Molecular Markers

Article · December 2013

DOI: 10.1007/978-81-322-1554-7_3

2 authors:

Pavan Kumar Agrawal
Govind Ballabh Pant Engineering College
53 PUBLICATIONS **1,056** CITATIONS

SEE PROFILE

Rahul Shrivastava
Jaypee University of Information Technology
44 PUBLICATIONS **448** CITATIONS

SEE PROFILE

Indu Ravi · Mamta Baunthiyal
Jyoti Saxena
Editors

Advances in Biotechnology

Springer

Advances in Biotechnology

Indu Ravi · Mamta Baunthiyal
Jyoti Saxena
Editors

Advances in Biotechnology

Springer

Editors
Indu Ravi
Indira Gandhi National Open University
Regional Centre Jaipur
Mansarovar, Jaipur, Rajasthan
India

Mamta Baunthiyal
Department of Biotechnology
Govind Ballabh Pant Engineering
College
Ghurdauri, Pauri, Uttarakhand
India

Jyoti Saxena
Department of Biochemical Engineering
Bipin Tripathi Kumaon Institute of
Technology
Dwarahat, Uttarakhand
India

ISBN 978-81-322-1553-0 ISBN 978-81-322-1554-7 (eBook)
DOI 10.1007/978-81-322-1554-7
Springer New Delhi Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013946949

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

The twenty-first century has been nicknamed as the era of biotechnology. It has grown and evolved to such an extent over the past few years that increasing numbers of professionals work in areas directly impacted by it. It has been turned into a high science topic to our everyday vocabulary over a short period of time.

It is quite remarkable to note how different branches of biotechnology have emerged to have both substantial academic and industrial impact in the not so distant future. The opportunities become wider and the hopes brighter. Modern biotechnology has opened up many opportunities in various sectors such as agriculture, food, forestry, waste treatment, medicine, and pharmaceutical production. Covering even the most important aspects of biotechnology in a single book that reaches readers ranging from students to active researchers in academia and industry is an enormous challenge. To prepare such a wide-ranging book on biotechnology, editors have harnessed their own knowledge and experience, gained in several departments and universities, and has assembled experts to write chapters covering a wide array of biotechnology topics, including the latest advances. *Advances in Biotechnology* is an important book that provides the information and insight to enable readers to participate in the biotechnology debate. This book is intended to serve both as a textbook for university courses as well as a reference for researchers. It is increasingly important that scientists and engineers, whatever their specialty, have a solid grounding in the fundamentals and potential applications of biotechnology. The editors and their team are to be warmly congratulated for bringing together this exclusive, timely, and useful biotechnology book.

D. S. Chauhan Vice Chancellor
Uttarakhand Technical University, Dehradun

Preface

The twenty-first century looks to Biotechnology as the world's fastest growing and most rapidly changing technology that can improve the human conditions. Modern biotechnology enables an organism to produce a totally new product which the organism does not or cannot produce in its normal course of life. The book *Advances in Biotechnology* is a collection of topics on recent advances in certain ongoing biotechnological applications. Fourteen authoritative chapters on current developments and future trends in biotechnology are empathized. The book aims to cover a wide range of topics under all specialized domains of microbial, plant, animal, and industrial biotechnology.

Chapter 1 provides a detailed account of high capacity vectors used for various applications of genetic engineering. Chapter 2 is devoted to the modern era DNA sequencing dealing with next generation sequencing. Up-to-date methodological approaches such as use of molecular markers (Chap. 3), DNA microarray technology (Chap. 6) and proteomics (Chap. 8) have revolutionized biotechnology with a wide array of applications in studies related to cancer biology, microbiology, plant science, environmental science, etc. Proteomics has recently been of interest to scientists because it gives a better understanding of an organism than genomics.

Chapters 4, 5, 9, 11, and 12 are focused on the crucial role of biotechnology in health care through gene therapy, gene silencing, stem cell technology, monoclonal antibodies, and edible vaccines. Gene therapy is being used for correcting defective genes that are responsible for disease development; RNAi is a valuable research tool not only for functional genomics, but also for gene-specific therapeutic activities. Monoclonal antibodies are widely used for immunodiagnostic, immunotherapy, and in biological and biochemical research. Key aspects of edible vaccines like host plants, mechanism of action, advantages, limitations, and different regulatory issues are contemplated upon.

In today's world where products of microbial origin have proved their utility in almost every sphere of life, metagenomic studies (Chap. 7) have become highly important as they give a clue to the hidden wealth of microbial world. Chapter 10 describes the utilization of biosensors in various industries for monitoring food quality control, medical research, clinical diagnosis, environmental monitoring, agriculture, bioprocesses,

and control. Genes from microbes, plants, and animals are being used successfully to enhance the ability of plants. Though improvement of plants by genetic engineering opens up new possibilities to tolerate, remove, and degrade pollutants, it is still in its research and development phase with many technical issues needing to be addressed as explained in Chap. 13. Finally, in Chap. 14, the great market potential involved for biotechnological companies has been highlighted with suggestions that can be set up for harnessing the vast potential involved in biotech products all over the world.

This book is clearly a team effort, and many thanks are due. The authors of the individual chapters have been chosen for their recognized expertise and their contributions to the various fields of biotechnology. Their willingness to impart their knowledge to their colleagues forms the basis of this book and is gratefully acknowledged. Authors relied on various sources, which are identified in the individual chapters. The authors would also like to thank their colleague Ms. Shweta Ranghar whose help during the preparation of this book was commendable. Thanks are also due to Mr. Vikas Kumar for working on the illustrations. The editors wish to thank their respective head of the institutions/center for their encouragement and providing required ambience.

Moreover, this work would not have been brought to realization without the prudence and the constant and conscientious support of the publisher. We are grateful to Springer for publishing this book with their customary excellence. Finally, special thanks go to our families, who put up with longer hours, helpful suggestions, indispensable help, and encouragement.

