ability of experimental inbred lines. The most common tester used is an elite inbred line from the opposite heterotic group. Testers with low frequency for favorable alleles (e.g., susceptible to diseases) are also used. The level of inbreeding when testcross evaluation is conducted varies among breeders and depends on the traits under consideration and the effectiveness of visual selection. Two basic systems are used: late testing and early testing. In late testing, hybrid evaluation is delayed until advanced generations of selfing (e.g.,  $S_5$  or  $S_6$ ), assuming that selection for additively inherited traits, and seed yield during inbred line development, will assist in reducing the number of lines for testcross evaluation. In early testing, evaluation of hybrid performance is conducted in early generations of inbreeding ( $S_1$ – $S_3$ ). Approximately 60% of the maize breeders in the USA evaluate new lines in testcrosses in  $S_3$  and  $S_4$  (Bauman 1981). Characterization and selection of inbreds is a sequential testing with some lines discarded either through early testing or by performance *per se* in early generations, and others discarded later by general and specific combining ability in hybrid combinations.

Evaluation of hybrids in late stages emphasizes wide area evaluation with multiple environments. This extensive evaluation permits the selection of hybrids for adaptation and stability. Managed environments may include cultural practices such as different densities, planting dates, drought stress, fertilization levels, tillage, crop rotation, etc. Environments are representative of the most common farmer field conditions. More weight in decision-making is applied to those environments more correlated with the target environment. Traits that are emphasized during hybrid screening depend on the target environments. In the US, desired traits include high grain yield, low moisture content at harvest, good standability, ear retention, stability over years, drought tolerance, stay-green, disease and insect resistance, early vigor, and proper maturity (Troyer 2001). The hybrid seed industry has concentrated on extensive and efficient testing, improved mechanization, reliable data collection and analysis, and high selection intensities (Coors 1999). After all the processes described, only a very small percentage of the experimental hybrids tested make it to commercial hybrids.

### References

- Bauman, L.F. 1981. Review of methods used by breeders to develop superior corn inbreds. *Proc. Annu. Corn Sorghum Ind. Res. Conf.* 36:199–208.
- Bernardo, R. 1999. Prediction of single-cross performance. *In: The genetics and exploitation of heterosis in crops* (Coors, J.G., and S. Pandey, eds), pp. 269–276. American Society of Agronomy, Madison, WI.
- Betrán, F.J., M. Bänziger, and M. Menz. 2003. Corn breeding. *In: Corn: Origin, history, technology, and production* (Smith, W., F.J. Betrán, and E. Runge, eds). John Wiley & Sons, New York.
- Birchler, J.A. 1994. Practical aspects of haploid production. *In: The maize handbook* (Freeling, M., and V. Walbot, eds). Springer-Verlag, New York.
- Coors, J.G. 1999. Selection methodologies and heterosis. *In: The genetics and exploitation of heterosis in crops* (Coors, J.G., and S. Pandey, eds), pp. 225–246. American Society of Agronomy, Madison, WI.
- Duvick, D.N., and K.G. Cassman. 1999. Post-green revolution trends in yield potential of temperate maize in the north-central United States. *Crop Sci.* 39:1622–1630.
- Goodman, M.M., J. Moreno, F. Castillo, R.N. Holley, and M.L. Carson. 2000. Using tropical maize germplasm for temperate breeding. *Maydica* 45:221–234.
- Hallauer, A.R. 1990. Methods used in developing maize inbreds. *Maydica* 35:1–16.
- Hallauer, A.R., and J.B. Miranda, Jr. 1988. Quantitative genetics in maize breeding, 2nd edn. Iowa State University Press, Ames, IA.
- Lee, M. 1995. DNA markers and plant breeding programs. *Adv. Agron.* 55:265–344.
- Melchinger, A.E. 1999. Genetic diversity and heterosis. *In: The genetics and exploitation of heterosis in crops* (Coors, J.G., and S. Pandey, eds), pp. 99–118. American Society of Agronomy, Madison, WI.
- Shull, G.H. 1909. A pure line method of corn breeding. *Rep. Am. Breeders Assoc.* 5:51–59.
- Smith, O.S., K. Hoard, F. Shaw, and R. Shaw. 1999. Prediction of single-cross performance. *In: The genetics and exploitation of heterosis in crops* (Coors, J.G., and S. Pandey, eds). American Society of Agronomy, Madison, WI.
- Troyer, A.F. 2001. Temperate corn – Background, behavior, and breeding. *In: Specialty corn* (Hallauer, A.R., ed), pp. 393–466. CRC Press, New York.

Consequently, it is critical that the inbred lines from which the crop will be produced, are high yielding. Selection at each generation should also be based on superior performance regarding disease and insect resistance, grain quality, lodging resistance, and other traits that support the breeding objectives.

### Establishing a breeding nursery

#### Field nursery

A proper layout of the breeding nursery facilitates a breeding program. A comprehensive corn breeding program will have selfing nurseries for the maintenance

of inbred lines, increasing inbred lines, and inbred line selection and breeding. There will also be crossing blocks (for single, three-way, and double crosses, and topcrosses) and recurrent nurseries.

The plants should be planted in rows with alleys for easy movement of the breeder. Parents for crossing should be arranged in paired rows. Where several crosses have one row, the common parent may be used as the male parent and planted adjacent to a block of the female parent.

#### Other nurseries

As previously indicated, corn breeders use field nurseries exclusively for breeding. Greenhouses and growth

chambers may be used for experimental purposes, but they are not amenable to practical breeding of corn. Breeders in cooler climates may use winter nurseries for additional breeding work.

## Special environment

No special environmental conditions are needed if the breeding nursery is established in an area adapted for corn production. Temperature, soil moisture, and humidity are critical for proper flowering and pollination of corn. Temperatures above 30°C and low humidity adversely impact pollen viability. Similarly, drought causes delay in silk emergence relative to tassel development. It is important to have a source of supplemental irrigation to provide moisture at critical times to avoid tassel firing and male flower abortion.

## Artificial pollination

### Materials and equipment

Crossing corn requires the use of simple equipment, including a knife and stapler. Materials include tassel bags, ear shoot bags, paper clips, a pencil, tags, and a carrying case or apron with pockets. The bags should be water repellent.

### Emasculation

Corn is a monoecious plant. The female flower should be covered and protected from unwanted pollen. The ear shoot should be covered with an ear bag before the silks emerge from the husk tips. Once covered, artificial pollination should be undertaken within about 3 days, otherwise the silks should be trimmed to prevent them from growing out of the bag and becoming contaminated. Further, trimming the silks to within 2 cm of the husk tip is usually done the day before pollination.

### Pollination

Where controlled pollination is desired, the tassel is also covered with a tassel bag and held securely in place by stapling or using paper clips at least 1 day before pollination. This prevents contamination from unwanted sources, and also saves desirable pollen that would otherwise have been lost to the wind. Copious amounts of pollen are available on the second or third day following pollen dehiscence.

When selfing, the pollen may be taken directly from the tassel and deposited on the silk. In cross-pollination, the pollen may be collected in a bag by shaking the tassel and dusting the pollen on the ears. A pollen gun may also be used for multiple pollinations.

The optimum time for pollination depends on temperature and humidity, but is usually about 3 hours after sunrise. The ear bag should be replaced immediately after pollination to avoid contamination. However, under some conditions, covering the ear shoot for a long time may cause ear tip rot. Pollination information, including parents involved in the cross and date of pollination, is written on the envelopes.

## Natural pollination

Controlled artificial pollination (and sometimes natural pollination) is used to develop commercial hybrid cultivars. Once developed, hybrid seed (or synthetic seed) production is conducted in the field using the wind as the agent of pollination. To maintain the genetic purity of the new cultivar, and to meet certification standards for corn, seed production must be conducted in isolation from other corn cultivars to keep contamination below the mandated 1%. It is also mandated that an isolation block for seed increase be located no less than 200 m from corn of a different color or texture.

## Common breeding objectives

The breeding objectives presented below are not in order of importance.

**1 Grain yield.** Grain yield of corn is a major breeding objective for this crop, which is among the major crops that feed the world. It is estimated that corn yield in the US has increased about 340% between 1940 and 1990. This increase is attributed to the development of cultivars with high yield potential. Over the period indicated, producers switched from open-pollinated cultivars to higher yielding cultivars. Yield is a complex trait and depends on the genes that are associated with basic physiological processes, plant structure, and morphology. Yield components of corn include number of ears, kernel rows and kernels per row, kernel test weight, and shelling percentage.

Days to maturity, standability, and resistance to environmental stress (biotic and abiotic) also affect grain yield of corn. Features of modern corn hybrid cultivars compared to older cultivars were summarized

by W. A. Russel as: longer grain filling period (i.e., early flowering, delayed senescence), rapid grain filling, increased sink size (i.e., more kernels per unit area, larger kernels, induced barrenness), higher harvest index and shelling percentage, short plant and tassel, upright leaves, shorter anthesis to silking interval, better standability, and better tolerance to abiotic stress. Further, modern-day cultivars are more efficient at exploiting high soil nitrogen, and also are amenable to cultivation under high plant density.

Another approach to yield breeding corn is to select for prolificacy (plant has more than one ear). Multiple ears per plant are known to improve standability under adverse conditions.

**2 Yield stability.** It is desirable for cultivars to have stable yield in the production region.

### 3 Agromorphological traits:

- (a) **Lodging resistance.** To accomplish lodging resistance or good standability in corn, the specific focus of corn breeders is improved stalk quality. Lodging in corn may be at the root level or the stalk level. Rind thickness and crushing strength are indicators used to select stalk strength. Tall genotypes will have ears too high on the stem and become prone to lodging. Stalk breakage is caused by boring insects, such as the European corn borer (*Ostrinia nubilalis*). One of the most successful applications of genetic engineering in plant breeding is the development and deployment of *Bt* corn varieties. These genetically modified (GM) varieties are resistant to attack by the borer. As previously indicated, the use of the *brachytic 2* gene in breeding is limited.
- (b) **Resistance to ear dropping.** When harvesting is delayed, mature corn may break off and drop to the ground before harvesting. Sometimes, the impact of the combine harvester may cause such ear drop to occur, leading to significant field losses. Disease may promote ear dropping. Genotypes with large, heavy ears are more susceptible to ear dropping. Long ear shanks are more susceptible to corn borer attack and are structurally weak and more prone to ear dropping. Plant breeders hence select for shorter and stronger shanks. However, it is desirable for the ear to bend downward when mature to accelerate drying.
- (c) **Husk covering.** The ear of corn is covered by a husk that has a protective role (against pests and weather) and also promotes rapid drying. Breeders select for a long husk or complete covering of the tip.

- (d) **Dry-down.** The husk and grain must dry before harvesting of corn. The more rapid the drying process, the better, as it allows for early harvesting of the crop.

### 4 Adaptation:

- (a) **Early maturity.** Maturity in maize is a qualitative trait. Corn is frost insensitive. Hence, early maturity is an important objective in temperate climates. Early maturing cultivars are usable in areas of shorter growing season and higher altitudes. Maturity traits used in breeding corn includes days to silk, tassel, and brown husk. The black layer is an indicator of physiological maturity.
- (b) **Drought resistance.** Corn is adapted to both temperate and tropical climates. However, excessive temperature and drought at flowering promotes reduced seed set and barrenness. Drought-resistance traits in maize include small male tassel, small leaf area, prolificacy, leaf elongation, heat tolerance, high abscisic acid, and low temperature.
- (c) **Cold tolerance.** Cold tolerance is important for germination under unseasonably cold soil temperatures, and also for quick dry-down in cold weather.

**5 Disease resistance.** Corn is plagued by over 100 diseases and over 100 insect pests worldwide, some of which have global importance.

- (a) **Seed rots and seedling blights.** Corn is plagued by many pathogens that cause rots and blights, the common ones including *Diplodia maydis*, *Fusarium moniliforme*, and *Penicillium* spp.
- (b) **Root, stalk, and ear rots.** Agents of common root rots include *Fusarium* spp. and *Pythium* spp. Common stalk rots include *Diplodia* rots and *Fusarium* rots. Corn is also attacked by ear rots caused by *Diplodia* and *Aspergillus*.
- (c) **Leaf blight or spots.** Leaf blights or spots of economic importance include bacterial wilt (caused by *Bacterium stewartii*) and *Helminthosporium* leaf diseases like *H. maydis* (causes southern corn leaf blight), *H. carbonum*, and *H. turcicum* (causes northern corn leaf blight). Common rusts include common corn rust (caused by *Puccinia sorghi*) and southern corn rust (caused by *P. polysora*). Important viral diseases of corn include maize dwarf mosaic virus and maize chlorotic dwarf virus.
- (d) **Smuts.** Important smuts include common smut caused by the fungus *Ustilago maydis* or *U. zeae*, and head smut caused by *Sphacelotheca reiliana*.

**6 Insect resistance.** Most of the research emphasis has been on breeding resistance to the European corn

borer. A chemical found in corn with the acronym DIMBOA, is reported to impart resistance to this pest to plants. Insect pests of economic importance include the following:

- (a) **Soil-inhabiting insects.** Important soil-inhabiting insect pests of corn include corn rootworms (*Diabrotica longicornis*), black cutworms (*Agrotis ypsilon*), wireworms such as *Melanotus cibulosus*, and the sugarcane beetle (*Euetheola rugiceps*).
- (b) **Leaf, stalk, and ear insects.** An important pest of corn is the corn earworm (*Heliothis armigera* or *H. zea*) and the European corn borer (*Pyrausta nubilalis*).
- (c) **Stored corn grain insects.** Major storage pests of corn include the rice or black weevil

(*Sitophilus oryzae*), and the Angoumois grain moth (*Sitotroga cerealella*).

- 7 **Product quality.** One of the major achievements in cereal grain protein improvement is the breeding of quality protein maize (QPM), a high protein corn bred by incorporating the *opaque-2* gene. Also, high oil content is a goal of some breeding programs. High oil corn is of high industrial value. Selection in the long-duration Illinois High Oil Strain Experiment has increased oil content from the initial 4.7% to over 21% over several decades. Other end-use quality breeding goals include improvement of milling quality, and enhancing the sugar content of sweet corn (by incorporating the *shrunken-2* (*sh2*) gene in breeding).

## References and suggested reading

- Acquaah, G. 2005. Principles of crop production. Theory, techniques and technology. Prentice Hall, Upper Saddle River, NJ.
- Kehra, A.S., B.S. Dhillon, and D.S. Multani. 1997. Corn. In: Genetics, cytogenetics, and breeding of crop plants. Vol. 2. Cereal and commercial crops (Bhal, P.N., P.M. Salimath, and A.K. Mandal, eds). Science Publishers, Inc., Enfield, NH.
- Russell, W.A. 1991. Genetic improvement of maize yields. Adv. Agron. 46:245–298.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Corn is a dioecious plant.
- 2 E. M. East founded the Hi-Bred Corn Company of Iowa.
- 3 Corn belongs to the family Poaceae.
- 4 B-chromosomes occur in corn.

### Part B

Please answer the following questions:

- 1 What is xenia?
- 2 Give three of the states in the Corn Belt of the USA.
- 3 Discuss the types of corn.
- 4 What does QPM stand for?

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the major objectives in breeding pest resistance in corn.
- 2 Discuss the major agromorphological traits of importance in corn breeding.
- 3 Discuss the main methods of breeding corn.



# Breeding rice

## Taxonomy

Kingdom	Plantae
Subkingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Liliopsida
Subclass	Commelinidae
Order	Cyperales
Family	Poaceae
Genus	<i>Oryza</i> L.
Species	<i>Oryza sativa</i> L.

## Economic importance

Rice accounts for about 20% of the world's total grain production, second only to wheat. It is the primary staple for more than 50% of the world's population. An estimated 90% of production and consumption of rice occurs in Asia where the per capita consumption is about 45 kg, compared to a world average of about 27 kg and a US average of about 9 kg. World consumption in 1993/94 was 358.5 million metric tons and has increased steadily to an estimated 408.8 million tons in 2002/03. Similarly, the area devoted to rice increased during the same period from 145.2 to 155.1 ha in 1990/2000, with a slight drop thereafter. Production trends mirror the area cultivated. Rice production in the USA is dominated by six states – Arkansas, California, Louisiana, Texas, Mississippi, and Missouri – together accounting for about 99% of the total US production. Other minor producers in the USA include Florida, Tennessee, Illinois, South Carolina, and Kentucky. Arkansas leads the production in the US, with about 100 cwt of rice produced in 2001 on about 1.61 million acres (see Appendix 2 for conversion rates for units).

The world's major producers of rice are China, India, Indonesia, and Bangladesh, which together account for

more than 70% of the world's total production. China, the world's leading producer regularly accounts for about 36% of the world's total production. The expansion in rice production is attributed largely to the impact of the Green Revolution, which was implemented in the 1960s and 1970s in Asia and other parts of the tropical world. Other major producers in Asia include Vietnam, Thailand, Japan, and Burma, together accounting for about 12% of the total world production. Outside of Asia, rice is produced in substantial amounts in countries including the USA, Italy, Spain, Egypt, Australia, and countries in the West African region.

Thailand leads the world in rice exports, accounting for about 25% of the total world exports (about 4 million metric tons). The USA is the second largest exporter of rice, exporting about 2.7 million metric tons. Latin America is the largest market for US rice export. Other major exporters of rice include Vietnam, Pakistan, China, Australia, Burma, Italy, India, and Uruguay.

## Origin and history

The origin of rice is not exactly known. Wild species of rice occur across South and East Asia, including India and south China. Rice was domesticated in the 5th millennium BC. The cultivated species of rice *Oryza sativa* is believed to have derived from annual progenitors found in a wide area extending from the Gangetic Plains, through Burma, northern Thailand, northern Vietnam, and southern China. Rice trade occurred between Egypt, India, and China. The Moors are credited with bringing rice to Spain, from where it was introduced to Italy (in the 15th century) and subsequently to Central America. The African cultigen (*O. glaberrima*) originated in the Niger River delta of Mali in West Africa.

A British sea captain brought rice to the US from Madagascar. The first commercial planting occurred in South Carolina in about 1685. Early rice production concentrated in the southeast US where South Carolina, Georgia, Louisiana, North Carolina, Mississippi, Alabama, Kentucky, and Tennessee were major producers in the early 1800s. However, in the early 19th century, shifts in rice production started in the US, reaching its highest in the late 1900s. Production became more modernized and spread to Arkansas and the Mississippi River delta where the flatter lands permitted larger scale, high-tech production (mechanized, irrigated, etc.). In 1838, South Carolina produced about 75% of the US total rice crop. In 1903, Louisiana and Texas produced 99% of the US crop. However, by 1990, Arkansas, Louisiana, and Mississippi accounted for more than two-thirds of the US rice production, with Texas and California accounting for most of the balance. Rice production in the USA is currently concentrated in the Arkansas grand prairie, the Mississippi River delta, southwestern Louisiana, the coast prairie of Texas, and the Sacramento Valley of California.

## Adaptation

Rice may be described as a semiaquatic plant. Rice is adapted to very wide agroecological zones, ranging from dry to submerged root growing conditions. Four general ecosystems can be identified for commercial rice production around the world, based on elevation, rainfall pattern, depth of flooding, and drainage.

### Rainfed lowland rice ecosystems

This ecosystem is common in densely populated rural regions of the world, where producers face severe economic challenges in addition to burgeoning population. Rice production under these conditions account for about 25% of the world's harvested rice area and 17% of the world's production. Producers prepare the land by puddling the soil or dyking fields to hold water for a variable duration of flooding, according to the rainfall. The soils alternate between flooded to dry conditions during the growing season. Rice is direct-seeded or transplanted into the field.

### Upland rice ecosystems

Upland rice production occurs on well-drained, level to steeply sloping farmlands. These soils are frequently

moisture deficient. Upland production occurs in regions of the world where slash-and-burn agriculture is common. The removal of vegetation from these slopes predisposes the soil to physical deterioration and nutrient depletion. Crop yields are generally low. Upland rice production constitutes about 13% of the world's harvested area and only 4% of the total rice produced. Rice is direct-seeded to non-flooded soils.

### Flood-prone rice ecosystems

Rice production in certain areas occurs on flooded soils throughout the growing season, until harvest time. Rice is direct-seeded or transplanted into flooded fields (50–300 cm deep in water) during the rainy season. Flooded rice production occurs widely in South and South East Asia, and some parts of West Africa and Latin America. Problems of salinity and toxicity from various ions are common in this ecosystem. Crop yields are unpredictable and generally low.

### Irrigated rice ecosystems

The key feature of this system is that moisture is controlled in both the dry and wet seasons. Various methods are used to provide and regulate soil moisture. About 55% of the world's harvested area and 75% of total production occurs in irrigated ecosystem. Production involves the use of modern technology with high production inputs (e.g., fertilizers). Consequently, yields are high, reaching about 5 t/ha in the wet season to about 10 t/ha in the dry season.

### Other classifications

T. Chang of the International Rice Research Institute (IRRI) reported a classification of rice according to topohydrological, edaphic, cultural, and seasonal criteria. According to ecogeographic differentiation, rice may be classified into three races of *O. sativa* as **indica**, **javanica**, and **sinica** (or **japonica**). The japonica race has upland and lowland cultivars, whereas the indica has cultivars that span dryland to deep water and floating cultivars.

### Commercial classes

Rice varieties are primarily classified according to the length of the grain – short (5.5 mm), medium (6.6 mm), and long (7–8 mm). The shorter grained varieties are also called japonica types, and have short,

stiff, lodging-resistant stalks, making them more responsive to heavy fertilization. The longer grained varieties, called indica types, have taller, weaker stems that lodge under heavy fertilization. The US produces mainly indica rice (about 65% of the annual production) primarily in Arkansas, Mississippi, Louisiana, and Texas, while California produces mostly medium- and short-grained rice.

Rice may also be classified in terms of maturity – early maturing (about 120–129 days), midseason (about 130–139 days), or late maturing (about 140 days or more). Rice may be scented (aromatic) or unscented, the two most common scented types being basmati and jasmine. Basmati rice has a distinctive odor, doubles in grain size upon cooking, and is non-sticky (grains remain separate). It is cultivated mainly in the Punjab area of central Pakistan and northern India. Jasmine rice is grown mainly in Thailand and is more preferred by the Asian community in the US. Cooked jasmine rice is soft, moist, and sticky. The stickiness derives from the types of starch in the grain. The endosperm starch of rice may be glutinous or commonly non-glutinous (non-sticky). The glutinous property is conferred by amylopectin-type starch.

### **Germplasm resources**

Over 85,000 rice accessions are maintained at IRRI in the Philippines, representing the largest rice germplasm collection in the world. The Africa Rice Center (ARC) in the Ivory Coast maintains large numbers of accessions, especially *O. glaberrima* genotypes. The US Department of Agriculture (USDA) maintains about 16,000 accessions of rice, including a large number of breeding lines. Genetic testers, developed mostly by Japanese researchers, are available at the Rice Genetics Cooperative.

### **Cytogenetics**

Twenty species of the genus *Oryza* are known, with a basic chromosome number of 12. The genus has six genome groups: A, B, C, D, E, and F. The cultivated species *O. sativa* ( $2n = 2x = 24$ ) has the genome formula AA, whereas *O. glaberrima* ( $2n = 2x = 24$ ) is designated A<sup>a</sup>A<sup>a</sup> because it does not pair well with *O. sativa*. There are two major species of the cultivated rice – *O. sativa* and *O. glaberrima* – the latter being native to Africa and cultivated in West Africa and Central Africa. Other

wild species are *O. alta* ( $2n = 2x = 48$ ; CCDD), *O. australiensis* ( $2n = 2x = 24$ ; EE), *O. longistaminata* ( $2n = 2x = 48$ ; CCDD), *O. brachyantha* ( $2n = 2x = 24$ ; FF), and *O. punctata* ( $2n = 2x = 48$ ; BBCC).

The 22 species are divided into diploids ( $2n = 2x = 24$ ), e.g., *O. australiensis*, *O. barthii*, *O. glaberrima*, and *O. sativa*, and tetraploids ( $2n = 2x = 48$ ), e.g., *O. alta*, *O. punctata*, and *O. grandiglamis*. Rice chromosomes have a tendency to form secondary associations during meiosis of normal diploids. Asynaptic plants occur in nature, as well as haploids, triploids, and tetraploids. They may be induced artificially. Primary trisomics have also been developed by various researchers for rice. Also, chromosome maps have been developed by different workers.

Crossing between plants from different ecogeographic races, or even within a race, may exhibit some hybrid sterility, inviability, or weakness. In practice, plant breeders are able to overcome these sterility problems through rigorous selection to obtain fertile lines. Wild germplasm has proved a useful genetic resource for rice breeding. The wild abortive cytoplasmic male-sterility (CMS) cytoplasm that is commercially used was obtained from *O. sativa* f. *spontanea*.

### **Genetics**

Rice genetics is quite extensive. Color development in rice is complex. Anthocyanin pigmentation is controlled by a complementary genetic system, the CAP gene system. C is the basic gene for chromogen production, while A controls the conversion of chromogen to anthocyanin. P, which is variable, determines the site of pigment expression (e.g., Pg, Pm, Ps, Px, etc., according to the site). The CAP gene system is affected by modifying genes, and occasionally an inhibitor gene in some species. Also, multiple alleles and various dominance levels are associated with C, A, and Pl. Color development is also affected by light intensity, growth stage, and fading and leaching. Color of the different layers of seed coat may be affected by different genes or sets of genes. A smooth hull is desirable for mechanical harvesting and processing. Pubescence on the surface of the blades and hull is controlled by a dominant gene, *glabrous* (gl).

Semidwarfism is desirable in rice and is controlled by a pair of recessive alleles, *sdl*. However, studies indicate that semidwarfism is a complex, quantitative trait. One three-gene unit with cumulative effect controls awning in rice, fully awned being *An<sub>1</sub>An<sub>2</sub>An<sub>3</sub>* (or *An<sub>1</sub>An<sub>2</sub>an<sub>3</sub>*),

while awnless is conditioned by  $an_1an_2an_3$ . Photo-period sensitivity is controlled by one or two genes ( $Se_1$ ,  $Se_2$ ), while one or two dominant genes control shattering. Genes for host resistance to many diseases have been identified, including bacterial blight ( $Xa-1$ ,  $Xa-2$ , etc.), blast ( $Pl_1$ ,  $Pl_2$ , and others), yellow dwarf virus ( $YdV$ ), and brown spot ( $He$  or  $he$ ).

One of the important mutations in rice breeding is the  $sd$ , the recessive gene that conditions semidwarf stature. It was discovered in a Taiwanese cultivar "Dee-geo-woo-gen". Hybrid rice using CMS entails crossing a CMS line ( $rffr-S$ ) with a restorer line ( $RfRfF$ ) to obtain an  $F_1$  ( $Rffr-S$ ). The "wild abortive" or the "Chinsurah Bone II" are the preferred sources of CMS for rice hybrids.

### General botany

An annual grass, rice has erect culms that may reach 1.8 m in some varieties. It produces about five tillers. Rice inflorescence is a loose terminal panicle consisting of spikelets that are self-pollinated. The rice grain is enclosed by the lemma and palea (constitutes the hull) that may be straw yellow, red, brown, or black in color. Depending on the variety, the lemmas may be fully awned, partly awned, tip-awned, or awnless. Hulled kernels vary in length from 3.5 to 8 mm, and are 1.7–3 mm in breadth and 1.3–2.3 mm thick. Furthermore, the kernels may be hard, semihard, or soft-textured. The color of unmilled rice kernel is variable and may be white, brown, amber, red, or purple; lighter colors (white, light brown) tend to be preferred in the US.

### Reproductive biology

Rice has a loose terminal panicle with branches that arise singly or in whorls. It is predominantly self-pollinated with less than 1% outcrossing. A panicle may contain 75–150 spikelets, or even several hundred in some varieties. Rice has perfect flowers that are borne in single-flowered spikelets. The flower consists of two lodicules, six stamens (instead of three as in most cereals), and two plumose stigmas on two styles, surrounded by floral bracts. The floral bracts (lemma and palea) may be straw, yellowish-gold, red, brown, or purplish in color. Further, the lemma may be fully awned, partly awned, tip-awned, or awnless.

Rice is predominantly self-pollinated. Time of anthesis is significantly influenced by environmental factors,

especially temperature, and to a lesser extent by genotype. The entire panicle completes blooming in 4–7 days. Peak anther dehiscence occurs between mid to late morning in tropical regions and at about noon in temperate regions.

### Common breeding methods

Being a self-pollinated species, rice can be improved by any of the methods of breeding for self-pollinated species. Field tests can be conducted to evaluate introductions from various parts of the world to identify adapted varieties for commercial production. International breeding programs and non-profit organizations may engage in breeding genotypes for use in developing countries with few resources for embarking on elaborate breeding programs.

Hybridization is the principal procedure used to generate a segregating population for a breeding program in rice. Depending on the breeding objective and the underlying genetics of the trait of interest, breeders may use any of the breeding methods for self-pollinated species in rice breeding.

Commercial hybrid rice seed production was started in the People's Republic of China in the 1970s, where hybrid varieties occupy more than 50% of the 75 million acres of rice cultivated annually. The key factor in the success of the commercial hybrid program was the discovery of a source of CMS, called the "wild abortive", in a wild rice plant (*O. sativa f. spontanea*). Fertility restoring genes were discovered in indica cultivars. Hybrid rice production is practical in China largely because of the low cost of labor. To enhance flowering, cross-pollination, and seed set, seed producers implement practices such as cutting back the flag leaves of female plants to enhance pollination.

### Establishing a breeding nursery

#### Field nursery

Rice is a warm season species and is successful under a mean temperature of about 20°C or higher. High humidity encourages diseases and hence is undesirable. Heavy-textured soil with impervious subsoil for holding moisture is desirable for rice. The tolerable pH ranges between 4.5 and 7.5. Rice is also a short-day plant.

Rice seeding is accomplished by using ground equipment (broadcast seeders or grain drills). Rice is drilled to



## Industry highlights

### *Breeding rice*

Anna Myers McClung

USDA-ARS Dale Bumpers National Rice Research Center and Beaumont Rice Research Unit, 1509 Aggie Drive, Beaumont, TX 77713, USA

#### ***Introduction***

The US Department of Agriculture-Agricultural Research Service (USDA-ARS) rice breeding program located at the Texas A&M Research and Extension Center at Beaumont has been in operation since the early 1930s. The objective of this program has been to develop superior performing rice cultivars that are adapted to the southern rice growing region of the United States, which includes Texas, Louisiana, Arkansas, Mississippi, and Missouri. Some 80% of this region is dedicated to the production of long grain cultivars, while 20% produces medium grain cultivars, with a small percentage dedicated to specialty rice cultivars for niche markets. In the USA, the conventional market classes of rice are categorized according to grain dimensions and cooking quality, the latter being primarily determined by amylose content and starch gelatinization temperature (Webb 1985). The development of new rice cultivars must include selection for agronomic traits, resistance to disease and insect pests, and grain quality traits (McClung 2002). In addition, cultivars that are produced along the Gulf coast of Texas and Louisiana, where the growing season is relatively long, are also evaluated for second crop potential, which is called the ratoon crop. After the main crop is harvested, a second crop develops from the stubble of the first crop. About 60 days after the main crop harvest, the ratoon crop is cut, producing up to 50% of the first crop's yield.

Most public rice breeding programs use standard pedigree, bulk, and backcross breeding approaches. However, mutation breeding and recurrent selection methods are also used to a limited extent. Many of these programs are now using molecular markers that are associated with traits of economic importance to expedite the breeding process. In addition, there are private breeding programs that are developing rice hybrids and others that are using transgenic technology (McClung 2004).

The development of "Saber" rice (McClung et al. 2004) is an example of a recent breeding project conducted by the Beaumont Rice Variety Development Program. In 1989, a cross was made between "Gulfmont" and an experimental selection, RU8703196. At the time, "Gulfmont" was a new release from the breeding program that was characterized as an early maturing, semidwarf, long grain cultivar having excellent main and ratoon crop yield as well as good milling quality. It is rated moderately resistant to blast disease (caused by *Pyricularia grisea*) and very susceptible to sheath blight disease (caused by *Rhizoctonia solani*) (Bollich et al. 1990), the two most common yield-limiting diseases in the southern rice growing region. RU8703196 is a long grain germplasm source that was released as an improved source of resistance to blast and sheath blight diseases (Marchetti et al. 1995). The F<sub>1</sub> of "Gulfmont"/RU8703196 was then crossed with "TeQing" (PI 536047). This is a medium grain cultivar from China that is characterized as having high yield potential, intermediate height, relatively late maturity, and excellent resistance to blast and sheath blight diseases when grown in the southern US. The objective of the cross was to maintain the plant stature and grain quality of "Gulfmont" while improving its yield and disease resistance.

#### ***Timeline for the development of "Saber" rice***

##### **1989**

Make cross B8910 ("Gulfmont"/RU8703196//TeQing") in the greenhouse and produce 27 F<sub>1</sub> seed (Figure 1). Plant each F<sub>1</sub> seed in separate containers in the greenhouse. Harvest F<sub>2</sub> seed from each F<sub>1</sub> plant separately.

Plant bulk of each F<sub>1</sub>-derived F<sub>2</sub> population in the fall planted nursery in Puerto Rico. Three of the F<sub>1</sub>-derived populations appear to be selfs and are discarded. Strip bulk of F<sub>3</sub> seed from F<sub>2</sub> plants at harvest.

##### **1990**

Plant bulk of F<sub>1</sub>-derived F<sub>3</sub> populations in the winter planted nursery in Puerto Rico. Strip bulk of F<sub>4</sub> seed from F<sub>3</sub> plants.

Plant bulk of F<sub>1</sub>-derived F<sub>4</sub> populations in summer nursery in Beaumont, TX. Strip bulk of F<sub>5</sub> seed from F<sub>4</sub> plants.

##### **1991**

Plant bulk of F<sub>5</sub> population in the summer nursery. Select panicles from a total of 215 plants for the B8910 project based upon field observations of plant height, days to heading, good tillering capacity, and good appearing grain dimensions.



**Figure 1** Rice panicle being prepared for crossing to produce F<sub>1</sub> seed.



**Figure 2** Plant height is measured in yield plots after flowering has occurred.

Regional Nursery (URRN) as entry RU9603178. The URRN is an extensive replicated yield trial that includes a total of 200 entries from four southern rice breeding programs and is planted in Texas, Arkansas, Louisiana, and Mississippi. This study has two sections: advanced entries, which are replicated four times, and preliminary entries, which are replicated twice. RU9603178 is entered into the preliminary trial section based upon limited availability of seed. In addition to the standard yield, agronomic, milling, and cooking quality traits that are evaluated, several states screen all 200 entries in inoculated nurseries for resistance to blast and sheath blight diseases as well as the physiological disorder, straighthead. Two of the states evaluate all entries for ratoon crop yield and for reaction to nine individual races of blast using controlled greenhouse conditions. All of the entries from the Beaumont program that are in the URRN trial also undergo additional yield testing at two to three other Texas locations which are representative of the state's rice growing area. The other four entries from the B8910 cross are planted only in these Texas replicated yield trials.

