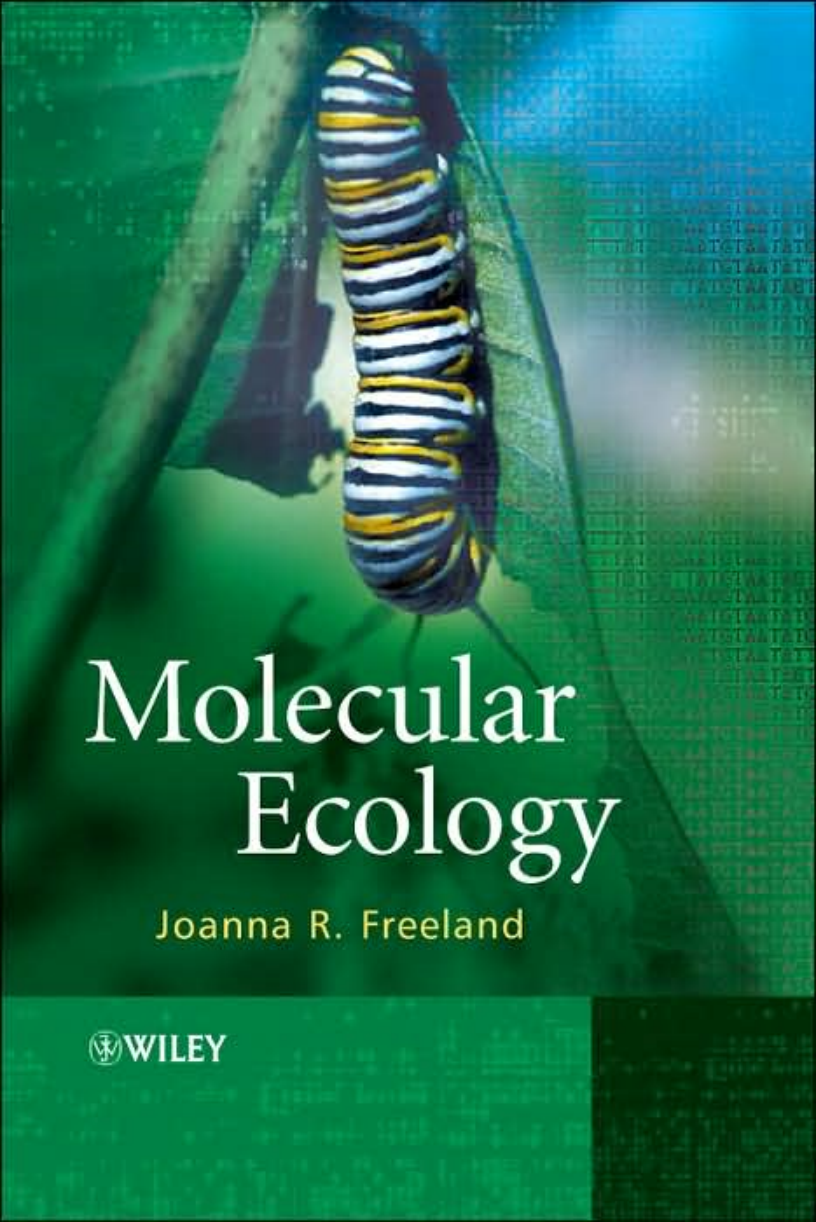
Molecular
Ecology
Joanna R. Freeland
WILEY

Molecular Ecology

Joanna R. Freeland
The Open University, Milton Keynes

John Wiley & Sons, Ltd

Molecular Ecology

Molecular Ecology

Joanna R. Freeland
The Open University, Milton Keynes

John Wiley & Sons, Ltd

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Freeland, Joanna.
Molecular ecology / Joanna Freeland.
p. cm.
Includes bibliographical references.
ISBN-13 978-0-470-09061-8 (HB) ISBN-13 978-0-470-09062-6 (PB)
ISBN-10 0-470-09061-8 (HB) ISBN-10 0-470-09062-6 (PB)
1. Molecular ecology. I. Title.
QH541.15.M63F74 2005
577–dc22 2005027904

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN-13 978-0-470-09061-8 (HB) ISBN-13 978-0-470-09062-6 (PB)
ISBN-10 0-470-09061-8 (HB) ISBN-10 0-470-09062-6 (PB)

Typeset in 10.5/13pt Minion-Regular by Thomson Press (India) Limited, New Delhi, India.
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire.
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Contents

Preface

The theory and practice of molecular ecology draw on a number of subjects, particularly genetics, ecology and evolutionary biology. Although the foundations of molecular ecology are not particularly new, it did not emerge until the 1980s as the discipline that we now recognize. Since that time the growth of molecular ecology has been explosive, in part because molecular data are becoming increasingly accessible and also because it is, by its very nature, a collaborative discipline. Molecular ecology is now a broad area of research that embraces topics as varied as population genetics, conservation genetics, molecular evolution, behavioural ecology and biogeography, and has added much to our understanding of ecology. Researchers in molecular ecology now routinely publish their work in a wide range of ecological and evolutionary journals (including *Molecular Ecology*, first published in 1992), and also in more general publications such as *Science* and *Nature*.