May 2013

Indu Ravi
Mamta Baunthiyal
Jyoti Saxena

Contents

Contributors

Pavan Kumar Agrawal Govind Ballabh Pant Engineering College, Ghurdauri, Pauri, Uttarakhand 246194, India, e-mail: p_k_agarwal@rediffmail.com

Rachana Arya Electronics and Communication Engineering Department, Bipin Tripathi Kumaon Institute of Technology, Dwarahat, Uttarakhand 263653, India, e-mail: rachna009@gmail.com

Bhakti Bajpai Biotechnology, ARIBAS, New Vallabh Vidya Nagar, Gujarat 388121, India, e-mail: bbajpai@yahoo.com

Mamta Baunthiyal Govind Ballabh Pant Engineering College, Ghurdauri, Pauri, Uttarakhand 246194, India , e-mail: mamtabaunthiyal@yahoo.co.in

J. P. Bhatt Department of Zoology and Biotechnology, HNB Garhwal University, Srinagar (Garhwal), Uttarakhand, India, e-mail: profjpbhatt@gmail.com

Pradeep Bhatnagar Life Sciences, The IIS University, Mansarovar, Jaipur, India, e-mail: Pradeepbhatnagar1947@yahoo.com

Satpal Singh Bisht Department of Biotechnology, School of Life Sciences, Mizoram University (A Central University), Tanhril, Aizawl, Mizoram 796004, India, e-mail: sps.bisht@gmail.com

Ashwini M. Charpe Plant Pathology Section, College of Agriculture, Dr. Panjabrao Deshmukh Krishi Vidyapeeth, Akola, Maharashtra 444104, India, e-mail: ashwinicharpe@yahoo.com

Sunita Chauhan Kumarappa National Handmade Paper Institute, Jaipur, Rajasthan, India, e-mail: itsneeru@yahoo.com

G. K. Joshi Department of Zoology and Biotechnology, HNB Garhwal University, Srinagar (Garhwal), Uttarakhand, India, e-mail: gkjoshi@rediffmail.com

J. Jugran Department of Zoology and Biotechnology, HNB Garhwal University, Srinagar (Garhwal), Uttarakhand, India, e-mail: jyotijugran28@gmail.com

Shweta Kulshreshtha Amity Institute of Biotechnology, Amity University of Rajasthan, Jaipur, India, e-mail: shweta_kul17@rediffmail.com

B. D. Lakhchaura Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, India, e-mail: lakhchaurabd@rediffmail.com

Mayank Biochemical Engineering Department, Bipin Tripathi Kumaon Institute of Technology, Dwarahat, Uttarakhand 263653, India, e-mail: findmayank12@yahoo.co.in

Amrita Kumari Panda Department of Biotechnology, Roland Institute of Pharmaceutical Sciences, Berhampur, Orissa 760010, India, e-mail: itu.linu@gmail.com

Indu Ravi IGNOU Regional Centre, Jaipur, Rajasthan, India, e-mail: induravi11@yahoo.com

Shweta Rawat Biochemical Engineering Department, Bipin Tripathi Kumaon Institute of Technology, Dwarahat, Uttarakhand 263653, India, e-mail: shweta.biotech24@gmail.com

K. Rohini Unit of Biochemistry, Faculty of Medicine, AIMST University, 3½ Bukit Air Nasi, Jalan Semeling, 08100 Bedong, Kedah Darul Aman, Malaysia, e-mail: rohinik23@gmail.com

Jyoti Saxena Biochemical Engineering Department, Bipin Tripathi Kumaon Institute of Technology, Dwarahat, Uttarakhand 263653, India, e-mail: saxenajyoti30@gmail.com

Rahul Shrivastava Department of Biotechnology, Jaypee University of Information Technology, Waknaghat, DumeharBani, Solan, Himachal Pradesh, India, e-mail: rahulmicro@rediffmail.com

Molecular Markers

3

Pavan Kumar Agrawal and Rahul Shrivastava

Abstract

Detection and analysis of genetic variation help in understanding the molecular basis of various biological phenomena in eukaryotes. Since the entire eukaryotes cannot be covered under sequencing projects, molecular markers and their correlation to phenotypes provide with requisite landmarks for elucidation of genetic variation. There are different types of DNA-based molecular markers. These DNA-based markers are differentiated into two types; first nonPCR-based (RFLP) and second is PCR-based markers (RAPD, AFLP, SSR, SNP etc.). Amongst others, the microsatellite DNA marker has been the most widely used in ecological, evolutionary, taxonomical, phylogenetic, and genetic studies due to its easy use by simple PCR, followed by a denaturing gel electrophoresis for allele size determination and high degree of information provided by its large number of alleles per locus. Despite this, a new marker type, named SNP, for Single Nucleotide Polymorphism, is now on the scene and has gained high popularity, even though it is only a bi-allelic type of marker. Day by day development of such new and specific types of markers makes their importance in understanding the genomic variability and the diversity between the same as well as different species of the plants. In this chapter, we have discussed types of molecular markers, their advantages, disadvantages, and the applications.

3.1 Introduction

The concept of genetic markers is not a new one; Gregor Mendel used phenotype-based genetic markers in his experiment in the nineteenth century. Later, phenotype-based genetic markers for *Drosophila* led to the establishment of the theory of genetic linkage. The limitations of phenotype-based genetic markers directed the development of more general and useful direct

P. K. Agrawal
Govind Ballabh Pant Engineering College, Ghurdauri, Pauri, Uttrakhand 246194, India
e-mail: p_k_agarwal@rediffmail.com

R. Shrivastava (✉)
Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh, India
e-mail: rahulmicro@gmail.com

I. Ravi et al. (eds.), *Advances in Biotechnology*,
DOI: 10.1007/978-81-322-1554-7_3,

DNA-based markers that became known as molecular markers. A molecular marker is defined as a particular segment of DNA that is representative of the differences at the genome level. They may or may not correlate with phenotypic expression of a trait. Molecular markers offer numerous advantages over conventional phenotype-based alternatives as they are stable and detectable in all tissues regardless of growth, differentiation, development, or defense status of the cell. Further, these are not confounded by the environment, pleiotropic and epistatic effects.

Due to the rapid developments in the field of molecular genetics, a variety of techniques have emerged to analyze genetic variation during the last few decades. These genetic markers may differ with respect to important features, such as genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements, and financial investment. No marker is superior to all others for a wide range of applications. The most appropriate genetic marker depends on the specific application, the presumed level of polymorphism, the presence of sufficient technical facilities and knowhow, time constraints, and financial limitations. The classifications of molecular marker technologies that have been widely applied during the last decades are summarized in Table 3.1.

Table 3.1 Some popular molecular markers, their characteristics and potential applications

Techniques	Marker type	Require prior molecular information	Mode of inheritance	Degree of polymorphism	Major applications	Discoverer
Non PCR-based techniques	Restriction fragment length polymorphism	Yes	Mendelian, codominant	Low	Linkage mapping, Physical mapping	Bostein et al. (1980)
PCR-based techniques	Random amplified polymorphic DNA	No	Mendelian, dominant	Intermediate	Fingerprinting for population study, Gene tagging, Hybrid identification	Williams et al. (1990)
	Amplified fragment length polymorphism	No	Mendelian, dominant	High	Linkage mapping, Gene tagging, population study	Vos et al. (1995)
	Microsatellite, SSR	Yes	Mendelian, codominant	High	Linkage mapping, population study, Genetic diversity, Paternity analysis	Litt and Lutty (1989)
	Minisatellite, VNTR	No	Mendelian, codominant	High	DNA fingerprinting for population study	Jeffrey (1985)
	Single nucleotide polymorphism	Yes	Mendelian, codominant	High	Linkage mapping, population study	Ching et al. (2002)

An ideal molecular marker technique should (1) be polymorphic and evenly distributed throughout the genome, (2) provide adequate resolution of genetic differences, (3) generate multiple, independent and reliable markers, (4) be simple, quick, and inexpensive, (5) need small amounts of tissue and DNA samples (6) have linkage to distinct phenotypes and (7) require no prior information about the genome of an organism.