#### 1992

Plant 215 F<sub>6</sub> panicle rows in the summer nursery. Select 63 rows based upon field observations of plant height, days to heading, good tillering capacity, and good appearing grain dimensions. Harvest five panicles per row.

#### 1993

Project not planted due to field space limitations.

#### 1994

Plant five F<sub>7</sub> panicle rows from each of 63 F<sub>6</sub>-derived families in the fall planted nursery in Puerto Rico. Select 29 rows based upon field observations of plant height, maturity relative to check varieties, good tillering capacity, and good appearing grain dimensions. In each selected row harvest five panicles and bulk harvest the row.

#### 1995

F<sub>8</sub> panicle rows are planted in the summer breeding nursery for each of the 29 families. Ten F<sub>9</sub> panicles are selected from one of the five rows for each of the 29 families.

In addition, the bulk harvested F<sub>8</sub> seed from Puerto Rico from each of the 29 families is planted in an unreplicated yield trial with some 600 other entries during the summer. In this trial, data is collected on height (Figure 2), heading, harvest maturity, yield, milling yield, and resistance to blast disease using the inoculated disease screening nursery (Figure 3). Fourteen of the 29 yield plots from the B8910 project are selected to undergo further evaluation during the winter for grain cooking quality traits (i.e., amylose content and alkali spreading value) (Figure 4) and grain chalkiness (Webb 1985). Using these data, five lines are selected for testing in the following year.

#### 1996

The 10 F<sub>9</sub> panicle rows from each of the five families selected based upon the yield trial are planted in the summer breeding nursery. One row is selected from each of the five families and 10 panicles are picked.

Using the bulked yield plot seed from 1995, one of the five families is entered into the Uniform Rice



**Figure 3** Rice cultivars are evaluated for their reaction to a mixture of several races of *Pyricularia grisea* in disease screening nurseries.



**Figure 4** Milled rice flour is used to determine grain amylose content.

rows are selected and 20 panicles per row are harvested. Twenty-two rows are removed that are earlier, later, or taller than the others and then the remaining rows are bulk harvested. In addition, detailed phenotypic data is collected from the Headrow 1 field as an objective description of the cultivar, which is required for application to the state department of agriculture certified seed program.

RU9603178 is evaluated in the 1998 URRN and Texas statewide trials using the bulked harvested  $F_{11}$  seed from the 1997 yield trial. Results demonstrate that RU9603178 has main crop and ratoon crop yields competitive with other cultivars in its maturity group, some lodging susceptibility at two locations, excellent milling yield, and good tolerance to sheath blight disease (rating

Results from the URRN are analyzed during the winter and demonstrate that RU9603178 has very good main and ratoon crop yield, good resistance to all races of blast disease (including an IB49 rating of 1 on a scale of 0 to 9), good tolerance to sheath blight (rating of 4 on a scale of 0 to 9) and straighthead, very clear, non-chalky grain, a cooking quality of typical US long grains, and excellent milling quality. It is advanced for testing in 1997 along with one of the other four entries (BPRE 18) that had been tested in the Texas statewide trials.

#### 1997

Ten  $F_9$ -derived  $F_{10}$  panicle rows of RU9603178 and of BPRE 18 are planted in the summer nursery and 20 panicles are selected from one row of the two entries.

Bulked  $F_{10}$  seed from the 1996 URRN yield trial in Beaumont of entry RU9603178 is used to be planted in the advanced section of the 1997 URRN as entry 23. The other B8910 project line (BPRE 18) is entered into the advanced URRN as entry 32 using seed harvested from one of the 1996 Texas replicated trials. The advanced entries of the URRN are tested using four replications at the four state locations and the Texas entries of the URRN are also evaluated in the multi-location replicated trials within Texas.

Results from these trials indicate that RU9603178 continues to demonstrate competitive yield potential, excellent milling quality, and has moderate tolerance (rating of 6) to the sheath blight pathogen. The other B8910 line, entry 32, does not perform as well as RU9603178 and is dropped from further testing. Moreover, RU9603178 demonstrates resistance to all races of the blast fungus. Based upon this spectrum of resistance and the pedigree of RU960178, this suggests that the line may possess a novel major resistance gene to *P. grisea*. Previous research had demonstrated success in developing genetic markers associated with major genes that control rice cooking quality (Ayres et al. 1997). This spawned further research to identify the novel resistance gene and develop closely linked DNA markers.

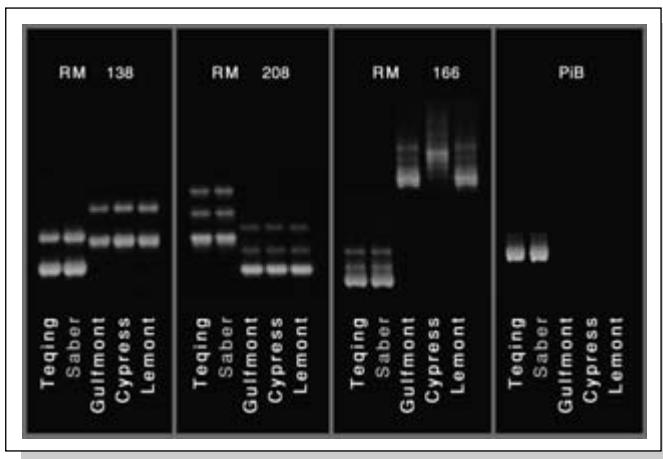
During the fall of 1997, 20  $F_{11}$  panicles from the summer breeding nursery selection of RU9603178 are planted in the Puerto Rico nursery; 287  $F_{12}$  panicles are harvested from the 20 panicle rows in Puerto Rico.

#### 1998

The 287  $F_{12}$  panicles are planted in an isolation block during the summer at Beaumont (Headrow 1). Thirty



**Figure 5** A foundation seed field of “Saber” rice.



**Figure 6** DNA markers on chromosome 2 that are associated with the *Pi-b* blast resistance gene are present in “Saber” and its parent, “TeQing” (Fjellstrom et al. 2004).

are confirmed to be resistant or moderately resistant and no susceptible plants are identified. This verifies the resistant/moderate resistant rating to IB49 infection of Headrow 2 of RU9603178 and indicates that resistance is not due to escapes nor is there heterogeneity for susceptibility. However, as a precaution, families rated as 3 are eliminated and only families that had an IB49 rating of 1 are provided to the foundation seed program. The objective description of the variety is provided to the state inspector to evaluate the foundation seed field (Figure 5).

Families that had a resistance rating of 2 to IB49 are bulked and used for further field testing during 2000. The same battery of field trials, disease screenings, and quality tests are conducted as before. DNA markers are developed that are associated with the *Pi-b* blast-resistance gene that is found in “TeQing” and explains the multirace resistance to blast that is observed in RU9603178 (Figure 6) (Fjellstrom et al. 2004).

By the end of the 2000 field season, RU9603178 has been evaluated at over 40 environments and it is decided to proceed with public release of the variety. The name “Saber” is selected in honor of the Texas A&M University Corps of Cadets. The complete set of data collected over the previous five field seasons is summarized to justify to the state seed board the release of “Saber”. In February 2001, the seed board accepts “Saber” into the state seed certification program, which allows the foundation seed that was produced in 2000 to be sold as certified ( $F_{15}$ ) seed for the 2001 planting season.

of 2). Its reaction to eight races of blast shows excellent resistance, however it appears to be segregating for resistance to race IB49 (rating of 2 and 4).

### 1999

In an effort to further clarify the level of resistance to the IB49 race of blast,  $F_{13}$  seed from the 265 bulk harvested 1998 Headrow 1 plants is evaluated for reaction to race IB49 during the 1999 spring greenhouse planting. Of some 300 seedlings that are screened, 37% are highly resistant (rating of 1–2), 61% are moderately resistant (rating of 3–4), and 2% are susceptible (rating of 5–6). This suggests that further selection in RU9603178 may allow for improvement in resistance to this pathotype.

During the summer, RU9603178 is evaluated in the same manner as before in the 1999 URRN and Texas statewide trials using the bulked harvested  $F_{13}$  seed from the 265 rows in Headrow 1. Agronomic and milling quality results from this trial are consistent with the previous year. Its reaction to eight of the blast races shows excellent resistance, however it is rated as susceptible (rating of 4) to IB49.

Headrow 2 is planted in the summer using 596  $F_{13}$  panicle rows derived from the 30 row selections from 1998. The information collected for the objective description of the variety in 1998 is verified in the Headrow 2 field. Each of the 596 rows is harvested separately for verification of reaction to blast race IB49 during the fall greenhouse planting. Over 99% of the families are rated as highly resistant (rating of 1–2) and the remaining families are rated as moderately resistant (rating of 3). This indicates that the 30 families that had been selected in the 1998 Headrow 1, and which constitute Headrow 2, are highly resistant to this pathotype.

### 2000

Approximately 20  $F_{14}$  families that had been rated as resistant (i.e., 1) and 20 families that had been rated as moderately resistant (i.e., 3) to IB49 in the fall 1999 greenhouse planting are re-evaluated in the spring 2000 greenhouse planting. All plants in the 40 families

### References

- Ayres, N.M., A.M. McClung, P.D. Larkin, H.F.J. Bligh, C.A. Jones, and W.D. Park. 1997. Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of U.S. rice germplasm. *Theor. Appl. Genet.* 94:773–781.
- Bollich, C.N., B.D. Webb, M.A. Marchetti, and J.E. Scott. 1990. Registration of Gulfmont rice. *Crop Sci.* 30:1159.
- Fjellstrom, R.G., C.A. Bormans, M.A. Marchetti, A.R. Shank, W.D. Park, and A.M. McClung. 2004. Development of DNA markers suitable for marker assisted selection of three *Pi* genes conferring resistance to multiple *Pyricularia grisea* pathotypes. *Crop Sci.* 44:1790–1798.
- Marchetti, M.A., C.N. Bollich, A.M. McClung, J.E. Scott, and B.D. Webb. 1995. Registration of RU87030196 disease resistant rice germplasm. *Crop Sci.* 35:600.
- McClung, A.M. 2002. Techniques for development of new cultivars. In: *Rice: Origin, history, technology, and production* (Smith, C.W., and R.H. Dilday, eds), pp. 177–202. John Wiley & Sons, New York.
- McClung, A.M. 2004. The rice plant: Growth, development and genetic improvement. In: *Rice chemistry and technology*, 3rd edn (Champagne, E.T., ed.), pp. 25–48. American Association of Cereal Chemists, St Paul, MN.
- McClung, A.M., R.G. Fjellstrom, C.J. Bergman, C.A. Bormans, W.D. Park, and M.A. Marchetti. 2004. Registration of "Saber" rice. *Crop Sci.* 44:693–694.
- Webb, B.D. 1985. Criteria of rice quality in the United States. In: *Rice chemistry and technology* (Juliano, B.O., ed.), pp. 403–442. American Association of Cereal Chemists, St Paul, MN.

about 25–50 mm in a good seedbed. Ground equipment may be used to apply fertilizer at the time of seeding. The rate of nitrogen application may vary between 30 and 100 lb/acre. High levels may cause lodging in some varieties. Where soils are deficient, application of moderate amounts of phosphorus and potassium may be beneficial.

A photoperiod of 10–12 hours is preferred for rice growth. The optimum temperature for growing rice is about 27°C, but the optimum temperature for flowering is dependent on photoperiod.

### Greenhouse nursery and growth chamber

Crossing indoors under controlled environments in a greenhouse or growth chamber is especially convenient for breeders working in the monsoon tropics or flooded paddies. Under these conditions, the parents may be grown in the field and dug up and placed in pots for crossing in the greenhouse. Working indoors allows the breeder to optimize photoperiod, temperature, and light level for optimal growth and flowering.

### Artificial pollination

#### Materials and equipment

Various methods are used by rice breeders for emasculation and pollination. The equipment and materials differ for the method used, the common ones being scissors, fine forceps, glassine bags, pot labels, paper

clips, wax pencil, tags, hot water bath, and vacuum emasculator.

#### Emasculation

The plant is ready for the preparation of the female for pollination when 50–60% of the panicle has emerged from the boot. Emasculation of individual flowers is done before anthesis and after they emerge from the boot. Emasculation in the tropics is best done after mid-afternoon, when anthesis has ceased for the day. In the temperate regions, preparation of the female can be done in the morning or late afternoon. Hot water emasculation is done by soaking the panicle for 5 minutes in water maintained at 43°C. This treatment must be followed by pollination within 30–60 minutes after emasculation. The hot water may be carried in a vacuum flask. A simpler and more efficient emasculation technique is to clip the spikelets and remove the anthers with a pair of forceps or using a vacuum unit.

Where plants have to be transplanted from the field into pots, the relocation must be completed at least 6 hours before emasculation, to allow plants to recover from any transplanting shock. Once a panicle has been identified for emasculation, it is separated from the others nearby to facilitate the emasculation process. The flag leaf is carefully removed. Florets at the top of the panicle that may have already self-pollinated and the young flowers at the bottom are cut with a pair of scissors. Next, about one-third to one-half of each floret is cut off at a slant to expose the anthers. If cut too low, the stigma may be damaged. If the anthers are to be

removed with forceps, the cut may be made across the anthers. The anthers can be extracted by using a vacuum pump or by forceps.

### Pollination

The stigma remains receptive for about 4–5 days, so pollination should be made sooner than later. Pollen remains viable for a shorter period (a few minutes to about half a day). Pollination should be conducted during the period of peak anther dehiscence, with pollen gathered just before anther dehiscence. The male panicles are cut and the flag leaf removed. These panicles are watched closely for anther extrusion, and used thereafter for pollination. The bag is taken from the female and the pollen shaken over it. The bag is replaced and clipped securely against the stem. Other techniques are also used for pollination.

### Natural pollination

Rice is highly self-pollinated. Where commercial F<sub>1</sub> hybrid seed production is undertaken, natural cross-pollination is often inadequate. Consequently, natural pollination is supplemented with hand pollination. The flag leaf may be cut and the leaf sheaths that enclose the panicle torn, to help the release of pollen. Also, a rope or pole may be dragged across the field at the level of the panicles each day during the flowering period to aid pollen dehiscence.

### Seed development and harvesting

The success rate in artificial pollination of rice is about 50% or higher depending on the technique used. Ovule swelling starts 3–4 days after pollination. The developing F<sub>1</sub> seed lacks a complete covering because the glumes were cut during the preparation of the female for pollination.

Rice does not mature uniformly in the head. Grain harvest moisture is critical to yield and produce quality.

The recommended grain moisture content is between 23 and 28%. At this stage the grains in the top portion of the head are ripe but those in the lower portion are in the hard-dough stage. Harvesting at this stage will include some immature grain but delaying harvesting increases the chances of shattering and checking of grains in susceptible varieties.

### Common breeding objectives

- 1 Grain yield.** Grain yield in rice depends on yield potential under the most favorable conditions, yield stability (across seasons), and crop productivity. The components of yield are panicle number per unit area, the number of filled grains per panicle, and grain weight. Some breeders use spikelet number per unit area and grain weight. The dramatic yield increases observed in the tropics in the 1960s were as a result of the development and use of semidwarf (*sd1* gene) cultivars that were environmentally responsive.
- 2 Grain quality.** Grain quality traits of interest vary from one region to another. They include grain size and shape, color of kernel, aroma, stickiness, and protein content.
- 3 Disease resistance.** Host resistances to many major diseases and insect pests have been identified in rice. Many of these are under oligogenic control and hence are susceptible to changes in the pest resulting in the pest becoming resistant to a resistant cultivar. The strategy of durable resistance is favored by many breeders. Resistance to insect-transmitted viral infections is complex to breed because the insect-resistance aspect (e.g., non-preference of the vector) can mask the disease-resistance component.
- 4 Resistance to environmental stresses.** Breeding resistance and tolerance to various environmental stresses is important in rice breeding. Drought and flooding frequently alternate with each other. Research indicates that regarding moisture stress, escape and avoidance mechanisms are important in dryland culture, while tolerance and recovery mechanisms are important under rainfed wetland culture. It should be pointed out that plant reactions to environmental stresses are site- and growth stage-specific.

### References and suggested further reading

- Chang, T. 1997. Rice. In: Genetics, cytogenetics and breeding of crop plants, Vol. 2 (Bahl, P.N., P.M. Salimath, and A.K. Mandel, eds). Science Publishers, Enfield, NH.
- Coffman, W.R., and R.M. Herrera. 1980. Rice. In: Hybridization of crop plants (Fehr, W.R., and H.H. Hadley, eds), pp. 511–522. American Society of Agronomy, Madison, WI.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Commercial classification of rice is based primarily on grain length.
- 2 The basic chromosome number of the genus *Oryza* is 12.
- 3 Rice is predominantly cross-pollinated.
- 4 Most of the world's rice production occurs in irrigated ecosystems.

### Part B

Please answer the following questions:

- 1 Give the three ecogeographic classes of rice.
- 2 List four major quality breeding objectives in rice breeding.
- 3 Describe the distinguishing features of the japonica class of rice.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the commercial classes of cultivated rice.
- 2 Discuss the common methods of breeding rice.
- 3 Discuss the genetics of grain color in rice.
- 4 Discuss the importance of the "wild abortive" gene in rice breeding.



# 30

## Breeding sorghum

### Taxonomy

Kingdom	Plantae
Subkingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Liliopsida
Subclass	Commelinidae
Order	Cyperales
Family	Poaceae
Genus	<i>Sorghum</i> Moench
Species	<i>Sorghum bicolor</i> (L.) Moench
Subspecies	<i>Sorghum bicolor</i> (L.) Moench subsp. <i>bicolor</i> – grain sorghum

### Economic importance

Sorghum was the world's fourth most important crop in 1995, accounting for 4% of the total cereal production, and a total harvest of 53 million metric tons. Sorghum and millet are two of the major world food crops that originate from Africa. Even though sorghum has become important in the agricultural production of developed countries, it is still primarily a developing country crop, with 90% of the world's acreage found in Africa and Asia.

In the USA, 470,525,000 bushels were produced in 2000 on 9,195,000 acres (see Appendix 2 for conversion rates of units). The leading producer was Kansas, with 3,500,000 acres and a production of 188,800,000 bushels, followed by Texas with an average of 3,000,000 acres and a production of 143,350,000 bushels. Other important producers were Nebraska, South Dakota, Colorado, Oklahoma, Missouri, Louisiana, New Mexico, and Arizona. The varieties grown in the US are primarily (78%) for grain. However, sorgho (sweet sorghum) is grown for forage, silage, and syrup (stalks crushed). The important sorgho states are Alabama, Mississippi, Georgia, Tennessee, and Iowa. Dual purpose cultivars

(grain and forage) are produced in states such as Texas, Kansas, Nebraska, and Oklahoma.

On the world scene, sorghum is produced principally in Africa, Asia, the Americas, and Australia. Important producers in Africa include Nigeria, Sudan, Burkina Faso, Cameroon, Chad, Mali, and Rwanda. About 74% of sorghum produced in Africa is used for food. Most developing countries use sorghum for food, unlike the USA where sorghum is mainly grown for feed. Even though most of the acreage for sorghum occurs in developing countries, the US leads the world in production, followed by India, Nigeria, China, Mexico, and Sudan.

### Origin and history

Sorghum originated in northeastern Africa (Ethiopia, Sudan, East Africa), where the greatest diversity of both wild and cultivated species occurs. It was domesticated in Ethiopia and parts of Congo between 5,000 and 7,000 AD with secondary centers of origin in India, Sudan, and Nigeria. It moved into East Africa from Ethiopia around 200 AD or earlier. It was distributed along trade and shipping routes throughout Africa, and through the Middle East to India at least 3,000 years ago. Sorghum was taken to India from eastern Africa during the first millennium BC. It arrived in China along the silk route. It was introduced into the Americas as guinea corn from West Africa through the slave trade, at around the middle of the 19th century.

### History of breeding in the USA

Concerned about the narrow genetic base of germplasm available to US breeders, J. C. Stephens and J. R. Quinby embarked upon a project to convert the tall, late maturing, tropical accessions in the world collection to

short, day-neutral genotypes for use in temperate regions, starting with 1,500 accessions. A large part of this material is also maintained in long-term storage at the National Seed Storage Laboratory, Fort Collins, Colorado.

### Germplasm resources

In the USA, sorghum germplasm is maintained at the Regional Plant Introduction Station, Experiment, Georgia, with duplicate samples stored at the National Seed Storage Laboratory, Fort Collins. A more comprehensive collection with accessions in excess of 25,000 is maintained at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India.

### Cytogenetics

Sorghum comprises three known ploidy levels:  $x = 5$ ,  $x = 10$ , and  $x = 20$ . Important sorghum species are: *S. bicolor* (L.) Moench ( $2n = 2x = 20$ ), *S. propinquum* (Kunth) Hitche ( $2n = 2x = 20$ ), and *S. halepense* (L.) Pers. ( $2n = 4x = 40$ ).

Of these three, the most important to crop production is *S. bicolor*. Grassy species such as *S. arundinaceum*, *S. verticilliflorum*, and *S. aethiopicum* have the same chromosome number ( $2n = 2x = 20$ ), and can be crossed with *S. bicolor*. Sorghum has also been successfully crossed with sugarcane and corn.

Cultivated races are readily crossable with each other to produce fertile hybrids. In sorghum production, natural interspecific crossing between *S. bicolor* and *S. halepense* is a common source of off-types in the field. Two types of products arise from this type of cross – sterile plants with 30 chromosomes and fertile plants with 40 chromosomes, the former often persisting in the field as perennial weeds.

Cytogenetic male sterility (CMS) and fertility restoration systems have been discovered in sorghum and are used in hybrid sorghum production. The CMS in sorghum genotypes was developed by backcrossing chromosomes of kafir into the cytoplasm of milo. Similarly, genetic male sterility (*ms*) has been discovered in sorghum male-sterile plants (*Msm*s).

### Genetics

Midrib color is controlled by a single dominant gene (*D*), whereas resistance of hydrogen cyanide (HCN)

appears to be controlled by more than one factor. Researchers have determined that grain color in sorghum is influenced by pericarp color, pericarp thickness, presence of testa, color of testa, endosperm color, glume color, and plant color. Each of these characteristics is determined by one or a few genes. For example, two genes, *R* and *Y*, determine whether the pericarp will be red (*RY*), colorless or white (*Ry*, *ry*), or lemon yellow (*rY*). The kernel starch is conditioned by a dominant allele (*Wx*); *wxwx* conditions a waxy endosperm. Similarly, sugary endosperm is controlled by a single locus, *sus*<sub>1</sub>.

Genetic male-sterility genes occur in sorghum, the most commonly used being the *ms*<sub>3</sub>. It has stable expression over different environments. CMS also occurs in sorghum, conditioned by an interaction between two major genes, *msc*<sub>1</sub> and *msc*<sub>2</sub>.

Important genes that have impacted sorghum breeding are those affecting maturity and plant height. Sorghum is a short-day plant. Maturity is influenced by photoperiod and temperature. Genes that influence sorghum maturity are designated *Ma*<sub>1</sub>, *Ma*<sub>2</sub>, *Ma*<sub>3</sub>, and *Ma*<sub>4</sub>; tropical cultivars being dominant at the *Ma*<sub>1</sub> locus. A genotype of *Ma*<sub>1</sub>*Ma*<sub>2</sub>*Ma*<sub>3</sub>*Ma*<sub>4</sub> takes about 90.5 days to anthesis, while *ma*<sub>1</sub>*ma*<sub>2</sub>*ma*<sub>3</sub>*Ma*<sub>4</sub> takes 55.3 days to anthesis. Similarly, four independent recessive genes, *dw*<sub>1</sub>, *dw*<sub>2</sub>, *dw*<sub>3</sub>, and *dw*<sub>4</sub>, reduce internode length of the sorghum stalk without affecting time of blooming and leaf size. A genotype *Dw*<sub>1</sub>*Dw*<sub>2</sub>*Dw*<sub>3</sub>*Dw*<sub>4</sub> has a flag leaf at 127 cm, while *dw*<sub>1</sub>*dw*<sub>2</sub>*dw*<sub>3</sub>*dw*<sub>4</sub> has a flag leaf located 43 cm above ground on the stalk. Sorghum cultivars containing dwarf genes are identified by the number of specific dwarf genes they contain as 1-dwarf (contains one dwarf gene, e.g., *dw*<sub>2</sub>), 2-dwarf (contains two dwarf genes, e.g., *dw*<sub>2</sub>*dw*<sub>3</sub>), and so on. US commercial sorghum hybrids are 3-dwarf.

### General botany

Sorghum (*Sorghum bicolor* (L.) Moench) is known by common names such as milo, kafir, and guinea corn. The annual sorghums have  $n = 10$  and include grain sorghum, sorgo, broomcorn, and sudangrass. *Sorghum halepense* (johnsongrass) is a perennial sorghum with  $2n = 20$ . The sorghum plant has culms that may stand 0.6–4.5 m tall, depending on the type and variety. It may produce two or more tillers. The stalk is solid. The center of the stem can be dry or juicy, insipid or sweet to taste. A dry stalked variety has leaves with a white or yellow midrib, while a juicy stalked variety has a dull

green midrib because of the presence of the juice instead of air spaces in the pithy tissues.

The number of leaves on the plant varies between 7 and 24 depending on the variety. The sorghum inflorescence is a panicle that may be loose or dense. It is usually erect but may curve to form a “gooseneck”. The panicle has a central rachis, with short or long primary, secondary, and mature tertiary branches, which bear groups of spikelets. The length and closeness of the panicle branches determine the shape of the panicle, which varies from densely packed conical or oval to spreading and lax. Sorghum is predominantly self-pollinated.

A fully developed panicle may contain 2,000 grains, each one usually partly covered by glumes. The grain is rounded and bluntly pointed, 4–8 mm in diameter, and of varying size, shape, and color according to the variety. Pigments occur in the pericarp, testa, or both. Cultivars with a pigmented pericarp have a yellow or red color. When the pericarp is white and a testa is present, the seed color may be buff or bluish-white. When a colored pericarp and a testa are present, the seeds tend to have a dark brown or reddish-brown color.

### Sorghum races

Five major races of sorghum are recognized – durra, kafir, guinea, bicolor, and caudatum. They differ in panicle morphology, grain size, and yield potential, among other characteristics. Durra sorghums developed primarily in Ethiopia and the Horn of Africa, from where they spread to Nigeria and the savanna region of West Africa. Kafir types originated in eastern and southern Africa. Guinea sorghums are grown mainly in West and Central Africa, while bicolor types are the least important to African production, and occur in East Africa. Caudatum varieties originated in Kenya or Ethiopia.

### Grain sorghum groups

Most of the grain sorghums in cultivation are hybrids, derived from kafir × milo crosses. The major commercial groups of grain sorghum are kafir, hegari, milo, feterita, durra, shallu, and kaoliang.

- 1 Kafir.** These have a thick juicy stalk, relatively large, flat, dark green leaves and awnless cylindrical heads. The seed color may be white, pink, or red.
- 2 Hegari.** These types have a more nearly oval head, more abundant leaves than kafir, and sweeter juice and hence are more desirable for forage.

**3 Milo.** This group has a less juicy stalk, curly light green leaves, and smaller leaves and stalks. The head is short, compact, and oval with large yellow or white seeds. The plant tillers more than kafir and is more drought tolerant.

**4 Feterita.** This group has few leaves, relatively dry stalks, and an oval compact head with very large, chalk-white seeds.

**5 Durra.** This group has dry stalks, flat seeds, and very pubescent glumes. The panicles are erect but may be compact or loose. The varieties are chiefly grown in North Africa, India, and the Near East.

**6 Shallu.** These are characterized by tall, slender, dry stalks, a loose head, and pearly white seeds. The varieties are late maturing.

**7 Kaoliang.** The varieties in this group have dry, stiff, slender stalks, an open bushy panicle, and small brown or white seeds. They are grown exclusively in China, Korea, Japan, and southeastern Siberia.

## Reproductive biology

### Floral biology

The sorghum inflorescence is a panicle ranging from 7.5 to 50 cm in length and 12.5 to 200 mm in width. Morphologically, the panicle (or head) ranges from compact to open. The spikelets are borne in pairs on branches arranged in whorls. One spikelet is sessile, bisexual, and fertile, whereas the other is pedicelled and male-sterile. The sessile spikelet contains two florets, one perfect and fertile, while the other is infertile. The fertile floret has a membranous lemma, a palea, two lodicules, three stamens, and an ovary with two long styles with plumose stigmas.

### Pollination

Blooming of sorghum starts within 3 days after the panicle emerges from the boot. Blooming starts at or near the panicle apex and proceeds downward, the process lasting 4–7 days. Blooming is hastened by short-day length and higher temperature. Optimal flowering occurs at temperatures of 21–35°C. Depending on the environment, the stigma may remain receptive for 5–16 days after anthesis if a flower is unpollinated. Anthesis usually occurs in the morning. The anthers dehisce as they are exserted or soon thereafter, usually becoming pendant. The pollen is most viable within the first 30 minutes; viability is negligible after 4 hours.



William Rooney

Texas A&amp;M University, College Station, TX 77843, USA

## Industry highlights

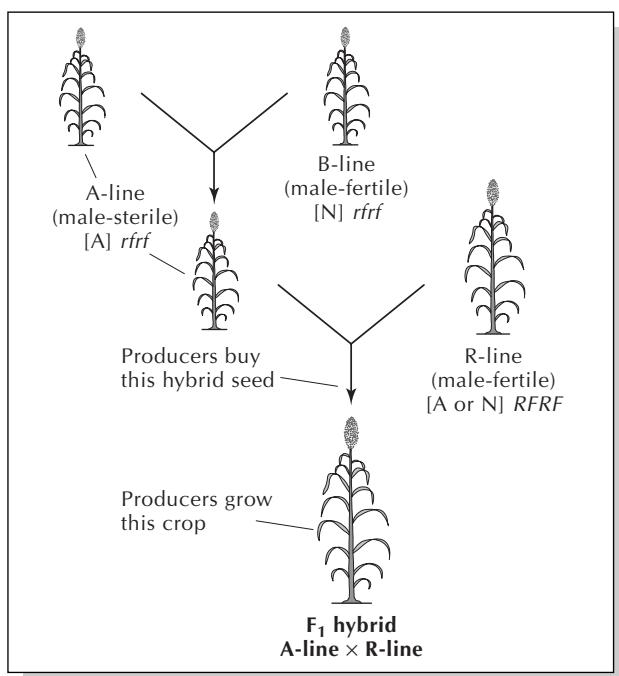
### Sorghum breeding

**Table 1** Genotypes and corresponding phenotypes for A-, B-, and R-lines in the A1 cytoplasmic-genetic male sterility system in sorghum.

Line	Cytoplasm*	Genotype†	Phenotype
A-line	[A]	<i>rfrf</i>	Male-sterile
B-line	[N]	<i>rfrf</i>	Male-fertile
R-line	[A] or [N]	<i>RFRF</i>	Male-fertile
Hybrid	[A]	<i>RFrf</i>	Male-fertile

\* Cytoplasm types: [A], sterility-inducing cytoplasm type; [N], normal cytoplasm.

† *RF* is the dominant allele for fertility restoration, and *rf* is the recessive allele for fertility restoration.



**Figure 1** The sorghum seed production process utilizing CMS. The genetics for each line are described in Table 1. The A-line parent is increased using pollen from the B-line. The  $F_1$  hybrid is produced by pollinating the A-line with an R-line pollinator. Both the B-line and R-line are maintained through self-pollination. [A], sterility-inducing cytoplasm; [N], normal cytoplasm.

### Introduction

Sorghum (*Sorghum bicolor* (L.) Moench) is one of the more important cereal grain crops in the world. In 2001, sorghum was produced on approximately 50 million hectares with an average yield of 1,280 kg/ha worldwide (FAO 2001). It is commonly grown in semiarid tropical, subtropical, and temperate regions of the world. The crop is used for many different purposes. The grain is used as a food grain, feed grain, and for industrial purposes. In many production systems, the vegetation is used as forage. The location of production often defines the ultimate end use and the specific types of sorghum that will be grown.

Sorghum is grown as a hybrid crop in the USA. Because it is a predominately self-pollinated crop, hybrid seed production requires the use of cytoplasmic genetic male sterility (CMS). Without this system, hybrid seed production would not be economically feasible. The CMS system is based on male-sterility-inducing cytoplasm that is complemented by alleles in the nuclear genome that either restore fertility or maintain sterility. In the CMS system, lines that have [A] cytoplasm must have a dominant allele present in the nuclear genome to restore male fertility (Table 1). If the line lacks the dominant allele for fertility restoration, the plant will be male-sterile.

Hybrid seed production requires maintenance of A-, B-, and R-lines (Figure 1). Seed of a male-sterile A-line is increased by pollination using the complementary B-line. The sole purpose of the B-line (also known as a maintainer) is to perpetuate or maintain the A. The A-line and B-line are genetically identical except that the A-line has a sterility-inducing cytoplasm while the B-line has normal fertile cytoplasm. Thus, A-line plants that are male-sterile can be pollinated with pollen from B-line plants to regenerate seed of the A-line. To produce hybrid seed, the male-sterile A-line is pollinated with pollen from the male-fertile R-line plants. The R-line (also known as a restorer line) is genetically very different to the A-line and carries the dominant fertility-restoration alleles needed to restore fertility in the progeny of the A-line. The seed that is produced on the A-line from this pollination is the seed that is planted by the producer for commercial grain production.

