

Contents

Dedication	v
Preface	vi
Acknowledgements	ix
List of TechBoxes	xvi
List of Case Studies	xvii

1 Introduction

The story in DNA

page 1



Reading the story in DNA	2
Individuals, families, and populations	3
Uncovering the evolutionary past	5
Tracing ancestors	5
About this book	10
What is this book for?	10
Who is this book for?	11
How the book is structured	12
How to use this book	13
❶ Conclusion	15
❷ Points to remember	16
❸ Ideas for discussion	16
❹ Examples used in this chapter	17
Heroes of the Genetic Revolution 1: Fred Sanger	18
TechBox 1.1: GenBank	19
TechBox 1.2: BLAST	23
Case Study 1.1: Identification: solving the mystery of the Chilean blob	25
Case Study 1.2: Forensics: using DNA to trace trade in whale meat	28

2 DNA

The immortal germline

page 31



Material basis of heredity	32
Principles of heredity	33

Discovery of the gene	36
Structure of DNA	38
The central dogma	41
DNA extraction	42
Ubiquity of DNA	42
Extracting DNA	42

❶ Conclusion	43
❷ Points to remember	44
❸ Ideas for discussion	45
❹ Sequences used in this chapter	45
❺ Examples used in this chapter	45

Heroes of the Genetic Revolution 2:

August Weismann	46
TechBox 2.1: DNA structure	47
TechBox 2.2: DNA extraction	52
Case Study 2.1: Environmental DNA: tracking changes in Arctic vegetation over time	54
Case Study 2.2: Ancient DNA: cataloguing diversity of extinct giant birds	57

3 Mutation

We are all mutants

page 61



Mutation	62
Point mutations	62
Single nucleotide polymorphisms (SNPs)	65
Identifying individuals using SNPs	65
Using SNPs to track inherited traits	67
Estimating the mutation rate	68
Mutation rates are a compromise	72
Is mutation random?	73

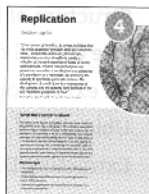
❶ Conclusion	75
❷ Points to remember	75
❸ Ideas for discussion	76
❹ Examples used in this chapter	76
Heroes of the Genetic Revolution 3: James Crow	78
TechBox 3.1: Single nucleotide polymorphisms (SNPs)	79

TechBox 3.2: Biobanking	84
Case Study 3.1: SNPs: genetic resources for crop development in a non-model species	88
Case Study 3.2: Mutation rate: mutation accumulation in a laboratory population	91

4 Replication

Endless copies

page 95



DNA replication	96
Template reproduction	97
Replication	98
DNA amplification	102
Polymerase chain reaction (PCR)	103
Contamination	106
Copy errors create hierarchies	106
Sequence evolution	107
Hierarchies reveal history	110
❖ Conclusion	113
⊙ Points to remember	114
⊗ Ideas for discussion	115
⊞ Sequences used in this chapter	116
⊛ Examples used in this chapter	116
Heroes of the Genetic Revolution 4:	
Rosalind Franklin	117
TechBox 4.1: DNA replication	118
TechBox 4.2: DNA amplification	121
Case Study 4.1: Replication: taller plants have lower rates of molecular evolution	124
Case Study 4.2: Hierarchies: genes and language track human settlement of Oceania	127

5 Genome

Accident and design

page 131



Genome size	132
The paradox of genome size	133
Doubling and dividing: genome replication	134
Recombination	136
Chromosomal rearrangement	138

Repeat sequences	139
Comparing repeat sequences	141
Transposable elements	145
LINEs, SINEs, and retroviruses	146
The genomic record of transposition	147

❖ Conclusion	149
⊙ Points to remember	150
⊗ Ideas for discussion	150
⊞ Sequences used in this chapter	151
⊛ Examples used in this chapter	151

Heroes of the Genetic Revolution 5:

Barbara McClintock	152
TechBox 5.1: Sanger sequencing	153
TechBox 5.2: High-throughput sequencing	157
Case Study 5.1: Duplication: gene copies and microsatellites help aphids tolerate nicotine	161
Case Study 5.2: Genome: a leukaemia virus becoming part of the koala genome	164

6 Gene

Making an organism

page 167



Genes	168
Gene structure and expression	170
Transcription and translation	171
Regulation of gene expression	173
Sequence evolution	175
Divided genes: exons and introns	175
Identifying genes	178
Regulatory sequences	181

❖ Conclusion	182
⊙ Points to remember	183
⊗ Ideas for discussion	184
⊞ Sequences used in this chapter	184
⊛ Examples used in this chapter	184

Heroes of the Genetic Revolution 6:

Francis Crick	185
TechBox 6.1: Genetic code	186
TechBox 6.2: Primers and probes	189
Case Study 6.1: Gene families: duplication and loss of taste receptor genes	191
Case Study 6.2: Genetic code: multiple proteins from a single human gene	194

7 Selection*Descent with modification*

page 197



Variation	198
Substitution	200
Natural selection	201
The power of selection	202
Fitness	205
Genetic background	206
Linkage	208
Are humans still evolving?	209
❖ Conclusion	214
❖ Points to remember	214
❖ Ideas for discussion	215
❖ Sequences used in this chapter	215
❖ Examples used in this chapter	215
Heroes of the Genetic Revolution 7: J. B. S. Haldane	216
TechBox 7.1: Fitness	217
TechBox 7.2: Detecting selection	220
Case Study 7.1: Selection: tracking insecticide resistance over time using museum specimens	222
Case Study 7.2: Variation: experiments on evolutionary rescue under environmental change	225

8 Drift*Chance and necessity*

page 229



Neutral theory	230
Genetic load	230
Genetic drift	232
Neutrality	233
Population size	236
Inbreeding	238
Patterns of substitution	239
Chance and complexity	242
❖ Conclusion	244
❖ Points to remember	245
❖ Ideas for discussion	246
❖ Sequences used in this chapter	247
❖ Examples used in this chapter	247

Heroes of the Genetic Revolution 8: Tomoko

Ohta (太田 朋子)	248
TechBox 8.1: Neutral theory	249
TechBox 8.2: Population size	252
Case Study 8.1: Substitution: mutation accumulation in asexual walking stick insects	255
Case Study 8.2: Population size: inbreeding depression in an endangered species	258

9 Species*Origin of species*

page 262



Taxonomy	263
Classification	264
Relatedness	266
Divergence	267
Speciation	267
Hybrid incompatibility	268
Species	269
Defining species	272
DNA taxonomy	274
❖ Conclusion	275
❖ Points to remember	276
❖ Ideas for discussion	277
❖ Examples used in this chapter	277
Heroes of the Genetic Revolution 9:	
Rosemary Gillespie	278
TechBox 9.1: Phylogeography	279
TechBox 9.2: Barcoding	284
Case Study 9.1: Speciation: evolution of novel flower colour through reinforcement	288
Case Study 9.2: Barcoding: cataloguing weevil diversity in New Guinea	291

10 Alignment*Same but different*

page 295



Homology	296
Systematic hierarchy	297
Shared by descent	300

DNA can reveal homology	301
Alignment establishes homology	302
Alignment	302
Alignment gaps	303
Homologous sites	304
Manual alignment	305
Automated alignment	306
How to make the best alignment	308
❖ Conclusion	310
⊙ Points to remember	310
⊙ Ideas for discussion	311
⊙ Sequences used in this chapter	311
⊙ Examples used in this chapter	311
Heroes of the Genetic Revolution 10:	
Margaret Oakley Dayhoff	312
TechBox 10.1: De-extinction	313
TechBox 10.2: Multiple sequence alignment	317
Case Study 10.1: Alignment: identifying insertions and deletions in endosymbiont genomes	321
Case Study 10.2: Horizontal gene transfer: genes in parasitic plants derived from their hosts	325