3.2 Molecular Marker Techniques

A vast array of DNA-based genetic markers for the detection of DNA polymorphism have been discovered since 1980 and new marker types are developed every year. There are several DNA marker types, which may be classified broadly into two groups based on their detection method, (1) **Non PCR based** or those based on DNA–DNA hybridization and (2) **PCR based** or those based on amplification of DNA sequences using the polymerase chain reaction (PCR) (Table 3.1).

3.2.1 Non PCR Based or DNA–DNA Hybridization

3.2.1.1 Restriction Fragment Length Polymorphism

Restriction fragment length polymorphism markers were the first DNA-based genetic markers developed (Botstein et al. 1980). Eukaryotic genomes are very large and there was no simple way to observe genetic polymorphisms of individual genes or sequences. The property of complementary base pairing where small piece of DNA could be used as probe allowed for methods to be developed to reveal polymorphisms in sequences homologous to the probe. The genetic system derived using this approach is called restriction fragment length polymorphism. RFLP (commonly pronounced "rif-lip") refers to a difference between two or more samples of homologous DNA molecules arising from differing locations of restriction sites. If two organisms differ in the distance between sites of cleavage of particular *Restriction endonucleases*, the length of the fragments produced will differ when the DNA is digested with a restriction enzyme. The dissimilarity of the patterns generated can be used to differentiate species (and even strains) from one another. In RFLP analysis, the DNA sample broken into pieces (digested) by restriction enzymes that are separated according to their lengths on agarose gel electrophoresis. It is possible to visualize DNA within such a gel by staining it with ethidium bromide, however, due to typically so many restriction fragments of all possible sizes, discrete fragments cannot be seen. To overcome this problem, the fractionated DNA is transferred and chemically bound to a nylon membrane by a process called Southern blotting, named after its inventor E. M. Southern (1975). Specific DNA fragments are visualized by hybridizing the DNA fragments bound to the nylon membrane with a radioactively or fluorescently labeled DNA probe. Different sizes or lengths of restriction fragments are typically produced when different individuals are tested. Such a polymorphism can be used to distinguish plant species, genotypes and, in some cases, individual plants. Labeling of the probe may be performed with a radioactive isotope or with alternative non-radioactive stains, such as digoxigenin or fluorescein. These probes are mostly species-specific single locus probes of about 0.5–3.0 kb in size, obtained from a cDNA library or a genomic library. Though genomic library probes may exhibit greater variability than probes from cDNA libraries, a few studies reveal the converse (Miller and Tanksley 1990). In addition to genetic fingerprinting, RFLP is an important tool in genome mapping, localization of genes for genetic disorders, and determination of risk for disease and paternity testing.

Advantages: The RFLP markers are relatively highly polymorphic, codominantly inherited, and highly reproducible. Because of their presence throughout the eukaryotic genome, high heritability and locus specificity of the RFLP markers are considered superior. The method

also provides opportunity to simultaneously screen numerous samples. DNA blots can be analyzed repeatedly by stripping and reprobing (usually 8 to 10 times) with different RFLP probes.

Disadvantages: The technique is not very widely used because it is time consuming, expensive, and only one out of several markers may be polymorphic, which is highly inconvenient especially for crosses between closely related species. The utility of RFLPs has been hampered due to the large quantities of purified, high molecular weight genomic DNA requirement for DNA digestion as well as Southern blotting. The use of radioactive isotope makes the analysis relatively expensive and hazardous. The requirement of prior sequence information for probe generation increases the complexity of the methodology. Their inability to detect single base changes restricts their use in detecting point mutations occurring within the regions at which they are detecting polymorphism.

Applications: Restriction Fragment length polymorphism is a powerful tool for the identification of organisms to the level of species, strains, varieties, or individuals. RFLPs have been widely used in population genetics to determine slight differences within populations, and in linkage studies to generate maps for quantitative trait locus (QTL) analysis. They also have been used to investigate relationships of closely related taxa or intraspecific level, as fingerprinting tools for diversity studies, and for studies of hybridization. They have been used in criminal and paternity tests, localization of genes for genetic disorders, and determination of risk for disease.

3.2.2 PCR-Based Marker

After the invention of polymerase chain reaction (PCR) technology (Mullis and Faloona 1987), a large number of approaches for generation of molecular markers based on PCR were detailed, primarily due to its apparent simplicity and high probability of success. Before PCR, the analysis of a specific DNA fragment generally required cloning of the fragment and amplification in a plasmid or compatible vector. PCR enables the production of a large amount of specific DNA sequence without cloning, starting with just a few molecules of the target sequence. Usage of random primers overcame the limitation of prior sequence knowledge for PCR analysis and facilitated the development of genetic markers for a variety of purposes.

3.2.2.1 Amplified Fragment Length Polymorphism

Amplified fragment length polymorphism PCR (or AFLP-PCR or just AFLP) is a PCR-based tool used in genetics research, DNA fingerprinting and in the practice of genetic engineering. Developed by Vos et al. (1995), AFLP is essentially an intermediate between RFLP and PCR. AFLP selectively amplifies a subset of restriction fragments from a complex mixture of DNA fragments obtained after digestion of genomic DNA with restriction endonucleases. Polymorphisms are detected from differences in the length of the amplified fragments by polyacrylamide gel electrophoresis (PAGE) or by capillary electrophoresis. AFLP uses restriction enzymes to cut genomic DNA, followed by ligation of adaptors to the sticky ends of the restriction fragments. A subset of the restriction fragments are then amplified using primers complementary to the adaptor and part of the restriction site fragments. The amplified fragments are visualized on denaturing polyacrylamide gels either through autoradiography or fluorescence methodologies (Vos et al. 1995). AFLPs are DNA fragments (80–500 bp) obtained from digestion with restriction enzymes, followed by ligation of oligonucleotide adapters to the digestion products and selective amplification by the PCR. AFLPs, therefore, involve both RFLP and PCR. The AFLP banding profiles are the result of variations in the restriction sites or in the intervening region. The AFLP technique simultaneously generates fragments from many genomic sites (usually 50–100 fragments per reaction) that are separated by polyacrylamide gel electrophoresis

and that are generally scored as dominant markers.

Advantages: There are many advantages of AFLP, they are produced in greater amount and have higher reproducibility, resolution, and sensitivity at the whole genome level and also has the capability to amplify between 50 and 100 fragments at one time. In addition, no prior sequence information is needed for amplification. As a result, AFLP has become extremely beneficial in the study of taxa including bacteria, fungi, and plants, where much is still unknown about the genomic makeup of various organisms. AFLPs can be analyzed on automatic sequencers, but software problems concerning the scoring of AFLPs are encountered on some systems. The use of AFLP in genetic marker technologies has become the main tool due to its capability to disclose a high number of polymorphic markers by single reaction (Vos et al. 1995).