### Private and public research

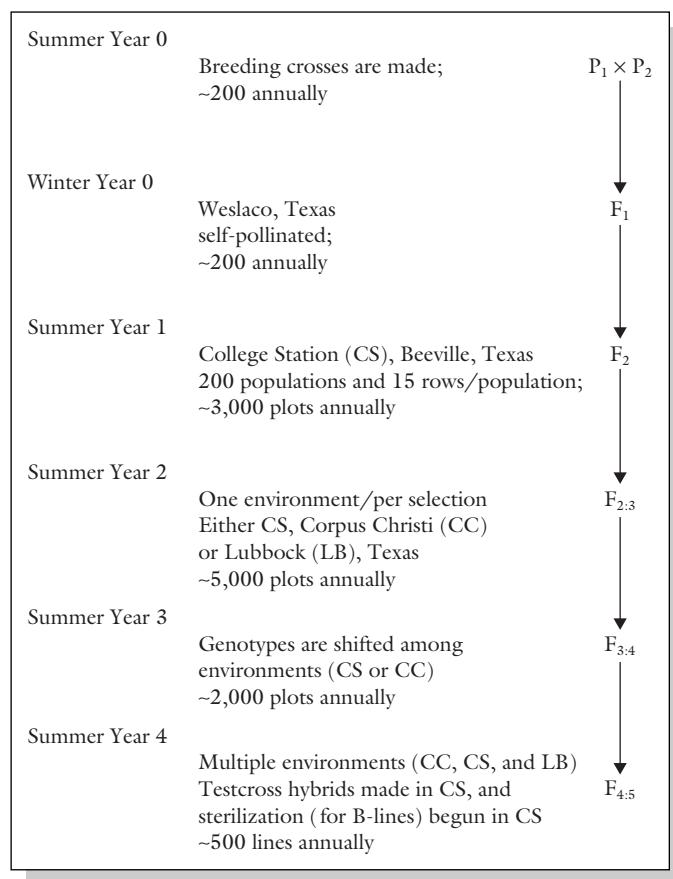
Because sorghum is grown as a hybrid, the hybrids that producers grow are produced and sold by private industry. Private industry also maintains a limited number of breeding programs for the production of new A/B- and R-lines for new hybrids. This work is supplemented and enhanced by researchers in public breeding programs, such as those at

Kansas State University, Texas A&M University, and USDA. These public programs do not produce or sell hybrids, but they develop parental lines and germplasms that are used by private industry in commercial hybrid production. In addition, public research programs in sorghum conduct research in long-term projects such as the introgression and development of new germplasm that may provide useful traits in the future.

The Texas Agricultural Experiment Station (TAES) sorghum breeding program located at Texas A&M University in College Station, Texas is part of a multiproject, multilocation sorghum improvement program supported by the TAES. Sorghum breeders work in conjunction with plant pathologists, entomologists, and grain quality and molecular geneticists to create an effective and important research and application oriented team. In terms of the sorghum breeding program at College Station, the breeding program has several objectives: (i) develop and release germplasm and parental lines with improved adaptability, yield, quality, and stress resistances; (ii) conduct research that increases our understanding and knowledge of sorghum breeding and genetics; and (iii) train undergraduate and graduate students in plant breeding.

#### *Methodology of the TAES sorghum breeding program at College Station*

For improved hybrids, new and improved parental lines must be developed. First, genetic variability must be developed through the selection and hybridization of parent material. This is a crucial step in the process. Usually elite germplasm is crossed to other material (elite lines, germplasm, genetic stocks) to correct a perceived deficiency in the elite material. For example, if an otherwise good A/B pair is susceptible to lodging, it will be hybridized with several different sources of lodging resistance with the goal of producing a new A/B pair with improved lodging resistance. In our program, the A/B program is managed separately from the R-line program to maintain heterosis between the two groups and keep the fertility restoration and maintenance genetics separate. Based on the considerations listed above, specific crosses are made using the methodology described by Rooney (2004). These  $F_1$  progeny are self-pollinated to produce an  $F_2$  population.

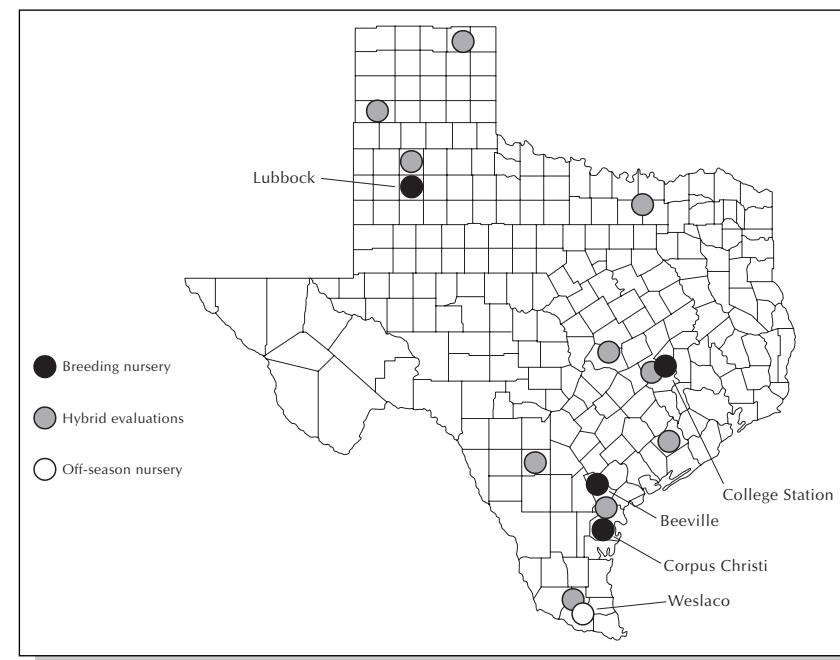


In our program, the  $F_2$  populations are created, our program utilizes a pedigree breeding approach for the development of inbred lines (Figure 2). From the  $F_2$  generation until the  $F_5$  generation (in which uniform lines are selected), the progeny rows are grown and panicles in the rows are visually selected on the basis of agronomic desirability, pest resistance, and abiotic stress tolerance.  $F_5$  lines that are phenotypically uniform are testcrossed to measure their general and specific combining ability and their suitability as parent lines in hybrid combinations.

Once  $F_5$  populations are created, our program utilizes a pedigree breeding approach for the development of inbred lines (Figure 2). From the  $F_2$  generation until the  $F_5$  generation (in which uniform lines are selected), the progeny rows are grown and panicles in the rows are visually selected on the basis of agronomic desirability, pest resistance, and abiotic stress tolerance.  $F_5$  lines that are phenotypically uniform are testcrossed to measure their general and specific combining ability and their suitability as parent lines in hybrid combinations.

The appropriate time for the selection of specific traits is dependent on the heritability of the trait and the environments in which the selection occurs. In our program, traits with higher heritability (maturity, height, grain color, etc.) are selected in the early generations while traits with lower heritability (yield, drought tolerance, disease and insect resistance) are selected in more advanced generations. These more complexly inherited traits must also be screened in multiple environments, because these traits may not be expressed in any given environment. Evaluation in multiple environments is crucial to the development of widely adapted sorghum genotypes. In our program, we use three basic regions for inbred selection: south Texas, central Texas, and the Texas high plains (Figure 3). These regions are each unique and force different selection pressures on the material grown therein. For example, our south Texas nurseries are rainfed and subject to drought stress and consistent disease pressure. In addition, this region is good for selecting genotypes that perform well in subtropical

**Figure 2** Pedigree breeding scheme used by the TAES sorghum breeding program at College Station, Texas. This scheme is used for the development of new B- and R-lines and germplasm. Initial crosses are made using either plastic bag crosses or hand emasculations. Open-pollinated selections are made in each generation until the  $F_5$  where the plot is self-pollinated and used to make testcross hybrids. At the  $F_5$  generation, new B-lines enter sterilization and testcrossing while new R-lines are evaluated in testcrosses.



**Figure 3** Map of Texas showing the locations used by the TAES sorghum breeding program for breeding, selection, and evaluation.

grain quality in sorghum has become more important. Until recently, sorghum in the US was used exclusively as feed grain, but the development of food quality sorghum hybrids has resulted in an increase in the sale and use of sorghum in food products. A major goal of our program is to produce sorghum lines that will make hybrids with grain quality suitable for both feed and food use.

Prior to testcrossing, the new B-lines must be male-sterilized. B-lines are sterilized using a backcross program in which a standard A-line with a similar pedigree to the new B-line is used as the source of the male-sterile cytoplasm (Figure 4). The new B-line is then used as a recurrent parent to produce an A-line that is genetically identical to the B-line (except that it is male-sterile). In each generation of backcrossing, plants and progenies that are fully male-sterile and are the most similar to the B-line are selected. The sterilization process usually requires a minimum of five backcrosses and most sorghum breeding programs utilize winter nurseries to reduce the amount of time required for sterilization.

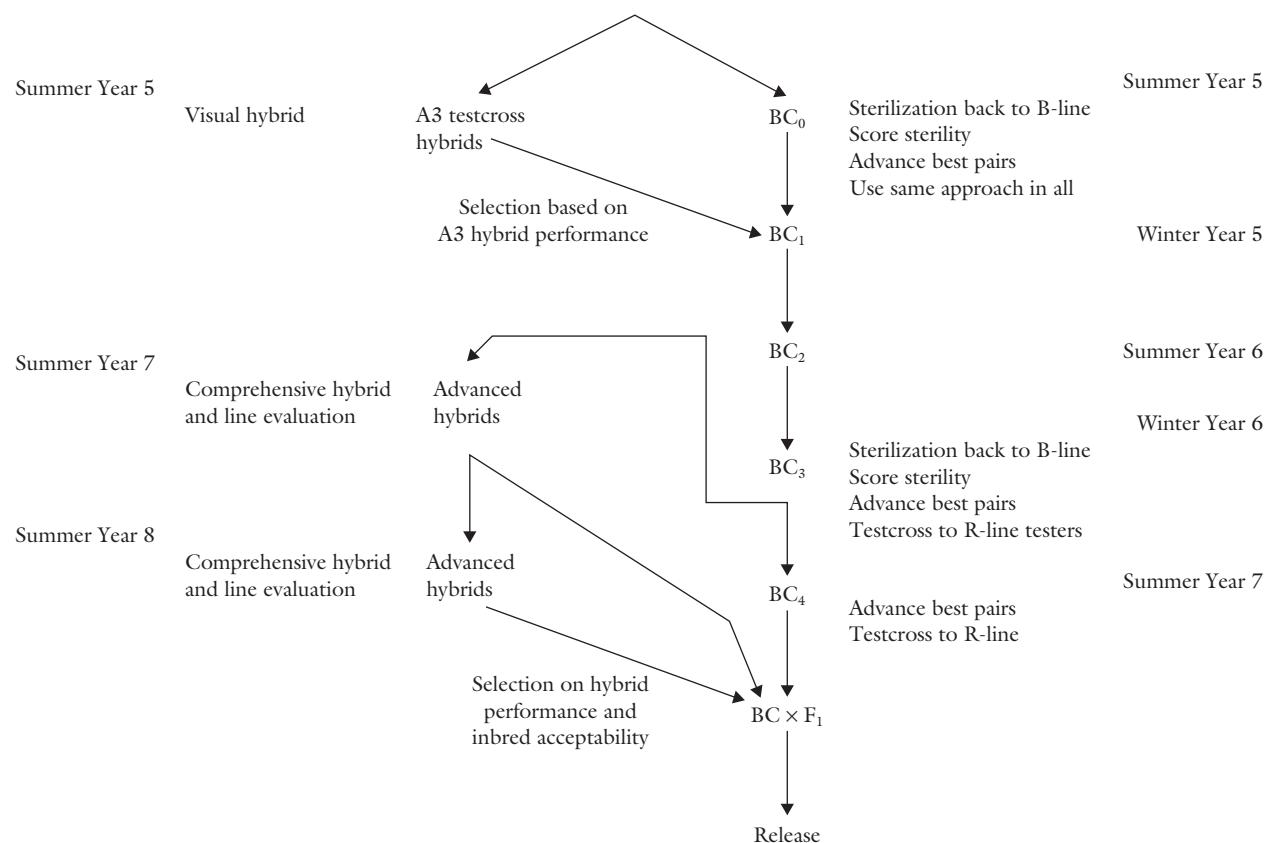
Our program begins testcrossing in the  $F_5$  for R-lines and in the  $BC_3$  of sterilization for A/B-lines. New R-lines are testcrossed to A-line testers and new A-lines are testcrossed to R-line testers to determine each new line's general combining ability and the fertility of the hybrids. Lines that produce high-yielding hybrids with appropriate agronomic parameters are advanced for additional testing. These lines are hybridized to several potential parental lines of the opposite group to identify those hybrids with good general and specific combining ability. These hybrids are tested in multiple locations and those lines that produce hybrids with high yield, good stability, good agronomic characteristics (height, maturity, etc.), and acceptable abiotic and biotic stress tolerances will be released. These releases are then tested by private industry to determine if they will be used in hybrid releases for producers (Rooney 2003a, 2003b).

In addition to hybrid performance, private companies must consider the parental line's performance with regard to seed production. Seed producers must be able to consistently coordinate flowering of the A-line and the R-line. If the two lines have a poor "nick", the A-line will have extremely low seed set and, consequently, poor yields of hybrid seed. The pollinator line should start shedding pollen prior to the emergence of stigmas in the A-line and the R-line must continue to shed pollen throughout the flowering of the A-line. In addition, the pollen shed from the R-line should be consistent and relatively unaffected by normal environmental conditions. Obviously, the R-line must consistently restore fertility to the hybrid and the A-line parent must produce seed yields high enough to justify seed production costs.

In addition to breeding for grain sorghums, there are directed breeding programs in sorghum to improve the crop for other uses. These include the improvement of sorghum for forages, such as silage, grazing, and hay. Other programs focus on the use of sorghum for sweeteners, in which the crop is used much like sugarcane. All of these uses have resulted in the development of numerous genotypes suitable for the production environments and purposes of the crop. The challenge for sorghum breeders is to continue future improvements given the relatively small number of researchers actively engaged in this field of research.

growing environments. In our high plains nurseries, the environment is typical of temperate production regions and different traits are of importance. The evaluation of material at these locations allows us to select widely adapted germplasm.

Many traits are important in our breeding program (both inbred and hybrids). Since grain sorghum producers are paid based on grain yield, it remains the most important trait. Most every breeding program measures the yield potential in hybrids, as the correlation between inbred yield and hybrid yield is rather poor (Rooney 2004). In addition to breeding for yield, any factor that reduces yield becomes important as well. Therefore, breeding for drought tolerance and disease and insect resistance have priority as well. Diseases of importance in US sorghum include stalk lodging, grain mold, anthracnose, downy mildew, head smut, sooty stripe, and leaf blights. Insects of economic importance include the sorghum midge, greenbug, and chinch bug. Finally,



**Figure 4** The advanced testing and sterilization scheme used by the TAES sorghum breeding program at College Station, Texas. From the  $F_5$  generation (see Figure 2), testcrossed hybrids and  $BC_0$  sterilization lines are grown for evaluation and continued backcrossing. If testcross performance is acceptable, sterilization is continued through backcrossing until the A-line is identical in phenotype to the B-line. As the A-line becomes available, additional hybrid evaluation is performed to confirm heterosis and line acceptability. The R-line testcross evaluation process is similar but eliminates the requirement of sterilization.

### References

- FAO. 2001. Available at [222.fao.org](http://222.fao.org).  
 Rooney, W.L. 2003a. Registration of Tx2912 to Tx2920 sorghum germplasm. *Crop Sci.* 43:442–443.  
 Rooney, W.L. 2003b. Registration of Tx2921 to Tx2928 sorghum germplasm. *Crop Sci.* 43:443–444.  
 Rooney, W.L. 2004. A overview of sorghum breeding methodology. *Adv. Agron.* 83:37–109.

### Common breeding methods

Early sorghum improvement depended on introductions. The introduction of short-day germplasm from the tropical regions into temperate regions with a longer photoperiod was unsuitable for commercial production because they usually matured too late or were too tall. Using backcrossing, US Department of Agriculture (USDA) and Texas A&M University researchers

embarked on the sorghum conversion program to convert tropical varieties to temperate climate-adapted cultivars by substituting two recessive alleles for the dominant height alleles in the tropical varieties, as well as the recessive *mal* maturity alleles for the *Mal* dominant counterparts.

Pure-line selections have been made in many countries in Africa and India. Pedigree selection is commonly used following hybridization. The population

improvement methods discussed in Part I of this book are applicable to sorghum breeding. To develop a random mating population the breeder starts by selecting 20–40 parents. The next step is to incorporate a male-sterility gene by crossing each parent individually to a male-sterile stock. The F<sub>2</sub> segregates for male sterility. The selected plants are backcrossed one or two times if the cytoplasmic male-sterile (CMS) stock lacks good agronomic qualities. An equal amount of F<sub>2</sub> seed from all crosses are bulked and grown in isolation for random mating.

Sorghum is also bred using population improvement methods for enhancing quantitative traits. The pure-line method of breeding has been used in sorghum improvement. With the discovery of CMS and fertility-restoring genes in the 1950s, hybrid sorghum breeding became practical and profitable using the A-, B-, R-line breeding systems as in corn.

The sorghum conversion program (adapting tropical sorghum to temperate climates) has been a significant part of the success of sorghum breeding in the USA. A conversion program starts with a cross of tropical × temperate lines at a tropical station. The resulting F<sub>1</sub> is grown at a temperate experimental station to obtain F<sub>2</sub> seed. The F<sub>2</sub> is planted at a tropical station. Selections are made and backcrossed (five cycles of backcrosses) to adapt the tropical line to temperate conditions.

### **Establishing a breeding nursery**

The specific layout of a nursery depends on the task to be performed and materials to be handled. A section of the nursery may be allocated to procedures such as selfing, crossing, population breeding, observations, forage breeding, and tropical conversion. To reduce walking, parents to be crossed are planted near each other. Plant density should be similar to that used by farmers in production. Spacing varies from 10 to 100 cm between rows and 5 to 30 cm within rows.

### **Artificial pollination**

#### **Materials and equipment**

The equipment used includes pollinating bags (6 × 12–40 cm), stapler, knife, marking pencils, scissors, hot water container, clips, string, and apron.

#### **Emasculation**

Sorghum breeders control pollen using one of four general methods – male sterility, hot water emasculation, hand emasculation, and control of anther dehiscence. The panicle of the male-sterile plant is bagged just before anthesis. Using male sterility enables the breeder to undertake large-scale hybrid seed production. Hot water emasculation is also used, but usually when male sterility is not available in the parents, and when complete pollen control (i.e., the presence of a few selfs) is not critical. To emasculate in the field, a suitable panicle (just beginning to flower) is selected and all opened spikelets removed. The panicle is enclosed in a rubber or plastic sleeve, tied tightly around the peduncle but open at the top. Water at 47–48°C is poured into the sleeve and held for 10 minutes. In the greenhouse, the particle may be emasculated by directly immersing the panicle in hot water by inverting the potted plant. The panicle is left to dry before covering it with a pollination bag.

Breeders may use hand emasculation when only a small quantity of seed is needed and complete pollen control is desired. This method of emasculation requires skill to succeed. Hand pollination in the field is undertaken in the afternoon when contamination from other plants is least. Anthers may be removed by using forceps, scissors, or other pointed instruments. The bag is placed over the emasculated panicle.

Anther dehiscence can be delayed by high humidity. In regions of high humidity, placing polythene or paper over a particle can delay anther dehiscence by about 30 minutes the next morning. However, under hot conditions, heat might build up under the bag and injure the flowers.

#### **Pollination**

Pollen collection is best done in the morning between 7 and 12 a.m. The center of the particle yields the most pollen. It may be necessary to cover the male plant, the day before pollination to keep out contaminants. When it is time to pollinate, the panicle is tilted so that the pollen can be shaken into the bag. Male-sterile plants and those emasculated by hot water are ready for pollination after about 5–8 days following bagging, when flowering is complete and the stigmas are extruded. The bag containing the pollen is inverted over the female panicle and shaken to pollinate. The operator may also remove a branch of the panicle and brush pollen on the stigma. The pollinated panicle is

covered with the bag. Pollination success varies with the method of pollination, operator's skill, environment, age of the stigma, and amount of pollen, among other factors.

### Natural pollination

CMS is used for large-scale hybridization in an isolation block where random mating populations are grown. A fertility-restorer pollinator is planted with several male-sterile females.

### Seed development and harvesting

Successful pollination is evidenced by the seventh day following pollination. The grains are physiologically mature when the black layer forms at the base of the seed. This stage occurs between 35 and 45 days after pollination.

Grain sorghum is a perennial plant. Consequently, the plant remains green and alive, even after the grain is matured, until killed by tillage or freezing temperatures. The grain dries slowly. At the hard-dough stage, the grain contains about 18–20% moisture. For effective combining, the grain moisture content should be about 13% or less. Waiting for the grain to dry in the field delays harvesting and increases the risk of damage by weather factors and birds. Furthermore, delayed harvesting also delays the rotation of sorghum with a winter crop (e.g., wheat). Desiccants (e.g., diquat, 28% nitrogen urea-ammonium nitrate) are applied as a preharvest treatment on grain sorghum. Roundup Ultra® may be used as a defoliate when the grain is for feed.

### Common breeding objectives

- 1 Grain yield.** Grain yield is one of the principal objectives of sorghum breeding. Grain yield improvement in sorghum has made significant strides over the years. The success mainly stems from access to additional germplasm from tropical accessions through the sorghum conversion program, which has augmented existing genetic variability and the development of hybrid cultivars. Also, disease and insect resistance protect grains in the field, reducing harvest losses.
- 2 Yield stability.** Sorghum is more adapted to marginal production environments than cereal crops such as corn and rice. Breeders are interested in stable per-

formance over variable environmental conditions in the production region.

### 3 Agromorphological traits:

- (a) **Lodging resistance.** Tall cultivars are prone to lodging. Breeders use dwarfing genes to attain short stature. Sources of stiff stalk include the kafir and hegari genotypes. Another way lodging resistance is enhanced is through the development of resistance to stalk and root diseases.
- (b) **Short stature.** As previously discussed, breeders use the four recessive dwarfing genes  $dw_1$ ,  $dw_2$ ,  $dw_3$ , and  $dw_4$  to develop short-statured cultivars. Most US cultivars are 3-dwarfs (have three recessive dwarfing genes). It should be mentioned that taller cultivars are preferred for the production of forage and silage, and also in dry production systems. Short stature is desirable for mechanized harvesting. In addition to short stature, crosses between milo and kafir produce recombinants with an erect head and stout stalk.

### 4 Adaptation:

- (a) **Early maturity.** Early maturing cultivars are advantageous in low rainfall regions by allowing the crop to escape damage from drought. These cultivars also allow an expansion of sorghum production to regions of high altitudes and short growing season. Genes for early maturity tends to reduce plant stature.
- (b) **Photoperiod insensitivity.** Photoperiod insensitivity adapts the crop to regions of shorter growing season.
- (c) **Drought resistance.** Sorghum is more resistant to heat and drought than corn. In breeding for drought and heat resistance, breeders select for extensive root systems that promote more extensive exploration of the soil for moisture, as well as plant features that reduce moisture loss from the leaves (e.g., fewer stomata).
- (d) **Tolerance of aluminum.** High levels of soil aluminum reduce root development and predispose plants to drought injury. Genotypes with aluminum tolerances have been identified for breeding.

### 5 Disease resistance.

Sorghum diseases of economic importance include the following:

- (a) **Rots.** Sorghum is known to be attacked by both root and stalk rots, the common ones including *Fusarium* root rot (caused by *Fusarium moniliforme*) and stalk rot, charcoal rot (caused by *Macrophomina phaeolina*), and *Periconia* root and crown rot.
- (b) **Blight.** An important blight is the northern leaf blight (caused by *Exserohilum turcicum*), a

disease that is common in humid production regions. Sources of resistance include grain and sweet sorghum types.

- (c) **Smuts.** An important smut disease is the head smut (caused by *Sporisorium reidianum*).
- (d) **Rust.** *Puccinia purpurea* is responsible for sorghum rust, a disease that is not economically very important in the US.
- (e) **Viral diseases.** One of the most economically important viral infections of sorghum is the maize dwarf mosaic, a disease that is spread by the greenbug and various species of aphids.

**6 Insect resistance.** Major insect pests of sorghum include the following:

- (a) **Greenbug.** The greenbug (*Schizaphis graminum*) is a sucking insect pest with biological races. Breeders, hence, have to continue to breed new cultivars as new races evolve.
- (b) **Sorghum midge.** The sorghum midge (*Contarinia sorghicola*) is of economic importance worldwide. Resistant germplasm has been identified among accessions from Ethiopia and

Brazil. Resistance is quantitatively inherited. Consequently, breeders should use resistant A- and R-lines in developing resistant hybrids.

- (c) **Stalk borers.** The European stalk borer (*Ostrinia nubilis*) and the southwestern corn borer (*Diatraea grandiosella*) are the most economically important.
- (d) **Shoot fly.** The shoot fly (*Antherigona socata*) is of economic importance primarily in the tropics.

**7 Product quality.** Sorghum is produced for grain for food, especially in developing countries. For food purposes, endosperm qualities are important in breeding objectives. Sorghum endosperm may be floury or corneous, the latter trait being important in dry milling. The pericarp may be pigmented or without pigment. To improve the nutritional quality of sorghum for feed, one objective is to reduce the tannin content. Sorghum is also grown for industrial uses, including flour, syrup, and malt production. For feed, an objective is to reduce the cynogenetic glucoside content of the plant and improve its fodder, forage, and silage qualities.

### Reference and suggested reading

Acquaah, G. 2005. Principles of crop production. Theory, techniques and technology. Prentice Hall, Upper Saddle River, NJ.

### Outcomes assessment

#### Part A

Please answer the following questions true or false:

- 1 Sorghum is a self-pollinated species.
- 2 Sorghum originated in Africa.
- 3 *Sorghum bicolor* is the most important species for crop production.
- 4 Most US sorghum cultivars are 2-dwarf.

#### Part B

Please answer the following questions:

- 1 Give the five major races of sorghum.
- 2 Give four examples of grain sorghum groups.
- 3 Give two of the top sorghum-producing states in the USA.

#### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the breeding of plant stature in sorghum.
- 2 Discuss the important diseases and insect pests of sorghum.
- 3 Discuss the pollination process in sorghum breeding.

# Breeding soybean



## Taxonomy

Kingdom	Plantae
Subkingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Magnoliopsida
Subclass	Rosidae
Order	Fabales
Family	Fabaceae
Genus	<i>Glycine</i> Willd
Species	<i>Glycine max</i> (L.) Merr

## Economic importance

Soybean (*Glycine max* (L.) Merrill) is the major world oil seed. It is also a major source of meal used for livestock feed. It consists of about 35–40% protein and less than 20% oil. The USA is the world's leading producer of soybean, accounting for nearly 50% of the world's total production, and about 40% of the cultivated acreage. The average yield of US soybean was only about 11 bushels/acre in 1924, but increased to 25.4 bushels/acre in 1966 (see Appendix 2 for conversion rates of units). In 1994, the production reached a new high of 41.9 bushels/acre. In that year, producers in Iowa recorded a yield of 50.5 bushels/acre. Similarly, the acreage of soybean in 1924 was 1.8 million acres, 18.9 million acres in 1954, and 63.4 million acres in 1996. It rose to 72.7 million acres in 2000.

Soybean production in the US occurs primarily in the north central states, which overlap the Corn Belt states. The major producing states are Iowa, Illinois, Minnesota, Indiana, Ohio, Missouri, and Nebraska, these states together accounting for 72% of the US total production in 2000. Iowa and Illinois produce more than 30% of the US crop. Soybean is important in the southern

and southeastern states of Arkansas, Mississippi, North Carolina, Kentucky, Tennessee, Louisiana, and Alabama, these states together producing about 10% of the US total crop. Other producers are South Dakota, Kansas, Missouri, Wisconsin, and North Dakota.

On the world scene, the US has dominated production since the 1950s, growing more than 75% of the world crop by the 1970s. By 2000, the USA, though still the world's leader, had a share of 45%, with Argentina and Brazil taking up about 15% and 21%, respectively. China ranks fourth in the world, with about 12% of the world's total production. Other producers in the world include Japan, Indonesia, and the former Soviet Union.

## History and origin

The soybean is considered among the oldest cultivated crops. The first record of the crop is contained in a 2838 BC Chinese book in which Emperor Cheng-Nung described the plant. Soybean was a "Wu Ku", one of the sacred five grains (the others being rice, wheat, barley, and millet) considered essential for the existence of Chinese civilization. Cultivated soybean is believed to have derived from a wild progenitor, *Glycine ussuriensis*, which occurs in eastern Asia (Korea, Taiwan, Japan, Yangtze Valley of central China, northeastern provinces of China, adjacent areas of Russia). The plant was first domesticated in the eastern half of north China in the 11th century BC. It was introduced into Korea from this region and then into Japan between 200 BC and 300 AD.

Soybean was known to be grown in Europe in the 17th century. Its first introduction into the USA is traced to a Samuel Bowen, an employee of East India Company, a seaman, who brought it to Savannah, Georgia, from China via England. Between 1804 and 1890, numerous soybean introductions were made into

the US from China, India, Manchuria, Korea, Taiwan, and Japan. In 1852, J. J. Jackson is reported to have first planted soybean as an ornamental plant in Davenport, Iowa. Most of the production of soybean in the US prior to the 1920s occurred in the southern states, mostly for hay, and then spread to the Corn Belt after about 1924.

Soybean is a subtropical plant. However, it is grown over a wide range of ecological zones, ranging from the tropics to 52°N. Its climatic requirements are similar to those of corn.

### History of breeding in the USA

Soybean was introduced into the USA in about the late 1700s. Initially, the crop was grown primarily as a forage species. The early introductions had weedy characteristics. They also shattered profusely, making them unsuitable for mechanized production. Producing soybean for seed started in the 1920s. Formal soybean research was initiated as a joint activity between the US Department of Agriculture (USDA) research wing and the Agricultural Experiment Stations of land grant universities in the midwest and southern states. In 1936, a cooperative state and federal program was established at the University of Illinois, Champaign/Urbana, to conduct soybean research. This program is now known as the US Regional Soybean Laboratory.

Seed companies became involved in soybean breeding as a result of the United States Plant Variety Act of 1970. With this protection against unauthorized use of proprietary material, seed companies could make money developing and selling improved seed. Modern cultivars were developed for seed, shifting the emphasis away from soybean as a forage crop to soybean as an oilseed crop.

### Cultivars

Soybean may be produced for forage, the varieties for this purpose generally being small seeded, finer stemmed, and very leafy. There are also cultivars for edible, dry, or green-shelled beans. These cultivars usually have straw-yellow or olive-yellow seed and a yellow, brown, or black hilum. Soybean grown for grain is grouped into 13 maturity classes, ranging from 000 to X. The 000 group consists of the earliest maturing cultivars while the X group consists of the latest maturing cultivars. Groups 000–IV are considered indeterminate, while

groups V–X are determinate cultivars. Further, early maturing cultivars (000–IV) are adapted to the northern production regions, while those with high maturity class designation are adapted to southern production regions. The 000 cultivars are adapted to the short summer growing seasons of northwest USA and Canada. Cultivars in groups II and III are best adapted to the midwest growing area. Maturity groups VIII and higher are grown in the southern or coastal plain counties.

### Germplasm resources

The USDA maintains about 15,000 accessions of *G. max* and a smaller number of other *Glycine* species at Urbana, Illinois, and Stoneville, Mississippi. Soybean accessions are also held by the Applied Genetics of Korea Atomic Energy Research Institute in Seoul, Korea, the National Institute of Agricultural Sciences at Hiratsuka, Japan, and the Asian Vegetable Research and Development Center, Taiwan.

### Cytogenetics

The genus *Glycine* has two subgenera – *Soja* and *Gycine*. The subgenus *Soja* consists of two species: *G. max* ( $2n = 2x = 40$ ), the cultivated species, and *G. soja* (L.) Sieb or *G. ussuriensis* ( $2n = 2x = 40$ ), a wild species. These two species are cross-fertile. There are 15 wild species of soybean of which *G. tabacinia* and *G. tomentella* have polyploid forms (including  $2n = 4x = 80$ ).

### Genetics

Several hundreds of genes for qualitative traits have been identified for soybean. Among them are four recessive genes for genetic male sterility, designated *ms<sub>1</sub>*, *ms<sub>2</sub>*, *ms<sub>3</sub>*, and *ms<sub>4</sub>*. Linkage groups have been identified for 13 of the 20 soybean chromosomes. Soybean has pigmentation on various parts of the plant. The black and brown seed coat and hilum color are conditioned by two gene pairs, *Tr* and *Rr*. The expression of these genes is modified in some conditions by brow seed coat (*TrO*), which is dominant over reddish-brown color (*Tr*), and green seed coat (*G*), which is dominant over yellow seed coat (*y*).

Soybean may have either brown (tawny) or gray pubescence, the trait being conditioned by a single gene

and brown being dominant over gray. Stem termination is controlled by two genes,  $Dt_1$  and  $Dt_2$ , with  $Dt_1$  conferring intermediate stem while  $Dt_2$  conditions semi-determinacy, with  $dt_1$  conditioning determinacy.  $Dt_1$  and  $Dt_2$  are dominant to  $dt_1$  and  $dt_2$ , respectively;  $dt_1$  is epistatic to  $Dt_2$  and  $dt_2$ . Pod color at maturity is controlled by two genes:  $L_1L_2$  and  $L_1l_2$  produce black pods, while  $l_1L_2$  produce brown color. Tan pods are conditioned by  $l_1l_2$ . Six independent recessive genes  $df_1$  to  $df_6$  are known to condition dwarfism in soybean, while five major and independent genes,  $E_1$  to  $E_5$ , condition flowering and maturity, with  $E_1$  and  $E_2$  delaying flowering and maturity.

### General botany

Soybean (*Glycine max* (L.) Merrill) is an annual summer legume. Cultivated soybean is usually erect with a well-defined main stem and branches, and numerous leaves. Both determinate and indeterminate cultivars are used in production. The leaves and stems are usually pubescent. The flowers are either purple or white, and are borne in axillary racemes on peduncles at the nodes. The plant produces a large number of flowers, but only about two-thirds to three-quarters of them produce pods. The pods are also pubescent. They range in color from light straw to black, containing 1–4 seeds (occasionally five). The seeds are usually unicolored and may be straw-yellow, greenish-yellow, green, brown, or black. Bicolored seeds exist, such as yellow with a saddle of black or brown. The hilum is also colored in various patterns – yellow, buff, brown, and black. Soybean is primarily self-pollinated. In the proper soil environment, soybean is infected by the bacterium *Rhizobium*, resulting in roundish nodules on the roots in which the nitrogen-fixing bacteria live.

### Reproductive biology

#### Floral biology

Soybean flowers are borne in the axil of a branch in clusters of 3–15. Soybean has the floral features of the family Papillionidae, comprising a large standard petal, two wing petals, and two keel petals. The five petals enclose a pistil surrounded by 10 stamens, nine of which are fused into a tube, the one nearest the standard being

free. The style curves towards the standard petal bearing a club-shaped stigma.

#### Pollination

The soybean flower is cleistogamous, with self-pollination occurring shortly before or after the flower opens. Pollen shed depends mainly on temperature. Under warmer conditions (30°C or higher) pollen shed may occur before 7–9 a.m. Anther dehiscence under moderate temperature starts later in the morning and continues for much of the day. Cool temperature reduces pollen shed and often causes flowers to prematurely self-pollinate before the breeder can emasculate the stem.

