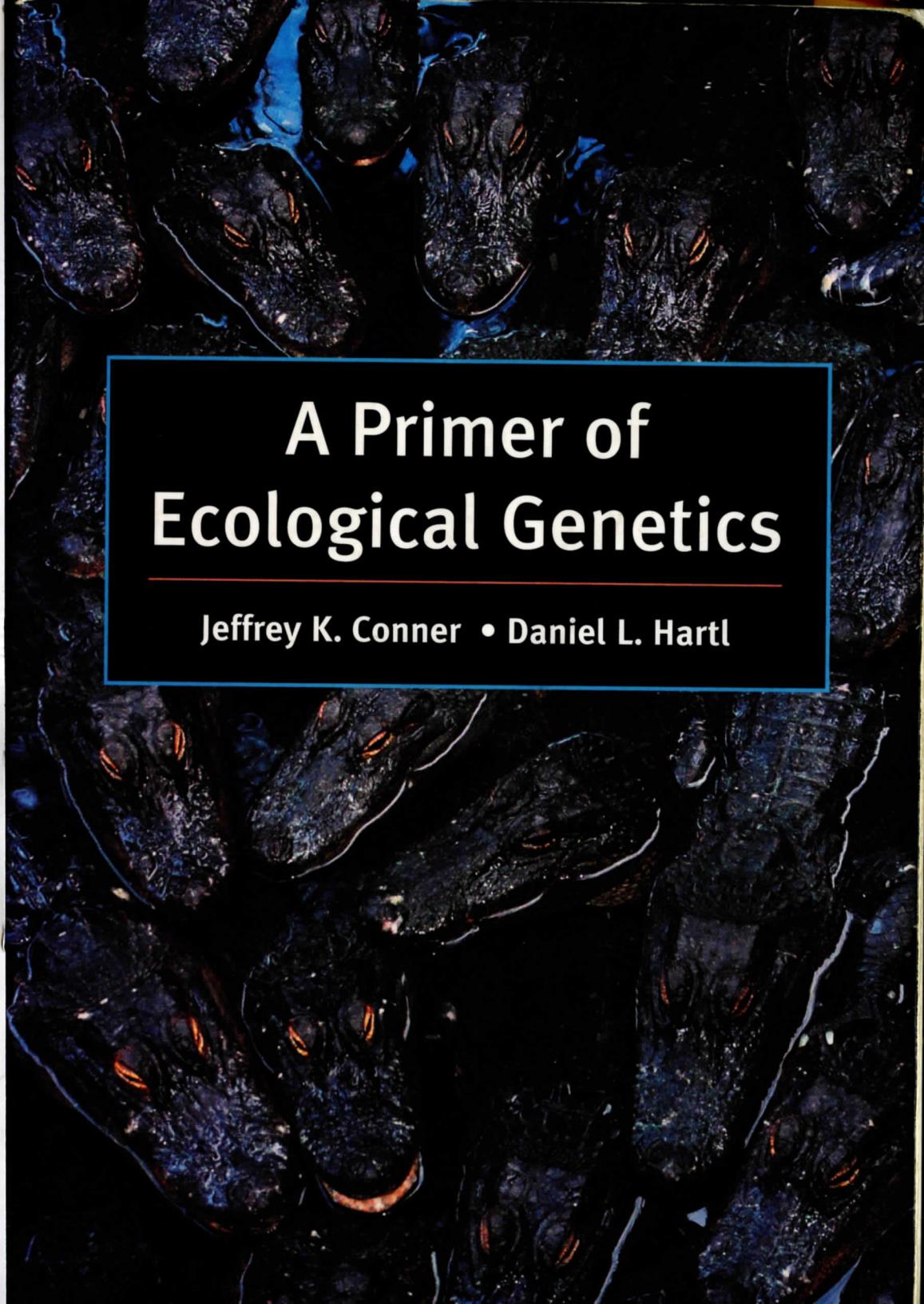


6317
Tuxford - 28028
ext.

A Primer of Ecological Genetics

Jeffrey K. Conner • Daniel L. Hartl



Cover photograph © Jacob Silverman Conner

Contents

A PRIMER OF ECOLOGICAL GENETICS

Copyright © 2004 by Sinauer Associates, Inc. All rights reserved. This book may not be reproduced in whole or in part for any purpose. For information address Sinauer Associates, Inc., 23 Plumtree Road, Sunderland, MA 01375 U.S.A.

Fax: 413-549-1118

Email: publish@sinauer.com

www.sinauer.com

Library of Congress Cataloging-in-Publication Data

Conner, Jeffrey K., 1957-

A primer of ecological genetics / Jeffrey K. Conner,
Daniel L. Hartl.— 1st ed.
p. ; cm.

Includes bibliographical references and index.

ISBN 0-87893-202-X (pbk.)