Disadvantages: Disadvantages of this technique are that alleles are not easily recognized, have medium reproducibility, labor intensive, and have high operational and developmental costs. AFLP needs purified, high molecular weight DNA, the dominance of alleles and the possible non-homology of co-migrating fragments belonging to different loci.

Applications: Most AFLP fragments correspond to unique positions on the genome and have the capability to detect various polymorphisms in different genomic regions simultaneously and hence can be exploited as landmarks in genetic and physical mapping. The technique can be used to distinguish closely related individuals at the subspecies level (Althoff et al. 2007) and also for gene mapping. Applications for AFLP in plant mapping include establishing linkage groups in crosses, saturating regions with markers for gene landing efforts (Yin et al. 1999), and assessing the degree of relatedness or variability among cultivars (Mian et al. 2002). Molecular markers are more reliable for genetic studies than morphological characteristics, because the environment does not affect them. AFLP is considered more applicable to intraspecific than to interspecific studies due to frequent null alleles. AFLP markers are useful in genetic studies, such as biodiversity evaluation, analysis of germplasm collections, genotyping of individuals, and genetic distance analysis. The availability of many different restriction enzymes and corresponding primer combinations provides a great deal of flexibility, enabling the direct manipulation of AFLP fragment generation for defined applications (e.g., polymorphism screening, QTL analysis, and genetic mapping).

3.2.2.2 Random Amplified Polymorphic DNA

The basis of randomly amplified polymorphic DNA technique is differential PCR amplification of genomic DNA. It deduces DNA polymorphisms produced by "rearrangements or deletions at or between oligonucleotide primer binding sites in the genome" using short random oligonucleotide sequences (mostly 10 bases long) (Williams et al. 1990). As the approach requires no prior knowledge of the genome that is being analyzed, it can be employed across species using universal primers. In the RAPD marker system, a PCR reaction is conducted using a very small amount of template DNA (even less than 10 ng is sufficient) and a single RAPD primer. Primers are usually just 10 base pairs long (10 mers) and are of random sequence. The analysis of RAPD is based on the PCR using short (about 10 bases) randomly chosen primers singly which anneal as reverted repeats to the complementary sites in the genome. The DNA between the two opposite sites with the primers as starting and end points is amplified by PCR. The amplification products are separated on agarose gels in the presence of ethidium bromide and viewed under ultraviolet light. The banding patterns distinguish organisms according to the presence or absence of bands (polymorphism). These polymorphisms are considered to be primarily due to variation in the primer annealing sites, but they can also be generated by length differences in the amplified sequence between primer annealing sites. Each product is derived from a region of the genome

that contains two short segments in inverted orientation, on opposite strands that are complementary to the primer. It is a peculiarity of RAPD analysis that it discriminates at different taxonomical level, viz., isolates and species, depending on the organism investigated and the primer used

Advantages: The main advantage of RAPDs is that they are quick and easy to assay. Because PCR is involved, only low quantities of template DNA are required, usually 5–50 ng per reaction. Since random primers are commercially available, no sequence data for primer construction are needed. Moreover, RAPDs have a very high genomic abundance and are randomly distributed throughout the genome. They are dominant markers; hence have limitations in their use as markers for mapping, which can be overcome to some extent by selecting those markers that are linked in coupling (Williams et al. 1993). RAPD assay has been used by several groups as efficient tools for identification of markers linked to agronomically important traits, which are introgressed during the development of near isogenic lines. The RAPD analysis of NILs (nonisogenic lines) has been successful in identifying markers linked to disease resistance genes in tomato (*Lycopersicon* sp.), lettuce (*Lactuca* sp.), and common bean (*Phaseolus vulgaris*). Due to the speed and efficiency of RAPD analysis, high-density genetic mapping in many plant species was developed in a relatively short time.

Disadvantages: The major drawback of the method is that the profiling is dependent on the reaction conditions so it may vary within two different laboratories and as several discrete loci in the genome are amplified by each primer, profiles are not able to distinguish heterozygous from homozygous individuals (Bardakci 2001).

Applications: The application of RAPDs and their related modified markers in variability analysis and individual-specific genotyping has largely been carried out, but is less popular due to problems such as poor reproducibility, faint or fuzzy products, and difficulty in scoring bands, which lead to inappropriate inferences. RAPDs have been used for many purposes, ranging from studies at the individual level (e.g., genetic identity) to studies involving closely related species and determination of genetic diversity. RAPDs have also been applied in gene mapping studies to fill gaps not covered by other markers.

3.2.2.3 Sequence Characterized Amplified Region

Sequence characterized amplified regions (SCARs) markers were developed by Michelmore et al. (1991) and Martin et al. (1991). SCARs are DNA fragments amplified by the use of PCR using specific 20–30 bp primers. The primers are designed from terminal ends of a RAPD marker. RAPD marker fragments associated with a phenotypic condition of interest are cloned and sequenced. This nucleotide sequence is then used for designing of unique primers for specific amplification of a particular locus. Use of unique primers decreases site competition for specific regions among primers, making the results less sensitive to reaction conditions. Use of longer primers also increases the reproducibility of the results by increasing specificity of template binding by the primer. PCR products obtained after amplification are analyzed for the presence of length polymorphism by gel electrophoresis. Conversion of RAPDs into SCARs may have additional advantage of obtaining a codominant marker, although dominance may be exhibited by SCARs when one or both primers partially overlap the site of sequence variation.

Advantages: Due to the application of PCR, low amounts of genomic or template DNA is required (10–50 ng per reaction) for analysis of the marker. The major advantage of SCARs is the ease of their use, and requirement of relatively less time for the analysis. SCAR markers are locus specific and the results obtained from analysis of SCARs are highly reproducible.

Disadvantages: Major limitation lies in the fact that analysis of SCARs can be done only for organism, species, or DNA fragment with known sequence; as knowledge of sequence data is required for designing of the PCR primers.

Applications: Major application of SCAR makers has been found in gene mapping studies as they are locus specific. They are also used in

marker assisted selection, where presence of a specific genotype and expression of its corresponding phenotype can be correlated to the presence or absence of a SCAR marker (Paran and Michelmore 1993).