11 Phylogeny

Tree of life

page 329



History	330
Life as a tree	332
Molecular phylogenetics	335
Phylogeny reconstruction	340
Splits	342
Similarity	345
Distance methods	345
Conflict	347
Networks	349
❖ Conclusion	350
⊙ Points to remember	350
⊙ Ideas for discussion	351
⊙ Examples used in this chapter	352
Heroes of the Genetic Revolution 11:	
Joseph Felsenstein	353
TechBox 11.1: Distance methods	354
TechBox 11.2: Phylogenetic networks	358

Case Study 11.1: Networks: bobtail squid and their symbiotic bioluminescent bacteria	362
Case Study 11.2: Distance: tracing the origins and evolution of the domesticated apple	365

12 Hypotheses

Seeing the wood for the trees

page 369



Comparing trees	370
Possible histories	371
Evaluating alternative trees	372
Multiple hits	374
Statistical inference of phylogeny	376
Models of molecular evolution	377
Likelihood	378
Testing phylogenetic hypotheses	379
Phylogenetic support	380
Which phylogenetic method should I choose?	384
❖ Conclusion	385
⊙ Points to remember	386
⊙ Ideas for discussion	387
⊙ Examples used in this chapter	387
Heroes of the Genetic Revolution 12:	
Hélène Morlon	388
TechBox 12.1: Maximum likelihood	389
TechBox 12.2: Bootstrap	392
Case Study 12.1: Epidemiology: using ancient DNA to identify the origins of a plague	396
Case Study 12.2: Prediction: relating ethnobotanical resources across different cultures	400

13 Rates

Tempo and mode

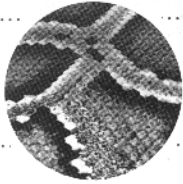
page 403



Rate of evolutionary change	404
Variation in the rate of evolution	405
Estimating branch lengths	406
Variation in rates across sites	408

Comparing rates	409	Divergence	443
Mutation rate variation	409	Divergence dates	444
Accounting for descent	411	Estimating number of substitutions	447
Generation time effect	413	Accuracy and precision	450
Lineage-specific rates	415	Dating with confidence	453
Comparing substitution rates	415	Molecular dating	455
Testing for rate variation	418	Modelling rate change	456
❶ Conclusion	419	Challenging assumptions	457
❷ Points to remember	420	❶ Conclusion	460
❸ Ideas for discussion	421	❷ Points to remember	461
❹ Sequences used in this chapter	421	❸ Ideas for discussion	461
❺ Examples used in this chapter	422	❹ Sequences used in this chapter	462
Heroes of the Genetic Revolution 13:		❺ Examples used in this chapter	462
Xia Hua (华夏)	423	Heroes of the Genetic Revolution 14:	
TechBox 13.1: Substitution models	424	Andrew Rambaut	464
TechBox 13.2: Bayesian phylogenetics	428	TechBox 14.1: Calibration	465
Case Study 13.1: Rates: flightless insects		TechBox 14.2: Molecular dating	471
have faster substitution rates	432	Case Study 14.1: Calibration: did kauri survive	
Case Study 13.2: Diversification: mutation		the Oligocene drowning of New Zealand?	476
rates are linked to species richness in plants	434	Case Study 14.2: Dates: using phylogenies	
		to trace the source of disease outbreaks	480
14 Dates		Coda: You are a scientist	
Telling the time		What do I do now?	
page 438		page 484	
Confidence	439	Glossary	487
Fossilization: a chance at immortality	441	Index	505
Alternative hypotheses	442		





Preface

This book is designed to give a from-the-ground-up introduction to the use of molecular phylogenetic analysis in evolution and ecology. It describes how you can use analysis of DNA sequence data to reconstruct evolutionary history and processes. To do this, you need to have a basic understanding of molecular evolution: how mutations occur that change the DNA sequence, how some of these mutations become substitutions that create differences between populations or species, and how we can use the patterns of substitutions to reconstruct evolutionary history. The aims of this book, and the features it contains, are covered more fully in Chapter 1. Here is a brief summary with references to pages where each point is discussed in more detail.