### Common breeding methods

As previously indicated, plant introductions played a significant role in early soybean breeding programs. Selections were made from introductions to develop commercial cultivars. Modern soybean breeders use a wide variety of methods in their programs, including backcrossing to transfer qualitative traits and single-seed descent to accelerate breeding programs. Pedigree selection is also used because the plants can be well spaced for individual observations.

Hybridization is most commonly used in soybean breeding programs for gene transfer. Both qualitative and quantitative traits have been improved in breeding programs using hybridization to create new variability for selection. Recurrent selection is possible if male sterility is incorporated into the breeding program.

### Establishing a breeding nursery

#### Field nursery

To facilitate crossing in the field, the two parents to be crossed are planted in opposite rows with wide spacing (65–100 cm) between them for the operator to move freely. Having the two parents to be crossed planted side by side facilitates that crossing process and reduces the chance of crossing the wrong parents.

#### Greenhouse nursery

Soybeans can be readily hybridized in the greenhouse. Potted plants can be moved around and positioned at levels that are convenient for the crossing operation.



## Industry highlights

### *Estimating inheritance factors and developing cultivars for tolerance to charcoal rot in soybean*

James R. Smith

USDA-ARS Crop Genetics and Production Research Unit, Stoneville, MS 38776, USA



**Figure 1** Soybean plants infected with *M. phaseolina* that have wilted and died prematurely. Note the dead leaves and petioles attached to the plants. (Courtesy of A. Mengistu.)



**Figure 2** Microsclerotia on the lower interior (vascular, cortical, and pith tissues) of a split soybean stem that was inoculated with *M. phaseolina* at the time of planting and harvested at physiological maturity.

The following is a summary of work in progress for the purpose of understanding the inheritance of soybean (*Glycine max* (L.) Merr.) tolerance to charcoal rot and for developing improved varieties with tolerance to it. The summary demonstrates the kind of thinking and planning that needs to take place to successfully achieve these objectives.

Charcoal rot is a disease of soybean caused by the fungus *Macrophomina phaseolina* (Tassi) Goidanich and occurs throughout the world, where it can cause severe yield losses to soybean. Charcoal rot symptoms appear during midsummer under conditions of high temperature (28–35°C) and low soil moisture availability. Diseased plants may wilt and prematurely die, with dead leaflets and petioles remaining attached to the plant (Figure 1). The most diagnostic symptom of charcoal rot is the black speckled (charcoal-like) appearance of microsclerotia on the exterior and interior (vascular, cortical, and pith tissues) of the lower stem of affected plants (Figure 2). Disease severity can be affected by heat, drought, and the presence of other pathogens. Attempts to control charcoal rot in soybean have included crop rotation, application of fungicides, biological controls, and adjustments to plant population, planting date, and irrigation protocols. Because these strategies have largely failed, researchers have become increasingly interested in the potential of varietal tolerance as a means of reducing yield losses. However, little progress has been made in developing charcoal rot-tolerant varieties, and the inheritance of varietal tolerance to charcoal rot is unknown.

In order to understand the inheritance of tolerance to charcoal rot and utilize it in breeding programs, it is necessary to generate segregating families (populations) of plants. Therefore, highly tolerant and highly susceptible soybean lines were chosen as parents, controlled pollinations between them were successfully made, and the appropriate segregating populations were generated.

Various options are available with segregating populations for estimating the importance of genetic factors in tolerance to charcoal rot. One way is by estimating heritability. Heritability can be expressed as the ratio of genetic variation over the total variation (genetic and non-genetic) influencing the trait. Traits with high heritability can be easily selected by plant breeders, whereas those with low heritability are not

easily selected. High heritability, in the narrow sense ( $\Phi_A^2/\Phi_P^2$ ), means that a high proportion of the total observed phenotypic variability ( $\Phi_P^2$ ) is accounted for by what is called additive genetic variance ( $\Phi_A^2$ ). Additive genetic variance is the variance due to the average effects of alleles (Bernardo 2002, p. 91). It measures the variation in the effects that are transmitted from one generation to the next. It does not imply that the alleles act in a purely additive manner. Rather, segregating loci with dominance, partial dominance, and overdominance gene actions, as well as additive gene action, can all contribute to  $\Phi_A^2$  (Bernardo 2002, p. 92). Epistatic gene action can also contribute to  $\Phi_A^2$  as well as to  $\Phi_D^2$  (dominance genetic variance) and  $\Phi_I^2$  (epistatic genetic variance) (Bernardo 2002, p. 96).  $\Phi_D^2$  is the variance of dominance deviations and indicates dominance gene action, which describes the genotypic interaction of alleles within a locus.  $\Phi_D^2$  is not considered useful for making progress from selection because intralocus interactions are not passed on to progeny. Rather, meiosis determines that only one allele from each diploid intralocus interaction will be passed on to progeny through gametes. Hence, an individual's dominance genotypic interactions are not passed on to its progeny and that is why  $\Phi_D^2$  is not considered useful for making progress from selection. When  $\Phi_D^2 = 0$ , dominance gene action is absent and the intralocus variance is comprised solely of  $\Phi_A^2$ , indicating purely additive gene action within the locus, but not necessarily between loci (Bernardo 2002, p. 92). Some components of  $\Phi_P^2$ , such as  $\Phi_{AA}^2$  (additive by additive epistatic genetic variance), are useful in selection (Hanson 1963, p. 133), because such interlocus interactions can be passed on to progeny. A broad-sense estimate of heritability ( $\Phi_G^2/\Phi_P^2$ ) includes the variance components of all types of gene action ( $\Phi_G^2$ ), some of which (dominance, etc.) are not passed on from parent to offspring. Exceptions to the above, where dominance gene action can be useful in selection, are when selecting among asexually propagated clones or among single-cross hybrids (Bernardo 2002, p. 109).

If tolerance were controlled by one or a few major genes, with only minimal environmental influence, then heritability would be expected to be high and selection for charcoal rot tolerance would be effective in the  $F_2$  generation among single plants or in the  $F_3$  generation among  $F_{2:3}$  progenies. If tolerance was controlled by multiple genes, but the environment had only a minor influence, then heritability might still be high and selection in early generations still effective. However, if heritability for tolerance to charcoal rot was low, then selection might only be effective among advanced generation breeding lines grown in replicated experiments.

Creating segregating populations of multiple generations ( $F_2$  and backcrosses to each parent) from two parents can serve multiple purposes in inheritance studies. Warner (1952) proposed utilizing the above generations ( $P_1$ ,  $P_2$ ,  $F_1$ ,  $F_2$ ,  $BC_1P_1$ ,  $BC_1P_2$ ) to estimate a narrow-sense heritability ( $\Phi_A^2/\Phi_P^2$ ), but suggested the need to test at least several hundred  $F_2$ ,  $BC_1P_1$ , and  $BC_1P_2$  individuals in order to reduce potential sampling error. Reinert and Eason (2000) provide an example of estimating a narrow-sense heritability in a self-pollinated species using the above generations. Bernardo (2002, pp. 110, 146) listed three disadvantages of estimating  $\Phi_A^2$  and  $\Phi_D^2$  using the above generations:

- 1 As single  $F_2$  and  $BC_1$  plants cannot be replicated, the lack of replication across environments causes these estimates of genetic variance to be confounded with the variance for genotype by environment.
- 2 Any linkage disequilibrium in the non-random-mated  $F_2$  and  $BC_1$  populations will cause the relationship between genotypic values at two loci to be confounded with  $\Phi_A^2$  and  $\Phi_D^2$ .
- 3 Individual plant measurements of quantitative traits are prone to large, non-genetic effects.

Another option for utilizing the above generations could be to determine the number of genes and their modes of action for simply inherited (one or two genes) traits. If all generations ( $P_1$ ,  $P_2$ ,  $F_1$ ,  $F_2$ ,  $BC_1P_1$ ,  $BC_1P_2$ ) are assayed together for tolerance to charcoal rot, segregation ratios of the  $F_2$ ,  $BC_1P_1$ , and  $BC_1P_2$ , along with the assay values of  $P_1$ ,  $P_2$ , and  $F_1$  individuals, can determine if inheritance is simple and if dominance gene action is evident. The segregating  $F_2$  generation can provide one estimate for an inheritance model, while the two backcross generations can provide a second and confirming estimate. Velez et al. (1998) and Singh and Westermann (2002) provide examples in dry bean (*Phaseolus vulgaris* L.) of utilizing the above generations in this way to determine the qualitative inheritance of resistance to bean disorders. However, when inheritance is affected by many genes and influenced greatly by environment, the above estimates may be inadequate.

An additional purpose for creating and utilizing the above generations could be to conduct an analysis of generation means. Generation mean analysis provides information on the relative importance of additive and dominance effects in populations created from two inbreds. It involves measuring the means of different generations ( $P_1$ ,  $P_2$ ,  $F_1$ ,  $F_2$ ,  $BC_1P_1$ ,  $BC_1P_2$ , etc.) derived from two inbreds and interpreting the means in terms of the different genetic effects (Bernardo 2002, p. 144). Because the actual means of single loci are unobservable, generation means estimate the pooled genetic effects across loci. Generation mean analysis is most useful when the two parents differ greatly in favorable alleles; that is when one parent has most, if not all, of the favorable alleles and the other parent has few, if any, favorable alleles. The pooled estimates of effects are summed across all loci for which  $P_1$  and  $P_2$  differ. Generation mean analysis has been commonly used to study disease resistance, where one parent has high resistance and the other has high susceptibility (Bernardo 2002, p. 145). Useful examples of generation mean analysis have been provided by Reinert and Eason (2000) in snap bean (*P. vulgaris* L.), Mansur et al. (1993) in soybean, and Campbell and White (1995) and Campbell et al. (1997) in maize (*Zea mays* L.).

Generation mean analysis can also be used to estimate effects due to epistasis, environment, genotype  $\times$  environment interactions, and linkage (Mather & Jinks 1971, pp. 83–119). However, the experiments can become much more complex and are unnecessary if the simpler additive-dominance model accounts for the variability present.

An important consideration for valid estimates of generation means is that there is sufficient sampling of segregating generations (Hallauer & Miranda 1981, p. 109). Bernardo (2002, p. 138) recommends that sampled breeding populations have a minimum of 50–100 progenies.

The advantages of generation mean analysis are:

- 1 It is relatively simple and statistically reliable (Mather & Jinks 1971, p. 126). Sampling errors are inherently smaller when working with means than with variances for estimating inheritance. Smaller experiments can therefore be used to obtain the same level of precision (Hallauer & Miranda 1981, p. 111; Campbell et al. 1997).
- 2 The estimation and interpretation of non-allelic interactions (epistasis) is more progressive for generation mean analysis than for variance estimates because mean effects are less confounded with one another and because the kinds of experiments required for analysis of means are smaller and easier to carry out than are those for variances (Mather & Jinks 1971, p. 126).
- 3 Populations evaluated in generation mean analysis can be used in applied breeding programs (Campbell et al. 1997).
- 4 It is equally applicable to both self- and cross-pollinated species (Hallauer & Miranda 1981, p. 111).

However, generation mean analysis also has several weaknesses:

- 1 It has limited value for quantitative traits whose parents have comparable mean performance (Bernardo 2002, p. 146).
- 2 As the information derived from the analysis is relevant to only a specific pair of parents, it has little application to other populations (Hallauer & Miranda 1981, p. 111).
- 3 Negative effects at some loci can cancel out positive effects at other loci, causing true genetic effects to be underestimated. Generation mean analysis does not reveal opposing effects (Bernardo 2002, p. 145). For example, a mean dominance effect of zero due to the cancellation of opposing effects does not mean that there are no dominance effects. But it does mean that there are no evident dominance effects and that a determination of the degree of dominance using variance components may be necessary.
- 4 Generation mean analysis does not provide estimates of heritability, which are essential for estimating predicted gain from selection (Hallauer & Miranda 1981, p. 111).
- 5 Finally, if epistatic effects are present, additive and dominance effects can be biased by the epistatic effects and by linkage disequilibrium (Hallauer & Miranda 1981, p. 110).

The use of generation mean analysis should be considered as complementary, rather than as an alternative, to variance component analyses (Mather & Jinks 1971, p. 126). However, if one estimates additive and dominance genetic effects using generation mean analysis, and also estimates  $\Phi_A^2$  and  $\Phi_D^2$ , there may be little relation in the magnitude of the two sets of estimates (Hallauer & Miranda 1981, p. 110). This might be expected because generation means estimate the sum of the genetic effects, whereas variances are the squares of the genetic effects. Further, the magnitude of  $\Phi_A^2$ ,  $\Phi_D^2$ , and  $\Phi_E^2$  may be poor indicators of the underlying gene actions for quantitative traits (Bernardo 2002, p. 144). For example, Moll et al. (1963) found that generation mean analysis detected epistatic effects that were not evident from estimates of  $\Phi_A^2$ ,  $\Phi_D^2$ , and  $\Phi_E^2$  (environmental variance), but noted that variance components may detect genetic variation not detected by generation mean analysis due to cancellation of mean effects.

Because of the weaknesses of single-plant data and the complexities of multigene quantitative traits, it may be advisable to generate  $F_3$  progenies and selfed progenies of both the  $BC_1P_1$  and  $BC_1P_2$  generations (Bernardo 2002, pp. 175–177). These generations are useful for estimating genetic variances, provided that they are developed without selection. The  $F_3$  progenies can be grown in replicated trials, which can provide an estimate of environmental effects and genotype  $\times$  environment interactions. Hamblin and White (2000) and Walker and White (2001) provide examples in maize of using the ANOVA (analysis of variance) of  $F_3$  progeny means to estimate  $\Phi_A^2$  heritability, and predicted gain from selection.

Hallauer and Miranda (1981, p. 91) recommended the use of  $F_3$  families for estimating  $\Phi_A^2$  in maize. Bernardo (2002, pp. 179–181) noted that an increase in either the number of environments or replications reduces the variance of an  $F_3$  family mean and in turn increases heritability. Hence, selection for quantitative traits among  $F_3$  families can be effective if each family is grown in extensive performance tests (Bernardo 2002, p. 181).

A further advantage of developing  $F_3$  progenies is that the segregation ratios within each  $F_3$  progeny can be used to confirm applicable  $F_2$  qualitative inheritance models. Thompson et al. (1997) used  $F_3$  progeny ratios to help accurately categorize each  $F_2$  plant genotype.

An additional genetic relationship that could be used to estimate heritability and predicted gain from selection from the above generations is that of the parent–offspring relationship between  $F_2$  individuals and  $F_{2:3}$  progeny means (Hallauer & Miranda 1981, p. 110). Using least squares regression of  $F_3$  offspring means onto individual  $F_2$  parental values, the slope ( $b$ ) is equal to  $(\Phi_A^2 + \frac{1}{2}\Phi_D^2)/\Phi_P^2$  and can be interpreted as the change in breeding value per change in phenotypic value (Bernardo 2002, p. 110). As were all the estimates of heritability previously discussed, this parent–offspring estimate is referenced to an  $F_2$  population assumed to be in Hardy–Weinberg equilibrium and considered to be non-inbred (Bernardo 2002, p. 34). This estimate of heritability may be biased upward by the presence of some potential amount of  $\Phi_D^2$  and any epistatic variance component involving  $\Phi_D^2$  ( $\Phi_{AD}^2$ , etc.).

An advantage of using parent–offspring regression to estimate heritability is that it is straightforward (Hanson 1963, p. 129). However,  $F_2$  individuals and derived  $F_3$  progenies are grown in separate environments (years), with no estimate of environmental variance and no estimate of genotype  $\times$  environment interaction. Depending on the variability of the trait measured, biases can be significant (Hanson 1963, p. 129).

In summary, the development and utilization of multiple generations ( $P_1$ ,  $P_2$ ,  $F_1$ ,  $F_2$ ,  $F_3$ ,  $BC_1P_1$ ,  $BC_1P_2$ ,  $BC_1P_1S_1$ ,  $BC_1P_2S_1$ , etc.) can provide flexibility and opportunities for multiple estimates of inheritance factors, both qualitative and quantitative. Such populations are also complementary for developing breeding populations.

In conjunction with the above early generation studies,  $F_2$  plants were grown for the purpose of producing recombinant inbred lines (RILs) by single-seed descent. The use of RILs in genetic studies requires that the finished inbred lines be an unselected sample of the  $F_2$  population; ideally, an inbred line for each  $F_2$  plant grown (Hallauer & Miranda 1981, p. 89). If inbreeding depression is a problem, resulting in the death of some plants before becoming inbred, a random non-selected RIL population may be difficult to obtain. However, this is usually not the case with self-pollinated crops such as soybean. The net result of using RILs to estimate inheritance, as opposed to an  $F_2$  population, is that the  $\Phi_A^2$  among RILs is twice as large as the  $\Phi_A^2$  among  $F_3$  families or among  $F_2$  individuals. Because each RIL is theoretically completely inbred,  $\Phi_D^2$  is zero (there are no intralocus genotypic interactions), which also means that there are no unusable  $\Phi_{AD}^2$  interactions. Hence, selection among RILs is expected to be more effective than selection among  $F_{2,3}$  families (Bernardo 2002, p. 180).

However, Hallauer and Miranda (1981, p. 91) noted that the use of RILs has two serious handicaps. First, as already noted, RILs require the development of a set of unselected inbred lines that are representative of genotypes of a reference population (the  $F_2$ ). This can be difficult if, during the selfing process, inbreeding produces weak plants that die and if high disease pressure kills plants. Second, the time required to develop RILs is much greater than that for developing and testing  $F_3$  progeny rows. Hence, it is wise to make early generation estimates of predicted gain from selection so as to determine if early generation selection can be profitable. Early generation selection can be highly desirable, if it is possible, enabling a greater allocation of resources to the most promising families (Bernardo 2002, pp. 119, 180, and 181).

However, where heritabilities are low (due to environment, genotype  $\times$  environment interactions,  $\Phi_D^2$ , etc.), RILs can provide useful estimates of genetic parameters and can be complementary to the breeding program. Replicated field trials across multiple locations and years with RILs will likely result in better estimates of heritability, better estimates of genetic gain from selection, and in the potential development of improved lines.

The development and utilization of RILs can facilitate the construction of a genetic map based on molecular marker linkages in the RIL population. If sufficient markers are located at enough strategic points, markers can be detected that are linked to genetic factors controlling the expression of quantitative traits. This was an important consideration in the development of RILs segregating for tolerance to charcoal rot and will allow for the development and release of charcoal rot-tolerant varieties, along with the release of molecular markers tightly linked to genetic factors affecting tolerance. Breeders can then create populations and test them in their own unique environments, while selecting for charcoal rot tolerance using molecular markers.

The most cost-effective way to reduce losses from plant diseases and stresses is through the use of cultivars with tolerance to the appropriate stresses. It is anticipated that this will be the case with charcoal rot of soybean. However, at this time, insufficient information is available to determine how tolerance to charcoal rot is inherited: as a single gene, as multiple genes with high heritability, or as multiple genes with low heritability. But, because of proper planning and execution, sufficient quantities of the appropriate segregating populations have been developed to effectively determine the inheritance of tolerance to charcoal rot and to maximize the available genetic potential for developing improved soybean varieties with tolerance to charcoal rot.

## References

- Bernardo, R. 2002. Breeding for quantitative traits in plants. Stemma Press, Woodbury, MN.
- Campbell, K.W., A.M. Hamblin, and D.G. White. 1997. Inheritance of resistance to aflatoxin production in the cross between corn inbreds B73 and LB31. *Phytopathology* 87:1144–1147.
- Campbell, K.W., and D.G. White. 1995. Inheritance of resistance to *Aspergillus* ear rot and aflatoxin in corn genotypes. *Phytopathology* 85:886–896.
- Hallauer, A.R., and J.B. Miranda. 1981. Quantitative genetics in maize breeding. Iowa State University Press, Ames, IA.
- Hamblin, A.M., and D.G. White. 2000. Inheritance of resistance to *Aspergillus* ear rot and aflatoxin production in corn from Tex6. *Phytopathology* 90:292–296.
- Hanson, W.D. 1963. Heritability. In: Statistical genetics and plant breeding (Hanson, W.D., and H.F. Robinson, eds), pp. 125–140. National Academy of Sciences and National Research Council, Washington, DC.
- Mansur, L.M., A.L. Carriquiry, and A.P. Rao-Arelli. 1993. Generation mean analysis of resistance to race 3 of soybean cyst nematode. *Crop Sci.* 33:1249–1253.
- Mather, K., and J. Jinks. 1971. Biometrical genetics, the study of continuous variation, 2nd edn. Cornell University Press, Ithaca, NY.
- Moll, R.H., D.L. Thompson, and P.H. Harvey. 1963. A quantitative genetic study of the inheritance of resistance to brown spot (*Physoderma maydis*) of corn. *Crop Sci.* 3:389–391.

- Reinert, R.A., and G. Eason. 2000. Genetic control of O<sub>3</sub> sensitivity in a cross between two cultivars of snap bean. *J. Am. Soc. Hort. Sci.* 125(2):222–227.
- Singh, S.P., and D.T. Westermann. 2002. A single dominant gene controlling resistance to soil zinc deficiency in common bean. *Crop Sci.* 42:1071–1074.
- Thompson, J.A., R.L. Bernard, and R.L. Nelson. 1997. A third allele at the soybean dt1 locus. *Crop Sci.* 37:757–762.
- Walker, R.D., and D.G. White. 2001. Inheritance of resistance to *Aspergillus* ear rot and aflatoxin production in corn from CI2. *Plant Dis.* 85:322–327.
- Warner, J.N. 1952. A method for estimating heritability. *Agron. J.* 44:427–430.
- Velez, J.J., M.J. Bassett, J.S. Beaver, and A. Molina. 1998. Inheritance of resistance to bean golden mosaic virus in common bean. *J. Am. Soc. Hort. Sci.* 123(4):628–631.

## Artificial pollination

### Equipment and materials

The equipment and materials used in soybean hybridization include fine-tip forceps, tags, a bottle with alcohol, Petri dish, pencil, and headband-mounted magnifiers.

### Emasculation

Soybean is emasculated in the bud stage. The suitable buds are those ready to open the following day. At this stage the corolla is visible at the tip of the bud. To emasculate, the flower is grasped between the thumb and index finger, keeping the hand steady to avoid snapping the pedicel. The calyx is removed with a pair of forceps by grasping the sepal and pulling with the forceps, and carefully pulling up while gently wiggling the forceps. It is important to remove all unemasculated flowers from the leaf axil. The leaf axil is tagged. If more than two cultivars are involved in the crossing program, the forceps should be sterilized between emasculations by dipping the tip in a bottle of alcohol.

### Pollination

Pollination immediately follows emasculation. Recently opened flowers have fresh pollen, and suitable flowers from the pollen source are picked. The stamen column and the style are together removed with a forceps, and used as a brush to dust pollen onto the stigma of the female flower. When a flower yields copious amounts of pollen, more than one stigma may be pollinated with one male staminal column.

Under humid conditions, as often occurs in the southern US, anther dehiscence occurs in the morning

when the female parts are less mature and more difficult to emasculate and the flowers may also be damp from the morning dew. The male flowers may be picked into Petri dishes and stored in desiccators until the afternoon. Pollen remains viable for about 2 days when stored in a cool dry place.

Pollinated flowers are tagged. The success rate of pollination depends on the skill of the operator and can range from zero to about 40%. The failure of a cross may be due to factors including immature female, injury to the stigma, injury to the pedicel, high temperature, and inadequate pollen.

## Natural pollination

Natural pollination is facilitated by the use of genetic male sterility. Breeders using recurrent selection method may benefit from this pollination method.

## Seed development and harvesting

Signs of successful cross are visible within 7 days of pollination. Sometimes, new pods may develop in the axis where the artificially pollinated flowers occur. These should be removed. A successful cross can be identified by the presence of the calyx scar resulting from the emasculation process.

Mature soybean should be harvested on time to avoid field loses. The crop is ready to harvest when the seeds are at the hard-dough stage. The moisture content of the seed should be 12–14%. Drier seed (less than 12%) increases the incidence of seed coat cracking and splitting and shattering of the seed. The crop may be harvested at high seed moisture content (17–18%) provided postharvest drying is available.

## Common breeding objectives

**1 Grain yield.** Soybean grain yield is a major breeding objective. Progress has been made, but not at the rate achieved in cereal crops such as corn. The major yield components are numbers of nodes per plant, number of pods per node, number of seeds per pod, and seed size. Adapting soybean to new production environments is followed by improvement in yield potential for cultivars in those new regions.

**2 Agromorphological traits:**

- (a) **Lodging resistance.** Lodging resistance is critical to mechanized production. An attempt to decrease plant height to improve lodging resistance by using the *dt<sub>1</sub>* gene for determinate growth, produced genotypes that were shorter but lower in plant yield.
- (b) **Shattering resistance.** Shattering resistance is important for production to allow the crops to attain full maturity before harvesting. Sometimes, inclement weather may force harvesting to be delayed even after the crop is mature. The mechanical harvesting process involves physical contact with moving parts of the combine, which may promote shattering.

**3 Adaptation:**

- (a) **Maturity.** Thirteen maturity groups have been identified for soybean. Late maturity in soybean has been found to be conditioned by three dominant genes, *E<sub>1</sub>*, *E<sub>2</sub>*, and *E<sub>3</sub>*. Another gene, *E<sub>4</sub>*, has been found to condition sensitivity to long photoperiods.
- (b) **Drought and other environmental stresses.** Drought resistance is important in areas where production is rainfed. Nutrient stress and heat stress may be important in areas where the production environment is not ideal.
- (c) **Herbicide resistance.** The modern technology of genetic engineering has enabled herbicide-

tolerant cultivars to be developed for commercial production (e.g., Roundup Ready®).

**4 Disease resistance.** Soybean is plagued by numerous diseases, the major ones including the following:

- (a) **Bacterial blight.** Caused by *Pseudomonas syringae*, this pathogen occurs worldwide. Resistance to the disease has been incorporated into various breeding programs.
- (b) **Rots.** *Phytophthora* root rot is the most important of the rots affecting soybean. Caused by *P. megasperma*, this problem affects all stages of development. Resistance is conditioned by six dominant genes, *Rps<sub>1</sub>–Rps<sub>6</sub>*.
- (c) **Viral disease.** The main viral disease of soybean is the soybean mosaic virus (SMV). It is spread by aphids. Genes for resistance, *Rsv<sub>1</sub>* and *Rsv<sub>2</sub>*, have been discovered for SMV.
- (d) **Nematodes.** The major species of nematodes of importance to soybean producers are the cyst nematode (*Heterodera glycines*) and root knot nematode (*Meloidogyne* spp.). Three recessive genes, *rhg<sub>1</sub>*, *rhg<sub>2</sub>*, and *rhg<sub>3</sub>*, and one dominant gene, *Rhg<sub>4</sub>*, have been identified and incorporated into breeding programs to develop resistant cultivars.

**5 Insect resistance.** Major insect pests of soybean include the southern green stink bug and beanflies, which are common in Asia and Africa.

**6 Seed compositional traits and quality:**

- (a) **Oil quality.** Soybean is the world's leading source of vegetable oil, accounting for more than 75% of the market share. Breeding objectives include an increase in oil content as well as improvement in oleic acid and reduced linolenic acid for high oil quality.
- (b) **Seed protein.** Soybean is also the leading source of protein meal. Because seed oil and seed protein are negatively correlated, breeding has tended to focus on developing cultivars with a high protein and low oil content.

## References and suggested reading

- Carlson, J.B., and N.R. Lersten. 1987. Reproductive morphology. In: Soybeans: Improvement, production, and uses, 2nd edn (J.R. Wilcox, ed.), pp. 95–134. American Society of Agronomy, Madison, WI.
- Chen, Y., and R.L. Nelson. 2004. Genetic variation and relationship among cultivated, wild, and semiwild soybean. Crop Sci. 44:316–325.

- Kisha, T., C.H. Sneller, and B.W. Diers. 1997. Relationship between genetic distance among parents and genetic variation in populations of soybean. Crop Sci. 33:1313–1325.
- Specht, J.E., D.J. Hume, and S.V. Kumudini. 1999. Soybean yield potential – A genetic and physiological perspectives. Crop Sci. 39:1560–1570.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 Soybean is a subtropical plant.
- 2 Soybean maturity group VI is produced in the northern parts of the USA.
- 3 Soybean is long-day plant.

### Part B

Please answer the following questions:

- 1 Give the scientific name of soybean.
- 2 Is cultivated soybean is tetraploid?
- 3 Describe the reproductive biology of soybean.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss emasculation for crossing in soybean.
- 2 Discuss the breeding of agromorphological traits in soybean.
- 3 Discuss the genetic resources available for soybean breeding.
- 4 Discuss the history of soybean breeding in the USA.



32

# Breeding peanut

## Taxonomy

Kingdom	Plantae
Subkingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Magnoliopsida
Subclass	Rosidae
Order	Fabales
Family	Fabaceae
Genus	<i>Arachis</i> L.
Species	<i>Arachis hypogaea</i> L.

## Economic importance

Peanuts are an important legume crop in the warm climates of the world. The USA produces about 10% of the world's peanut crop on about 3% of the total world peanut acreage. This disproportionate share of the world's production is attributable to the high average yield per acre in the US (2,800–3,000 lb/acre) compared to the world's average yield (800–1,000 lb/acre) (see Appendix 2 for conversion rates of units). Nine states in the USA account for 99% of the US peanut crop: Georgia, Texas, Alabama, North Carolina, Florida, Oklahoma, Virginia, South Carolina, and New Mexico. Georgia alone produces 39% of the US total production. In 2001, the US harvested acreage was 1,411,900 acres, a total production of 4,276,704,000 lb, and an average yield of 3,029 lb/acre. There are three main peanuts regions in the US, Georgia–Florida–Alabama (southeast region), Texas–Oklahoma–New Mexico (southwest region), and the Virginia–South Carolina–North Carolina region, with the southeast region accounting for about 55% of all US production.

On the world scene, peanuts are produced in Asia, Africa, Australia, and the Americas. India and China

together account for more than 50% of the world's total production. Other substantial peanut producing nations include Senegal, Sudan, Brazil, Argentina, South Africa, Malawi, and Nigeria.

## Origin and history

The peanut is native to the Western Hemisphere. It probably originated in South America, the center of origin most likely being Brazil, where about 15 wild species are found. The Spanish explorers are credited with its spread throughout the New World. They introduced it to Europe from where traders spread it to Asia and Africa. Peanuts reached North America via the slave trade. Commercial production of peanuts in the USA began in about 1876. The demand for the crop increased after the Civil War, transforming it from a regional (southern) food to a national food. Production came to the Cotton Belt after 1900. The expansion of the peanut industry was driven by advances in technology that resulted in the development of equipment and machinery for planting, harvesting, and processing the crop.

## Market types

There are four basic market types – runner, Virginia, Spanish, and Valencia.

### Runner

Runners have become the predominant peanut type in the USA following the introduction of the cultivar "Florunner", which was responsible for the dramatic yield increase of the crop in the US. They have uniform size and are grown mainly in Georgia, Alabama, Florida,

Texas, and Oklahoma. About 54% of the crop is used for making peanut butter. Runners mature in about 130–150 days, depending on the cultivar. The seeds are medium sized (900–1,000 seeds/lb).

### *Virginia*

Virginia cultivars have dark green foliage and large pods. They have the largest seeds of all the types (about 500 seeds/lb). The pods usually have two seeds (occasionally 3–4), which have a russet testa. They are grown mainly in Virginia and North Carolina. They mature in about 135–140 days, and may have runner or bunch types. Large seeds are sold as snack peanuts.

### *Spanish*

The Spanish group of peanuts comprises bunch types with erect, light green foliage. The pods rarely contain more than two seeds, which are short with a tan testa. The seeds are small sized (1,000–1,400 seeds/lb). They have a higher oil content compared with other types. They are grown mainly in Oklahoma and Texas and are used mainly for making peanut candies, and also snack nuts and peanut butter. They mature earlier than the runner types (about 140 days).

### *Valencia*

The Valencia types typically bear many pods with 3–4 seeds and a bright-red testa color. They are erect and sparsely branching with dark green foliage. They are very sweet peanuts and are usually roasted and sold as in-the-shell or boiled peanuts. Valencias are grown mainly in New Mexico.

## Germplasm resources

Over 10,000 peanut accessions are held in germplasm banks in various countries. In the USA, the Southern Regional Plant Introduction Station at Experiment, Georgia, maintains about 4,000 accessions. The International Crops Research Institute for Semi-Arid Tropics (ICRISAT) also maintains thousands of accessions.

## Cytogenetics

The genus *Arachis* is classified into seven sections. The cultivated peanut belongs to the section *Arachis*, which

consists of three series – annuals, perennials, and amphidiploids. There are four annual diploids ( $2n = 20$ ), several (at least five) perennial diploids, and two annual tetraploids ( $2n = 40$ ). The series amphidiploide has two tetraploids – *A. hypogaea* and *A. monticola*. Some intersectional crosses yield fertile hybrids. Others produce meiotic problems leading to embryo abortions. The cultivated peanut is a tetraploid with two subspecies and four interfertile varieties:

- 1 Subspecies *A. hypogaea hypogaea* var. *hypogaea* (US market types Virginia and runner) and var. *hirsute*.
- 2 Subspecies *A. hypogaea fastigiata* var. *fastigiata* (US market type Valencia) and var. *vulgaris* (US market type Spanish)

## General botany

The peanut (*A. hypogaea*), also called groundnut and earthnut, among other names, is technically a pea not a nut. It is an unusual plant in the sense that it flowers above ground, but fruits below ground. The cultivated plant is an annual with a central upright stem that may stand up to about 45 cm tall. It bears numerous branches that vary from prostrate to nearly erect. It has pinnately compound leaves. The cultivars in cultivation may be grouped into two, based on the arrangement of the nuts at the base of the stem. In **bunch types**, the nuts are closely clustered about the base, while in **runner types** the nuts are scattered along prostrate branches that radiate from the base of the plant to the top. Peanuts have a strong taproot system; most roots are nodulated for biological nitrogen fixation. The branching pattern in the peanut may be alternate or sequential. Virginia types have alternate branching, a trait that is generally believed to be dominant over the sequential branching of Spanish and Valencia peanuts. Aerial podding is conditioned by dominance or partial dominance.