3.2.2.4 Minisatellites, Variable Number of Tandem Repeats

The term 'minisatellite' was introduced by Jeffrey et al. (1985). These loci contain tandem repeats that vary in the number of repeat units between genotypes and are referred to as variable number of tandem repeats (VNTRs) (i.e., a single locus that contains variable number of tandem repeats between individuals) or hypervariable regions (HVRs) (i.e., numerous loci containing tandem repeats within a genome generating high levels of polymorphism between individuals). Minisatellites are a conceptually very different class of marker. They consist of chromosomal regions containing tandem repeat units of a 10–50 base motif, flanked by conserved DNA restriction sites. A minisatellite profile consisting of many bands, usually within a 4–20 kb size range, is generated by using common multi locus probes that are able to hybridize to minisatellite sequences in different species. Locus specific probes can be developed by molecular cloning of DNA restriction fragments, subsequent screening with a multi locus minisatellite probe and isolation of specific fragments. Variation in the number of repeat units, due to unequal crossing over or gene conversion, is considered to be the main cause of length polymorphisms. Due to the high mutation rate of minisatellites, the level of polymorphism is substantial, generally resulting in unique multilocus profiles for different individuals within a population.

Advantages: The main advantages of minisatellites are their high level of polymorphism and high reproducibility.

Disadvantages: Disadvantages of minisatellites are similar to RFLPs due to the high similarity in methodological procedures. If multilocus probes are used, highly informative profiles are generally observed due to the generation of many informative bands per reaction. In that case, band profiles cannot be interpreted in terms of loci and alleles and similar sized fragments may be nonhomologous. In addition, the random distribution of minisatellites across the genome has been questioned (Schlötterer 2004).

Applications: The term DNA fingerprinting was introduced for minisatellites, though DNA fingerprinting is now used in a more general way to refer to a DNA-based assay to uniquely identify individuals. Minisatellites are particularly useful in studies involving genetic identity, parentage, clonal growth and structure, and identification of varieties and cultivars and for population-level studies. Minisatellites are of reduced value for taxonomic studies because of hyper variability.

3.2.2.5 Simple Sequence Repeat or Microsatellites

The simple sequence repeats (SSR) are also referred to as microsatellites. 'Microsatellites', coined by Litt and Lutty (1989), consist of tandemly repeating units of DNA ubiquitous in prokaryotes and eukaryotes, scattered throughout most eukaryotic genomes (Powell et al. 1996). Microsatellites can comprise repetition of mono-, di-, tri-, tetra- or penta-nucleotide units. The variation in number of repeat units of a microsatellite resulting in length polymorphisms is mainly attributed to Polymerase slippage during DNA replication, or slipped strand mispairing. Function of such microsatellites is largely unknown, though they can occur in protein-coding as well as noncoding regions of the genome. Microsatellite sequences are especially exploited to differentiate closely related species or genotypes. Due to the high degree of variability present in microsatellites, they are preferred in population studies (Smith and Devey 1994) and to distinguish closely related plant cultivars.

Simple sequence repeat polymorphism is detected by the use of PCR, in conditions when nucleotide sequences of the flanking regions of the microsatellite are known, or by Southern

hybridization using labeled probes. Unlike minisatellites, which also represent tandem repeats, microsatellites consist of short tandem repeat motifs of 1–6 base pair (bp) long units. Knowledge of nucleotide sequence is used to design specific primers (generally 20–25 bp) for amplification of the microsatellite region along with flanking sequences by PCR. Such amplified amplicon is then identified by construction of a small-insert genomic library, followed by screening the library with a synthetically labeled oligonucleotide probe. Positive clones thus obtained are sequenced for detection and confirmation of polymorphism.

Microsatellite polymorphism of organism whose genome is not sequenced may also be identified by screening sequence databases of closely related species for microsatellite sequence motifs. Primers designed from flanking regions after such comparative study are then used to amplify the SSR region followed by analysis.

Advantages: High number of alleles and their co-dominance, along with widespread distribution throughout the genome, are the most significant features of microsatellites which make them important molecular markers. The SSRs are present in high genomic abundance in eukaryotic organisms and are mostly localized in low-copy regions (Morgante et al. 2002). As the analysis of SSRs is PCR based, low quantities of template DNA (10–50 ng per reaction) are required. In comparison to other markers, such as RAPD, longer primers can be designed and used for microsatellites, thus, increasing reproducibility of the results obtained. The analysis of microsatellite can be semi-automated and does not necessitate the use of radioactive isotopes.

Microsatellite analysis allows identification of multiple alleles at a single locus. Although it is done mostly for single locus on the genome, in conditions where the size ranges of the alleles on different loci are markedly different, gel electrophoresis analysis or multiplex PCR can easily be used to simultaneously study multiple microsatellites (Ghislain et al. 2004). This would help in decreasing the analytical cost significantly.

Individual microsatellite loci can be converted into PCR-based markers for analysis as monolocus, codominant SSRs. In conditions where sequence of the microsatellite region is unknown, primers are designed from genomic regions adjacent to the SSR sequences. The amplified genomic fragment thus obtained is cloned and sequenced to obtain specific sequence of the microsatellite region. Primers are designed using this microsatellite sequence for specific amplification using PCR. This approach is termed as sequence-tagged microsatellite site (STMS).

Disadvantages: Principal drawback of microsatellites lies in their high development costs. Sequence information of the template or genomic DNA is required for designing and synthesis of primers. Thus, study or analysis of microsatellites is extremely difficult to apply for any unstudied group of species or organism of interest. Although SSRs are codominant in nature, mutations in the primer binding sites may result in false-negative result because of nonamplification of the intended PCR product (occurrence of null alleles), which may lead to errors in genotype analysis.

Such null alleles may lead to underestimation of heterozygosity and biased estimate of allelic and genotypic frequencies. Similarly, underestimation of genetic divergence may occur due to homoplasy at microsatellite loci, which is a result of different forward and backward mutations.

Interpretation of bands obtained on electrophoretic gel, after PCR amplification of microsatellite markers may be difficult sometimes. Appearances of artifacts in form of stutter bands have been seen to influence proper visualization and size determination of the fragments, thus interpretation of bands. Formation of such bands occurs due to DNA slippage during PCR amplification. Appropriate reference genotypes of known DNA fragments of specific band sizes can be run along with test samples to overcome the complications in interpretation of such results.

Applications: Microsatellites have high degree of mutability, hence are thought to play a

significant role in genome evolution by creation and maintenance of quantitative genetic variation. Thus, high level of polymorphism shown by microsatellites makes them informative markers for population genetics studies, ranging from the individual level (e.g., strain identification) to that of closely related species level. Conversely, due to their high degree of mutability, microsatellites markers are not suitable for analysis and correlation studies of higher taxonomic groups.

SSR markers have been useful for genetic variation studies in germplasm collections and are considered to be ideal molecular markers for gene mapping studies (Hearne et al. 1992; Jarne and Lagoda 1996). SSRs are molecular markers used for recombination mapping studies, to verify parental relationships and population genetic studies. They are the only molecular marker to provide information about presence or absence of closely related alleles. Increase or decrease in the number of SSR repeats in genes of known function can be correlated with changes in biological functions or phenotypic variation in the organism (Ayers et al. 1997). SSRs have been shown to be useful for study of functional diversity among closely related species, in relation to adaptive changes (Eujay et al. 2001).