1. Ecological genetics. 2. Population genetics. 3. Quantitative genetics.

[DNLM: 1. Genetics, Population. 2. Ecology. 3. Selection (Genetics) 4. Variation (Genetics) QH 456 C752p 2004] I. Hartl, Daniel L. II. Title.

QH456.C665 2004

576.5'8—dc22

2003022664

Printed in U.S.A.

Preface xi

Acronyms, Abbreviations, and Symbols xiii

1 Introduction 1

What is Ecological Genetics? 1

Overview of the Book 3

Basic Genetic Terms 4

2 Population genetics I: Genetic variation, random and nonrandom mating 9

What is Population Genetics? 9

What are populations and why are they important? 9

Genetic Variation 11

Measuring Genetic Variation: Genetic Markers 12

Visible polymorphisms 13

Molecular markers 13

Protein electrophoresis 14

DNA markers 16

Measuring Genetic Variation: Simple Summary Statistics 22

Organization of Genetic Variation within Populations 24

Random mating 25

<i>Assumptions of the Hardy-Weinberg model</i>	25
Box 2.1 WHAT IS A MODEL?	26
<i>Deriving the Hardy-Weinberg equations</i>	27
Uses of the Hardy-Weinberg Model: Tests for Departure from HWE	31
Box 2.2 A BRIEF INTRODUCTION TO STATISTICS	32
<i>Recessive alleles hidden in heterozygotes</i>	35
Nonrandom Mating	36
<i>Assortative mating</i>	36
<i>Inbreeding</i>	37
Problems	43
Suggested Readings	44
Chapter References	44
3 Population genetics II: Changes in allele frequency	47
Mutation	47
Migration	48
Genetic Drift	52
<i>Random genetic drift</i>	52
<i>Population differentiation</i>	55
<i>F-statistics</i>	57
Box 3.1 CALCULATION OF F-STATISTICS	60
<i>Effective population size (N_e)</i>	62
<i>Inbreeding with random mating in small populations</i>	65
Natural Selection: Selection on Genotypes	66
Box 3.2 EXAMPLE OF INBREEDING DEPRESSION DUE TO GENETIC DRIFT	66
<i>Fitness</i>	67
<i>How does selection at one locus change allele frequencies at that locus?</i>	69
<i>Overdominance or heterozygote advantage</i>	71
<i>Frequency dependent selection</i>	76
Population Genetics: Synthesis	77
Interactions Between the Four Evolutionary Forces	79
<i>Mutation-selection balance</i>	79
<i>Selection vs. migration</i>	81
<i>Selection vs. genetic drift</i>	81
<i>Migration vs. genetic drift</i>	84
<i>Methods of measuring gene flow</i>	85
<i>Selection, genetic drift, and gene flow: Wright's shifting balance theory</i>	86

Problems	89
Suggested Readings	92
Suggested Readings Questions	92
Chapter References	94

4 Quantitative genetics I: Genetic variation 97

Mendelian Basis of Continuous Traits	97
<i>Different types of gene action</i>	103
<i>How can we quantify gene action?</i>	104
<i>Population mean</i>	106
<i>Population variance</i>	108
<i>Breeding value</i>	111
Heritability	112
Estimating Additive Variance and Heritability	115
<i>Offspring-parent regression</i>	116
Box 4.1 REGRESSION DETAILS	118
<i>Maternal and paternal effects</i>	120
<i>Sibling analyses</i>	121
Box 4.2 ANALYSIS OF VARIANCE (ANOVA)	122
<i>Comparison of methods</i>	131
Problems	133
Suggested Readings	134
Suggested Readings Questions	134
Chapter References	135

5 Quantitative genetics II: Advanced topics 137

Phenotypic Plasticity and Genotype-by-Environment Interaction	138
Box 5.1 TWO-WAY ANOVA	141
Multiple Subpopulations—Differentiation vs. Adaptive Plasticity	145
Correlations Among Traits	150
<i>Sources of genetic covariance among traits</i>	155
<i>What can cause linkage disequilibrium?</i>	160
<i>Mechanisms of genetic correlations in nature</i>	160
Artificial Selection	163
<i>Advantages and disadvantages of artificial selection</i>	169
QTL Mapping	170
Problems	180
Suggested Readings	183

Suggested Readings Questions 184
Chapter References 186

6 Natural selection on phenotypes 189

The Chicago School Approach to Phenotypic Evolution 191
Selective Agents and Targets 196
Multiple traits; direct and indirect selection 199
Selection gradients 200
Experimental manipulation 205
Correlational selection 208
How do we study selective agents? 211
Notes on the Study of Adaptation 215
Synthesis: Predicting Short-Term Phenotypic Evolution 216
Box 6.1 SUMMARY OF BIVARIATE RELATIONSHIPS IN ECOLOGICAL GENETICS 217
The Future of Ecological Genetics 223
Problems 224
Suggested Readings 226
Suggested Readings Questions 227
Chapter References 229