Who is this book for?

This book is aimed at biology students who need to learn how to use DNA sequence analysis to understand evolution and ecology. The book assumes no more than a basic grounding in biology and genetics, so should be suitable for entry-level undergraduate students (see Chapter 1, 'Who is this book for?', page 11). However, this book may also be useful to bioinformaticians with a background in maths, statistics, or computation, who need to learn evolutionary and genetic principles in order to interpret patterns in biological data.

What is in this book?

This book is designed to be read in a variety of ways, depending on your interests or background knowledge (see Chapter 1, 'How to use this book', page 13). To facilitate you getting what you need from this book, each chapter is divided into several sections.

The main text provides the background to key issues in molecular evolution and phylogenetics, and should be accessible to entry-level students with no previous knowledge of the subject.

- **Points to remember** provides a bullet-point summary of the key issues covered in the chapter, as a revision aid.
- **Ideas for discussion** draw on the subjects covered in the chapter, but asks you to step beyond the material covered to think more broadly about the concepts.

• **Sequences used in this chapter** gives the Genbank description and accession number for any sequence used for illustrative purposes in the chapter.

• **Examples used in this chapter** is not a comprehensive bibliography but the publication references of specific examples mentioned in the text.

There are three kinds of boxes, which provide optional additions to the main text, as follows.

TechBoxes go into more detail on specific methods or ideas. The language is more technical and they may be challenging to entry-level students, but more informative for students with some previous knowledge of molecular genetic analysis.

TECHBOX 10.1		De-extinction
- RELATED TECHBOXES - TB 2.1: DNA extraction - TB 3.2: Sequencing	Summary New genetic technologies offer hope of bringing recently extinct species back. Keywords cloning, gene bank, high-throughput sequencing, somatic nuclear cell transfer, genome editing, endangered, reproductive technology, frozen zoo	
RELATED CASE STUDIES		

Case Studies illustrate recent scientific research related to the principles covered in the chapter, cross-referenced to TechBoxes that explain the techniques used. There will be many terms and concepts that will be unfamiliar to beginners, so they may be better suited to more advanced students. There are three '**Check your understanding**' questions which are based on material from the box. The '**What do you think?**' section aims to get you to think more deeply about the implications of the topics discussed in the box, and could be used for class discussions. '**Delve deeper**' encourages you to go beyond the material presented in the case study and do some independent research on a related topic. **References** cited in the box are listed at the end, as are cross-references to other useful boxes.

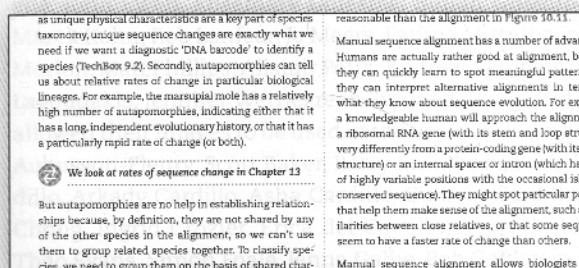
CASE STUDY 1.2		Forensics: using DNA to trace trade in whale meat
- RELATED TECHBOXES - TB 1.2: Genbank - TB 2.2: Sequencing	Baker C.S., Steel D., Choi Y., Lee H., Kim K.S., Choi S.K., Ma Y.H., Hanisatom C., Palocz L., Brownell R.L., Fujiwara H. (2010) Genetic evidence of illegal trade in protected whales from Japan with the US and South Korea. <i>Biology Letters</i>, Volume 6, page 647. These results confirmed the power of molecular methods in monitoring	

Heroes of the Genetic Revolution illustrate scientists who have worked in the area.



Other features of the text

Cross-references to information in other chapters, to help you navigate through the book and build up a knowledge base.



Glossary contains some basic biological terms and explains some tricky concepts in evolution and genetics.