Although somewhat varied, the areas of research within molecular ecology are united by the fact that they all use molecular genetic data to help us understand the ecology and evolution of organisms in the wild. Although there are many excellent texts that cover general ecology and evolution, there is currently a shortage of books that provide a comprehensive overview of molecular ecology. The most important goal of this book, therefore, has been to present molecular genetics, population genetics and applied molecular ecology in a logical and uncomplicated – but not oversimplified – manner, using up-to-date examples from a wide range of taxa. This text is aimed at upper-level undergraduate and postgraduate students, as well as at researchers who may be relatively new to molecular ecology or are thinking about different ways to address their research questions using molecular data.

Each chapter may be read in isolation, but there is a structure to the book that should be particularly useful to students who read the text in its entirety. The first two chapters provide a brief history of molecular ecology and a review of genetics, followed by an overview of molecular markers and the types of data they generate. Chapters 3 and 4 then build on this foundation by looking at ways in which molecular data can be used to characterize single and multiple populations. Having read Chapters 1–4, readers should have a good understanding of the relevant theory and practice behind molecular markers and population genetics.

Chapter 5 then builds on this by adding an explicit evolutionary component within the context of phylogeography. Chapters 6 and 7 then focus on two additional, specific applications of molecular ecology, namely behavioural ecology and conservation genetics. Finally, chapter 8 provides a more general overview of the practical applications of molecular ecology, paying particular attention to questions surrounding law enforcement, agriculture and fishing, which will be of interest to biologists and non-biologists alike.

As an aid to the reader, each chapter is followed by a summary, a list of useful websites and software and some recommended further reading. Suggestions for further reading also can, of course, come from the extensive reference list at the end of the book. There are review questions after each chapter that students can use to identify key points and test their knowledge. There is also a glossary at the end of the book, and glossary words are highlighted in bold when they first appear in the text. An ongoing website (www.wiley.com/go/freeland) will be maintained upon which corrections and new developments will be reported, and from which figures that may be used as teaching aids, can be downloaded.

Joanna Freeland

Acknowledgements

Many thanks to James Austin, Amanda Callaghan Colin Ferris, Hélène Fréville, Trevor Hodkinson and Steve Lougheed for reading part or all of this book and providing helpful comments. The cover photo and design concept were by Kelvin Conrad. Thanks also to James Austin, Spencer Barrett, P.G. Bentz, David Bilton, Kelvin Conrad, Mike Dodd, Claude Gascon, Beth Okamura, Kate Orr and Jon Slate for providing photos. Kelvin Conrad also helped with some figures and provided essential technical advice. This book is dedicated to Eva and William.

Joanna Freeland

1

Molecular Genetics in Ecology

What is Molecular Ecology?

Over the past 20 years, molecular biology has revolutionized ecological research. During that time, methods for genetically characterizing individuals, populations and species have become almost routine, and have provided us with a wealth of novel data and fascinating new insights into the ecology and evolution of plants, animals, fungi, algae and bacteria. Molecular markers allow us, among other things, to quantify genetic diversity, track the movements of individuals, measure inbreeding, identify the remains of individuals, characterize new species and retrace historical patterns of dispersal. These applications are of great academic interest and are used frequently to address practical ecological questions such as which endangered populations are most at risk, from inbreeding, and how much hybridization has occurred between genetically modified crops and their wild relatives. Every year it becomes easier and more cost-effective to acquire molecular genetic data and, as a result, laboratories around the world can now regularly accomplish previously unthinkable tasks such as identifying the geographic source of invasive species from only a few samples, or monitoring populations of elusive species such as jaguar or bears based on little more than hair or scat samples.

In later chapters we will take a detailed look at many of the applications of molecular ecology, but before reaching that stage we must first understand just why molecular markers are such a tremendous source of information. The simplest answer to this is that they generate data from the infinitely variable **deoxyribonucleic acid** (**DNA**) molecules that can be found in almost all living things. The extraordinarily high levels of genetic variation that can be found in most species, together with some of the methods that allow us to tap into the goldmine of information that is stored within DNA, will therefore provide the focus of this chapter. We will start, however, with a retrospective look at how

Molecular Ecology Joanna Freeland

the characterization of proteins from fruitfly populations changed forever our understanding of ecology and evolution.