7 Applied ecological genetics 231

Conservation Genetics 232
What is the unit to be conserved? 233
How do genetic factors directly affect extinction risk? 235
Adaptation to environmental change in the future 241
Conservation genetics in the future 244
Evolution of Invasive Species 245
Transgene Escape 248
Establishment of transgenes in wild populations 250
Fitness effects of transgenes 250
Strategies to reduce transgene escape risk 252
Evolution of Resistance to Pesticides and Antibiotics 253
Selection 254
Genetic variance and covariance 257
Gene flow and population structure 259
Resistance management 260

The Future of Applied Ecological Genetics 262
Suggested Readings 263
Suggested Readings Questions 263
Chapter References 265

Appendix 269

Glossary 271

Index 284

Preface

This book covers basic concepts in population and quantitative genetics, including measuring selection on phenotypic characters, with a focus on methods applicable to field studies of ecologically important traits. It is designed for a broad audience of advanced undergraduates and graduate students, and is also aimed at providing professionals outside the field with an accessible introduction. The concepts included are critical for training students in ecology, evolution, conservation biology, agriculture, forestry, and wildlife management. The book should be useful as a textbook for courses on ecological and evolutionary genetics, conservation biology, and population biology, as well as a supplement to readings from the original literature in graduate-level courses in evolution.

The guiding principle of the book is to focus on clear explanations of the key concepts in the evolution of natural and managed populations. For example, we discuss the similarities and differences between inbreeding and random genetic drift, describe how dominance and epistasis represent interactions among alleles within and between loci, respectively, examine the most common misconceptions about heritability, and show the relationships among different methods of measuring selection. Mathematics and conceptual material are integrated and fully explained. The mathematics is used as a tool to improve understanding of the biological principles, not an end in itself. For most concepts, examples from the literature, mainly from studies of natural populations, are briefly described. These examples are not necessarily classic studies (although many are) but rather were chosen

because they provide clear and simple illustrations of the concept under discussion. The solutions to these problems, worked out in full, can be found at the website www.sinauer.com/connerhartl.

We have also striven to synthesize and integrate the different chapters and sections to give students a clearer idea of the “big picture” of how different concepts relate to each other. Our main goal is to enable students to understand the concepts well enough so they can gain entry into the primary literature. For this reason, most chapters have sets of discussion questions keyed to a few original papers in ecological genetics, each of which represents an important contribution to the literature and is written clearly and simply enough to be accessible to a beginning student.

We welcome any comments or suggestions, particularly those that point out any errors or unclear concepts or descriptions. Please direct these remarks to connerj@msu.edu.

Acknowledgments

The authors thank Butch Brodie, Richard Frankham, and Mike Wade for reviewing one or more chapters, and especially Mike Whitlock for reviewing the entire manuscript. The book was also improved by comments from students in Conner’s laboratory and in the graduate evolution course at Michigan State, especially Meghan Duffy, Frances Knapczyk, Angela Roles, Heather Sahli, James Sobel, Christy Stewart, and Rachel Williams. Nancy Haver drew the pen and ink drawings, Michele Ruschaupt created the graphic art, and Roberta Lewis copyedited most of the book. The staff at Sinauer Associates did truly excellent work, including the production team led by Chris Small and Joan Gemme whose skill in design and layout are evident in this book. We particularly thank Sydney Carroll for overseeing the editing and production of the book and Andy Sinauer for guidance, encouragement, and extraordinary patience. Jeff Conner would like to thank his wife, Buffy Silverman, for trying to teach him to write well, and his children Jake and Emma for putting up with hearing about this book at too many dinners. Dan Hartl thanks Christine and Christopher for their enduring support, and all of the people in the Hartl laboratory for their continuing enthusiasm, creativity, and hard work.

Acronyms, Abbreviations, and Symbols

Note: Numbers in brackets refer to the chapter in which each symbol is introduced.

β	Linear selection gradient [6]
χ^2	Chi-square value [2]
γ	Non-linear (variance) selection gradient [6]
μ	Mutation rate at a locus per generation [3]
σ^2	Theoretical expectation of a variance [4]
a	Additive effect, which equals the genotypic value for the homozygotes [4]
A	Adenine, the purine base, or a nucleotide containing adenine [1]
AFLP	Amplified fragment length polymorphism [2]
ANOVA	Analysis of variance [4]
BLUP	Best linear unbiased prediction [4]
bp	Base pair, a unit of length in nucleic acids equal to 1 base pair [5]
Bt	The bacterium <i>Bacillus thuringiensis</i> , which produces insecticidal toxins [7]
c	Frequency of recombination between two gene loci, also denoted r [5]
C	Cytosine, the pyrimidine base, or a nucleotide containing cytosine [1]