3.2.2.6 Inter-Simple Sequence Repeat

AFLP involves high cost of analysis, while RAPD is low in reproducibility; analysis of SSR markers also requires prior knowledge of DNA sequence of flanking regions for designing of species-specific primers. Thus, analysis of polymorphism by application of any of these techniques has its limitations, which pose major obstacle in regular use of these markers. These limitations can be overcome to some extent by the use of ISSR-PCR technique.

SSRs or microsatellites are short segments of DNA (1–4 base pairs) universally present in eukaryotic genomes. The number of repeats of SSRs varies in different organisms of same or different species amounting to polymorphism. ISSR is a PCR-based technique, in which SSRs are used as primers to primarily amplify the inter-SSR regions. The method involves PCR amplification of DNA fragments present at a suitable amplifiable distance in between two identical microsatellite regions. The ISSRs are regions between adjacent, oppositely oriented microsatellite sequences.

Primers are designed by using microsatellite repeat regions. As the repeat sequences flanking an inter-SSR region are same, a single primer of 17–25 bp length is generally used both as forward and reverse primers. The primers used for amplification of ISSRs may consist of di-, tri-, tetra-, or penta-nucleotide repeat units.

The designed primers are usually anchored at 3′ or 5′ end with 1–4 bases extended into the nonrepeat adjacent regions of SSRs (Zietkiewicz et al. 1994). Unanchored primers can also be used for amplification of ISSRs.

Using the primer, about 10–50 inter- SSR sequences of different sizes (about 200–3,000 bp) are amplified simultaneously from multiple genomic loci by a single PCR reaction. These amplicons are separated by gel electrophoresis and analyzed for the presence or absence of DNA fragments of particular length.

Some methods related to analysis of ISSRs are:

1. Single primer amplification reaction (SPAR)—that uses a single primer designed only from the core motif of microsatellite region (unanchored primer), and
2. Directed amplification of minisatellite region DNA (DAMD)—that employs a single primer designed only from the core motif of a minisatellite.

Advantages: Principle advantage of ISSRs is that no sequence data of the fragment to be amplified is required for designing of the primers, as primers are designed from the repeat units of flanking microsatellite region. The analysis requires very small quantities of template DNA (10–50 ng per reaction). ISSRs are mostly exhibited as dominant marker, though it might also be present in codominant state. In comparison to RAPD, which is also a PCR-based technique using a single primer, use of longer primers and corresponding higher annealing

temperature (45–60 °C) leads to higher stringency of PCR amplification. Thus, better reproducibility in results is obtained from analysis of ISSRs as molecular markers.

Disadvantages: As ISSR is a multilocus technique, fragments with the same size and mobility originating from nonhomologous regions, may lead to misinterpretation of genetic similarity estimates. Further, under certain conditions ISSRs can also show low reproducibility of results similar to RAPDs.

Applications: ISSR analysis leads to multilocus fingerprinting profiles, thus it is useful in areas of study of genetic diversity, construction and study of phylogenetic tree, and evolutionary biology studies in a wide range of species. ISSR analysis can be applied in studies involving parentage establishment, clone and strain identification, and taxonomic studies of closely related species. ISSRs are also considered useful in genome mapping studies. Though ISSRs are present mostly as dominant markers, in some cases they may also segregate as codominant markers enabling distinction between homozygotes and heterozygotes.

3.2.2.7 Cleaved Amplified Polymorphic Sequence

Similar to RFLP, CAPS is based on the principle that any change or mutation in the genome of an individual could lead to creation or deletion of restriction sites. CAPS markers are generally developed by sequencing of RFLP probes used for hybridization. Moreover, 20–25 bp specific primers are designed for amplification of 800–2,000 bp DNA fragments. To increase the chance of finding a region containing polymorphism introns or 3′ untranslated regions are generally used for the amplification. Sequences of DNA from target genotypes are amplified using PCR and the resulting amplicons are digested individually with one or more restriction enzymes. The digested products after gel electrophoresis show readily distinguishable pattern. Thus, length polymorphisms resulting from variation in the occurrence of restriction sites are identified. CAPS have also been referred to as PCR-Restriction fragment length polymorphism (PCR–RFLP).

Advantages: Being a PCR-based technique, analysis of CAPS marker requires very low quantities of template DNA (50–100 ng per reaction). Compared to RFLPs, analysis of CAPS genotypes does not require the technically demanding and time taking steps of Southern hybridization and use of radioactive isotopes. CAPS markers are codominant in nature and locus specific. Most CAPS markers are easily scored and interpreted. Results obtained are highly reproducible.

Disadvantages: Sequence data of the genome to be analyzed are required for synthesis of primers. A limited size of DNA fragment can be amplified by PCR reaction, which would be used for restriction digestion. Hence, in comparison to RFLP analysis, specific regions with CAPS polymorphisms are more difficult to locate.

Applications: CAPS markers have been applied predominantly in gene mapping studies, also helpful in identification of mutation in the genome of an organism.

3.2.2.8 Single Nucleotide Polymorphism

A novel class of DNA markers namely single nucleotide polymorphism in genome (SNPs) has recently become highly applicable in genomic studies. SNP describes polymorphisms caused by point mutations that give rise to different alleles containing alternative bases at a given nucleotide position within a locus. Such sequence differences due to base substitutions have been well characterized since the beginning of DNA sequencing in 1977, but the ability to genotype SNPs rapidly in large numbers of samples was not possible until the application of gene chip technology in the late 1990s. SNPs are again becoming a focal point in molecular marker development since they are the most abundant polymorphism in any organism, adaptable to automation, and reveal hidden polymorphism not detected with other markers and methods. Theoretically, a SNP within a locus can produce as many as four alleles, each

containing one of four bases at the SNP site: A, T, C, and G. Practically, however, most SNPs are usually restricted to one of two alleles (most often either the two pyrimidines C/T or the two purines A/G) and have been regarded as bi-allelic.

Analytical procedures require sequence information for the design of allele specific PCR primers or oligonucleotide probes. SNPs and flanking sequences can be found by library construction and sequencing or through the screening of readily available sequence databases. Once the location of SNPs is identified and appropriate primers designed, one of the advantages they offer is the possibility of high-throughput automation. To achieve high sample throughput, multiplex PCR and hybridization to oligonucleotide microarrays or analysis on automated sequencers are often used to interrogate the presence of SNPs. SNP analysis may be useful for cultivar discrimination in crops where it is difficult to find polymorphisms, such as in the cultivated tomato. SNPs may also be used to saturate linkage maps in order to locate relevant traits in the genome. For instance, in *Arabidopsis thaliana* a high density linkage map for easy to score DNA markers was lacking until SNPs became available (Cho et al. 2000; Ching et al. 2002). To date, SNP markers are not yet routinely applied in gene banks, in particular because of the high costs involved.