The flowers arise in the leaf axils above the ground and are self-pollinated. Upon fertilization, the ovary begins to enlarge while the section behind it, called the **peg** or **gynophore**, elongates to push the ovary into the soil for fruit development. The fruit is an indehiscent pod that may contain 1–6 (but usually 1–3) seeds. The pods form predominantly underground. However, mutants with aerial podding have been developed. Consequently, it is critical that the pegs reach the soil. The seed has a thin papery testa that varies in color – brick red, russet, light tan, purple, white, black, or



## Industry highlights

### Peanut (*Arachis hypogaea L.*) breeding and root-knot nematode resistance

Charles Simpson

Texas A&M University, College Station, TX 77843, USA

#### Program objectives

The Texas peanut breeding program originated in 1939 with the principal objective of improving yield stability of the crop. A later objective included the development of disease (pest) resistance, and, more recently, an objective to improve the edible quality of the product has been implemented. Pedigree selection with and without single-seed descent has been used extensively in the program, but the mainstay of the disease-resistance transfer has been accomplished via the backcross technique. Diseases and nematodes have been recognized as major constraints to production since the very early stages of the program. After the discovery of moderate resistance in some plant introduction material in the late 1950s, a project was initiated to develop leaf spot resistance. To enhance the available gene pool, a concentrated effort began in 1970 to introduce leaf spot resistance genes from the wild species. In 1976 an international effort was initiated to collect, preserve, evaluate, and utilize wild *Arachis* from the area of origin of the genus in South America. Through this effort more than 1,500 new wild *Arachis* accessions were added to the germplasm collection, including more than 60 new species and additional representatives of the 22 previously described species. We also introduced more than 3,900 landraces of *A. hypogaea*.

#### Introgression program

##### Leaf spot

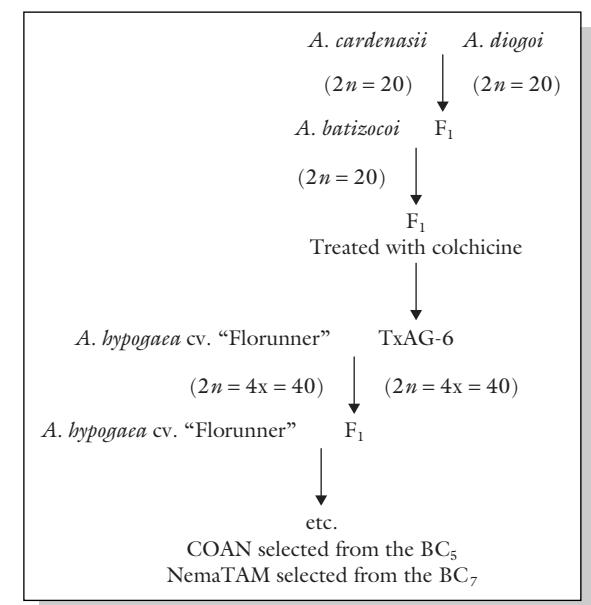
Our introgression program began in earnest in 1972. Abdou et al. (1974) had identified near immunity to two leaf spot pathogens (*Cercosporidium personatum* (Beck and Curtis) Deighton and *Cercospora arachidicola* Hori) in *A. cardenasii* Krapov. and W.C. Gregory and *A. chacoensis* nom. nud. (now classified as *A. diogoi* (Krapovickas & Gregory 1994)), respectively. Most of the wild *Arachis* species are  $2n=20$ , but the cultigen is  $2n=4x=40$ . We had used both *A. cardenasii* and *A. diogoi* species individually in attempts to cross with *A. hypogaea* and form fertile hexaploids from interspecific triploid progenies, followed by crossing of the  $6x$  progeny back to  $4x$  individuals, then further backcrossing of progenies to  $4x$  to accomplish chromosome segregation and to eventually develop tetraploid progenies that were cross-compatible with *A. hypogaea*, and possessing the leaf spot resistance(s) (Simpson 1990). With both species, widespread sterility occurred after one or two generations of backcrossing to the *A. hypogaea* recurrent parent and in several lines sterility prevented the second backcross. In 1973 we attempted to combine the two resistances into one diploid hybrid and use that hybrid to attempt the hexaploid route. A successful hybridization of *A. diogoi*  $\times$  *A. cardenasii* was achieved only once in 3,500 pollinations, resulting in a partially fertile hybrid with 50% pollen stain. However, that was the end of that pathway's success. Major sterility problems ensued as we doubled chromosome numbers of the triploid and attempted crosses/backcrosses to *A. hypogaea* (Simpson 1990).

The next attempt consisted of doubling the chromosome number of the *A. diogoi*  $\times$  *A. cardenasii* hybrid, then crossing with *A. hypogaea*. Even though this was a tetraploid  $\times$  tetraploid cross, the process met with equal, if not greater, sterility problems than previous attempts.

In this timeframe, Smartt et al. (1978a, 1978b) published their theory that the cultivated peanut was comprised of two genomes, made up of an A-genome which is represented by *A. cardenasii* and *A. diogoi*, and a B-genome represented by *A. batizocoi*. Based on this hypothesis, we crossed the *A. diogoi*  $\times$  *A. cardenasii* hybrid onto *A. batizocoi* and obtained a sterile (pollen stain = 0.01%) hybrid. We doubled the chromosomes of this plant to yield a highly fertile complex amphiploid (tetraploid) hybrid with normal meiosis and a pollen stain of >90%. From this point we initiated our crossing/backcrossing effort to establish fertile progenies with leaf spot resistance (Figure 1). Progress was slow primarily because we had not established a reliable laboratory technique for screening large numbers of plants, therefore each cycle required more than 18 months when selections were made based on leaf spot resistance. However, we carried out two sections of the program, with the hope that selection based on fertility and agronomic traits other than resistance would yield fertile plants with leaf spot resistance at an earlier date. Without selection for resistance, the program moved rapidly, with a cycle being made in 10 months or less. The initial complex amphiploid from this program was later released as the germplasm line TxAG-6 (Simpson et al. 1993).

##### Nematodes

Root-knot nematodes (*Meloidogyne arenaria* (Neal) Chitwood) became a serious pest of peanut in the major growing area of central Texas (as well as many other places in the USA and the world) by the early 1970s. However, no useful level of resistance had



**Figure 1** Tetraploid pathway for introgression of root-knot nematode resistance in *Arachis hypogaea*.

### Backcross program

Our backcross efforts (Simpson 1990) consisted of making crosses in the fall of any one year at the Texas A&M University Research and Extension Center in Stephenville; harvest and then plant the  $F_1$  seeds as soon as they were mature enough to germinate (usually mid-December); harvest seeds ( $F_2$  or  $BC_nF_1$ ) from these plants in late April; and then send the seeds to the Plant Nematology Laboratory in College Station for resistance screening. Seeds were planted in 10 cm pots and the seedlings inoculated with 10,000 root-knot nematode eggs per pot. Plants were harvested 8 weeks after inoculation and highly resistant plants were identified based on low levels of nematode reproduction (Starr et al. 1990, 1995). Stem cuttings from resistant individuals were then sent back to Stephenville to be used as male parents in the next backcross generation, which was initiated in September and October. In each cycle, the primary selection criterion was nematode resistance, especially in the earlier generations. After the  $BC_3$  generation we also began making selections based on the plant characters of pod shape and seed per pod. We continued this process through seven backcrosses (Figure 1) (Simpson 1990).

At about the third backcross generation we determined that at least two resistance genes were present in the early generation materials, but at this stage it was not certain that more than one dominant gene controlled the trait (Starr et al. 1990) in our more advanced generation lines. At this time we also initiated a collaboration with A. H. Paterson and M. D. Burow to develop molecular markers linked to the resistance loci (Burow et al. 1996).

### Releases

After the fifth backcross, we also initiated seed increases of lines that appeared to have most of the plant, pod, and seed characteristics of "Florunner" as well as nematode resistance. Field tests for yield potential began in 1995 and from these tests we identified five lines with reasonable yield and excellent resistance. Following a second year of testing we made a winter increase of numerous sister seed rows in Puerto Rico from one line, and from this material we released the world's first root-knot-resistant peanut cultivar, "COAN", in 1999 (Simpson & Starr 2001). This release was rushed through the process and had one major flaw – the plant size was too small, limiting seed production of the plants. Even though the yields under heavy nematode pressure were 150–200% better than susceptible cultivars, the total yields were still too low to be profitable for grower acceptance.

During this time the molecular marker work was progressing and preliminary tests with several RFLP (restriction fragment length polymorphism) loci indicated that we would be able to use this technology for more efficient screening for resistance, and even be able to identify plants homozygous for the one major dominant gene that appeared to control resistance (Burow et al. 1996).

After the  $BC_7$  was made, selection and seed increase resulted in lines that performed much better than "COAN" in yield tests and still contained the resistance gene. After 1 year of testing we decided to do a Puerto Rico winter increase of three of the best

been identified in the cultigen, so we began testing for resistance in the wild species. Banks (1969) had demonstrated some resistance to *M. hapla* in the *Rhizomatosae*, but since the rhizomatous taxa would not cross with the cultivated peanut or its sectional relatives (Gregory & Gregory 1979) these genes were unavailable. In 1989 numerous accessions of wild peanut were identified with resistance (near immunity) to the root-knot nematode (Nelson et al. 1989). Fortunately one of those accessions was *A. cardenasi*, the primary A-genome parent we had been working with in the leaf spot resistance program. Our next effort was to test the advanced backcross generation materials from the leaf spot program in hopes that the gene(s) for nematode resistance might have been retained. Unfortunately, the only materials that proved to be resistant to the root-knot nematode were TxAG-6 and the  $F_1$  from the first backcross to *A. hypogaea*. Even so, availability of a high level of resistance to the root-knot nematode in genotypes that were readily cross-compatible with the cultigen was a tremendous benefit to the effort to develop a nematode-resistant cultivar (Simpson 1990).

We then initiated an intense backcross breeding program with a selection from the  $BC_1F_1$  materials, subsequently released as the germplasm line TxAG-7 (Simpson et al. 1993), to transfer the gene(s) to the most desirable cultivated peanut available, which at that time was "Florunner". "Florunner", a revolutionary peanut because of its large increase in yield, was released from the University of Florida in 1969 (Norden et al. 1969) and by 1986 had become the most successful peanut variety ever developed, worldwide.

lines in hopes of being able to select a line for release as a cultivar with greater yield potential. In the second testing year we space planted 300 individual seeds from the winter nursery of the three lines in the field at Stephenville. We collected tissue for DNA analyses from the 900 plants and at harvest we had molecular data to identify the homozygous resistant plants that had the dominant (*RR*) gene (Church et al. 2000). These data were utilized when we were evaluating the individual plants for plant, pod, and seed characters and selecting individual plants for a breeder seed increase. We selected numerous desirable plants that were homozygous resistant from each line. The seeds from these were planted as plant rows in a Puerto Rico winter nursery to gain another generation. After a third year of yield testing and extensive other evaluations, the cultivar "NemaTAM" was released in 2001 (Simpson et al. 2003). "NemaTAM" has c. 30% greater yield potential than "COAN" and the same high level of root-knot resistance. Two major benefits of these resistant cultivars is that: (i) the resistance will eliminate the need for use of nematicides even at very high nematode population densities; and (ii) the inhibition of nematode reproduction due to the resistance results in lower nematode population densities such that a susceptible crop plant in rotation with the resistant cultivar will be subjected to less nematode disease pressure (Starr et al. 2002).

The resistance in "COAN" and "NemaTAM" to *M. arenaria* is controlled by a single dominant gene (Burow et al. 1996; Choi et al. 1999; Church et al. 2000), but evidence for additional genes in the same species used to form TxAG-6 indicates that we have the opportunity to pyramid genes for more stable resistance (Burow et al. 1996; Garcia et al. 1996; Choi et al. 1999).

Further testing indicates that the resistance in TxAG-6 is conditioned by at least two genes, one dominant and one recessive (Church 2002). We also discovered that "COAN" and "NemaTAM" are resistant to *M. javanica*, which is also parasitic on peanut and especially prevalent in India and northern Africa. This resistance to *M. javanica* has been confirmed in an independent study (Timper et al. 2003). At present we can only assume the resistance is conditioned by the same gene; we have not tested this hypothesis.

### *The future*

The program continues to try to identify, characterize, and locate flanking molecular markers for the second resistance gene so we can pyramid the genes. It would be desirable to move from the current RFLP marker-assisted selection system to one based on the polymerase chain reaction, which would increase the efficiency of the system. We are also continuing our efforts to identify genes and markers resistance to *M. hapla* as well as *M. javanica*. However, now our major efforts are to combine the nematode resistance gene(s) with other characters to develop cultivars with multiple traits, including high O/L (ratio of oleic free fatty acid to linoleic free fatty acid), tomato spotted wilt virus resistance, sclerotinia resistance (*Sclerotinia minor* Jagger), and leaf spot resistance.

### *References*

- Abdou, Y.A.-M., W.C. Gregory, and W.E. Cooper. 1974. Sources and nature of resistance to *Cercospora arachidicola* Hori and *Cercosporidium personatum* (Beck and Curtis) Deighton in *Arachis* species. Peanut Sci. 1:6–11.
- Banks, D.J. 1969. Breeding for northern root-knot nematode, *Meloidogyne hapla*, resistance in peanuts. J. Am. Peanut Res. Educ. Assoc. 1:23–28.
- Burow, M.D., C.E. Simpson, A.H. Paterson, and J.L. Starr. 1996. Identification of peanut (*Arachis hypogaea* L.) RAPD markers diagnostic of root-knot nematode (*Meloidogyne arenaria* (Neal) Chitwood) resistance. Mol. Breed. 2:369–379.
- Choi, K., M.D. Burow, G. Church, G. Burow, A.H. Paterson, C.E. Simpson, and J.L. Starr. 1999. Genetics and mechanism of resistance to *Meloidogyne arenaria* in peanut germplasm. J. Nematol. 31:283–290.
- Church, G.T. 2002. Resistance to *Meloidogyne arenaria* in peanut: Gene identification and molecular markers. PhD dissertation, Texas A&M University, College Station, TX.
- Church, G.T., C.E. Simpson, M.D. Burow, A.H. Paterson, and J.L. Starr. 2000. Use of RFLP markers for identification of individuals homozygous for resistance to *Meloidogyne arenaria* in peanut. Nematology 2:575–580.
- Garcia, G.M., H.T. Stalker, E. Shroeder, and G. Kochert. 1996. Identification of RAPD, SCAR, and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* into *Arachis hypogaea*. Genome 39:836–845.
- Gregory, M.P., and W.C. Gregory. 1979. Exotic germ plasm of *Arachis* L. interspecific hybrids. J. Hered. 70:185–193.
- Krapovickas, A., and W.C. Gregory. 1994. Taxonomia del genero *Arachis* (Leguminosae). Bonplandia 8(1–4):1–186.
- Nelson, S.C., C.E. Simpson, and J.L. Starr. 1989. Resistance to *Meloidogyne arenaria* in *Arachis* spp. germplasm. J. Appl. Nematol. Suppl. 21:654–660.
- Norden, A.J., R.W. Lipscomb, and W.A. Carver. 1969. Registration of Florunner peanuts. Crop Sci. 9:850.
- Simpson, C.E. 1990. Pathways for introgression of pest resistance into *Arachis hypogaea* L. Peanut Sci. 18:22–26.
- Simpson, C.E., and J.L. Starr. 2001. Registration of "COAN" peanut. Crop Sci. 41:918.
- Simpson, C.E., J.L. Starr, G.T. Church, M.D. Burow, and A.H. Paterson. 2003. Registration of "NemaTAM" peanut. Crop Sci. 43:1561.
- Simpson, C.E., J.L. Starr, S.C. Nelson, K.E. Woodard, and O.D. Smith. 1993. Registration of TxAG-6 and TxAG-7 peanut germplasm. Crop Sci. 33:1418.

- Smartt, J., W.C. Gregory, and M.P. Gregory. 1978a. The genomes of *Arachis hypogaea* I. Cytogenetic studies of putative genome donors. *Euphytica* 27:665–675.
- Smartt, J., W.C. Gregory, and M.P. Gregory. 1978b. The genomes of *Arachis hypogaea* II. The implications in interspecific breeding. *Euphytica* 27:677–680.
- Starr, J.L., E. Morgan, and C.E. Simpson. 2002. Management of the peanut root-knot nematode, *Meloidogyne arenaria*, with host resistance. Available at Plant Health Progress doi:10.1094/PHP-2002-1121-01-HM.
- Starr, J.L., G.L. Schuster, and C.E. Simpson. 1990. Characterization of the resistance to *Meloidogyne arenaria* in an interspecific *Arachis* spp. hybrid. *Peanut Sci.* 17:106–108.
- Starr, J.L., C.E. Simpson, and T.A. Lee, Jr. 1995. Resistance to *Meloidogyne arenaria* in advanced generation breeding lines of peanut. *Peanut Sci.* 22:59–61.
- Timper, P., C.C. Holbrook, and W.F. Anderson. 2003. Reproduction of *Meloidogyne* spp. on resistant peanut genotypes from three breeding programs. *J. Nematol.* 35:417–421.

multicolored. Some cultivars exhibit seed dormancy. Well-developed nuts have a shelling percentage of 70–80%.

## Reproductive biology

### Floral morphology

The peanut inflorescences occur in clusters of three or more flowers in the axil of the cataphyll of foliage leaves. Inflorescences may occur on either the main stem or lateral branches. Spanish and Valencia cultivars bear their inflorescences on the main stem, whereas the inflorescences of Virginia cultivars occur on the lateral branches. A flower is subtended by a bract and occurs on a minute branch of the inflorescence, which arises in the axil of a second bract. The calyx and corolla are borne at the top of the hypanthium, which surrounds the staminal column. The calyx has five lobes, whereas the staminal column is usually composed of 10 filaments, eight of which are normally anther bearing. About 50–75% of the bottom parts of the filaments are fused.

### Pollination

Normally, only one of the flowers in an inflorescence matures to anthesis. Anthesis generally occurs before flower opening. Bud opening occurs at the beginning of the light period. The stigma remains receptive for about 12–24 hours after flower opening, and a flower only stays open for about 5–6 hours.

## Common breeding methods

The common approaches to breeding peanuts include the use of plant introductions. Introductions of peanut

germplasm can be a beginning point for crop improvement. Selections from introductions may provide a parental material for breeding. Selection may also be used to isolate pure lines from hybrid populations.

A widely used method of breeding peanuts is hybridization of superior parents to create opportunities for transgressive segregation to occur. Pedigree selection may be used to advance generations. However, the use of single-seed-descent is more rapid when used in conjunction with winter nurseries. Backcross breeding may be used to incorporate specific genes of interest.

## Establishing a breeding nursery

Peanut crossing is often done in the greenhouse using potted plants. However, wild species are more successfully hybridized in the field than in the greenhouse.

Success of hybridization, whether under field or greenhouse conditions, depends on proper humidity. Drought causes low success. Breeders may emasculate the flowers in the evening, and pollinate the next morning.

## Artificial pollination

### Materials and equipment

The equipment needed includes forceps, sharp knife, scalpel, razor blade, magnifier (2–3×) attached to a head band, camel hair brush, and a bottle of alcohol.

### Emasculation

Flowers near the main stem are preferred for emasculation. Further, one flower in each inflorescence is selected

for emasculation, which is done in the bud stage. To emasculate, the bud is grasped between the thumb and index finger. Next, the petal in front of the keel and the sepal on the side of the standard are folded down. The standard petal is opened with the forceps and the wing petals are pulled out and down. The standard is held back with the thumb and index finger while the operator pulls the keel free of the stigma and anthers. Alternatively, the keel and wing petals may be removed all together. The anther and the stamens are removed to complete the emasculation. If an emasculated flower will not be immediately pollinated, the hypanthium is tagged (e.g., with a small thread).

### Pollination

Pollen collected between 5 and 7 a.m. is most viable. Flowers are artificially pollinated in the morning to early afternoon. The keel may be detached and used as a brush to directly deposit the pollen onto the stigma. Some operators transfer the pollen with a camel hair brush or the tip of the forceps. The environment may be humidified, but care should be taken not to dislodge the pollen. Pollen may be collected and stored under desiccated condition in a cool place (6°C) for up to about a week without losing viability.

Fertilization occurs between 12 and 16 hours after pollination, after which the ovary elongates as the intercalary meristem at its base grows. The peg eventually penetrates the soil where it develops into a mature peanut fruit. An identification wire may be tied to the peg from the emasculated flower before it penetrates the soil.

### Seed development and harvesting

Seed is mature for harvesting after about 55–65 days, depending on the environment and cultivar. Peanuts are ready to be harvested when 65–70% of them are mature. Early harvesting results in shriveled kernels. At proper maturity, the kernels display the distinct texture and color of the variety, the inside of the shell beginning to color and show dark veins. Harvesting is easier in sandy soils, and occurs with less loss of pods with small-podded than large-podded varieties.

### Common breeding objectives

Some of the major objectives in peanut breeding are as follow:

- 1 Yield potential and stability.** Breeders are interested in high crop yield *per se*, but also a stable yield with adaptation to various agroecological zones. High shelling percentage is important. Early maturity is desired in some production areas.
- 2 Disease resistance.** Important diseases of peanut include foliar ones such as leaf spot (caused by *Cercospora* spp.). Stem and peg rots (caused by *Sclerotium rolfsii*) and charcoal rot (caused by *Rizoctonia* spp.) are economic diseases in some production area.
- 3 Insect resistance.** Important insect pests include leafhoppers, corn earworm, cutworms, and tobacco thrips.
- 4 Product quality.** Peanuts consist of 40–48% oil and 25–30% protein, making the crop a major source of vegetable oil and plant protein. Breeding high oil content is an important objective.

### References and further reading

- 
- Boerma, H.R., and R.S. Hussey. 1992. Breeding plants for resistance to nematodes. *J. Nematol.* 24:242–252.
- Hyman, B.C. 1990. Molecular diagnosis of *Meloidogyne* species. *J. Nematol.* 22:24–30.
- 
- Isleib, G., and G. Kochert. 1994. Introgression from *A. cardenasi* germplasm lines. *Crop Sci.* 33:1418.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- 1 The cultivated peanut is a diploid.
- 2 Peanut is self-pollinated.
- 3 The Virginia variety of peanuts is characterized by small pods.

### Part B

Please answer the following questions:

- 1 Give the four basic market types of peanut.
- 2 Give the scientific name of peanut.
- 3 Distinguish between runner and bunch types of peanut.

### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the reproductive biology of peanut.
- 2 Discuss the common breeding methods for peanut.
- 3 Discuss the emasculation of peanut flowers.

# Breeding potato



## Taxonomy

Kingdom	Plantae
Subkingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Magnoliopsida
Subclass	Asteridae
Order	Solanales
Family	Solanaceae
Genus	<i>Solanum</i> L.
Species	<i>Solanum tuberosum</i> L.

## Economic importance

Potato is among the top five crops that feed the world, the others being wheat, corn, sorghum, and rice. In 1998, potatoes ranked as the fourth most important food crop in the USA. The total harvest in 1998 was 1.388 million acres and a total yield of 475.8 million cwt (see Appendix 2 for conversion rates of units). About 35% of the production was processed into frozen products (primarily fries). Per capita consumption of potatoes in the US in 1999 was 144.7 lb. The top five producing states in 1998 were Idaho, Washington, Oregon, Wisconsin, and North Dakota, with Idaho leading all production with about 450 million cwt, followed by Washington.

On the world scene, 293 million tons of potatoes were produced on 18 million ha worldwide in 1998. With the breakup of the former Soviet Union, China has become the world's leading producer of potatoes. Developing countries accounted for 36% of total production in 1998. In Asia, the key producers include China, India, Indonesia, and Nepal; in Africa, Egypt, South Africa, Algeria, and Morocco account for 80% of the region's total production. Latin American countries

that produce substantial amount of potato include Ecuador, Peru, Brazil, and Mexico. In Europe, the leading countries include the former Soviet Union, Poland, West Germany, and France.

## Origin and history

The potato originates in the Andean mountains of Peru and Bolivia where the plant has been cultivated for over 2,400 years. The Aymara Indians developed numerous varieties on the Titicaca Plateau, some 3,000 m above sea level. The first written account of potato was made in 1553 by the Spanish Conquistador Pedro Cieza de Leon in his journal "Chronicle of Peru". The Spanish introduced the potato to Europe between 1565 and 1580. It was taken from England to Bermuda and later to Virginia, USA, in 1621. It was also introduced into Germany in the 1620s where it became a part of the Prussian diet by the time of the Seven Year War (1756–1763). Antoine Parmentier, a prisoner of war in Prussia, introduced the crop to France after the war. Potato was introduced into North America when Irish immigrants (hence the name "Irish potato") brought it to Londonderry, New Hampshire, where large-scale production occurred in 1719.

By the 1840s, devastation by the fungus *Phytophthora infestans* (causes late blight) and heavy rains brought untold hunger and starvation to Ireland. This sparked a mass immigration of about 2 million Irish, mostly to North America. In America, Luther Burbank was the first to undertake improvement of the plant, subsequently releasing the "Burbank" potato to west coast states in the late 1800s. A mutation of the "Burbank" potato was discovered in Colorado that was disease resistant. It had reticulated skin and became known as the "Russet Burbank" and is grown on most farms in Idaho.

## Adaptation

Potato is a cool season crop. The optimum temperature for shoot growth and development is 22°C. In the early stages, a soil temperature of about 24°C is ideal. However, in later stages, a cooler temperature of about 18°C is desired for good tuberization. Tuber formation and development is slowed when soil temperature rises above 20°C and ceases at 29°C. Above this temperature, the effect of respiration exceeds the rate of accumulation of assimilates from photosynthesis.

Potato is sensitive to the hard frosts that may occur in the fall, winter, or early spring. Tubers will freeze at about -2°C and loose quality upon thawing. Tuberization is best under conditions of short photoperiod, cool temperature, and low nitrogen, while vegetative shoot production is favored by long days, high temperature, low light intensity, and high amounts of nitrogen. However, tuberization can occur at 12°C night air temperatures. The best potato production occurs in regions with daily growing season temperatures averaging between 15.5 and 18°C. High soil temperatures result in knobbly and malformed tubers. Potato flowers and sets seed best when long days and cool temperatures prevail. Consequently, potatoes set seed when grown in the northern states but not the southern states.

## Cultivars

Four cultivars account for about 75% of the potato acreage in the US – “Russet Burbank”, “Katahdin”, “Kennebeck”, and “Red Pontiac”. These major varieties represent the most common shapes of potato – the long cylindrical and russet skin of the “Russet Burbank”, the red and short rounded shape of “Red Pontiac”.

Many new modern cultivars have been developed for various markets – baking, frying, cooking, canning, creaming, dehydrating, and chipping. These cultivars have certain specific characteristics that make them suitable for their specific uses. The round, smooth-skinned, white eastern cultivars are used for chipping (potato chips) and cooking (boiling), while the mutated western types are used for baking and frozen products (mainly French fries). The russet potato varieties have higher dry matter content than the eastern types. Dry matter is measured by the specific gravity of the tuber. High dry matter (1.085 specific gravity or higher) is desired for baking and processing. For frying, boiling, or mashing, the tuber should have a specific gravity of 1.080 or higher and at least 19.8% solids plus 14% starch.

## Germplasm resources

Potato diversity is maintained in a number of germplasm banks in various parts of the world. The major repositories include the International Potato Center (CIP) in Peru (with the best collection of cultivated potatoes, especially Andigena, Phureja, and “bitter potatoes”), the Commonwealth Potato Collection at Pentlandfield, Scotland, and the German-Dutch Potato Collection at Braunschweig, Germany. In the USA, the IR-1 project at Sturgeon Bay, Wisconsin is important to US breeders.

## Cytogenetics

The genus *Solanum* contains about 2,000 species, of which only about 150 are tuber-bearing. The cultivated potato, *S. tuberosum*, is a tetraploid ( $2n = 4x = 48$ ). Five cytological groups of potato have been identified, with somatic numbers of 24, 36, 48, 60, and 72. About 70% of tuber-bearing potatoes are diploids, while 5% and 8% are tetraploids and hexaploids, respectively. Most of the diploids are self-incompatible, producing seed only when fertilized by pollen containing a different *S* allele.

A cultivated diploid that is used in South American production is *S. phureja*. It is used in bridgecrossing and other genetic studies. Triploid potato is sterile, and only a few are cultivated. Producing triploid potato by crossing  $2n \times 4n$  is seldom successful because of the so-called “triploid block”. Hexaploids are self-fertile. A widely used hexaploid is *S. demissum*. It is the source of the major *R* gene that confers resistance to late blight.

## Genetics

Potato genetics is complex, because of its autotetraploid origin. There can be four different alleles at a locus. Intralocus interactions (heterozygosity) and interlocus interaction (epistasis) occur and can be exploited by using the appropriate breeding procedure (see Chapter 13).

## General botany

Potato (*S. tuberosum*) is an annual plant with short (300–600 mm), erect, and branched stems. Its compound leaves can be 300–600 mm long, with a terminal leaflet. The flowers are borne in compound, terminal cymes with long peduncles. The flower color may be white, rose, lilac, or purple. The plant bears fruits



## Industry highlights

### *The breeding of potato*

John E. Bradshaw

Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK

#### *Evolution of the modern potato crop*

The domestication of one or a few wild, tuber-bearing, diploid *Solanum* species took place in the Andes of South America between 7,000 and 10,000 years ago, probably in what is now the Lake Titicaca to Lake Poopo region of north Bolivia (Hawkes 1990). The result was diploid *S. stenotomum*, also referred to as a form of *S. tuberosum* (group Stenotomum), from which other cultivated species were derived, including diploid *S. phureja* (or group Phureja), tetraploid *S. tuberosum* subsp. *andigena* (or group Andigena) and tetraploid *S. tuberosum* subsp. *tuberosum* (or group Tuberousum). Andigena potatoes became the most widely grown form in South America. Tuberousum potatoes were selected from Andigena types for tuber production in long-days conditions in coastal Chile and are referred to as Chilean Tuberousum potato. Phureja potatoes were selected from Stenotomum for lack of tuber dormancy so that up to three crops per year could be grown in the lower, warmer, eastern valleys of the Andes. Andigena potatoes were introduced into the Canary Isles around 1562 and from there to mainland Europe in the 1570s (Hawkes & Francisco-Ortega 1993). As the growing of potato spread northeastwards across Europe, the potato became adapted to the long summer days of northern Europe and evolved sufficiently to be classified as subspecies *tuberosum* (European Tuberousum). There may also have been early introductions from Chile as well as the Andes. Starting in the 17th century, the potato was taken from Europe and cultivated in many other parts of the world. Today, potato is grown in 149 countries from latitudes 65°N to 50°S and at altitudes from sea level to 4,000 m (Hijmans 2001), and potato is the fourth most important food crop after wheat, maize, and rice (Lang 2001). As well as being a staple food, potato is grown as a vegetable for table use, is processed into French fries and crisps (chips), and is used for dried products and starch production.

#### *Potato breeding and the need for new cultivars*

##### Potato breeding

The reproductive biology of potato is ideal for creating and maintaining variation. Potato flower and set true seed in berries following natural pollination by insects, particularly bumble bees. Outcrossing is enforced in cultivated (and most wild) diploid species by a gametophytic self-incompatibility system. Whilst this system does not operate in *S. tuberosum* because it is a tetraploid, 20% natural cross-pollination was estimated to occur in an artificially constructed Andigena population (Glendinning 1976). This sexual reproduction creates an abundance of diversity by recombining the variants of genes which arose by mutation, and potatoes are highly heterozygous individuals that display inbreeding depression on selfing. The genetically unique seedlings that grow from true seeds produce tubers which can be replanted as seed tubers, and hence distinct clones can either be established and maintained by asexual vegetative reproduction or discarded. Domestication involved selection of less bitter, and hence less toxic, tubers and Andean farmers certainly retained a much wider variety of tuber shapes and skin and flesh colours than seen in wild species (Simmonds 1995).

Potato breeding in the modern sense began in 1807 in England when Knight made deliberate hybridizations between varieties by artificial pollination (Knight 1807). It flourished in Britain and elsewhere in Europe and North America during the second half of the 19th century when many new cultivars were produced by farmers, hobby breeders, and seedsmen. A single Chilean Tuberousum cultivar, "Rough Purple Chili", was introduced into the USA in 1851 (Goodrich 1863), and North America's most popular potato cultivar, "Russet Burbank", was derived from it by three generations of open-pollination with selection, and released in 1914 (Ortiz 2001). The descendants of "Rough Purple Chili" were widely employed as female parents in crosses with European Tuberousum at the end of the 19th century. Modern potato breeding started later in the 1930s in China and India but these countries are now two of the leading potato producers in the world. By 2003, the *World catalogue of potato varieties* (Hamester & Hils 2003) was able to list 3,200 cultivars from 102 countries, a remarkable achievement since the genetic base of this potato breeding must be considered narrow, despite some introgressions of disease-resistance genes from wild and cultivated relatives of Tuberousum potato and some base broadening with Andigena and Phureja/Stenotomum potato that had been selected to tuber in long days (Bradshaw & Mackay 1994).

##### Need for new cultivars

Despite the large number of cultivars currently available there is a continuing need for new ones. At least two contrasting scenarios can be seen. In the European Union, the potato industry is trying to increase potato usage in an economically and environmentally

sustainable way. New cultivars must give more yield of saleable product at less cost of production. They must have inbuilt resistances to pests and diseases, and increased water and mineral use efficiency, that will allow a reduced use of pesticides and fungicides and better use of water and fertilizers. Finally, they must help meet consumer demands for convenience foods, improved nutritional and health benefits, improved flavour, and novel products. In contrast, in Asia and Africa there is a need for increased and stable potato production to meet increased demand for food. New cultivars must deliver higher yields under low inputs, disease and pest attacks, and environmental stresses such as heat, cold, drought, and salinity. If possible, they should also have improved nutritional and health properties, but the greatest need is to raise fresh weight yields from a world average of 17 t/ha to European and North American levels of 45 t/ha (Lang 2001).

### Breeding finished cultivars

#### Parents

Potato breeding worldwide has traditionally involved making crosses between pairs of parents with complementary features and this is still the main route to new cultivars. The aim has been to generate genetic variation on which to practice phenotypic selection over a number of vegetative generations, for clones with as many desirable characteristics as possible for release as new cultivars. The choice of parents is all important as breeding can never simply be a numbers game. Crossing the 3,200 cultivars in the world catalog in all possible combinations would generate 5,118,400 progenies, and raising 500 seedlings of each would give a staggering total of 2,559,200,000 for evaluation – an impossible task. In contrast, a phenotypic assessment of 3,200 cultivars is feasible, and so is a genotypic assessment of diversity with molecular markers. Hence breeders can now think in terms of capturing allelic diversity in a smaller core set of parents and of using association (linkage disequilibrium) genetics to choose parents genotypically as well as phenotypically (Simko 2004). They can also use genetic distance based on molecular markers to complement co-ancestry/pedigree analysis (Sun et al. 2003) in order to avoid closely related parents, and hence inbreeding depression, and to ensure genetic variation for continued progress.