CAPS	Cleaved amplified polymorphic site [2]
Cov	Covariance [5]
cM	Centimorgan, a unit of length in a genetic map equal to 1% recombination [5]
cpDNA	Chloroplast DNA
d	Genotypic value for the heterozygote [4]
$d.f.$	Degrees of freedom [2]
DNA	Deoxyribonucleic acid [1]
E	Environmental deviation [5]
e^2	Fraction of the phenotypic variance attributable to the environment; the complement of heritability [5]
ESU	Evolutionarily significant unit [7]
F_1	First offspring generation of a cross [5]
F_2	Second offspring generation usually resulting from self-fertilization of the F_1 generation [5]
F or F_{IS}	Inbreeding coefficient of individuals relative to the subpopulation [2, 3]
F_{ST}	Fixation index of a subpopulation relative to the total population [3]
F_{IT}	Metapopulation increase in homozygosity due to inbreeding and population substructure [3]
5' end	The end of a nucleic acid strand containing a free 5' phosphate group on the sugar
G	Guanine, the purine base, or a nucleotide containing guanine [1]
G	Genotypic value [4]
G	Additive genetic variance/covariance matrix [6]
$G \times E$	Genotype-by-environment interaction [5]
g-e	Genotype-by-environment interaction [5]
GE	Genetically engineered [7]
GM	Genetically modified [7]
GMO	Genetically modified organism [7]
G_{ST}	Fixation index (similar to F_{ST}) [5]
H	Heterozygosity, the frequency of heterozygous genotypes in a population [2]
H_0 or H_S	Heterozygosity expected at Hardy-Weinberg equilibrium ($2pq$) [2, 3]
H_O	Observed heterozygosity within subpopulations [3]
H_T	Expected metapopulation heterozygosity assuming random mating ($2p_0q_0 = 2\bar{p}\bar{q}$) [3]

<i>h</i>	Degree of dominance of a recessive allele [3]	r_P	Phenotypic correlation [5]
h^2	Usual symbol for narrow-sense heritability [4]	r_{IJ}	Coefficient of relatedness [2]
h^2_B	Broad-sense heritability (V_G/V_P) [4]	<i>r</i>	Frequency of recombination between two gene loci; also denoted <i>c</i> [5]
h^2_N	Unambiguous symbol for narrow-sense heritability (V_A/V_P) [4]	<i>R</i>	Allele for resistance to a pesticide [7]
HWE	Hardy-Weinberg equilibrium (genotypic frequencies p^2 , $2pq$, q^2) [2]	RAPD	Random amplified polymorphic DNA [2]
IBD	Identical by descent [2]	REML	Restricted maximum likelihood [4]
indel	Insertion or deletion [3]	RFLP	Restriction fragment length polymorphism [2]
kb	Kilobase, a unit of length in nucleic acids equal to 10^3 bases or 10^3 base pairs [5]	RNA	Ribonucleic acid [1]
LD	Linkage disequilibrium [5]	<i>s</i>	Selection coefficient [3]
LE	Linkage equilibrium [5]	<i>S</i>	Allele for susceptibility to a pesticide [7]
lod or LOD	The log odds score; logarithm of the likelihood ratio [5]	SCAR	Sequence-characterized amplified region [2]
<i>m</i>	Migration rate [3]	SNP	Single-nucleotide polymorphism [2]
MHC	Major histocompatibility complex [7]	SPAR	Single-primer amplified region [2]
ML	Maximum likelihood [4]	SS	Sum of squares [4]
MS	Mean square (variance) [4]	SSCP	Single-stranded conformational polymorphism [2]
mRNA	Messenger RNA [1]	SSR	Simple sequence repeat [2]
mtDNA	Mitochondrial DNA [7]	SSRP	Simple sequence repeat polymorphism [2]
<i>n</i>	Haploid chromosome number [2]	STR	Simple tandem repeat [2]
<i>n</i> or <i>N</i>	Sample size [2]	STS	Sequence-tagged site [2]
N_a	Actual population size [3]	<i>T</i>	Thymine, the pyrimidine base, or a nucleotide containing thymine [1]
N_e	Effective population size [3]	3' end	The end of a nucleic acid strand containing a free 3' hydroxyl on the sugar [1]
<i>ns</i>	In a statistical test, no significant difference from that expected with random sampling	<i>U</i>	The whole-genome mutation rate, especially for loci affecting fitness [7]
<i>p</i>	Frequency of one allele, often the wildtype, dominant, or selectively favored allele [2]	<i>U</i>	Uracil, the pyrimidine base, or a nucleotide containing uracil [1]
<i>P</i>	Frequency of one homozygote [2]	V_A	Additive genetic variance [4]
<i>P</i>	Probability value in a statistical test (<i>P</i> -value) [2]	V_D	Dominance variance [4]
<i>P</i>	Phenotypic value [4]	V_E	Environmental variance [4]
PCR	Polymerase chain reaction [2]	V_G	Genotypic variance [4]
<i>q</i>	Frequency of one allele, often the recessive, rare mutant, or deleterious allele [2]	V_I	Epistatic variance [4]
<i>Q</i>	Frequency of one homozygote [2]	V_P	Phenotypic variance [4]
Q_{ST}	Quantitative trait differentiation [5]	Var	Variance [4]
QTL	Quantitative trait locus [5]	VNTR	Variable number of tandem repeats polymorphism [2]
r_A	Additive genetic correlation [5]	<i>w</i>	Relative fitness [3]
r_E	Environmental correlation [5]	<i>z</i>	Phenotypic trait value [6]
		*	Statistical significance at the 5% probability level [5]
		**	Statistical significance at the 1% probability level [5]

1

Introduction

What is Ecological Genetics?