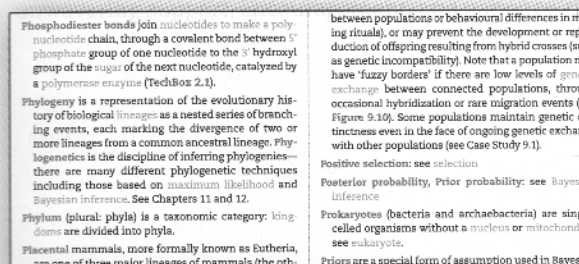
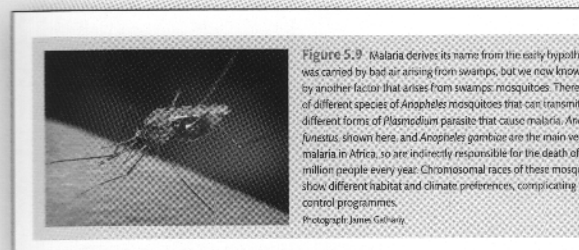


Figure legends sometimes have interesting facts not directly related to the subject of the chapter, because that's what makes biology fun.



What is not in this book?

Instructions for computer programs: The aim of this book is to give you an understanding of the evolutionary principles which are the basis of molecular phylogenetic analysis, rather than to give you training in the use of particular programs. The reason that this book sticks to general principles and avoids specific methods is that new methods are constantly being developed. Any programs or methods I describe today are likely to be out of date within a few years. But by giving you the general principles underlying those methods, I hope to equip you with the intellectual tools you need to learn how to use particular methods, and to adapt your analysis when new methods become available.

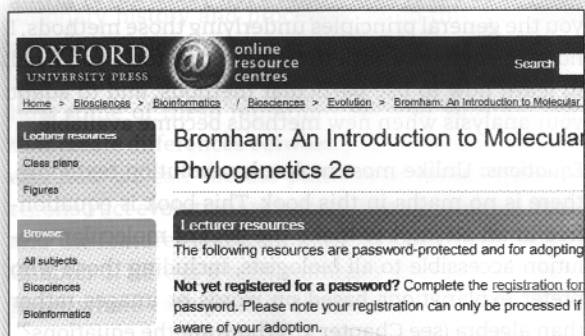
Equations: Unlike most molecular evolution textbooks, there is no maths in this book. This book is equation-free in an attempt to make the field of molecular evolution accessible to all biologists, including those who prefer explanations based on words or images rather than algebra (see Chapter 1, 'Where are the equations?', page 15). You might choose to read this book alongside a traditional molecular evolution text which will provide all the population genetic equations for the principles discussed here.

What has changed since the first edition?

The first edition of this book was titled *Reading the Story in DNA: a beginner's guide to molecular evolution*. This new edition contains some of the same material, but has been considerably expanded, updated, and remodelled in response to feedback from readers. The book has increased from 8 chapters to 14, which allows the material to be divided more evenly into key concepts and may make the book better fit a semester teaching structure. There are more features designed to help students and teachers come to grips with the material, such as summary points at the end of chapters, and discussion questions in the chapters and case studies. There are now nearly twice as many case studies, and many of the previous case studies have been replaced with more recent examples. I have discovered that the problem with writing a book based on molecular genetics is that the field changes so fast that much of the first edition had been superseded only five years after its publication. For example, high-throughput (next-gen) sequencing, which took up only a paragraph in the first edition, is on its way to eclipsing Sanger sequencing as the method of choice in molecular phylogenetic analysis, so the chapters and case studies have been updated

to reflect this fundamental change in data generation and analysis. It's inevitable that by the time this edition is in print, the field will have moved on even further. The best approach is to use this book as a jumping-off point for an examination of the most recent techniques in the scientific literature.

What is in the Online Resource Centre?



For lecturers: downloadable figures, tutorial exercises, practical projects.

For students: searchable glossary, links to relevant resources, scientific updates.

For everyone: the chance to give feedback on the book.

Go to www.oxfordtextbooks.co.uk/orc/bromham2e