As genetic knowledge accumulates, it will be possible to choose parents for use in pair crosses so that one or both parents have the desired major genes and alleles of large effect at quantitative trait loci (QTLs). Major genes have been mapped for flesh and skin color, for tuber shape and eye depth, and for resistance to late blight, nematodes, potato viruses X, Y, and A, and wart. QTLs of large effect have been mapped for maturity and resistance to late blight, potato cyst nematodes, and potato leaf roll virus (PLRV). In contrast, many economically important traits still appear to be complex polygenic traits and these include tuber dormancy, dry matter and starch content, fry color, resistance to *Erwinia* blackleg and tuber soft rot, tuberization, and yield. For these traits, breeders will still have to rely on phenotypic data and use knowledge of offspring–midparent regressions to determine crossing strategy. A statistically significant regression is evidence of heritable variation, and the slope of the regression line is a measure of heritability. With a highly heritable trait like fry color, the midparent value is a good predictor of the mean performance of the offspring and a few carefully chosen crosses can be made (Bradshaw et al. 2000). In contrast, with only a moderately heritable trait such as yield, offspring mean is less predictable and more crosses need to be made to ensure that they include the best possible.

#### Early generations

The program at the Scottish Crop Research Institute (SCRI) before 1982 was typical in its handling of the early generations (Bradshaw & Mackay 1994). Visual selection reduced the number of potential cultivars from 100,000 (200 crosses × 500 seedlings) in the seedling generation in the glasshouse to 40,000 spaced plants at a high grade seed site in the first clonal generation, then to 4,000 four-plant plots at the seed site in the second clonal generation, and finally to 1,000 clones in replicated yield trials at a ware site in the third clonal generation. Several independent reviews concluded that such intense early generation visual selection was very ineffective (Bradshaw & Mackay 1994).

A potato breeding strategy (Table 1) has been developed at SCRI that avoids intense early generation visual selection between seedlings in a glasshouse and spaced plants at a seed site (Bradshaw et al. 2003). Once pair crosses have been made, progeny tests are used to discard whole progenies before starting conventional within-progeny selection at the unreplicated small-plot stage. Clones are also visually selected from the best progenies for use as parents in the next cycle of crosses whilst they are multiplied to provide enough tubers for assessment of their yield and quality. Midparent values, as well as progeny tests, are then used to select between the resultant crosses. Material from other breeding programs can be included in the parental assessments and used in the next cycle of crosses if superior. Finally, in seeking new cultivars, the number of clones on which to practice selection can be increased by sowing more true seed of the best progenies, but without selection until the small-plot stage. The theoretical superiority of this strategy lies in being able to practice between-cross selection for a number of economically important traits within 1 or 2 years of making crosses, something that is not possible on individuals as seedlings in the glasshouse or spaced plants at the seed site. At SCRI, seedling progeny tests are used for resistance to late blight, resistance to the white potato cyst nematode (*Globodera pallida* Stone), and tuber yield and appearance, as visually assessed by breeders. Tuber progeny tests are used for fry color and a second visual assessment of tuber yield and appearance. The use of progeny tests for key traits also means that full-sib family selection can be operated on a 3-year cycle for these traits, an improvement on the practice of clonal selection after a further six vegetative generations (i.e., not using potential cultivars as new parents until they are entered into official National List Trials).

**Table 1** The Scottish Crop Research Institute (SCRI) strategy for breeding finished cultivars.

Year	Strategy	
<b>0</b>	<b>Decide objectives and evaluate potential parents</b> Midparent values used to predict mean performance of crosses for quantitative traits + other genetic information	
	GLASSHOUSE	
<b>1</b>	Choose parents and make 200 crosses	
<b>2</b>	Seedling progeny tests on 200 progenies $\times$ 2 replicates $\times$ 25 seedlings (late blight, potato cyst nematodes, visual assessment of tubers) Select best 40 progenies (after first cycle will also have midparent values for fry color and other traits from year 5 of previous cycle)	
	SEED SITE	
<b>3</b>	Tuber progeny tests on 40 progenies $\times$ 2 replicates as 2,000 spaced plants (visual assessment of tubers and fry color). Select 500 spaced plants at harvest (four tubers of each plant) Sow more seed of 10 best progenies in the glasshouse to provide a further $10 \times 250 = 2,500$ (four tubers of each) clones for year 4 Select clones for use as parents in next cycle of crosses at random from those (500) advancing to year 4	
<b>4</b>	3,000 unreplicated four-plant plots (including parents of next cycle of crosses)	
	<b>Assessment for yield and quality and special disease tests</b>	
Seed site (number of plants)	Clones (number)	Ware site (number of plots and plants per plot)
5	6	1,000 $2 \times 5$ ; include parents used in crosses
6	20	360 $2 \times 10$
7	100	120 Two harvests or sites of $2 \times 10$
8	300	40 Ware trials at a number of sites
9	700	20 Ware trials at a number of sites
10	2,000	2 Official trials
11	2,000	1 Official trials

**Multiplication and commercialization from virus-free stock****Intermediate and later generations**

The SCRI can again be considered typical in being able to handle 1,000 clones in the first year of replicated yield trials at a ware site. However, the relatively slow rate of natural vegetative reproduction, together with the complicated logistics of accurately assessing 1,000 or more clones for a very large number of traits, meant that another 5 years elapsed before one or a few potential cultivars could be confidently entered into official statutory trials. During this period, decreasing numbers of selected clones were grown in increasingly sophisticated trials over as wide a geographic range as economics permitted and breeding objectives demanded. During these intermediate and final stages of selection, the production of seed tubers was separated from the trials which were grown under ware conditions, designed as far as possible to approximate to those of good commercial practice. In addition to yield and agronomic performance, clones undergoing selection were assessed for their cooking and processing characteristics and tested for their resistances to numerous pests and diseases. The selection criteria and testing procedures were largely governed by practical considerations and experience of the reliability of the various tests used, rather than by genetic knowledge of, for example, herabilities or genetic correlations between traits. It would therefore be worthwhile trying to develop a more robust decision-making process based on multtrait, multistage selection theory and estimated genetic parameters, or at least to assess the extent to which current practice is suboptimal (Bradshaw & Mackay 1994).

**Genetic knowledge and molecular marker-assisted selection**

As knowledge increases about the number and chromosomal locations of genes affecting economically important traits, breeders should be able to design better breeding programs. As well as selecting parents that complement one another genotypically, they will be able to determine the seedling population size required for certainty of finding the desired genotype, and, more realistically,

the number of cycles of crossing and selection required before this is achievable in practice in the size of population they can handle. A big impact on the efficiency and rate of progress would be the identification of superior clones genotypically as seedlings in the glasshouse, and the use of modern methods of rapid multiplication to progress them to commercialization. This will require molecular marker-assisted selection or preferably direct recognition of the desired allele at a genetic locus.

### ***Widening the genetic base for future potato breeding***

In future, for many traits, greater use can be expected of the world collection of 3,527 potato cultivars native to Latin America which is maintained by the International Potato Centre (CIP) in Peru (Huaman et al. 1997). In addition, further improvements in resistance to abiotic and biotic stresses should come from a greater use of wild species, given the wide range of habitats in which they have evolved. The Inter-genebank Potato Database (IPD) contains 7,112 different accessions of 188 taxa (species, subspecies, varieties, and forms) out of the 247 tuber-bearing wild potato taxa recognized by Hawkes (Huaman et al. 2000) and data are available for more than 33,000 evaluations covering 55 traits. The species form a polyploid series from diploid ( $2n = 2x = 24$ ) to hexaploid ( $2n = 6x = 72$ ). By manipulation of ploidy, with due regard to endosperm balance number, virtually any potato species can be utilized for the introgression of desirable genes into *S. tuberosum* (Ortiz 1998, 2001). In the past, it took up to five backcross generations and 30 years to transfer a major dominant resistance gene from a wild species into a successful cultivar, but today, molecular marker-assisted introgression offers the possibility of faster progress. As potato is a heterozygous outbreeder, use of the same recurrent parent during introgression would result in a self of the recurrent parent and hence inbreeding depression. This can be avoided by using different *Tuberosum* parents for each backcross but would result in an entirely new cultivar, which may or may not be the desired outcome. The only way to introduce a gene into a known cultivar is by the transgenic route. Hence the molecular cloning of natural resistance genes and their transfer by *Agrobacterium*-mediated transformation into well-adapted but susceptible cultivars is being pursued in a number of laboratories worldwide. Given the timescale of conventional breeding, the genetic improvement of popular potato cultivars such as "Russet Burbank" by transformation is an attractive proposition, despite public concerns about genetically modified (GM) crops in some countries. The interested reader can consult a recent review by Davies (2002) of current and future prospects for a whole range of traits including insect, virus, and herbicide resistance and antibruise tubers.

### ***References***

- Bradshaw, J.E., M.F.B. Dale, and G.R. Mackay. 2003. Use of mid-parent values and progeny tests to increase the efficiency of potato breeding for combined processing quality and disease and pest resistance. *Theor. Appl. Genet.* 107:36–42.
- Bradshaw, J.E., and G.R. Mackay. 1994. Breeding strategies for clonally propagated potatoes. In: *Potato genetics* (Bradshaw, J.E., and G.R. Mackay, eds), pp. 467–497. CAB International, Wallingford, UK.
- Bradshaw, J.E., D. Todd, and R.N. Wilson. 2000. Use of tuber progeny tests for genetical studies as part of a potato (*Solanum tuberosum* subsp. *tuberosum*) breeding programme. *Theor. Appl. Genet.* 100:772–781.
- Davies, H.V. 2002. Commercial developments with transgenic potato. In: *Fruit and vegetable biotechnology* (Valpuesta, V., ed.), pp. 222–249. Woodhead Publishing Ltd, Cambridge, UK.
- Glendinning, D.R. 1976. Neo-*Tuberosum*: New potato breeding material. 4. The breeding system of neo-*Tuberosum*, and the structure and composition of the neo-*Tuberosum* gene pool. *Potato Res.* 19:27–36.
- Goodrich, C.E. 1863. The origination and test culture of seedling potatoes. *Trans. New York State Agric. Soc.* 23:89–134.
- Hamester, W., and U. Hils. 2003. World catalogue of potato varieties. Agrimedia GmbH, Bergen/Dumme, Germany.
- Hawkes, J.G. 1990. The potato: Evolution, biodiversity, and genetic resources. Belhaven Press, London.
- Hawkes, J.G., and J. Francisco-Ortega. 1993. The early history of the potato in Europe. *Euphytica* 70:1–7.
- Hijmans, R.J. 2001. Global distribution of the potato crop. *Am. J. Potato Res.* 78:403–412.
- Huaman, Z., A. Golmirzaie, and W. Amoros. 1997. The potato. In: *Biodiversity in trust: conservation and use of plant genetic resources in CGIAR centres* (Fuccillo, D., L. Sears, and P. Stapleton, eds), pp. 21–28. Cambridge University Press, Cambridge, UK.
- Huaman, Z., R. Hoekstra, and J.B. Bamberg. 2000. The inter-genebank potato database and the dimensions of available wild potato germplasm. *Am. J. Potato Res.* 77:353–362.
- Knight, T.A. 1807. On raising of new and early varieties of the potato (*Solanum tuberosum*). *Trans. Hort. Soc. Lond.* 1:57–59.
- Lang, J. 2001. Notes of a potato watcher. Texas A&M University Press, College Station, TX.
- Ortiz, R. 1998. Potato breeding via ploidy manipulations. In: *Plant breeding reviews*, Vol. 16 (Janick, J., ed.), pp. 15–86. John Wiley & Sons, New York.
- Ortiz, R. 2001. The state of the use of potato genetic diversity. In: *Broadening the genetic base of crop production* (Cooper, H.D., C. Spillane, and T. Hodgkin, eds), pp. 181–200. CABI Publishing, Wallingford, UK.
- Simko, I. 2004. One potato, two potato: Haplotype association mapping in autotetraploids. *Trends Plant Sci.* 9:441–448.
- Simmonds, N.W. 1995. Potatoes. In: *Evolution of crop plants*, 2nd edn (Smartt, J., and N.W. Simmonds, eds), pp. 466–471. Longman Scientific & Technical, London.
- Sun, G., G. Wang-Pruski, M. Mayich, and H. Jong. 2003. RAPD and pedigree-based genetic diversity estimates in cultivated diploid hybrids. *Theor. Appl. Genet.* 107:110–115.

(berries) called **potato balls**. The underground commercial part is a modified stem (or tuber) that is borne at the end of a stolon. The “eyes” on the tuber are actually rudimentary leaf scars favored by lateral branches. Each eye contains at least three buds protected by scales. When potatoes sprout, the sprouts are lateral branches with several buds. A section across a tuber reveals a pithy central core with branches leading to each of the eyes.

## Reproductive biology

### Floral biology

Potato has a terminal inflorescence consisting of 1–30 (but usually 7–15) flowers, depending on the cultivar. The five petals give an open flower a star shape. A flower also has a stigma that protrudes above a cluster of five large, bright yellow anthers. The corolla color varies from white to a complex range of blue, red, and purple. Flowers open, starting with those nearest the base of the inflorescence and proceeding upwards, at the rate of about 2–3 each day. At the peak bloom, there are usually 5–10 open flowers. Flowers stay open for only 2–4 days, and the receptivity of the stigma and duration of pollen production is about 2 days.

### Pollination

Potato is predominantly self-pollinated. The peak time of pollination is early morning. Pollen can be collected ahead of the time of use, and kept in a cool dry place (e.g., in a desiccator for longevity). Pollination is most successful when temperatures are not high. Some breeders collect the desirable flowers to be used a day ahead and lay them out to dry. The pollen is then shaken out over a sieve. The pollen is collected in tubes for use.

## Common breeding methods

Potato has a wide array of wild germplasm that easily crosses with cultivated types. Hybridization is the principal procedure for gene transfer. Selection is conducted in the  $F_1$ , because the parents are also widely used in modern potato improvement. Protoplast fusion techniques may be used to fuse monoploid (1x) to form dihaploids (2x). A cross of  $4x \times 2x$  using a particular accession of *S. phureja* as the male, is a technique for generating haploids at a high frequency. Hybridization

can be used to increase the frequency of tetra-allelic loci and thereby to increase the intra- and interlocus interactions for increased vigor. The techniques of unilateral sexual polyploidization ( $4x \times 2x$ ) or bilateral sexual polyploidization ( $2x \times 2x$ ) may be used in potato breeding. They are practical procedures because many diploid potatoes hybridize among themselves or with tetraploid species.

Genetic engineering procedures have been used to achieve the development of *Bt* resistance to Colorado potato beetle and viral coat protein-based resistance to several viral disease (e.g., the potato leaf roll virus – PLRV). The cloning and use of the *AGPase* gene has enabled cultivars with high solids to be developed.

## Establishing a breeding nursery

### Field nursery

Breeders commonly cross potatoes in the greenhouse because pollination in the field does not yield good seed. Instead, some breeders produce the pollen source in the field and cut stems containing large inflorescences for use in the greenhouse. Parents for pollen sources should be free from virus infection and should be properly managed, with protection against insect pests, and fertilizing and irrigating the plants for healthy flowers to be produced.

### Greenhouse nursery

Potato is a long-day plant. Hence, at least 16 hours of sunlight (or 20 klux of artificial light) is needed to successfully grow potato to flowering and maturity. The greenhouse temperature should be maintained at about 19°C. Seed potato or stem cuttings can be used for planting. Plants may be raised in a ground bed or in pots placed on raised benches. Crossing is often done in the winter. The vegetative growth is controlled by pruning and staking the plants to make the flowers more accessible.

## Artificial pollination

### Materials and equipment

A mechanical vibrator can be used to aid pollen collection, which is then used to pollinate a large number of plants. Pollen that is shed is collected into test tubes. A

blunt scalpel may be used to scrape pollen from the anthers for direct deposit on the anthers.

### Emasculation

Mature unopened buds are selected for emasculation. At this stage, the petals appear ready to open. Emasculation may be done in the afternoon for pollination the next morning. Greenhouse-grown plants usually fail to set fruit, unless hand-pollinated. Consequently, in the absence of air currents (e.g., from ventilators, open doors) that may agitate the flowers to cause pollen shed, emasculation is not necessary. However, it should be noted that there is always a possibility of some selfing occurring under such circumstances.

### Pollination

The best time to pollinate is in the morning, soon after the flowers are fully open. This is the time when pollen is most abundant. Flowers with plump, bright yellow anthers capped with brownish tips, give the best quality pollen. Because flowers in inflorescences do not all open at once, pollinating 3 days per week will allow most of the suitable flowers to be pollinated. It is best to pollinate flowers in the same inflorescence with the same pollen source, to reduce contamination. A blunt scalpel may be used to deposit pollen on the stigma, or the stigma may be dipped into the tube containing pollen. To authenticate hybridity, anthocyanin pigment markers may be incorporated into the breeding program.

### Natural pollination

Potato has a significant amount of self-pollination. Crossing blocks may be used if one of the parents to be crossed is self-incompatible. This will make sure that all fruits produced on the plant are hybrids.

### Seed development

Successful crosses develop into small fruits. It is recommended that each set of fruits is inserted in a paper bag, about 4 weeks after pollination, to prevent loosing fruits (and seed) to fruit drop. Softened fruits are picked for seed extraction. The seeds are squeezed into a beaker of water and then strained through a cheese cloth. The seeds of tetraploid species have seed dormancy ranging from about 6 months to about 2 years.

### Common breeding objectives

Some of the major breeding objectives in potato breeding are discussed next. The order of presentation is arbitrary.

**1 Tuber yield.** Increased tuber yield is the primary objective of potato breeding. Tuber yield and shape are influenced by photoperiod. Responsiveness to photoperiod is quantitatively inherited.

**2 Adaptation:**

- (a) **Heat tolerance.** Temperature variation is critical in potato production. While germination and growth are favored by warm temperatures, tuberization is favored by cool temperatures. Tuberization is inhibited at temperatures above 29°C. Heat tolerance is desirable for tuberization when unseasonable weather occurs during the production season.
- (b) **Frost resistance.** This trait is desirable for areas where fall potatoes are grown.
- (c) **Drought resistance.** This trait is necessary for production under rainfed conditions.

**3 Disease resistance.** Some of the key diseases of economic importance to potato production include the following:

- (a) **Late blight.** Caused by *Phytophthora infestans*, this fungal disease is the most economic in potato production, causing both foliar and tuber decay. Breeding for resistance is complicated by the fact that foliage resistance and tuber resistance may differ in the same plant. Resistance conditioned by major genes (designated  $R_1$ ,  $R_2$ ,  $R_3$ , etc.) has been discovered.
- (b) **Charcoal rot.** Caused by *Macrophomina phaseoli*, this disease occurs when temperatures are high. It is a common storage disease.
- (c) **Viral disease.** Several viral diseases occur in potato fields, the most economically important one being the virus X. Plant response to this virus is varied including resistance to infection, hypersensitivity, and immunity. Viral coat protein-based resistance to PLRV has been developed.

**4 Root-knot nematodes.** The root-knot nematode (*Meloidogyne incognita*), along with other nematodes, causes economic damage to potato tubers.

**5 Insect resistance:**

- (a) **Aphids.** Pubescent cultivars have resistance to aphids.
- (b) **Colorado potato beetle.** The Colorado potato beetle (*Leptinotarsa decemlineata*) can cause over 50% loss of a potato crop by feeding on the leaves. *Bt* resistance to this insect has been developed.

**6 Potato tuber quality improvement.** Potatoes are sold for the fresh market (for baking or cooking) or processed (chipped, frozen, starch, alcohol). The quality standards differ according to the product or end use. Long-term storage is a major aspect of marketing potatoes for chip processing. The reducing sugar content of the tuber should be low to avoid browning of chips (caramelization) during

preparation. Many cultivars have a low temperature sweetening potential, making them unsuitable for chip processing. Cultivars are being developed that will not accumulate reducing sugars at colder storage. High solids are also desired in potato breeding. Cultivars with high solids have been developed by genetic engineering procedures using the *AGPase* gene.

### References and suggested reading

- Jewell, L.D. 1988. Agricultural statistics 1987. US Department of Agriculture, Washington, DC.
- Page, O.T. 1982. Utilization of the genetic resources of the potato, III. International Potato Center, Lima, Peru.
- Ross, H. 1986. Potato breeding – Problems and perspectives: Advances in plant breeding. Journal of Plant Breeding Supplement 13. Paul Parey, Berlin and Hamburg.
- Talburt, W.F. 1987. History of potato processing. In: Potato Processing, 4th edn (Talburt, W.F., and O. Smith, eds), pp. 1–10. Van Nostrand, New York.

### Outcomes assessment

#### Part A

Please answer the following questions true or false:

- 1 Potato originated in India.
- 2 Potato is cool season crop.
- 3 Cultivated potatoes are tetraploids.
- 4 High reducing sugar levels are desired for chipping potatoes.

#### Part B

Please answer the following questions:

- 1 Exposure of potato tubers to light promotes the formation of a toxic alkaloid called .....
- 2 Potato bears fruits called .....
- 3 Give three major diseases and insect pests of potato.

#### Part C

Please write a brief essay on each of the following topics:

- 1 Discuss the artificial pollination of potato.
- 2 Discuss the importance of temperature adaptation in potato production.
- 3 Discuss the common potato breeding methods.

# Breeding cotton



## Taxonomy

Kingdom	Plantae
Subkingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Magnoliopsida
Subclass	Dilleniidae
Order	Malvales
Family	Malvaceae
Genus	<i>Gossypium</i> L.
Species	<i>Gossypium hirsutum</i> L.

## Economic importance

Cotton is the most important natural fiber in the world for textile manufacture, accounting for about 50% of all fibers used in the textile industry. It is more important than the various synthetic fibers, even though its use is gradually reducing. It is grown all over the world in about 80 countries. The USA is the second largest producer of cotton after China, producing 16.52 million bales on 14.6 million acres in 1999. Most of the cotton is produced in the Cotton Belt, and currently occurs west of the Mississippi river, due to the spread of the devastating boll weevil attack in the eastern states of the USA. The leading production states in 2000 were Texas, Mississippi, Arkansas, and California, in this order of decreasing importance. Other important cotton producing states are Alabama, Arizona, Tennessee, Georgia, Louisiana, South Carolina, North Carolina, Missouri, Oklahoma, and Minnesota.

On the world scene, 88.7 million bales of cotton were harvested from 32.2 million hectares in 2000. The leading producing countries in the world are China (mainland), USA, India, Pakistan, Uzbekistan, and Turkey. China and the US produce nearly 50% of the total world production. Other major producers are Mexico, Brazil, Egypt, Greece, and Columbia.

## Origin and history

The exact origin of cotton is not conclusive. However, two general centers of origin appear to have been identified – Indo-China and tropical Africa in the Old World (Mohenjo Daro, Indus valley of Asia in 2500 BC), and South and Central America in the New World (Huaca Prieta, Peru in 8000 BC). Cotton was cultivated in India and Pakistan, and in Mexico and Peru about 5,000 years ago. Cotton was grown in the Mediterranean countries in the 14th century and shipped from there to mills in the Netherlands in Western Europe for spinning and weaving. Wool manufacturers resisted its introduction into Europe until the law banning the manufacturing of cotton was repealed in 1736. Similarly, the English resisted the introduction of cotton mills into the USA until Samuel Slater, who previously had worked in a cotton mill in England, built one in 1790. Three types of cotton are grown in the US: Sea Island, American-Egyptian, and upland. These types probably originated in America, the first two are believed to have come from South America. Upland cotton may have descended from Mexican cotton or from crosses of Mexican and South American species.

World production of cotton occurs between latitudes 45°N and 30°S, where the average temperature in summer is at least 25°C. Cotton requires a frost-free production period (175–225 days), availability of moisture, and abundant sunshine for good plant growth, development, and ripening.

## Cultivars

The varieties of cotton in cultivation are derived from four species that produce seed fibers (lint) that is of economic value.

1 *Gossypium hirsutum*. About 87% of all cultivated cotton is derived from this species. The varieties are

grown in America, Africa, Asia, and Australia. About 99% of all US cotton is of this type, which is also classified as upland cotton. The varieties in this type produce fiber of variable length and fineness. The plant can attain a height of 2 m.

- 2 *Gossypium barbadense*. This type accounts for 8% of the world's cotton production and is grown in America, Africa, and Asia. The plant can reach a height of 2.5 m and has yellow flowers and small bolls. It is also classified as Egyptian cotton and has long, fine, and strong fibers. The cotton is used for manufacturing sewing threads. This type is also called Pima cotton. New Pima cultivars are identified by a number (e.g., "Pima S-1", "Pima S-6").
- 3 *Gossypium arboreum*. This species constitutes about 5% of the total production and is grown mainly in East Africa and South East Asia. The plant can attain a height of 2 m and has red flowers. The varieties from this species are also classified as Asiatic cotton.
- 4 *Gossypium herbaceum*. This species is also classified as Asiatic cotton. The fibers produced are short (less than 25 mm) and of poor quality. This type of cotton fiber is used for manufacturing surgical supplies.

Another classification of cotton based on fibers, from longest to shortest, is Sea Island > Egyptian > American upland long staple >, American upland short staple >, Asiatic.

### American upland cotton

Four types of *G. hirsutum* (called American upland cotton) are grown in the USA.

- 1 **Eastern upland**. This type has medium-sized open bolls, and medium length staple. It is resistant to *Fusarium* wilt.
- 2 **Delta highland**. This type is grown mainly in the Mississippi delta, Texas, and Arizona. It has small to medium open bolls, and medium staple.
- 3 **Plains or storm-proof**. This type of cotton is so-called because the bolls are closed or only partially open at maturity, and hence the fiber resists being blown away by the wind. The bolls are large with short staple and medium fiber length. It is grown in Oklahoma and Texas.
- 4 **Acala**. Grown widely in California, this type of cotton has medium to large bolls, medium to long staple, strong fibers, and a long fruiting period.

The Pima cottons are also grown in the US in places like Arizona, New Mexico, Texas, and California, and make up only about 1% of the total US cotton acreage.

### Germplasm resources

Germplasm collections are held in banks in the US and other parts of the world. In the USA, obsolete cultivars of *Gossypium hirsutum* are maintained at the US Cotton Physiology and Genetics Laboratory, Stoneville, Mississippi, whereas the US Cotton Research Laboratory in Phoenix, Arizona maintains accessions of *G. barbadense*. Texas A&M University, College Station maintains a collection of diploid cottons and race stocks of *G. hirsutum*, whereas the National Seed Storage Facility, Fort Collins, Colorado maintains seeds of various cotton types.

### Cytogenetics

Seven genomes of *Gossypium* species, designated *A*, *B*, *C*, *D*, *E*, *F*, and *G*, have been identified according to chromosomal size and affinity at meiosis. The basic chromosome number of *Gossypium* is 13. Most (45) of the 60 known species are diploids ( $2n = 2x = 26$ ). Cotton may be divided into two major groups:

- 1 **Old World cotton** ( $2n = 26$ ). The diploids in this group have *A*, *B*, *E*, or *F* genomes. The cultivated types have the *AA* genome and comprise: (i) *G. herbaceum*, which has five races that originated in Africa and Asia; and (ii) *G. arboreum*, which has six races of tree cotton and is found in India.
- 2 **New World cotton** ( $2n = 52$ ). These are tetraploids with the genome *AADD* (13 pairs of each of large and small chromosomes). The dominant species are: (i) *G. barbadense* (Sea Island and Egyptian cotton); and (ii). *G. hirsutum* (upland cotton).

Sixty-two translocations that identify chromosomes 1 through 25 of the 26 chromosomes of the diploid cotton are known. They have been used to locate genes on chromosomes and the production of aneuploids, among other uses.

### Genetics

A genetic linkage map of *G. barbadense* and *G. hirsutum* is available. Homeologous gene pairs related to several homeologous chromosomes have been discovered. For example, there is homeology between the loci for anthocyanin pigmentation ( $R_1R_2$ ) and cluster fruiting habit ( $cl_1cl_2$ ) of linkage groups II and III. Isogenic lines

of a number of allotetraploid mutants (e.g., glandlessness, plant hairiness, okrleaf) have been developed.

Anthocyanin pigmentation is controlled by a multiple allelic series, while monopodial and sympodial branching habits are influenced by several major genes and minor genes. Flower color is controlled by two duplicate genes,  $Y_1 Y_2$ .  $Y_1$  conditions yellow petals in all allotetraploids (except *G. darwinii*) in which  $Y_2$  conditions petal color. Yellow petal color in *G. barbadense* is controlled by  $Y_1$ . *G. hirsutum* is mostly cream petalled ( $Y_1 Y_1 y_2 y_2$ ). However, both yellow and cream petals occur in wild forms. Eleven male-sterility loci have been identified, 10 of which are located in *G. hirsutum* and one in *G. barbadense*.

Because of high cross-pollination in cotton, a cultivar can have a high level of heterozygosity that is uncharacteristic of self-pollinated cultivars. Further, the genetic constitution of a cultivar can change from year to year. Both genetic and cytoplasmic male sterility (CMS) have been identified in tetraploid cotton. The  $M_{s4}$  dominant gene confers complete sterility on plants.

### General botany

Cotton (*Gossypium* spp.) belongs to the family Malvaceae, the mallow family. There are about 40 species in this genus, but only four species are grown for their economic importance as fiber plants. It is considered an annual plant but it grows as a perennial in tropical areas where the average temperature for the coldest months stays above 18°C. The plant has a central stem that attains a height of 0.6–1.5 m. It has a deep taproot system. The leaves are arranged spirally around the stem. They are petioled and lobed (3–7 lobes). The stem and leaves are pubescent.

The flowers have five separate petals with the stamens fused into a column surrounding the style. Three large leaf-like bracts occur at the base of the flower. The ovary develops into a capsule or boll (the fruit). The fruit bud (young fruit) is called a **square**. When dry, the capsule splits open along the four or five lines. Average boll length is about 40–50 mm. Only about 45% of the bolls produced are retained and develop to maturity. The plant is predominantly self-pollinated. When pollinated by foreign pollen, the phenomenon of **xenia** causes a reduction in fiber length.

An open, mature cotton boll reveals the economic product, a fluffy mass of fibers surrounding the seeds. Each fiber is a single-cell hair that grows from the epidermis of the seed coat. The long hair is called **lint** and

the short hair, **fuzz**. Most wild species of cotton do not have lint.

## Reproductive biology

### Floral biology

Generally, cotton has perfect flowers. It takes 21–25 days for the square to reach anthesis. While in the bud form, a flower is enclosed by three bractoles (occasionally four) forming the epicalyx. The base of the flower is occupied by a five-lobed calyx. The cotton flower is large, the petal length reaching 9 cm in some cultivars. In most species, the corolla opens widely, but *G. barbadense* and *G. raimondii* have tubular flowers. The corolla is commonly white (in *G. hirsutum* and *G. arboreum*), while it is yellow (and various shades of yellow) in *G. barbadense* and *G. herbaceum*.

Each flower has an anther column, which may bear about 100 anthers in *G. hirsutum*. The tip of the stigma is exserted above the column. The pistil is compound and has 3–5 carpels. The ovary develops into a three- to five-loculed capsule (boll) containing 7–9 seeds that are covered with lint.

### Pollination

Cotton is predominantly self-pollinated, but up to about 30% and sometimes higher cross-pollination occurs. Once in bloom, the flower is usually receptive to pollination for no more than 8 hours. Pollination is predominantly by insects.

## Common breeding methods

Like other self-pollinated species, cotton improvement follows three general approaches – introduction, selection, and hybridization. The most common breeding procedure used in cotton breeding is hybridization. It is used for generating recombinants, followed commonly by pedigree selection to identify superior genotypes. The goal of breeding cultivars in cotton is to achieve sufficient uniformity for major traits (e.g., plant type, fiber properties, disease resistance), while retaining some heterozygosity for vigor. Pure-line selection is rarely if ever practiced in cotton breeding because of the floral biology. Some breeders use recurrent selection to concentrate genes of quantitatively inherited traits. Introgressive hybridization to incorporate many



## Industry highlights

### Cotton breeding

Don L. Keim

Delta and Pine Land Company, One Cotton Row, PO Box 157, Scott, MS 38772, USA

#### Introduction

Modern upland cotton, *Gossypium hirsutum*, is an allotetraploid which, though primarily self-pollinated, is readily cross-pollinated by insects. Cotton varieties are developed primarily by pure-line breeding techniques. But many varieties have traditionally been mixtures of closely related genotypes as a result of cross-pollination, cotton morphology, and breeding procedures. Uniquely, the primary economic value is not the seed itself, but in the fiber produced as an extension of the seed coat cells.

Some terminology unique to cotton is explained below:

- 1 **Lint**: the economic fiber that is separated from the seed during ginning. After ginning and cleaning, the lint is compacted into a bale. In the USA, each bale is identified and sampled for quality determination and marketed in bale units of approximately 480 pounds net weight of lint.
- 2 **Cotton boll**: a single fruiting body that contains at harvest the seed and attached lint fibers, and the dried ovary wall (burs).
- 3 **Lint percent**: the proportion of the lint to the harvested seed cotton (seed with lint attached).
- 4 **Fiber quality**: the characteristics that affect the processing of lint into yarn and finished textile goods. Four primary characteristics that affect the marketability and pricing of fiber are grade, length, micronaire, and strength. Many other characteristics affect the usability of fiber, some of which are uniformity, elongation, and short fiber content. Many fiber characteristics have a genetic component, but a major portion of the variation is due to environmental, harvest, and ginning effects. Fiber characteristics are determined in the commercial channels, and in most breeding programs, through the use of an HVI (high volume instrumentation) machine.
- 5 **Grade**: the combination of the amount of leaf trash and color of a lint sample.
- 6 **Length of staple**: the length of the fibers in a sample (2.5% span length).
- 7 **Micronaire**: the combination of relative fineness and fiber maturity of a sample.
- 8 **Strength**: the amount of force required to break a specified bundle of fibers.

Historically, a major factor contributing to the high cost of cotton production has been in the area of insect management. The insects most affecting US production have been the bollworm/budworm complex, *Heliothis virescens* and *Helicoverpa zea*, and the boll weevil, *Anthonomus grandis*. Recent adoption of new practices has dramatically reduced control costs associated with these two pests. The government-sponsored boll weevil eradication program has practically eliminated boll weevil as a pest in major areas of the US Cotton Belt. Additionally, in 1996, farmers began adopting the use of transgenic cotton which contains a gene that confers resistance to the bollworm/budworm complex. This gene was developed by Monsanto and was given the trade-name Bollgard® in cotton.

Cotton breeding as compared to breeding in other field crops is affected by two key factors. First, the fiber must be removed from the seed (ginning). Second, the remaining short fibers (seed fuzz) must also be removed with acid (delinting) so that the seed is "flowable" and can be planted with modern equipment. These processes, ginning and delinting, are costly in resources (time, effort, money). Additionally, cotton breeding is expensive due to much hand harvesting (hand picking), the limited number of seed per boll, and the need for hand selfing.