Ecological genetics is at the interface of ecology, evolution, and genetics, and thus includes important elements from each of these fields. We can use two closely related definitions to help describe the scope of ecological genetics:

1. Ecological genetics is concerned with the *genetics of ecologically important traits*, that is, those traits related to fitness such as survival and reproduction. Ecology is the study of the distribution and abundance of organisms—in other words, how many individuals there are, where they live, and why. Distribution and abundance are determined by birth rates and death rates, which in turn are determined by interactions with the organism's biotic and abiotic environment. These interactions include predation, competition, and the ability to find mates, food, and shelter. Consider traits that would help an organism deal with each of these interactions. Cryptic coloration could help a beetle avoid being eaten, growing tall could help a plant compete with other plants for light, and a thick coat of fur might help a mouse survive winter cold. These are examples of *ecologically important traits*: those traits that are closely tied to fitness or, in other words, are important in determining an organism's adaptation to its natural environment, both biotic and abiotic.
2. Ecological genetics can also be defined as the *study of the process of phenotypic evolution occurring in present-day natural populations*. Phenotypic evolution can be defined as a change in the mean or variance of a trait

across generations due to changes in allele frequencies. The four processes that can cause evolution are **mutation**, **genetic drift**, **migration**, and **natural selection**. All of these processes are described in Chapter 3, and the last three in particular are closely related to ecology and therefore appear throughout the book. Ecological factors can cause population size to decline, and the resulting small population size causes genetic drift. Migration is clearly ecological, but how is natural selection related to ecology? Selection is caused by differences in fitness among organisms in a population, and these fitness differences are caused in part by interactions with the environment as previously mentioned.

Our two definitions are tied together by the concept of *adaptation*, which is the central theme of ecological genetics. An **adaptation** is a phenotypic trait that has evolved to help an organism deal with something in its environment. Like most ecologically important traits, the examples given above are adaptations. Natural selection is special among the four evolutionary processes because it is the only one that leads to adaptation. Mutation, genetic drift, and migration can either speed up or constrain the development of adaptations, but they cannot cause adaptation.

An overview of these ideas is shown in Figure 1.1, which summarizes much of what will be covered in this book. Beginning at the top, ecological factors, both biotic and abiotic, can cause fitness differences among organisms with different phenotypes within the population; this is natural selection. If mutation and recombination create genetic variation for these phenotypic traits, then the selection can act on this variation to change the

genetic composition of the population. The genetics of the population can also be affected by gene flow from other populations with different genetic composition, or by genetic drift if the population size is small. All these changes in genetic composition are likely to feed back and affect the genetic variation for the phenotypic traits, as well as change the average phenotype in the population across generations. These phenotypic changes can lead to an improvement in the ability of the population to survive and reproduce in its biotic and abiotic environment; that is, it can lead to adaptation.

As an example, a deer mouse is part of a beetle's biotic environment, and may cause beetles with increased defensive secretions to have higher fitness than those with less secretions, if the secretions deter deer mouse predation. If this phenotypic variation is caused at least in part by underlying genetic variation, then this will cause an increase in the frequency of alleles that increase defensive secretions. (An **allele** is a particular type of a given gene.) This increase in allele frequency across generations may be slowed by random genetic drift, or by gene flow from other beetle populations with low frequencies of the high-secretion alleles (perhaps because there are fewer mice coexisting with those other populations). Selection and drift may decrease genetic variation for secretion quantity, also slowing future evolution of secretions. If the average secretion quantity in the population increases in spite of these constraints, then this may reduce the impact of mouse predation in subsequent generations, increasing adaptation of the beetles to this environmental factor.

Overview of the Book

Chapters 2 and 3 cover the field of population genetics. **Population genetics** is the study of genetic variation within and among populations, focusing on the processes that affect genotypic and allele frequencies at one or a few gene loci. These processes include inbreeding, mutation, migration, drift, and selection; the genotypic and allele frequencies are revealed mainly through molecular markers. Population genetics for the most part does not focus on phenotypes, since the genes and alleles underlying most phenotypic traits are unknown, especially in natural populations. This is because most phenotypic traits are complex, being affected by several to many gene loci and by the environment.