SNP markers are inherited as codominant markers. Several approaches have been used for SNP discovery including single-strand conformational polymorphism assays analysis, heteroduplex analysis, and direct DNA sequencing. DNA sequencing has been the most accurate and most used approach for SNP discovery. Random shotgun sequencing, amplicon sequencing using PCR, and comparative EST analysis are among the most popular sequencing methods for SNP discovery.

Advantages: Despite technological advances, SNP genotyping is still a challenging endeavor and requires specialized equipment. Traditional methods available for SNP genotyping include: direct sequencing, single base sequencing, allele specific oligonucleotide, denaturing gradient gel electrophoresis (DGGE), SSCP, and ligation chain reaction (LCR). Each approach has its advantages and limitations, but all are still useful for SNP genotyping, especially in small laboratories limited by budget and labor constraints. Large-scale analysis of SNP markers, however, depends on the availability of expensive, cutting-edge equipment. Several options are available for efficient genotyping using state-of-the-art equipment.

Particularly, popular are methods involving matrix-assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometry, pyrosequencing, Taqman allelic discrimination, real-time (quantitative) PCR, and the use of microarray or gene chips. Mass spectrometry and microarray technologies require a large investment in equipment.

Applications: In plants, SNPs are rapidly replacing simple sequence repeats as the DNA marker of choice for applications in plant breeding and genetics because they are more abundant, stable, amenable to automation, efficient, and increasingly cost-effective (Duran et al. 2009). Generally, SNPs are the most abundant form of genetic variation in eukaryotic genomes. Moreover, they occur in both coding and noncoding regions of nuclear and plastid DNA. As in the case of human genome, SNP-based resources are being developed and made publicly available for broad application in rice research. These resources include large SNP datasets, tools for identifying informative SNPs for targeted applications, and a suite of custom-designed SNP assays for use in marker assisted and genomic selection. SNPs are widely used in breeding programs for several applications such as, (1) marker assisted and genomic selection, (2) association and QTL mapping, positional cloning, (3) haplotype and pedigree analysis, (4) seed purity testing, (5) variety identification, and (6) monitoring the combinations of alleles that perform well in target environments (Kim et al. 2010).

The advantage and disadvantage of common molecular marker techniques are summarized in Table 3.2.

Table 3.2 Advantages and disadvantage of commonly used genetic markers

Marker	Advantages	Disadvantages
Amplified fragment length polymorphism	Sequence information not required	Very tricky due to changes in patterns with respect to materials used
	Can be used across species	
	Work with smaller RFLP fragments	Low reproducibility
	Useful in preparing contig maps	Need to have very good primers
	High genomic abundance	
	High polymorphism	
Restriction fragment length polymorphism	Sequence information not required	Large amount of good quality DNA required
	High genomic abundance	
	Codominant markers	
	High reproducibility	Labour intensive (in comparison to RAPD)
	Can use filters many times	
	Good genome coverage	Difficult to automate
	Can be used across species	Need radioactive labeling
	Needed for map-based cloning	Cloning and characterization of probe are required
Randomly amplified polymorphic DNA	Sequence information not required	No probe or primer information
		Dominant markers
		Very low reproducibility
	Good genome coverage	Cannot be used across species
	Ideal for automation	Not well established
	Less amount and poor quality DNA acceptable	
	No radioactive labeling	
	Relatively faster	
	High genomic abundance	

3.3 Array-Based Platforms

Several different types of molecular markers have been developed over the past three decades (Gupta and Rustgi 2004), motivated by requirements for increased throughput, decreased cost per data point, and greater map resolution. Recently, oligonucleotide-based gene expression microarrays have been used to identify DNA sequence polymorphisms using genomic DNA as the target.

3.3.1 Diversity Arrays Technology

Diversity arrays technology (DArT) is microarray hybridization-based technique that permits simultaneous screening of thousands of polymorphic loci without any prior sequence information. The DArT methodology offers a high multiplexing level, being able to simultaneously type several thousand loci per assay, while being independent of sequence information. DArT assays generate whole genome fingerprints by scoring the presence versus absence of DNA fragments in genomic representations generated from genomic DNA samples through the process of complexity reduction. DArT has been developed as a hybridization-based alternative to the majority of gel-based marker technologies currently in use. It can provide hundreds to tens of thousands of highly reliable markers for any species as it does not require any precise information about the genome sequence

(Jaccoud et al. 2001). Moreover, DArT was found to provide good genome coverage in wheat and barley (Akbari et al. 2006). An important step of this technology is a step called "genome complexity reduction" which increases genomic representation by reducing repetitive sequences that are abundant in eukaryotes. With DArT platform, comprehensive genome profiles are becoming affordable for virtually any crop, genome profiles which can be used in management of bio-diversity, for example in germplasm collections. DArT genome profiles enable breeders to map QTL in 1 week.

DArT profiles accelerate the introgression of a selected genomic region into an elite genetic background (for example, by marker assisted backcrossing). In addition, DArT profiles can be used to guide the assembly of many different regions into improved varieties (marker assisted breeding). The number of markers DArT detects is determined primarily by the level of DNA sequence variation in the material subjected to analysis and by the complexity reduction method deployed. Another advantage of DArT markers is that their sequence is easily accessible compared to amplified fragment length polymorphisms making DArT a method of choice for non-model species (James et al. 2008). DArT has also been applied to a number of animal species and microorganisms.

3.3.2 Restriction Site-Associated DNA

Another high-throughput method is restriction site-associated DNA (RAD) procedure which involves digestion of DNA with a particular restriction enzyme, ligating biotinylated adapters to the overhangs, randomly shearing the DNA into fragments much smaller than the average distance between restriction sites, and isolating the biotinylated fragments using streptavidin beads (Miller et al. 2007a). RAD specifically isolates DNA tags directly flanking the restriction sites of a particular restriction enzyme throughout the genome. More recently, the RAD tag isolation procedure has been modified for use with high-throughput sequencing on the Illumina platform. In addition, Miller et al. (2007b) demonstrated that RAD markers, using microarray platform, allowed high-throughput, high-resolution genotyping in both model and non-model systems.