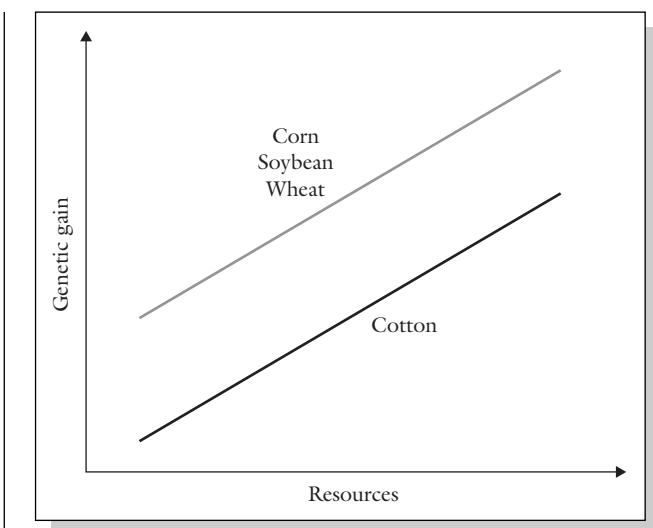
These factors, along with the need for expensive fiber quality evaluations, make the resource cost per unit of genetic gain much higher with cotton when compared to several other crops (Figure 1). Given fixed resources, the genetic gain will generally be less with cotton.

Until recently, cotton could have been considered to be relatively unbred when compared to the "state of the art" in crops like corn, soybean, and wheat. In the past, resource limits had forced cotton breeders to have fewer populations with small population sizes and to advance limited numbers of relatively unselected strains.

The recent widespread adoption and use of transgenic varieties has greatly increased the value of seed. Seed companies reaping this value have been able to place more resources in the hands of breeders. The recent gains in lint yield and fiber quality have been a direct result of these increased resources and the efficiencies gained from modernization of breeding procedures.

#### Breeding program

A major effort in this breeding program has been to maximize the gain enabled by the added resources made available. Several technique and procedural changes have been adopted to increase the efficiency of the utilized resources. This has resulted in an increased number of populations, larger population sizes, heavy selection pressure, and the early identification of superior strains.



**Figure 1** A comparison of resource cost per unit genetic gain between cotton and selected field crops.

for each row are compared against the average values of four surrounding check rows. Selected lines are advanced to replicated tests and breeder increase (Table 1).

The testing advancement schedule includes breeder tests within region for 2 years and an advanced testing program that involves uniform testing across all regions. Experimental lines advanced to the fourth year of testing (stage 4) are additionally entered into independently conducted trials managed by the Technical Services Department.

#### Transgene introgression program

At the beginning of the second year of testing, conventional lines are crossed to transgene donor parents to begin backcross introgression. Typically, three backcrosses are made prior to deriving lines homozygous for the desired transgene(s). Once  $F_{2:3}$  rows

**Table 1** Cotton breeding program.

Year	Location	Generation	Procedure/nursery	Harvest	Selection
1	Scott, MS	$F_0$	Crosses made		None
1	Costa Rica winter	$F_1$	Single row	Bulk	None
2	Scott, MS	$F_2$	Bulk population, $F_2$ yield trial	Modified bulk (single lock from 300+ plants)	Light visual, some populations discarded
2	Costa Rica	$F_3$	Bulk population	Single selfed boll from 200 plants	Light visual
3	Scott, MS	$F_{3:4}$	Progeny rows (7 ft)	Two plants per selected row	Heavy visual, fiber quality
4	Scott, MS	$F_{4:5}$	Progeny rows (27 ft) two replicates	Bulk row	Heavy visual, lint yield, fiber quality
5	Scott and Winterville, MS	$F_{4:6}$	Replicated tests	Bulk harvest, breeder increase	Lint yield, fiber quality, yield consistency
6	Three midsouth locations	$F_{4:7}$	Replicated tests, begin transgenic integration	Bulk harvest, breeder increase	Lint yield, fiber quality, yield consistency
7	Advanced Testing Program	$F_{4:8}$	Replicated tests	Bulk harvest, breeder increase	Lint yield, fiber quality, yield consistency

Such changes have included: utilization of winter nurseries (time savings); modified single-seed descent procedures (cost and labor savings); single boll to  $F_{3:4}$  progeny row advancement (cost, labor, and time savings);  $F_{4:5}$  progeny row yield testing (higher selection pressure); and yield trial technique changes (higher selection pressure). Additionally, changes in harvest mechanization and data collection have enabled better utilization of resources in the advanced stages where development costs are much higher.

A key aspect addressed is the degree of outcrossing due to pollinator activity. At Scott, the main breeding station, very little outcrossing occurs and the plants are treated as self-pollinated. In Costa Rica, a high degree of outcrossing occurs, therefore selfing is required in the  $F_3$  generation. A single selfed boll is harvested from each plant.

For each population, two opportunities are given for intensive selection: in the  $F_{3:4}$  single boll-derived rows and in the  $F_{4:5}$  progeny rows. In the  $F_{3:4}$ , intensive visual selection is practiced, along with postharvest selection based on lint percent (a component of lint yield) and fiber quality. The  $F_{4:5}$  progeny are planted in single rows that are replicated twice. Visual preselection at harvest is followed by harvest for lint yield, lint percent, and fiber quality. Values

are grown in transgene isolation, the breeder of the recurrent parent becomes responsible for the selection, evaluation, and advancement of lines. In many instances, lines showing close similarity are combined to form a bulk for testing and increase.

In all phases of transgenic development, the plants or seeds are evaluated for transgene presence, purity, and zygosity. All seed lots through to commercialization are evaluated for transgene purity.

Key transgenic introgression lines or bulks are entered directly into the advanced testing program as prior performance history exists for the conventional recurrent parents. Lines must go through gene equivalency evaluation prior to release approval by the technology provider. This is done to ensure the transgene is effective in the specific genotype.

### *Program successes*

One of the dilemmas faced in cotton improvement is the negative association of lint yield and fiber quality. In the early 1990s, as new higher yielding varieties were released, the fiber fineness (higher micronaire) and length tended to decrease. As fiber quality is determined and reported on each bale, growers suffered penalties in the marketplace. Early on, the higher yield overcame the lower price, thus increasing the growers' incomes. In recent years the penalties for less than desirable fiber have become more severe, negating many of the advantages of higher yielding varieties. Thus, a major goal of this program has been to develop higher yielding varieties with improved fiber traits.

One result has been the release of DP 491, a conventional, high-yielding variety with the outstanding fiber quality traits of lower micronaire, very long fiber, and high strength (Table 2). DP 491 was developed from a cross of DP 5415 × DP 2156. Apparently, transgressive segregation was captured in DP 491 for micronaire, length, and strength. The fiber characteristics of the DP 5415 parent are average with a tendency to have high micronaire. The other parent, DP 2156, is a stripper-type variety with very poor fiber characteristics.

DP 491 was converted to a "stacked gene" variety (containing the Bollgard® and Roundup Ready® transgenes) by backcrossing. The resulting recent release, DP 488 BG/RR, has been shown to have similar advantages in lint yield and fiber quality over DP 458 B/RR, the DP 5415 backcross-derived stacked gene variety (Table 3).

Another recent release from this program is DP 432 RR, a high-yielding early Roundup Ready® variety. DP 432 RR has shown in tests, distinct yield and fiber quality advantages over the popular ST 4793 RR, a backcross derivative of ST 474 (Table 4). DP 432 RR was developed from a straight cross of ST 474 × DP 5415 RR, rather than from the more typical backcross introgression of transgenes.

The examples presented demonstrate that the yield-quality barrier is not insurmountable. In fact, numerous experimental lines, developed by Delta and Pine Line breeders, exist that demonstrate the ability to combine high yield and high fiber quality.

### *The future*

The future holds great promise for the genetic improvement of cotton. In addition to continued lint yield and fiber quality improvement through intensive conventional breeding, opportunities exist in the areas of new transgenes, marker-assisted selection, and in increasing the useful diversity in germplasm base.

**Table 2** Comparison between new release DP 491 and DP 5415.

Variety	Value (\$/acre)	Lint yield (lb/acre)	Lint (%)	Micronaire	Length (2.5% span)	Strength (g/tex)
DP 491	725	1096	40.4	4.3	1.17	31.1
DP 5415	621	948	37.7	4.6	1.12	29.5
No. tests	96	105	105	101	100	96
% wins	86%	87%				
Significance <sup>1</sup>	***	***	***	***	***	***

<sup>1</sup> Paired *t*-test across tests.

**Table 3** Comparison between new release DP 488 BG/RR and DP 458 B/RR.

Variety	Value (\$/acre)	Lint yield (lb/acre)	Lint (%)	Micronaire	Length (2.5% span)	Strength (g/tex)
DP 488 BG/RR	808	1201	37.6	4.3	1.16	31.4
DP 458 B/RR	732	1094	36.3	4.4	1.11	30.7
No. tests	87	90	90	87	87	87
% wins	76%	76%				
Significance <sup>1</sup>	***	***	***	***	***	***

<sup>1</sup> Paired *t*-test across tests.

**Table 4** Comparison between new release DP 432 RR and ST 4793 RR.

Variety	Value (\$/acre)	Lint yield (lb/acre)	Lint (%)	Micronaire	Length (2.5% span)	Strength (g/tex)
DP 432 RR	805	1208	38.1	4.4	1.13	31.2
ST 4793 RR	766	1164	38.8	4.6	1.10	30.5
No. tests	121	123	123	121	121	121
% wins	74%	65%				
Significance <sup>1</sup>	***	***	***	***	***	***

<sup>1</sup> Paired *t*-test across tests.

Several new transgenes are being evaluated in upland cotton. Monsanto has developed a second *Bt* gene to be used in combination with the initial *Bt* gene (Bollgard II®) in order to increase effectiveness against lepidopteron pests and to prevent or greatly delay the advent of insect resistance. Initial varieties were released in 2003.

Roundup Ready® Flex is a new glyphosate-resistance gene developed by Monsanto that is designed to give a longer window of application than the original Roundup Ready® trait in cotton. Several companies are introgressing the gene into their varieties in anticipation of regulatory approval.

Because of the great success of Bollgard® and Roundup Ready® in the marketplace (78% of US cotton acreage in 2003 was planted to transgenic varieties), other technology providers are developing genes for herbicide and insect resistance in cotton. Liberty Link® cotton, which contains resistance to glufosinate herbicide, has been approved, and varieties were released for the 2004 plantings. Syngenta has developed vegetative insecticidal protein (VIP) technology in cotton providing resistance to lepidopteron pests and is seeking regulatory approval. Dow has developed Widestrike® technology (lepidopteron resistance) in cotton.

Marker-assisted selection offers in cotton improvement the potential to enhance screening for traits that are not easily evaluated. One example may be root-knot nematode resistance screening. Screening for this pest involves expensive and time-consuming greenhouse procedures. A marker or set of markers closely linked with the resistance gene(s) would greatly aid the breeder in identifying root-knot nematode resistant genotypes. Initial work in identifying markers associated with similarly useful traits is a major focus of the D&PL molecular breeding program. Once established, marker-assisted selection for specific goals will be integrated into the commercial breeding programs.

The crossing program has extensively involved the use of diverse upland varieties from all over the world. A key criteria for selecting parent lines is that they be high yielding in the regions for which they were developed. These lines were crossed with germplasm adapted primarily to the US midsouth. In several instances, three-way crosses were made utilizing an additional midsouth-adapted parent.

However, upland cotton germplasm derives from a very narrow base relative to other crops. Marker studies have indicated little polymorphism (5–17%). A strong need exists to reach beyond this pool into such diverse sources as upland race stocks and related species. Short-day sensitivity is a major barrier to the utilization of race stocks of upland cotton (modern upland varieties are day length insensitive). Barriers brought about by chromosomal incompatibilities have limited the broad use of related species crosses. Identification of desired traits and associated markers will enable the rapid and focused utilization of exotic germplasm.

### Conclusion

Cotton improvement has enjoyed several recent successes that have greatly benefited growers. The widespread utilization of transgenic cotton varieties has enabled growers to experience great savings in insect control and weed control costs. The convenience factor (limited tillage, reduce equipment/labor needs) with the Roundup Ready® technology has contributed greatly to its popularity.

The recent release of high-yielding varieties that have better fiber quality has contributed to more income in the farmer's pocket. It has been a great experience to have been a contributor to this ongoing revolution in cotton agriculture.

### Further reading

- Bowman, D.T. 2000. Attributes of public and private cotton breeding programs. *J. Cotton Sci.* 4:130–136.
- Cotton Varieties Planted 2003. Crop. USDA-AMS – Cotton Program, Memphis, TN.
- Dugger, P., and D.A. Richter (eds). Proceedings of the Beltwide Cotton Conference, National Cotton Council, Memphis, TN. Published annually. Available at <http://www.cotton.org/beltwide/proceedings.cfm>.
- Kohel, R.J., and C.F. Lewis (eds). 1984. Cotton. Agronomy Series No. 24. American Society of Agronomy, Madison, WI.

economically important traits has been achieved in cotton. This includes genes for disease resistance, insect resistance, and fiber quality.

Hybrid cotton seed is produced by hand emasculation and pollination, or hand pollination with male-sterility, in places where labor is inexpensive, such as in India. Where labor is expensive, the A-, B-, R-line male-sterility system is used to produce hybrid seed.

### Establishing a breeding nursery

Spacing in rows may be 75–200 mm in 100 cm rows. A final stand of 2–3 plants per 30 cm of row is optimum. Cotton responds to moderate amounts of balanced fertilization. Cotton breeders commonly use the greenhouse for crossing and selfing. The advantage of this practice is the control the breeder has over the environment to ensure proper flowering and fruiting, especially in photoperiod-sensitive cottons.

### Artificial pollination

#### Materials and equipment

Some of the materials and equipment used by breeders include fine-tipped tweezers, soda straw, and wired paper tags.

#### Emasculation

Flowers are emasculated at the whitebud stage. Usually, emasculation is done in the afternoon before anthesis. The flower is large enough to allow emasculation using bare fingers to be possible (but requires skill). Commonly, tweezers are used to strip off the anthers. Cultivars of the *G. barbadense* species do not tolerate a loss of petals during emasculation. Should it occur, a drop of gibberellic acid solution (100 ppm) may be applied at the base of the anther to increase the chance of fruit set. A section of soda straw may be used to protect the stigma prior to pollination.

#### Pollination

The day before use, the flower to be used as a pollen source should be protected from contamination from the visitation of insects, by sealing the corolla by tying it with copper wire or securing it with a paper clip (or some other techniques). It is advisable to inspect the

flower for intrusion by insects before using the flower. The stigma is fully receptive on the morning of anthesis. Pollen is deposited directly on the stigma using the anthers. Pollen remains viable in the field for only a few hours after anthesis. Some breeders use the soda straw to scoop the pollen and slip it over the stigma.

To authenticate hybridity, various markers may be incorporated into the breeding program. For example, the dominant  $R_1$  allele in *G. hirsutum* conditions a wine-red plant color. Pollinated flowers are tagged with a wired tag.

### Natural pollination

As previously indicated, cotton is capable of cross-pollination to a great extent. Genetic male sterility is used to aid mass crossing in the field. The crossing block may have to be isolated for this purpose.

### Seed development

Date to maturity varies among species and cultivars. Small crossing blocks are harvested by hand and delinted using a micro-gin. The fuzz is removed by chemical treatment with sulfuric acid and then dried at 43°C before storage.

### Common breeding objectives

Some of the major breeding objectives in cotton are as follows:

- 1 **Lint fiber yield.** The primary components of fiber yield are number of bolls per plant, size of bolls, and percent lint, of which the number of bolls per plant is the most important. Plant breeders select plants that are prolific. Bolls with five locks yield higher than those with four locks. Boll number and size are negatively correlated, making it difficult to improve both traits simultaneously. Recurrent selection may be used to break this undesirable association.
- 2 **Agromorphological traits.** Harvesting is mechanized in cotton production in the US. Traits that facilitate mechanized harvesting include lodging resistance, bolls set high on the plant, bolls borne singly, and natural leaf shedding at maturity. Early maturing and rapid fruiting are also desirable traits.
- 3 **Adaptation.** Cotton is produced in dry regions of the world. Consequently, drought resistance is an important breeding objective for this crop.

- 4 Disease resistance.** Some of the major diseases of cotton are as follows:
- Seedling diseases.** Some of the major diseases of cotton are caused by *Fusarium* spp., *Pythium* spp., and *Rhizoctonia solani*. These soil-borne fungi cause diseases to seedlings in wet soil, including damping-off and seed rot. The consequence of these diseases is reduced crop stand.
  - Nematodes.** The root-knot nematode, *Meloidogyne incognita*, is a destructive pest in some growing regions. Resistance to the disease is quantitatively inherited.
  - Verticillium wilt.** This disease occurs widely in the US. Some genotypes with resistance occur in the *G. barbadense* species.
  - Bacterial blight.** Also called angular leaf spot, bacterial blight (*Xanthomonas malvacearum*) is widespread throughout all cotton production regions. Resistance to the disease is conditioned by two or more major genes with minor or modifier genes.
- 5 Insect resistance.** The major insect economic insect pests of cotton are the boll weevil and bollworm.
- Boll weevil.** Caused by *Anthonomus grandis*, this insect pest causes cotton squares to drop. Early maturing cultivars tend to escape the pest.
  - Cotton bollworms.** Various lepidopteran insects belong to this group of devastating insect pests of cotton. These include the cotton bollworm (*Helicoverpa zea*), tobacco budworm (*Heliothis virescens*), and pink bollworm (*Pectinophora gossypiella*). Genetic engineering procedures have been used to address the attack by these pests through the development of *Bt* cultivars.
- 6 Fiber quality.** The quality traits of importance in the cotton industry include fiber length, fiber strength, and fiber fineness. Breeding objectives include improving fiber length and uniformity for improved spinning performance and utility of cotton. The fiber strength determines yarn strength, while fiber fineness affects the texture or feel of the fiber. Fiber strength is important for the current technology of open-end (rotor) spinning, which demands stronger fibers.
- 7 Seed quality.** In addition to lint, cotton is also grown for its seed oil. A major goal in breeding seed quality in cotton is to reduce the pigmentation that discolors the seed oil. The use of glandless cotton cultivars (the glands produce gossypol, a terpenoid compound responsible for the discoloration of the seed oil) helps to improve seed oil quality. However, glandless cultivars are more susceptible to insect attack.

## References and suggested reading

- Basu, A.K., and S.S. Narayanan. 1997. Cotton. In: Genetics, cytogenetics, and breeding of crop plants, Vol. 2. Cereal and commercial crops (Bahl, P.N., P.M. Salimath, and A.K. Mandal, eds). Science Publishers, Inc., Enfield, NH.
- Cotton Information. 2002. North Carolina Cooperative Extension Service. North Carolina State University, Raleigh, NC.
- Jiang, C.X., T.Z. Zhang, J.J. Pan, and R.J. Kohel. 1998. Polyploid formation created unique avenues for response to selection in *Gossypium* cotton. Proc. Natl. Acad. Sci. USA 95:4419–4424.
- Meredith, W.R., Jr. 2005. Minimum number of genes controlling cotton fiber strength in backcross population. Crop Sci. 45:1114–1119.
- Rahman, M., D. Hussain, and Y. Zafar. 2002. Estimation of genetic divergence among elite cotton cultivars – genotypes by DNA fingerprinting technology. Crop Sci. 42:2137–2144.
- Stiller, W.N., J.J. Reed, G.A. Constable, and P.E. Reid. 2005. Selection for water use efficiency traits in a cotton breeding program: Cultivar differences. Crop Sci. 45:1107–1113.

## Outcomes assessment

### Part A

Please answer the following questions true or false:

- The fruit bud of the cotton plant is called a boll.
- New World cotton species have a genomic formula of *AA*.
- The long fiber of cotton seed is called the lint.
- The economically important species of cotton is *G. barbadense*.

**Part B**

Please answer the following questions:

- 1 Give the four species of cotton that produce economic seed fibers.
- 2 Distinguish between Old and New World cotton groups.
- 3 Discuss the importance of *G. hirsutum* in cotton production.

**Part C**

Please write a brief essay on each of the following topics:

- 1 Discuss the American upland cotton types grown in the US.
- 2 Discuss pollination of cotton in breeding.
- 3 Discuss the key breeding objectives of cotton relating to lint yield and quality.

# Glossary of terms

**Accession:** A distinct, uniquely identified sample of seeds, plants, or other germplasm materials that is maintained as an integral part of a germplasm collection.

**Adaptedness:** The degree or capacity of an individual to survive in a local environment and to transmit its genotype to the next generation.

**Additive gene effect:** The effect of an allele expected after it has replaced another allele at a locus.

**Agrobacterium:** A type of soil-inhabiting bacteria that is capable of introducing DNA from plasmids in the bacteria into the genome of plant cells. Often used in the genetic transformation of plants.

**Allele:** One of several alternate forms (DNA sequences) that resides at the same locus on the chromosome and controls the same phenotype (although with potentially differing effects).

**Allogamy:** Alternative term for cross-pollination.

**Allopolloid (or allotetraploid):** An individual with somatic cells that contain more than two sets of chromosomes, each of which derives from a different species.

**Amino acid:** A building block of proteins. Each protein consists of a specific sequence of amino acids (with the sequence of amino acids determined by the sequence of the underlying DNA). There are 20 types of amino acid molecules that make up proteins.

**Amphidiploid (or amphiploid):** An allopolloid with the complete chromosome complements of two diploid species.

**Aneuploid:** An individual with a chromosome number that is not the exact multiple of the basic number for the species.

**Antisense:** The complementary strand of a coding sequence (gene); often an expressed copy of an antisense sequence is transformed into a cell or organism to shut off the expression of the corresponding gene.

**Apomixis:** Asexual reproduction in plants through the formation of seeds without fertilization (agamospermy).

**Asexual reproduction:** The reproduction process that does not involve the union of gametes.

**Autopolloid (or autoploid):** An individual with more than two complete sets of the basic number of chromosomes for the species.

**Average effect of a gene:** The change in mean value of the population produced by combining a gene with a random sample of gametes from the original population.

**Backcross:** A cross of an  $F_1$  to either parent used to generate it.

**Base collection:** A comprehensive collection of germplasm accessions held for the purpose of long-term conservation.

**Base pair (bp):** Two nitrogenous bases (adenine and thymine or guanine and cytosine) held together by weak bonds. Two strands of DNA are held together in the shape of a double helix by the bonds between base pairs.

**Bioinformatics:** A broad term to describe applications of computer technology and information science to organize, interpret, and predict biological structure and function. Bioinformatics is usually applied in the context of analyzing DNA sequence data.

**Biopharming:** The use of genetically transformed crop plants and livestock animals to produce valuable compounds, especially pharmaceuticals. Also called pharming.

**Bioremediation:** The use of biological organisms to render hazardous wastes non-hazardous or less hazardous.

**Biotechnology:** A set of biological techniques developed through basic research and now applied to research and product development.

**Breeding:** The science and art of manipulating the heredity of an organism for a specific purpose.

**Breeding line:** A genetic group that has been selected and bred for its special combinations of traits.

**Breeding value:** The mean genotypic value or the progeny of an individual expressed as a deviation from the population mean.

**Bt (*Bacillus thuringiensis*):** A naturally occurring bacterium that produces a protein toxic to certain lepidopteran insects.

**Callus:** A cluster of undifferentiated plant cells that have the capacity to regenerate a whole plant in some species.

**Cell:** The fundamental level of structural organization in complex organisms. Cells contain a nucleus (with chromosomes) and cytoplasm with the protein synthesis machinery, bounded by a membrane.

**Cell culture:** A technique for growing cells under laboratory conditions.

**Cell fusion:** The formation of a hybrid cell produced by fusing two different cells.

**Centimorgan (cM):** A unit of measure of recombination frequency. One centimorgan is equal to a 1% chance that a marker at one genetic locus will be separated from a marker at a second locus due to crossing over in a single generation.

**Central dogma:** The underlying model for describing gene structure and function. It states that genes are transcribed in the nucleus into messenger RNA molecules, which are then translated into proteins on ribosomes.

**Certified seed:** The progeny or increase from a breeder or foundation seed and approved by a certifying agency.

**Chimera:** An individual consisting of cells of two or more types.

**Chromosome:** A condensed structure found in the cell nucleus that contains the genes of that cell.

**Clonal propagation:** The reproduction of plants through asexual means, such as cuttings, grafts, or tissue culture.

**Cloning:** Asexually producing multiple copies of genetically identical cells or organisms descended from a common ancestor.

**Codon:** A triplet of nucleotides in a DNA or RNA molecule that codes for one of the 20 amino acids in proteins, or for a signal to start or stop protein production. Each gene that codes for protein is a series of codons that gives the instructions for building that protein.

**Combining ability:** The performance of a line with others in a cross.

**Complementary:** The opposite or “mirror” image of a DNA sequence. A complementary DNA sequence has an A for every T, and a C for every G. Two complementary strands of single-stranded DNA will join to form a double-stranded molecule.

**Complementary DNA (cDNA):** A single-stranded DNA molecule that is complementary to a specific RNA molecule and synthesized from it. Complementary DNAs are important laboratory tools as DNA probes and for isolating and studying individual genes.

**Conserved sequence:** A base sequence in a DNA molecule (or an amino acid sequence in a protein) that has remained essentially unchanged throughout evolution.

**Crossing over:** The breaking during meiosis of one maternal and one paternal chromosome, the exchange of corresponding sections of DNA, and the rejoining of the chromosomes.

**Cultivar:** A product of plant breeding that is released for access to producers.

**Deoxyribonucleic acid (DNA):** The molecule that encodes genetic information. DNA is a double-stranded molecule held together by weak bonds between base pairs of nucleotides. The four nucleotides in DNA contain the bases: adenine (A), guanine (G), cytosine (C), and thymine (T). In nature, base pairs form only between A and T and between G and C; thus the base sequence of each single strand can be deduced from that of its partner.

**Diploid:** A full set of genetic material consisting of paired chromosomes, one chromosome from each parental set.

**DNA chip:** A high density array of short DNA molecules bound to a solid surface for use in probing a biological sample to determine gene expression, marker pattern, or nucleotide sequence of DNA/RNA. See also **Microarray**.

**DNA probe:** A single-stranded DNA molecule used in laboratory experiments to detect the presence of a complementary sequence among a mixture of other single-stranded DNA molecules. Also called gene probe.

**DNA profile:** The distinctive pattern of DNA restriction fragments or PCR products that can be used to identify, with great certainty, any person, biological sample from a person, or organism from the environment.

**DNA replication:** The use of existing DNA as a template for the synthesis of new DNA strands. In humans and other eukaryotes, replication occurs in the cell nucleus.

**DNA sequencing:** Determining the order of nucleotides in a specific DNA molecule.

**Domestication:** The process of bringing wild plants under cultivation to produce crops under the supervision of humans.

**Dominant:** A phenotype that is expressed in an organism whose genotype may be either homozygous or heterozygous for the corresponding allele.

**Double helix:** The shape that two linear strands of DNA assume when bonded together.

**Doubled haploid:** An individual that is produced by doubling its gametic ( $n$ ) chromosome number into  $2n$ .

**Electrophoresis:** A method of separating substances, such as DNA fragments, by using an electric field to make them move through a “gel” at rates that correspond to their electric charge and size.

**Embryo rescue:** The removal and culture of an immature embryo to produce a plant, often following a wide cross.

**Enhancement:** The process of improving a germplasm accession by breeding while retaining the important genetic contributions of the accession.

**Enzyme:** A protein that acts as a catalyst, speeding the rate at which a biochemical reaction proceeds but not altering the direction or nature of the reaction.

**Epistasis:** The interaction of genes at different loci; the situation in which one gene affects the expression of another.

**Eukaryote:** Cell or organism with a membrane-bound, structurally discrete nucleus and other well-developed subcellular compartments.

**Functional genomics:** The field of study that attempts to determine the function of all genes (and gene products), largely based on knowing the entire DNA sequence of an organism.

**Gamete:** Mature male or female reproductive cell (sperm or ovum) with a haploid set of chromosomes.

**Gene:** The fundamental unit of heredity; a bundle of information for a specific biological structure or function.

**Gene cloning:** Isolating a gene and making many copies of it by inserting the DNA sequence into a vector, then into a cell, and allowing the cell to reproduce and make many copies of the gene.

**Gene expression:** The process in which a cell produces the protein, and thus the characteristic, that is specified by a gene's nucleotide sequence.

**Gene library:** A collection of DNA fragments (carried on vector molecules) that, taken together, represents the total DNA of a certain cell type or organism.

**Gene regulation:** The process of controlling the synthesis or suppression of gene products in specific cells or tissues.

**Gene splicing:** Joining pieces of DNA from different sources using recombinant DNA technology.

**Genetic code:** The language in which DNA's instructions are written. The code consists of triplets of nucleotides (codons), with each triplet corresponding to one amino acid in a protein structure or to a signal to start or stop protein production.

**Genetic engineering:** The manipulation of genes, composed of DNA, to create heritable changes in biological organisms and products that are useful to people, living things, or the environment.

**Genetic erosion:** The loss of genetic diversity caused by either natural or manmade processes.

**Genetic marker:** A genetic factor that can be identified and thus acts to determine the presence of genes or traits linked with it but not easily identified.

**Genetic stocks:** Accessions that typically possess one or more special genetic traits that make them of interest for research.

**Genetic vulnerability:** The condition that results when a crop or a plant species is genetically and uniformly susceptible to a pest, pathogen, or environmental hazard.

**Genetically modified (GM) organism:** An organism whose genetic makeup has been changed by any method including natural processes, genetic engineering, cloning, mutagenesis, or others.

**Genetics:** Study of the patterns of inheritance of specific traits.

**Genome:** The complete set of chromosomes found in each cell nucleus of an individual.

**Genomics:** The field of study that seeks to understand the structure and function of all genes in an organism based on knowing the organism's entire DNA sequence, with an extensive reliance on powerful computer technologies.

**Genotype:** The specific combination of alleles present at a single locus in the genome.

**Germ cells:** The sex cell(s) of an organism (sperm or egg, pollen or ovum). They differ from other cells (somatic) in that they contain only half the usual number of chromosomes. Germ cells fuse during fertilization to begin the next generation.

**Germplasm:** The sum total of all hereditary material in a single (interbreeding) species.

**Green Revolution:** An aggressive effort between 1950 and 1975 where agricultural scientists applied modern principles of genetics and breeding to improve crops grown primarily in less developed countries.

**Haploid:** A cell or organism with a single genome.

**Heterozygosity:** The presence of different alleles at one or more loci on homologous chromosomes.

**Heterozygous:** Situation where the two alleles at a specific genetic locus are not the same.

**Homologous:** Stretches of DNA that are very similar in sequence, so similar that they tend to stick together in hybridization experiments. Homologous can also be used to indicate related genes in separate organisms controlling similar phenotypes.

**Homologous chromosomes:** A pair of chromosomes containing the same linear gene sequences, each derived from one parent.

**Homozygous:** Situation where the two alleles at a specific genetic locus are identical to one another.

**Hybrid:** The progeny of a cross between two different species, races, cultivars, or breeding lines.

**Hybridization (or crossing):** The process of pollen transfer from the anther of the flower of one plant to the stigma of the flower of a different plant for the purpose of gene transfer to produce an offspring (hybrid) with a mixed parental genotype.

**Hybridization:** Bringing complementary single strands of nucleic acids together so that they stick and form a double strand. Hybridization is used in conjunction with DNA and RNA probes to detect the presence or absence of specific complementary nucleic acid sequences.

**In vitro:** Performed in a test tube or other laboratory apparatus.

**In vivo:** In the living organism.

**Inbreeding:** The breeding of individuals that are related.

**Isoenzyme (isozyme):** Different chemical forms of the same enzyme that can generally be distinguished from one another by electrophoresis.

**Landrace:** A population of plants, typically genetically heterogeneous, commonly developed in traditional agriculture from many years of farmer-directed selection, and which is specifically adapted to local conditions.

**Linkage:** The proximity of two or more markers (e.g., genes, RFLP markers) on a chromosome.

**Linkage map:** A map of the relative positions of genetic loci on a chromosome, determined on the basis of how often the loci are inherited together. Distance is measured in centimorgans (cM).

**Locus:** The position on a chromosome where the gene for a particular trait resides; a locus may be occupied by any one of several alleles (variants) for a given gene.

**Meiosis:** The process of two consecutive cell divisions in the diploid progenitors of sex cells. Meiosis results in four rather than two daughter cells, each with a haploid set of chromosomes.

**Messenger RNA (mRNA):** The ribonucleic acid molecule that transmits genetic information from the nucleus to the cytoplasm, where it directs protein synthesis.

**Microarray:** A large set of cloned DNA molecules spotted onto a solid matrix (such as a microscope slide) for use in probing a biological sample to determine the gene expression, marker pattern, or nucleotide sequence of DNA/RNA.

**Microsatellite:** A repeated motif of nucleotides, usually only two or three bases in length, where the number of repeats frequently differs between different members of a species.

**Mitosis:** The process of nuclear division in cells which produces daughter cells that are genetically identical to each other and to the parent cell.

**Molecular marker:** An identifiable physical location on a chromosome (e.g., restriction enzyme cutting site, gene) whose inheritance can be monitored.

**Multiline:** A mixture of isolines, each of which is different for a single gene conditioning different forms of the same trait.

**Mutagen:** A substance that induces mutations.

**Mutation:** A permanent change in the genetic material involving either a physical alteration in the chromosome or a biochemical change in the underlying DNA molecule.

**Nitrogenous base:** A nitrogen-containing molecule having the chemical properties of a base.

**Nucleic acid:** A large molecule composed of nucleotide subunits.

**Nucleotide:** A subunit of DNA or RNA consisting of a nitrogenous base (adenine, guanine, thymine, or cytosine in DNA).

**Nucleus:** Membrane-bound structure in the cell that contains the chromosomes (genetic material). The nucleus divides whenever the cell divides.

**Pathogen:** A specific biological causative agent of disease in plants or animals.

**Pedigree:** A record of the ancestry of an individual or family.

**Phenotype:** A biological characteristic or trait possessed by an organism that results from the expression of a specific gene.

**Physical map:** A map of the locations of identifiable landmarks on DNA (e.g., restriction enzyme cutting sites, genes), regardless of inheritance. Distance is measured in base pairs.

**Plasmid:** A small, self-replicating molecule of DNA that is separate from the main chromosome. Because plasmids are easily moved from cell to cell or to the test tube, scientists often cleave them with restriction enzymes and insert foreign DNA, and then transfer the recombinant DNA plasmid molecule (as a vector) into other cells.