Chapters 4 and 5 cover the field of **quantitative genetics**, which does focus on the phenotype, usually without knowing the genotypes underlying the traits. In the place of genotypic information, statistical abstractions such as variance, correlation, and heritability are used in quantitative genetics to help understand the genetics of complex phenotypic traits. QTL mapping (covered at the end of Chapter 5) is a marriage of molecular and statistical techniques for studying the genetics of complex phenotypic traits. QTL mapping

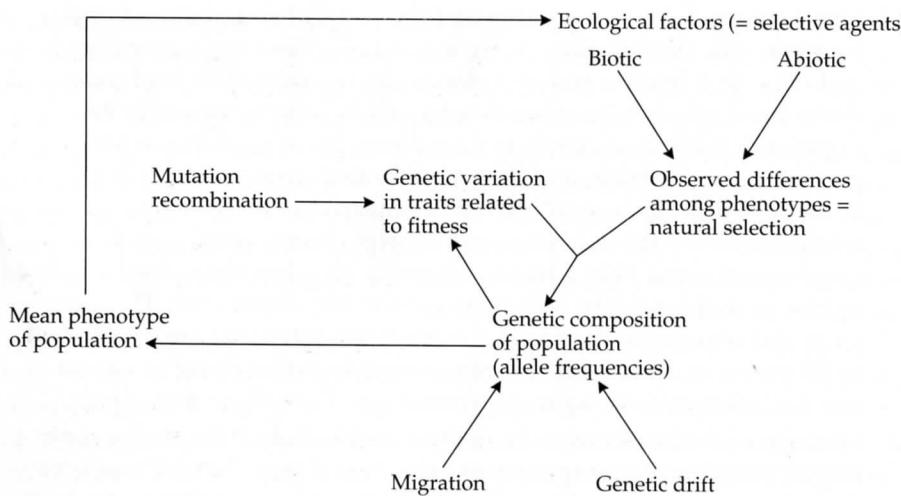


Figure 1.1 A schematic overview of key concepts in ecological genetics.

is a first step in discovering the genes underlying phenotypic traits in natural populations, bringing together the fields of population and quantitative genetics. This convergence is very likely to lead to fundamental new insights in ecological genetics.

Chapter 6 is on techniques developed from quantitative genetics for studying natural selection on phenotypic traits (rather than on genotypes as in population genetics). These techniques have allowed biologists to measure the strength and direction of selection in natural populations, as well as help determine the ecological causes of the selection. Chapter 6 also synthesizes the quantitative genetic material in Chapters 4 through 6, and shows how short-term evolution can be predicted in natural populations using knowledge of genetic variance and the strength of selection.

Since ecological genetics is at the interface between ecology, evolution and genetics, it is a critical component of all three fields, as well as essential for the study of some of society's problems. In Chapter 7 we will discuss the importance of ecological genetic principles in conservation, the spread of invasive species, the evolution of pesticide, herbicide, and antibiotic resistance, and the environmental effects of genetically modified organisms used in agriculture.

The focus of the book will be on diploid sexual organisms. Most of the concepts covered also apply to asexual and haploid organisms, but there are important differences. Most of our examples will come from studies of plants and animals, because the ecological genetics of most microorganisms and fungi are not as well known.

Basic Genetic Terms

A **gene** is a stretch of DNA (deoxyribonucleic acid) coding for a polypeptide chain; one or more polypeptides make up a protein. The genetic information in DNA is coded in the sequence of four nucleotides, abbreviated according to the identity of the nitrogenous **base** that each contains: A (adenine), G (guanine), T (thymine), or C (cytosine). DNA molecules normally consist of two complementary helical strands held together by pairing between the bases: A in one strand is paired with T in the other strand, and G in one strand is paired with C in the other.

The process of creating proteins from the genetic code in DNA is called **gene expression**. The essentials of gene expression in the cells of eukaryotes are outlined in Figure 1.2. The first step is **transcription**, in which the sequence of nucleotides present in one DNA strand of a gene is faithfully copied into the nucleotides of an **RNA** (ribonucleic acid) molecule. As the RNA transcript is synthesized, each base in the DNA undergoes pairing with a base in an RNA nucleotide, which is then added to the growing RNA strand. The base-pairing rules are the same as those in DNA, except that in RNA nucleotides the base U (uracil) is found instead of T (thymine). The second step of gene expression is **RNA processing**, in which intervening sequences or **introns** are removed