The Emergence of Molecular Ecology

Ecology is a branch of biology that is primarily interested in how organisms in the wild interact with one another and with their physical environment. Historically, these interactions were studied through field observations and experimental manipulations. These provided phenotypic data, which are based on one or more aspects of an organism's morphology, physiology, biochemistry or behaviour. What we may think of as traditional ecological studies have greatly enhanced our knowledge of many different species, and have made invaluable contributions to our understanding of the processes that maintain ecosystems.

At the same time, when used on their own, phenotypic data have some limitations. We may suspect that a dwindling butterfly population, for example, is suffering from low genetic diversity, which in turn may leave it particularly susceptible to pests and pathogens. If we have only phenotypic data then we may try to infer genetic diversity from a variable morphological character such as wing pattern, the idea being that morphologically diverse populations will also be genetically diverse. We may also use what appear to be population-specific wing patterns to track the movements of individuals, which can be important because immigrants will bring in new genes and therefore could increase the genetic diversity of a population. There is, however, a potential problem with using phenotypic data to infer the genetic variation of populations and the origins of individuals: although some physical characteristics are under strict genetic control, the influence of environmental conditions means that there is usually no overall one-to-one relationship between an organism's **genotype** (set of genes) and its **phenotype**. The wing patterns of African butterflies in the genus *Bicyclus*, for example, will vary depending on the amount of rainfall during their larval development period; as a result, the same genotype can give rise to either a wet season form or a dry season form (Roskam and Brakefield, 1999).

The potential for a single genotype to develop into multiple alternative phenotypes under different environmental conditions is known as **phenotypic plasticity**. A spectacular example of phenotypic plasticity is found in the oak caterpillar *Nemoria arizonaria* that lives in the southwest USA and feeds on a few species of oaks in the genus *Quercus*. The morphology of the caterpillars varies, depending on which part of the tree it feeds on. Caterpillars that eat catkins (inflorescences) camouflage themselves by developing into catkin-mimics, whereas those feeding on leaves will develop into twig mimics. Experiments have shown that it is diet alone that triggers this developmental response (Greene, 1996). The difference in morphology between twig-mimics and catkin-mimics is so pronounced that for many years they were believed to be two different species. There

Table 1.1 Some examples of how environmental factors can influence phenotypic traits, leading to phenotypic plasticity

Characteristic	Environmental influence	Example
Gender	Temperature during embryonic development	Eggs of the American snapping turtle *Chelydra serpentina* develop primarily into females at cool temperatures, primarily into males at moderate temperatures, and exclusively into females at warm temperatures (Ewert, Lang and Nelson, 2005)
Growth patterns in plants	Soil nutrients and water availability	Southern coastal violet (*Viola septemloba*) allocated a greater proportion of biomass to roots and rhizomes in poor-quality environments (Moriuchi and Winn, 2005)
Leaf size	Light intensity	Dandelions (*Taraxacum officinale*) produce larger leaves under conditions of relatively strong light intensity (Brock, Weinig and Galen, 2005)
Migration between host plants	Age and nutritional quality of host plants	Diamond-back moths (*Plutella xylostella*) are most likely to migrate as adults if the juvenile stage feed on mature plants (Campos, Schoereder and Sperber, 2004).
Feeding-related morphology	Food availability	Sea-urchin larvae (*Strongylocentrotus purpuratus* and *S. franciscanus*) produce longer food-gathering arms and smaller stomachs when food is scarce (Miner, 2005)
Plumage colouration	Carotenoids in diet	The plumage of male house finches (*Carpodacus mexicanus*) shows varying degrees of red, orange and yellow depending on the carotenoids in each bird's diet (Hill, Inouye and Montgomerie, 2002)

is also a behavioural component to these phenotypes, because if either is placed on a part of the tree that it does not normally frequent, the catkin-mimics will seek out catkins against which to disguise themselves, and the twig-mimics will seek out leaves or twigs. Some other examples of phenotypic plasticity are given in Table 1.1.

Phenotypic plasticity can lead to overestimates of genetic variation when these are based on morphological variation. In addition, phenotypic plasticity may obscure the movements of individuals and their genes between populations if it causes the offspring of immigrants to bear a closer resemblance to individuals in their natal population than to their parents. Complex interactions between genotype, phenotype and environment provided an important reason why biologists sought long and hard to find a reliable way to genotype wild organisms; genetic data would, at the very least, allow them to directly quantify genetic variation, and to track the movements of genes – and therefore individuals or **gametes** – between populations. The first milestone in this quest occurred around 40 years ago, when researchers discovered how to quantify individual genetic

variation by identifying structural differences in proteins (Harris, 1966; Lewontin and Hubby, 1966). This discovery is considered by many to mark the birth of molecular ecology.