**Pollination:** The transfer of pollen from the anthers to the stigma of a flower.

**Polymerase chain reaction (PCR):** A technique to amplify a specific DNA sequence *in vitro* using a DNA replicating enzyme, specific oligonucleotide primers, and repeated cycles of heating and cooling. PCR often amplifies the starting material many thousands or millions of times.

**Polymorphism:** The simultaneous occurrence of two or more distinct forms in a population in a frequency that cannot be accounted for by the balance of mutation and selection.

**Polypliody:** An individual with more than two sets of chromosomes characteristic of the species.

**Primer:** Short pre-existing polynucleotide chain to which new deoxyribonucleotides can be added by DNA polymerase.

**Probe:** Single-stranded DNA or RNA molecules of a specific base sequence, labeled either radioactively or immunologically, that are used to detect the complementary base sequence by hybridization.

**Prokaryotes:** Organisms whose genetic material is not enclosed by a nucleus.

**Promoter:** A DNA sequence preceding a gene that contains regulatory sequences controlling the rate of RNA transcription of that gene. In effect, promoters control when and in which cells a given gene will be expressed.

**Protein:** A molecule composed of amino acids arranged in a special order determined by the genetic code. Proteins are required for the structure and function of all living organisms.

**Pure line:** The progeny of a single homozygous individual produced by repeated selfing.

**Recessive:** A phenotype that is expressed in organisms only if it is homozygous for the corresponding allele.

**Recombinant DNA:** A hybrid DNA molecule produced in the laboratory by joining pieces of DNA from different sources.

**Recombinant DNA technologies:** Procedures used to join together DNA segments in a cell-free system (an environment outside a cell or organism). Under appropriate conditions, a recombinant DNA molecule can enter a cell and replicate there, either autonomously or after it has become integrated into a cellular chromosome.

**Recombination:** The process by which progeny derive a combination of genes different from that of either parent. In higher organisms, this can occur by crossing over.

**Recurrent selection:** A breeding method whereby plants are repeatedly selected and intercrossed to increase the frequency of favorable alleles.

**Regeneration:** The process of growing an entire plant from a single cell or group of cells.

**Reporter gene:** A gene sequence that is easily observed when it is expressed in a given tissue or at a certain stage of development.

**Restriction enzyme:** An enzyme that recognizes a specific nucleotide base sequence (usually four to six base pairs in length) in a double-stranded DNA molecule and cuts both strands of the DNA molecule at every place where this sequence occurs.

**Restriction fragment length polymorphism (RFLP):** The presence of two or more variants in the size of DNA fragments produced by a restriction enzyme. These different sized fragments result from an inherited variation in the presence of a restriction enzyme's target sequence. RFLPs are used for gene mapping and DNA profiling.

**Ribonucleic acid (RNA):** A molecule that translates the instructions encoded in DNA to build proteins.

**Ribosomes:** Small cellular components composed of specialized ribosomal RNA and protein; site of protein synthesis.

**Selection (field):** The process of discriminating among genetic variability to advance a fraction to the next generation or breeding cycle.

**Selection (*in vitro*):** A method to retain specific cells (or clones of cells) expressing a specific trait, such as antibiotic or herbicide resistance, while killing off all other cells that do not express that trait.

**Somatic cell:** Cells in the body that are not involved in sexual reproduction (that is, not germ cells).

**Southern blotting:** Transfer by absorption of DNA fragments separated in electrophoretic gels to membrane filters for the detection of specific base sequences by radiolabeled complementary probes.

**Tissue culture:** Growing cells, tissues, or tissue fragments (from complex, multicellular organisms) on a nutrient medium in a dish, test tube, or flask.

**Totipotent:** A cell that is capable of regenerating an entire adult organism by itself.

**Trait:** A distinguishing characteristic or quality of an organism.

**Transcription:** The transfer of information from specific

sequences in a DNA molecule to produce new strands of messenger RNA, which then carry this information from the nucleus to the cytoplasm (where the messenger RNA is translated into protein).

**Transformation:** Introduction of an exogenous DNA molecule into a cell, causing it to acquire a new phenotype (trait).

**Transgenic:** An organism that has been transformed with a foreign DNA sequence.

**Translation:** Synthesis of protein using information contained in a messenger RNA molecule.

**Vector:** A type of DNA molecule, usually a plasmid or virus, that is used to move recombinant DNA molecules from one cell to another.

# Appendix 1

## Internet resources

### Chapter 1

<http://www.foodfirst.org/media/opeds/2000/4-greenrev.html> Lessons from the Green Revolution.  
<http://www.arches.uga.edu/~wparks/ppt/green/> Biotechnology and the Green Revolution. Interview with Norman Borlaug.  
<http://cuke.hort.ncsu.edu/cucurbit/wehner/741/hs741hist.html> History of plant breeding.

### Chapter 2

<http://agronomy.ucdavis.edu/gepts/pb143/pb143.thm> Gepts, P. 2002. The evolution of crop plants.  
<http://cucurbitsvr.hort.ncsu.edu/breeding/usplantbreeding/uspbmain.html> Plant breeding in the USA. List of land grant institutions and seed companies.  
<http://pas.byu.edu/AgHrt100/evolutio.htm> Synopsis on plant breeding and evolution.

### Chapter 3

<http://molvis.sdsc.edu/atlas/atlas.htm#dnarna> Colorful animated figures of DNA and other macromolecules.  
<http://www-scf.usc.edu/~chem203/resources/DNA/doublehelix.html> DNA structure.  
[http://biog-101-104.bio.cornell.edu/BioG101\\_104/tutorials/cell\\_division.html](http://biog-101-104.bio.cornell.edu/BioG101_104/tutorials/cell_division.html) Well illustrated tutorial site on cell division.

### Chapter 4

<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/A/AsexualReproduction.html> Asexual reproduction in plants.  
<http://www.emc.maricopa.edu/faculty/farabee/BIOBK/BioBookflowers.html> Excellent illustrations and discussion of aspects of reproduction in flowering plants.  
<http://www.ukans.edu/~bio152/17/sld001.htm> Excellent slides on plant reproduction.

<http://billie.bnny.purdue.edu/apomixis/apomixis.html> Excellent overview of apomixis.

### Chapter 6

<http://www.barc.usda.gov/psi/ngrl/ngrl.html> Website of National Germplasm Resources Lab.  
<http://www.plantstress.com/admin/WRFFiles/germplasmwr.htm> List of websites for plant germplasm resource centers worldwide.  
<http://www.ciesin.org/docs/002-256a/002-256a.html> Paper on current status of biological diversity by E. O. Wilson of Harvard University.

### Chapter 7

[http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/H/Hardy\\_Weinberg.html](http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/H/Hardy_Weinberg.html) Excellent discussion of population genetics.

### Chapter 10

<http://csf.colorado.edu/perma/stse/isolate.htm> Basic isolation practices for reducing or eliminating natural cross-pollination in field crossing.  
[http://www.actahort.org/books/200/200\\_3.htm](http://www.actahort.org/books/200/200_3.htm) Application of wide crosses in tomato improvement.

### Chapter 11

<http://aggie-horticulture.tamu.edu/tisscult/microprop/microprop.html> Links to numerous aspects of plant micropropagation.  
<http://billie.bnny.purdue.edu/apomixis/apomixis.html> Excellent overview of apomixis.  
<http://www.sprrs.usda.gov/apomixis.htm> Comments from foremost scientists in field of apomixis.  
[http://www.blogontheweb.com/tissue\\_culture](http://www.blogontheweb.com/tissue_culture) Excellent discussion on tissue culture.

## **Chapter 12**

<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/M/Mutations.html> Detailed discussion of mutations in nature.  
<http://www.plantmutations.com/> Excellent discussion of mutagenesis in plant breeding.

## **Chapter 13**

<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/P/Polypliody.html> Good discussion on polypliody.  
<http://wheat.pw.usda.gov/ggpages/BarleyNewsletter/42/oral04.html> Application of doubled haploids in barley breeding.

## **Chapter 15**

[http://www.wipo.org/about-ip/en/about\\_patents.html](http://www.wipo.org/about-ip/en/about_patents.html) General information about patents.  
<http://www.3bsproject.com/html/ip.html> European perspectives on intellectual property.  
<http://www.cid.harvard.edu/cidbiotech/comments/comments117.htm> Biotechnology and morality debate.  
<http://www.aphis.usda.gov/biotech/OECD/usregs.htm> US biotechnology regulatory oversight.  
[http://www.agwest.sk.ca/saras\\_reg\\_int.shtml](http://www.agwest.sk.ca/saras_reg_int.shtml) Links to international regulations on biotechnology.  
<http://www.codexalimentarius.net/> Codex Alimentarius Commission.  
[http://www.enn.com/enn-features-archive/2000/03/03052000/gefood\\_5991.asp](http://www.enn.com/enn-features-archive/2000/03/03052000/gefood_5991.asp) Raging debate over biotechnology food.  
<http://www.cid.harvard.edu/cidbiotech/comments/comments117.htm> Debate over morals in biotechnology.  
<http://uk.fc.yahoo.com/g/genetic.html> GM foods debate.

## **Chapter 18**

<http://www.ontariocorn.org/ocpmag/dec99feat.html> Historical account of corn hybrids.

## **Chapter 22**

<http://www.ext.colostate.edu/PUBS/CROPS/00307.html> Brief overview of pharming in plants.  
<http://www.molecularfarming.com/molecular-farming-patents.html> Site to various products and companies engaged in pharming.

## **Chapter 25**

<http://www.cgiar.org>  
<http://impact.cgiar.org>  
<http://www.cgiar.org/centers/index.html>

## **Chapter 26**

<http://www.anth.org/ifgene/breed2.gif> A scheme for organic plant breeding.

## **Chapter 31**

<http://www.notrans.iastate.edu/research.html#breeding>  
Breeding for low linolenic acid.  
<http://www.ag.uiuc.edu/~stratsoy/research/ind10.html>  
Breeding for high seed protein.  
<http://www.indianasoybeanboard.com/Links.shtml>  
Soybean links.

## **Chapter 32**

<http://www.vaes.vt.edu/tidewater/peanut/> Peanut production guide.

## **Chapter 33**

<http://oregonstate.edu/potatoes/potliv.html> Potato links.  
<http://www.umaine.edu/paa/Breeding/B&Gsec31802.htm> Site of the Potato Association of America.  
<http://www.aphis.usda.gov/bbep/bp/potato.html> Brief overview of the crop.

# Appendix 2

## Conversion rates

Imperial unit	Metric conversion
<i>Volume/Capacity</i>	
Cubic inch	16.39 cubic centimeters
Bushel	0.036 cubic meters
Pint	0.57 liters
Quart	1.14 liters
Gallon	4.55 liters
<i>Area</i>	
Acre	4046.86 square meters
Acre	0.4 hectares
Square yard	0.8 square meters
Square feet	0.09 square meters
<i>Mass</i>	
Ounce (avoirdupois)	28.35 grams
Pound	0.45 kilogram
Hundredweight	50.80 kilograms
Ton	1.02 tonnes
<i>Length</i>	
Inch	2.54 centimeters
Foot	0.31 meter
Yard	0.91 meter
Mile	1.61 kilometers

# Index

- A-chromosomes, 229  
A-line, 344  
abiotic resistance, 385  
active collections, 99  
adaptation, 9, 427  
additive effect, 125  
additive gene action, 125  
additive variance, 127  
advanced backcross breeding, 96, 252  
advanced yield trial, 419  
adventitious shoot production, 183  
agamospermy, 64  
agriculture, 3, 6  
alien addition lines, 228  
alleles, 38, 40–43  
    multiple, 40  
    recessive, 40  
allogamous (cross-pollination), 56  
allopoloidy (allopolloid), 215, 221  
allozymes, 42, 248  
alternative hypothesis, 147  
alternative splicing, 51  
amphiploidy (amphiploid), 215  
amplified fragment length  
    polymorphisms (AFLP), 249  
androecium, 58  
androgenesis, 64  
angiosperms, 76  
annuals, 57  
ANOVA (analysis of variance), 154  
Anther culture, 186  
antibiotic selection, 235  
antiparallel, 48  
anueploidy, 227  
apomictic cultivar, 283  
apomixis, 63–67, 196  
apospory, 64  
artificial hybridization, 165  
    *see also* hybridization  
Association of Official Seed Analysis,  
    445  
Association of Official Seed Certifying  
    Agencies, 447  
autogamous (self-pollination), 56  
autopoloidy (autopoloids), 215, 220  
avirulent, 369  
avoidance, 369, 379  
axillary shoot production, 183  
B-chromosomes, 229  
B-line, 344  
backcross breeding, 303–308, 331  
backup collections, 99  
base collections, 99  
better-parent heterosis, 340  
biennials, 57  
binomial system of classification, 7,  
    76  
biochemistry, 26  
bioinformatics, 26, 238, 240  
biological variation, 75  
biological yield, 353  
biomass, 353  
biopharming, 4  
biotechnology, 231–255  
    cloning genes, 233  
    definition, 231  
    gene transfer, 234–237  
    general steps in rDNA technology,  
        231  
    genomics, 237  
    landmark discoveries, 232  
    legal risks of adoption of GM crops,  
        278  
    public perception and fears, 276  
    regulation of the industry, 269  
biotechnology in developing countries,  
    459  
biparental inheritance, 37  
blends (cultivar), 308  
boom and bust cycles, 375  
Borlaug, N., 10, 11–13  
botany, 25  
breeder seed, 436, 438  
breeders' equation, 131  
breeder's eye, 24  
breeders' rights, 258  
breeders' trials, 418  
breeding objectives, 495, 507, 517,  
    527, 535, 544, 553  
breeding potato, 537–545  
breeding value, 126  
bridge crosses, 178  
broad sense heritability, 128  
*bt* resistance, 377  
bulbosum method, 187  
bulk population breeding, 294–297  
Calyx, 58  
Camerarius, R.J., 7  
canonical correlation, 161  
Cartagena Protocol on biosafety, 274  
cDNA library, 233  
cell, 35  
central dogma of molecular biology, 49  
centralized plant breeding, 462–468  
certified seed, 436, 441  
CGIAR centers, 453  
chemical mutagens, 203  
chiasma, 80  
chimera, 200, 206  
chimerism, 194  
chi-square test, 153  
chromosomal mutation, 202  
chromosome deletion, 228  
chromosome elimination, 187  
chromosome mapping, 46  
chromosome substitution, 229  
chromosomes, 37  
classification, 75–78  
    operational system, 77  
    rules, 76  
cleistogamy, 60  
clonal cultivar, 283  
clonal propagation, 61  
clonally propagated species breeding,  
    193–196  
    breeding apomictic cultivars, 196  
    breeding approaches, 195  
    genetic issues, 194  
clones, 194, 284  
cloning vectors, 233  
cluster analysis, 161  
coat protein mediated viral resistance,  
    382  
codominance, 42  
codon, 50  
coefficient of inbreeding, 116  
coefficient of variation, 151  
cold stress breeding, 392  
combining ability, 139–141, 328  
comparative genomics, 237  
complementary DNA (cDNA), 49  
complete block designs, 430  
complete flower, 58  
complete linkage, 43

- complex inheritance, 42  
composite cultivar, 310  
composites, 310  
compositional traits, 9, 404  
congruency backcross, 308  
constitutive expression, 52  
constitutive promoters, 235  
Consultative Group of International Agricultural Research (CGIAR), 452–455  
conventional breeding, 29  
convergent crosses, 170  
cool season plants, 78  
copyrights, 257  
corn breeding, 485–497  
corolla, 58  
correlated response, 133  
correlation coefficient, 151  
cotton breeding, 546–554  
coupling phase, 45  
crop introduction, 20–22  
crop registration, 432  
cross-pollinated, 60  
crossing, 165  
crossing over, 38  
cryopreservation, 100  
cultivar, 4  
cultivar release process, 438  
cuttings, 62  
cybrids, 186  
cytoplasmic inheritance, 37  
cytoplasmic male sterility (CMS), 335, 345, 488
- Darwin, C., 17  
decentralized participatory plant breeding, 463  
delayed ripening, 412  
determinacy, 364  
developing countries (breeding), 451  
diallel cross, 143, 170  
diallel cross test, 328  
dichogamy, 61  
dicots, 76  
differential cultivars, 374  
differential expression, 52  
dihaploid, 214  
dioecious plants, 59  
dioecy, 61, 68, 349  
diploidization, 219  
direct organogenesis, 184  
directed selection, 193  
directional selection, 115  
disease, 368  
disease resistance, 367
- disease triangle, 369  
disruptive selection, 115  
divergent crossing, 170  
DNA, 8  
DNA amplification fingerprinting (DAF), 249  
DNA markers, 247  
DNA microarrays, 239, 243  
DNA sequencing, 232, 237  
domestication, 18–24  
    centers of diversity, 19  
    law of homologous series, 20  
    patterns, 19  
    syndrome, 23  
dominance, 39  
dominance effect, 12  
dominance gene action, 126  
dominance theory (of heterosis), 339  
dominant mutation, 201  
double cross, 170  
double dwarfs, 363  
double fertilization, 60  
double helix, 48  
doubled haploids, 187, 190  
drought stress, 387  
drought stress breeding, 387–392, 393–396
- ear-to-row selection, 317  
early generation testing, 139  
economic yield, 353  
ecosystem, 5  
emasculination, 26, 166, 481  
embryo rescue, 185  
end-use quality breeding, 412  
environmental index, 424  
environmental variation, 79  
epidemics, 383  
epistasis, 43  
ethics, 263  
evolution, 17  
*ex situ* conservation, 91, 98  
exons, 51  
experimental design, 427  
experimental unit, 148, 427  
explant, 183  
expressivity, 46
- factor analysis, 155  
Fairchild, T., 7  
female gametophyte, 59  
fertilization, 59  
field plot designs, 429  
field plot technique, 427  
filament, 58
- first division restitution (FDR), 216  
“FlavrSavr” tomato, 413  
flood stress, 402  
flowers, 58  
food allergy, 271  
forward mutation, 200  
foundation seed, 436, 440  
frame shift mutation, 202  
full-sib family selection, 318  
functional genomics, 238
- G × E* interaction, 386, 420–423  
gametic incompatibility, 172  
gametogenesis, 59  
gene action, 125  
gene banks, 91  
gene-for-gene hypothesis, 372  
gene frequency, 110, 113  
gene mutation, 200  
gene pool, 109  
gene pyramiding, 247, 374  
general worth, 139  
genes, 4  
genetic advance, 130  
genetic code, 50  
genetic determination, 128  
genetic engineering, 9  
genetic erosion, 96  
genetic gain, 130  
genetic linkage, 43–46  
genetic markers, 25  
genetic transformation, 80, 234  
genetic use restriction systems (GURT), 244–246  
genetic variation, 79  
genetic vulnerability, 90  
genetically modified (GM), 6  
genome, 36, 237  
genomic DNA library, 233  
genomic formula, 214  
genomic mutation, 200  
genomics, 477  
genotype, 40  
genotype × environment (*G × E*) interaction, 386, 420–423  
germplasm, 29, 87–107  
    collection, 98  
    conservation, 91, 98, 103  
    enhancement, 93  
    gene pools, 89  
    genetic vulnerability, 89, 95  
    importance, 87  
    sources, 88  
    storage technologies, 100  
germplasm collection, 98

- germplasm composites, 327  
 germplasm enhancement, 91, 102  
*gigas*, 216  
 "Golden Rice", 5, 410  
 green revolution, 10  
 growing degree days, 347  
 gymnosperms, 76  
 gynoecium, 58
- half-sib family selection, 315  
 haploid number, 214  
 haploids, 186–188  
   anther culture, 186  
   definition, 189  
   doubled haploids, 187, 191  
   ovary culture, 187  
 Hardy–Weinberg equilibrium, 110, 112  
 Harlan, J.R., 23, 89  
 harvest index, 355  
 harvest maturity, 365  
 heat stress breeding, 399  
 heat units, 347  
 herbicide resistance engineering, 382  
 heredity, 18  
 heritability, 127–130  
   applications, 130  
   computation, 129  
   definition, 128  
   types, 128  
 hermaphrodite, 55  
 heterosis, 339  
   biometrics of heterosis, 340  
   dominance theory, 340  
   genetic basis, 339  
   overdominance theory, 340  
 heterostyly, 61  
 heterotic relationship (groups), 341  
   definition, 342  
   heterotic groups and patterns, 342  
   methods for developing groups, 342  
 heterozygote, 40  
 high lysine, 405  
 homeologous chromosome, 222  
 homologous chromosomes, 37  
 homozygote, 40  
 horizontal resistance, 371  
 host, 369  
 hybrid, 7  
 hybrid breeding, 491–494  
 hybrid cultivar, 283, 334  
   definition, 335  
   genetic basis of heterosis, 339  
   hybrid vigor, 335  
   inbreeding depression, 336  
   types of hybrids, 343  
 hybrid cultivation, 336  
 hybrid seed production, 448  
 hybrid vigor, 335  
 hybrid weakness, 172  
 hybridization, 164–172  
   artificial hybridization, 165  
   genetic issues, 168  
   linkage drag, 169  
   types of population generated, 170  
 identity preservation, 275  
 ideotype concept, 354  
*in situ* conservation, 91, 98  
*in vitro* culture, 182  
*in vitro* selection, 188  
 inbred line, 285, 344, 346  
 inbreeding, 116–119  
 inbreeding depression, 61, 335  
 incomplete block designs, 431  
 incomplete dominance, 42  
 incomplete flower, 59  
 incomplete linkage, 45  
 independent curling, 133  
 indexing, 195  
 indirect organogenesis, 184  
 inflorescence, 59  
 insect resistance engineering, 377  
 intellectual property, 257–262  
   concept, 257  
   patents, 258–263  
 international agricultural research centers, 106, 452–455  
 International Board of Plant Nomenclature, 76  
 international centers, 452  
 International Crop Improvement Association, 447  
 international crop research centers, 453  
 International Seed Testing Association, 445  
 interspecific cross, 7  
 intrapopulation improvement, 315  
   ear-to-row selection, 317  
   full-sib family selection, 318  
   half-sib family selection, 316  
   half-sib reciprocal recurrent selection, 326  
   modified half-sib selection, 318  
 introns, 50  
 intuitive index, 135  
 ionizing radiations, 203  
 isolines (near isogenic lines, NIL), 283, 308  
 isozymes, 248
- Jones, D.F., 7  
 Koelreuter, J., 7  
 law of independent assortment, 39  
 law of segregation, 39  
 legal risk, 278  
 line cultivar, 288  
 linkage, genetic, 43  
 linkage block, 123  
 linkage drag, 169, 307  
 linkage group, 43  
 Linneaus, C., 8  
 living modified organisms (LMO), 274  
 local control, 148, 429  
 locus, 39  
 lodging resistance, 362  
 long-day plants, 364  
 low phytate breeding, 412  
 male gametophyte, 59  
 male sterility, 70–72  
 marker-assisted selection (MAS), 251–254, 478  
 mass selection, 286–288  
 maternal effect, 37  
 maternal inheritance, 37  
 Mather, C., 7  
 mating design  
   diallel cross, 143  
   North Carolina design I, 142  
   North Carolina design II, 142  
   polycross, 142  
 mating system, 60, 115, 119  
 maturity (early), 365  
 McClintock, B., 8  
 measures of central tendency, 149  
 measures of dispersion, 150  
 meiosis, 38  
 Mendel, G., 42  
 Mendelian concepts, 39, 42  
 Mendelian ratios, 43  
 Mendel's laws, 7  
 messenger RNA, 48  
 metric traits, 84  
 microarrays (DNA), 239  
 microprojectile bombardment (biolistics), 234  
 micropagation, 62, 183  
   *see also* tissue, culture  
 middling resistance, 374  
 midparent heterosis, 340  
 migration, 113  
 mineral deficiency stress, 401  
 mineral toxicity breeding, 400

- minor gene resistance, 371  
 minor genes (polygenes), 43, 84  
 missense mutation, 202  
 mitosis, 38  
 modified pedigree method, 297  
 modifying genes, 124  
 molecular breeding, 246, 231  
 molecular markers, 247–250  
 monocarps, 57  
 monocistronic gene, 52  
 monocots, 76  
 monoecious plants, 59  
 monoecy, 61, 68, 349  
 monoploid, 214  
 morals, 263  
 morphological markers, 247  
 multiline breeding, 308, 374  
 multilines, 283  
 multivariate analysis, 155  
 mutagenesis, 8, 199  
 mutation, 81, 113  
 mutation breeding, 196, 199–215  
     breeding clonally propagated species, 206  
     breeding seed-bearing plants, 205  
     factors affecting mutagenesis, 204  
     limitations of mutagenesis in breeding, 211  
     mutagenic agents, 203  
     types of materials used, 204  
     types of mutation, 200
- narrow sense heritability, 128  
 natural selection, 17  
 new world crops, 20  
 Nilsson, H., 7  
 non-disjunction, 227  
 non-vascular plants, 76  
 Norin 10 dwarfing genes, 475  
 normal distribution, 150  
 North Carolina designs, 142  
 novel traits (breeding), 415  
 nuclear division, 37–39  
 nucleic acids, 47–49  
     DNA, 47  
     RNA, 48  
 nucleoside, 47  
 nucleotide, 47  
 null hypothesis, 147
- obsolete cultivars, 97  
 official trial, 419  
 oil quality breeding, 411  
 old world crops, 20  
 oligogenes (major genes), 43  
 oligonucleotide, 47  
 open-pollinated cultivar, 283  
 open reading frame (ORF), 51  
 organic agriculture, 468  
 organic plant breeding, 468  
 orthodox seed, 98  
 overdominance gene action, 117, 126  
 overdominance theory (of heterosis), 339  
 ovule culture  
 oxidative stress, 401  
 oxygen enhancement ratio, 204
- panicle, 59  
 parent-offspring regression, 129  
 parthenogenesis, 188  
 particle gun, 234  
 passport data, 100  
 patents, 258–262  
 path analysis, 161  
 pathogen, 369  
 pathogeneticity, 369  
 peanut breeding, 529–535  
 Pearson correlation coefficient, 151  
 pedigree, 285  
 pedigree selection, 289–294  
 penetrance, 46  
 peptide bonds, 50  
 perennials, 57  
 perfect flower, 59  
 periclinal mutants, 206  
 phenology, 390  
 phenotype, 40  
 photoperiod, 364  
 photoperiod-insensitive, 9  
 physical mutagens, 203  
 physiological maturity, 365  
 physiological race, 373  
 physiological traits, 352  
 pin, 68  
 plant breeder, 3, 30  
 plant breeding, 14  
     art and science of, 24–27  
     conducting, 28–30  
     cost, 33  
     industry, 30–32  
     international, 32  
 plant exploration, 102  
 plant genetic resources, 96, 99  
 plant genomics, 36, 237  
 plant ideotype, 354  
 plant introduction, 102, 104  
 plant physiology, 25  
 plant taxonomy, 75  
 plant variety protection, 262, 442  
 pleiotropy, 43  
 ploidy modification, 81  
 pollination, 59–61, 167  
     cross-pollination, 61  
     self-pollination, 60  
 polycistronic gene, 52  
 polycross test, 328  
 polygenes, 123  
 polygenic inheritance, 43, 84  
 polymerase chain reaction (PCR), 232  
 polynucleotide, 47  
 polypeptide synthesis (protein synthesis), 49  
 polyploidy, 214–230  
     aneuploidy, 227  
     autopoloidy, 216  
     breeding allopoloids, 226  
     breeding autopoloids, 220  
     cytology, 217  
     effects on plants, 216  
     genetics of autopoloids, 218  
     origin, 216  
     terminology, 214  
     variation in chromosome number, 215  
 polyploidy series, 215  
 population genetics, 109  
 population improvement, 119  
 population mean, 149  
 population variance, 149  
 position effect, 47  
 prebreeding, 102  
 precautionary principle, 273  
 preliminary yield trial, 419  
 primary gene pool, 89  
 primary trisomics, 227  
 principal component analysis, 160  
 progeny test, 7  
 promoters, 235  
 protandry, 61  
 protein, 51–53  
     regulation, 52  
     structure, 52  
     synthesis, 51  
 protein content engineering, 409  
 protein markers, 247  
 proteomics, 238  
 protogyny, 61  
 protoplast, 185  
 pro-vitamin A, 5, 9  
 public sector breeding, 32  
 pure line, 282  
 pure-line selection, 288–290  
 pure-line theory, 7  
 purines, 47

- purity test, 446  
 pyrimidines, 47
- QTL mapping, 251  
 qualitative variation, 83  
   definition, 121  
   polygenes and polygenic inheritance, 123  
   quantitative traits, 122  
   *see also* Variation  
 quality protein maize (QPM), 406–409  
 quantitative genetics, 121–127  
 quantitative (polygenic) inheritance, 123  
 quantitative trait loci (QTLs), 250  
 quantitative traits, 122, 127  
 quantitative variation, 84  
   *see also* variation
- r-line, 346  
 random amplified polymorphic DNA (RAPD), 248  
 random mating, 111, 115  
 random sampling, 147  
 randomization, 148, 429  
 recalcitrant seed, 98  
 recessive (recessivity), 39  
 recessive mutation, 201  
 recognition sequence (site), 232  
 recombinant, 45  
 recombinant DNA, 29  
 recombination frequency, 45  
 recurrent mass selection, 136  
 recurrent selection, 136, 310  
   concept, 314  
   genetic basis, 315  
   types, 315  
 regeneration, 222  
 registered seed, 436, 441  
 regulation of biotechnology, 270–275  
   biosafety regulation at international level, 274  
   concept of substantial equivalence, 272  
   concept of precautionary principle, 273  
   labeling of biotech products, 274
- Reid, R., 7  
 replicated tests, 430  
 replication, 49, 148, 429  
 reporter genes, 235  
 reproduction, 55–74  
   alternation of generation, 56  
   asexual, 62  
   mode, 55  
   sexual, 56
- types, 56  
   types of flowers, 58  
 reproductive isolation barriers, 171–178  
   overcoming challenges, 173  
 resistance (disease), 368, 370  
   breeding strategies, 373  
   durable resistance, 371  
   gene pyramiding, 374  
   genetic resistance, 371  
   genetics, 372  
   horizontal resistance, 371, 375  
   pest resistance, 373  
   polygenic resistance, 374  
   vertical resistance, 371  
   vertifolia effect, 375  
 response to selection, 130  
 restriction enzymes (restriction endonucleases), 232  
 restriction fragment length polymorphism (RFLP), 248  
 reverse mutation, 200  
 ribosomes, 49  
 rice breeding, 498–507  
 risk analysis, 264–269
- salinity stress breeding, 398  
 salt tolerance, 399  
 sample, 147  
 sample mean, 149  
 second division restitution (SDR), 216  
 secondary gene pool, 89  
 sectorial mutants, 206  
 seed certification, 435, 441, 445, 447  
 seed germination test, 445  
 seed industry, 436  
 seed testing, 445  
 seedlessness (breeding), 414  
 segmental allopoloids, 216, 221  
 selection, 6, 18, 113  
   index, 133  
   intuitive index, 134  
   modes, 115  
   for multiple traits, 133  
   response in breeding, 130–134  
 selection differential, 131  
 selection index, 133  
 selective (preferential) pairing, 219  
 selfed family selection, 319  
 self-incompatibility, 68  
 self-pollinated, 60  
 semidwarfs, 363  
 sequence characterized amplified regions (SCAR), 249  
 sequence tagged sites (STS), 249
- sexual hybridization, 164  
 shattering resistance, 362  
 shoot culture, 183  
 short-day plants, 364  
 silent mutation, 202  
 Simmonds, N.W., 23  
 simple linear regression, 152  
 simple sequence repeats (SSR), 249  
 single-cell selection, 193  
 single cross, 170  
 single nucleotide polymorphisms (SNP), 249  
 single-seed descent, 297–303  
 site-directed mutagenesis, 83  
 somaclonal variation, 188, 193  
   *see also* variation  
 somatic embryogenesis, 183  
 somatic hybridization, 185  
 sorghum breeding, 509–518  
 southern hybridization, 232  
 soybean breeding, 519–527  
 spike, 59  
 spontaneous mutation, 200  
 stability analysis, 423–425  
 stabilizing selection, 115  
 stamen, 58  
 standard deviation, 150  
 standard error of the mean, 151  
 statistical hypothesis, 147  
 statistical significance, 147  
 statistics, 146–162  
   analysis of variance, 154  
   chi-square test, 153  
   coefficient of variation, 151  
   experimental design, 148  
   hypothesis, 147  
   measures of central tendency, 149  
   measures of dispersion, 150  
   multivariate statistics, 155–161  
   normal distribution, 150  
   population versus sample, 147  
   probability, 149  
   simple linear regression, 152  
   t-test, 153
- structural genes, 52  
 structural mutation, 200  
 substantial equivalence, 272  
 supernumerary chromosomes, 229  
 suspension culture, 182  
 synteny, 53  
 synthetic cultivars, 327  
 synthetic seed, 184
- t-test, 153  
 tagging commercial seed, 446  
 tandem selection, 133

- technology protection system (TPS), 244  
template strand, 50  
tertiary gene pool, 89  
testcross, 41  
three-way cross, 170  
thrum, 68  
tissue 181–188  
    applications, 184–186  
    culture, 181–185  
    culture medium, 182  
    directed selection, 193  
    *in vitro* selection, 188  
    microppropagation, 183  
    somatic embryogenesis, 184  
topcross test, 328  
totipotency, 181  
trademarks, 258  
traditional breeding, 29  
transcription, 49  
transfer RNA, 49  
transgenic plants, 9, 377  
transgressive segregation, 81, 123  
transient expression, 235  
translation, 49, 51  
transposable elements, 82, 488  
triangle of U, 221  
triticale, 226  
type II error, 147  
variable gene expressivity, 47  
variable gene penetrance, 46  
variation, 18, 79–85  
    environmental, 79  
    genetic, 79  
    qualitative, 83  
    quantitative, 84  
    scale, 83  
    somaclonal, 83  
    types, 79  
variety blend (composite), 310  
variety protection, 433, 442–444  
vascular plants, 76  
Vavilov, N., 7, 19  
vegetative propagation, 62  
vertical resistance, 371  
vertifolia effect, 375  
Vigor test, 446  
Vilmorin, L., 7  
viral resistance engineering, 382  
virulence, 369  
warm season plants, 78  
wheat breeding, 472–483  
wide crosses, 172–177  
bridge crosses, 178  
developing new species, 179  
objectives, 172  
successes, 173  
wild type, 200  
wobble hypothesis, 50  
yield, 353  
    biological pathways to economic  
        yield, 356  
    biological yield, 353  
    definition, 353  
    economic yield, 353  
    selecting for yield *per se*, 356  
    yield components, 356  
    yield plateau, 361  
    yield stability, 361  
    yield components, 356  
    yield gains, 435  
    yield plateau, 361  
    yield potential, 357, 386  
    yield stability, 361  
    yield trials, 418  
xenia, 488