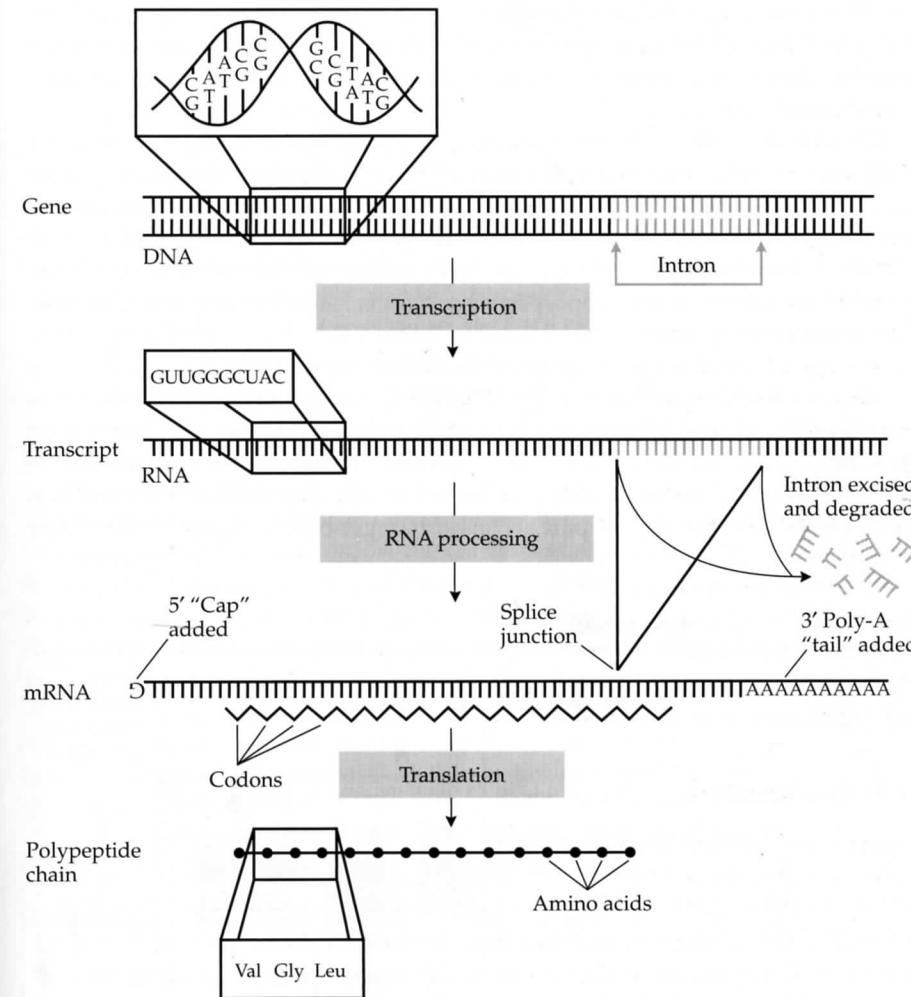


Figure 1.2 Principal processes in gene expression in eukaryotes.

from the RNA transcript by splicing and the ends of the transcript are modified. The regions between the introns that remain in the fully processed RNA are known as **exons**; these are the sequences that actually code for proteins. The fully processed RNA constitutes the **messenger RNA** (mRNA).

The messenger RNA undergoes **translation** on ribosomes in the cytoplasm to produce the polypeptide that is encoded in the sequence of nucleotides. In the translated part of the messenger RNA, each adjacent group of three nucleotides constitutes a coding group or **codon**, which specifies a corresponding amino acid subunit in the polypeptide chain. The stan-

dard **genetic code** showing which codons specify which amino acids is given in Table 1.1. After each three-letter codon are the three- and one-letter designations for the 20 amino acids. The three-letter and one-letter abbreviations are both established conventions. Note that in many cases changes in the third base in the codon do not change the amino acid that is specified; therefore, much variation at this position is not expressed (sometimes called the “silent” position). The codon AUG specifies methionine and also serves as the start codon for polypeptide synthesis. Any of three codons—UAA, UAG, or UGA—specify the end, or termination, of polypeptide synthesis, upon which the completed polypeptide chain is released from the ribosome. The start and stop codons are shaded in Table 1.1.

All the DNA in a cell is collectively called the **genome**. Genome size is typically expressed as the amount of DNA in a reproductive cell (sperm or egg), and it differs greatly among species. For example, the genome of *Arabidopsis thaliana*, a model plant for genetic studies, consists of about 120 million base pairs, whereas the genome of the lily *Fritillaria* is 1000 times as large, about 120 billion base pairs. The human genome is about 3 billion base

TABLE 1.1 *The standard genetic code*

First nucleotide in codon (5' end)	Second nucleotide in codon				Third nucleotide in codon (3' end)
	U	C	A	G	
U	UUU Phe/F	UCU Ser/S	UAU Tyr/Y	UGU Cys/C	U
	UUC Phe/F	UCC Ser/S	UAC Tyr/Y	UGC Cys/C	C
	UUA Leu/L	UCA Ser/S	UAA Stop	UGA Stop	A
	UUG Leu/L	UCG Ser/S	UAG Stop	UGG Trp/W	G
C	CUU Leu/L	CCU Pro/P	CAU His/H	CGU Arg/R	U
	CUC Leu/L	CCC Pro/P	CAC His/H	CGC Arg/R	C
	CUA Leu/L	CCA Pro/P	CAA Gln/Q	CGA Arg/R	A
	CUG Leu/L	CCG Pro/P	CAG Gln/Q	CGG Arg/R	G
A	AUU Ile/I	ACU Thr/T	AAU Asn/N	AGU Ser/S	U
	AUC Ile/I	ACC Thr/T	AAC Asn/N	AGC Ser/S	C
	AUA Ile/I	ACA Thr/T	AAA Lys/K	AGA Arg/R	A
	AUG Met/M	ACG Thr/T	AAG Lys/K	AGG Arg/R	G
G	GUU Val/V	GCU Ala/A	GAU Asp/D	GGU Gly/G	U
	GUC Val/V	GCC Ala/A	GAC Asp/D	GGC Gly/G	C
	GUA Val/V	GCA Ala/A	GAA Glu/E	GGG Gly/G	A
	GUG Val/V	GCG Ala/A	GAG Glu/E	GGG Gly/G	G