Protein allozymes

In the 1960s a method known as starch gel **electrophoresis** of allozymic proteins was an extremely important breakthrough that allowed biologists to obtain direct information on some of the genetic properties of individuals, populations, species and higher taxa. Note that we are not yet talking about DNA markers but about proteins that are encoded by DNA. This distinction is extremely important, and to eliminate any confusion we will take a minute to review the relationship between DNA, genes and proteins. **Prokaryotes**, which lack cell nuclei, have their DNA arranged in a closed double-stranded loop that lies free within the cell's cytoplasm. Most of the DNA within the cells of **eukaryotes**, on the other hand, is organized into **chromosomes** that can be found within the nucleus of each cell; these constitute the nuclear genome (also referred to as **nuclear DNA** or **nrDNA**). Each chromosome is made up of a single DNA molecule that is functionally divided into units called genes. The site that each gene occupies on a particular chromosome is referred to as its **locus** (plural **loci**). At each locus, different forms of the same gene may occur, and these are known as **alleles**.

Each allele is made up of a specific sequence of DNA. The DNA sequences are determined by the arrangement of four nucleotides, each of which has a different chemical constituent known as a base. The four DNA bases are adenine (A), thymine (T), guanine (G) and cytosine (C), and these are linked together by a sugar–phosphate backbone to form a strand of DNA. In its native state, DNA is arranged as two strands of complementary sequences that are held together by hydrogen bonds in a double-helix formation (Figure 1.1). No two alleles have exactly the same DNA sequence, although the similarity between two alleles from the same locus can be very high.

The function of many genes is to encode a particular protein, and the process in which genetic information is transferred from DNA into protein is known as **gene expression**. The sequence of a protein-coding gene will determine the structure of the protein that is synthesized. The first step of protein synthesis occurs when the coding region of DNA is transcribed into **ribonucleic acid** (**RNA**) through a process known as **transcription**. The RNA sequences, which are single stranded, are complementary to DNA sequences and have the same bases with the exception of uracil (U), which replaces thymine (T). After transcription, the **introns** (non-coding segments of DNA) are excised and the RNA sequences are translated into protein sequences following a process known as **translation**.

Translation is possible because each RNA molecule can be divided into triplets of bases (known as **codons**), most of which encode one of 20 different **amino acids**, which are the constituents of proteins (Table 1.2). Transcription and

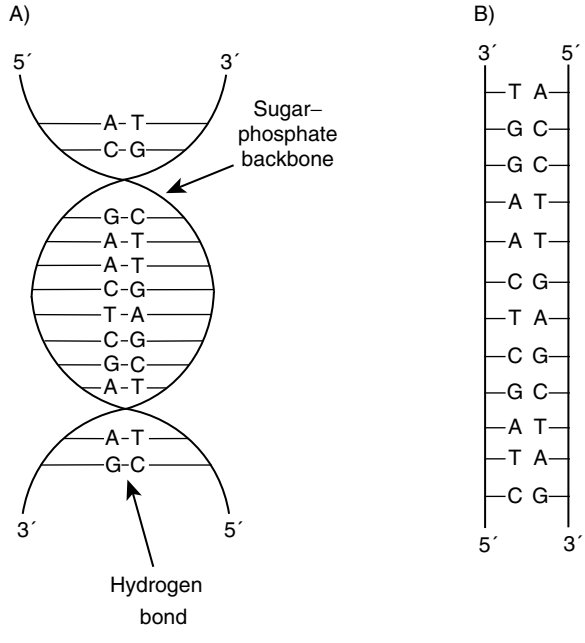

Figure 1.1 (A) A DNA double helix. Each sequence is linked together by a sugar–phosphate backbone, and complementary sequences are held together by hydrogen bonds; 3′ and 5′ refer to the orientation of the DNA: one end of a sequence has an unreacted 5′ phosphate group and the other end has an unreacted 3′ hydroxyl group. (B) Denatured (single-stranded) DNA showing the two complementary sequences. The DNA becomes denatured following the application of heat or certain chemicals

translation involve three types of RNA: **ribosomal RNA** (**rRNA**), **messenger RNA** (**mRNA**) and **transfer RNA** (**tRNA**). Ribosomal RNA is a major component of ribosomes, which are the **organelles** on which mRNA codons are translated into proteins, i.e. it is here that protein synthesis takes place. Messenger RNA molecules act as templates for protein synthesis by carrying the protein-coding information that was encoded in the relevant DNA sequence, and tRNA molecules incorporate particular amino acids into a growing protein by matching amino acids to mRNA codons (Figure 1.2).

Specific combinations of amino acids give rise to **polypeptides**, which may form either part or all of a particular protein or, in combination with other molecules, a protein complex. If the DNA sequences from two or more alleles at the same locus are sufficiently divergent, the corresponding RNA triplets will encode different amino acids and this will lead to multiple variants of the same protein. These variants are known as **allozymes**. However, not all changes in DNA sequences will result in different proteins. Table 1.2 shows that there is some redundancy in the **genetic code**, e.g. leucine is specified by six different codons. This redundancy means that it is possible for two different DNA sequences to produce the same polypetide product.