3.3.3 Single Feature Polymorphism

A third high-throughput method is single feature polymorphism (SFP) which is done by labeling genomic DNA (target) and hybridizing to arrayed oligonucleotide probes that are complementary to insert-delete (INDEL) loci. The SFPs can be discovered through sequence alignments or by hybridization of genomic DNA with whole genome microarrays. Each SFP is scored by the presence or absence of a hybridization signal with its corresponding oligonucleotide probe on the array. Both spotted oligonucleotides and Affymetrix-type arrays have been used in SFP. Borevitz et al. (2003) coined the term "single feature polymorphism" and demonstrated that this approach can be applied to organisms with somewhat larger genomes, specifically *A. thaliana* with a genome size of 140 Mb. Similarly, whole genome DNA-based SFP detection has been accomplished in rice (Kumar et al. 2007), with a genome size of 440 Mb, barley, which has a 5,300 Mb genome composed of more than 90 % repetitive DNA (Cui et al. 2005). Thus, SFPs have become an attractive marker system for various applications including parental polymorphism discovery. The development of DNA-based technologies such as SFP, DArT, and RAD which are based on microarray have the merits of SNP without going through sequencing. These technologies have provided us platforms for medium- to ultra-high-throughput genotyping to discover regions of the genome at a low cost, and have been shown to be particularly useful for genomes, where the level of polymorphism is low (Gupta et al. 2008). These array-based technologies are expected to play an important role in crop improvement and will be used for a variety of

studies including the development of high-density molecular maps, which may then be used for QTL interval mapping and for functional and evolutionary studies.

3.4 Conclusion

DNA-based molecular markers have acted as versatile tools and have found their own position in various fields like taxonomy, physiology, embryology, genetic engineering, etc. They are no longer looked upon as simple DNA fingerprinting markers in variability studies or as mere forensic tools. Ever since their development, they are constantly being modified to enhance their utility and to bring about automation in the process of genome analysis. The discovery of PCR was a landmark in this effort and proved to be a unique process that brought about a new class of DNA profiling markers. This facilitated the development of marker based gene tags, map-based cloning of agronomically important genes, variability studies, phylogenetic analysis, synteny mapping, marker assisted selection of desirable genotypes, etc. Thus, molecular marker provides new dimensions to concerted efforts of breeding and marker aided selection that can reduce the time span of developing new and better varieties and will make the dream of super varieties come true. These DNA markers offer several advantages over traditional phenotypic markers, as they provide data that can be analyzed objectively. With the advent of molecular markers, it is now possible to make direct inferences about genetic diversity and interrelationships among organisms at the DNA level without the confounding effects of the environment and/or faulty pedigree records. Genetic analyses of plant and animal populations and species for taxonomic, evolutionary, and ecological studies tremendously benefited from the development of various molecular marker techniques. Each molecular marker technique is based on different principles but their application is to bring out the genome-wide variability.

References

Akbari M, Wenzl P, Caig V, Carling J, Xia L, Yang S, Uszynski G, Mohler V, Lehmensiek A, Kuchel H, Hayden M, Howes N, Sharp P, Vaughan P, Rathmell B, Huttner E, Kilian A (2006) Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theor Appl Genet 113:1409–1420

Althoff DM, Gitzendanner MA, Segraves KA (2007) The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. Syst Biol 56:477–484

Ayers NM, McClung AM, Larkin PD, Bligh HFJ, Jones CA, Park WD (1997) Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. Theor Appl Genet 94:773–781

Bardakci F (2001) Random amplified polymorphic DNA (RAPD) markers. Turk J Biol 25:185–196

Borevitz JO, Liang D, Plouffe D, Chang HS, Zhu T, Weigel D, Berry CC, Winzeler E, Chory J (2003) Large-scale identification of single-feature polymorphisms in complex genomes. Genome Res 13:513–523

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32(314):331

Ching ADA, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski A (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet 3:19

Cho YG, Shii T, Temnyk S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (Oryza sativa L.). Theoretical and Applied Genetics 100:713–722

Cui XP, Xu J, Asghar R, Condamine P, Svensson JT, Wanamaker S, Stein N, Roose M, Close TJ (2005) Detecting single-feature polymorphisms using oligonucleotide arrays and robustified projection pursuit. Bioinformatics 21:3852–3858

Duran CN, Appleby T, Clark D, Wood M, Batley IJ, Edwards D (2009) Auto SNPdb: an annotated single nucleotide polymorphism database for crop plants. Nucleic Acids Res 37:D951–D953

Eujay I, Sorrells M, Baum M, Woltersand P, Powell W (2001) Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. Euphytica 119:39–43

Ghislain M, Spooner DM, Rodríguez F, Villamon F, Núñez C, Vásquez C, Bonierbale M (2004) Selection of highly informative and user-friendly microsatellites (SSRs) for genotyping of cultivated potato. Theor Appl Genet 108:881–890

Gupta PK, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. Funct Integr Genomics 4:139–162

Gupta PK, Rustgi S, Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. Heredity 101:5–18

Hearne CM, Ghosh S, Todd JA (1992) Microsatellites for linkage analysis of genetic traits. Trend Genet 8:288–294

Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res 29:e25

James KE, Schneider H, Ansell SW, Evers M, Robba L, Uszynski G, Pedersen N, Newton AE, Russell SJ, Vogel JC, Kilian A (2008) Diversity arrays technology (DArT) for pan-genomic evolutionary studies of non-model organisms. PLoS ONE 3:e1682

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. Trends Ecol Evol 11:424–429

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. Nature 314:67–73

Kim JS, Ahn SG, Kim CK, Shim CK (2010) Screening of rice blast resistance genes from aromatic rice germplasms with SNP markers. Plant Pathol J 26:70–79

Kumar R, Qiu J, Joshi T, Valliyodan B, Xu D, Nguyen HT (2007) Single feature polymorphism discovery in rice. PLoS ONE 2(3):e284

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. Amr J Hum Genet 44:397–401

Martin GB, Williams JGK, Tanksley SD (1991) Rapid identification of markers linked to a Pseudomonas resistance gene in tomato by using random primers and near-isogenic lines. Proc Natl Acad Sci 88:2336–2340

Mian MAR, Hopkins AA, Zwonitzer JC (2002) Determination of genetic diversity in tall fescue with AFLP markers. Crop Sci 42:944–950

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor Appl Genet 80:437–448

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007a) Rapid and cost effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res 17(2):240–248

Miller MR, Atwood TS, Eames BF, Eberhart JK, Yan YL, Postlethwait JH, Johnson EA (2007b) RAD marker microarrays enable rapid mapping of zebrafish mutations. Genome Biol 8(6):R105

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially present with non-repetitive DNA in plant genomes. Nat Genet 30:194–200

Mullis KB, Faloona F (1987) Specific synthesis of DNA *in vitro* via polymerase chain reaction. Methods Enzymol 155:350–355

Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985–993

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Tren Plant Sci 1:215–222

Schlötterer C (2004) The evolution of molecular markers—just a matter of fashion? Nat Rev Genet 5:63–69

Smith DN, Devey ME (1994) Occurrence and inheritance of microsatellites in *Pinus radia*. Genome 37:977–983

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Williams JGK, Hanafey MK, Rafalski JA, Tingey SV (1993) Genetic analysis using random amplified polymorphic DNA markers. Meth Enzymol 218:705–740

Yin X, Stam P, Dourleijn CJ, Kropff MJ (1999) AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. Theor Appl Genet 99:244–253

Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR) anchored polymerase chain reaction amplification. Genomics 20:176–183