pairs. Genes are arranged in linear order along microscopic threadlike bodies called **chromosomes**. Each human **gamete** (sperm or egg) contains one complete set of 23 chromosomes; this is the **haploid** chromosome number, designated as n . Chromosome number can vary greatly: $n = 2$ in some scorpions and 127 in a species of hermit crab! A typical chromosome contains several thousand genes, in humans averaging approximately 1500 genes per chromosome. The position of a gene along a chromosome is called the **locus** of the gene. Sometimes the words gene and locus are used interchangeably, which can lead to confusion. **Recombination** between loci can occur during meiosis, which creates new combinations of alleles at these different loci. Recombination is rarer between loci that are close together on the chromosome; these loci are said to be genetically **linked**.

In most multicellular organisms, each individual cell contains two copies of each type of chromosome, one inherited from its mother through the egg and one inherited from its father through the sperm (so the **diploid** chromosome number, $2n$, is 46 in humans and 254 in hermit crabs). Note that these two copies of the chromosome are not the two complementary strands of DNA; each chromosome consists of a double-stranded DNA molecule. At any locus, therefore, every diploid individual contains two copies of the gene—one at each corresponding (homologous) position in the maternal and paternal chromosome. These two copies are the alleles of the gene in that individual. If the two alleles at a locus are indistinguishable according to any particular experimental criterion, then the individual is **homozygous** at the locus under consideration. If the two alleles at a locus are distinguishable by means of this criterion, then the individual is **heterozygous** at the locus.

The **genotype** of an individual is the diploid pair of alleles present at a given locus. Therefore, homozygous and heterozygous are the two major categories of genotypes. Typographically, genes are indicated in italics, and alleles are typically distinguished by uppercase or lowercase letters (*A* versus *a*), subscripts (A_1 versus A_2), superscripts (a^+ versus a^-), or sometimes just + and -. Using these symbols, the genotype of homozygous individuals would be portrayed by any of these formulas: AA , aa , A_1A_1 , A_2A_2 , a^+a^+ , a^-a^- , $+/+$, or $-/-$. As in the last two examples, the slash is sometimes used to separate alleles present in homologous chromosomes to avoid ambiguity. The genotype of heterozygous individuals would be portrayed by any of the formulas Aa , A_1A_2 , a^+a^- , or $+/-$.

The outward appearance of an organism for a given characteristic is its **phenotype**. Phenotypic traits can be defined at a number of hierarchical levels, each one dependent on a number of traits at lower levels. For example, the form of an enzyme encoded by a gene is a phenotype, as is a physiological function like metabolic rate that depends on a number of enzymes. A number of different physiological functions affect morphological traits like

height, and physiology and morphology together can affect behavioral phenotypes such as courtship. Finally, all these lower level traits can affect life history traits like survival and reproduction, which determine the ultimate trait of individual fitness. The traits that are higher in this hierarchy are more complex and affected by more gene loci. The expression of most phenotypic traits, and especially the higher level ones, are also affected to varying degrees by the environment. This complexity means that the same genotype can produce different phenotypes, through the action of the environment. Conversely, the different genotypes can produce the same phenotypes, again due to the environment and also due to gene interactions. We will discuss complex phenotypic traits and fitness in more detail in Chapters 4 through 6.

2

Population genetics I: Genetic variation, random and nonrandom mating

What is Population Genetics?

In its broadest sense, **population genetics** is the study of naturally occurring genetic differences between organisms; these differences are called **genetic variation**. Genetic variation is important because it is the raw material for evolution. Genetic variation can occur at three hierarchical levels: within populations, between populations of the same species, and between different species. Therefore, to understand the purview of population genetics, we need to have a better understanding of populations and genetic variation.

What are populations and why are they important?

Populations are important because evolutionary processes occur primarily within populations; it is populations that evolve through changes in allele frequencies. In fact, most of the important concepts and processes in ecological genetics have meaning only at the level of the population, not the individual: These include genetic variation, allele and genotype frequencies, gene flow, drift, natural selection, heritability, and genetic correlation.

A species is rarely, if ever, a single interbreeding **panmictic** group, that is, one in which any member of the species can potentially mate with any other member of the opposite sex. Instead, species are divided into **populations**. A population can be defined as a local interbreeding (panmictic) group that has reduced gene flow with other groups of the same species. By reduced