

An Introduction to Molecular Ecology

Trevor J. C. Beebee
Graham Rowe

OXFORD
UNIVERSITY PRESS

Preface

What is molecular ecology? As a new scientific discipline, the subject has become increasingly well known in recent years but there have been few attempts at formal definition. Indeed there are those at both ends of the biological sciences spectrum, ecologists and molecular biologists alike, who remain sceptical that molecular ecology is a discrete discipline at all. To those familiar with the history of biochemistry there is a sense of *déjà vu* about this debate. A century or so ago, strongly held opinions were voiced by both physiologists and chemists that 'physiological chemistry' (later to become biochemistry) was a spurious hybrid science that should be strangled at birth. How wrong they were. With the benefit of hindsight we can appreciate how the emergence of biochemistry as a mainstream discipline revolutionized our understanding of the natural world. Only time will tell as to whether molecular ecology will enjoy comparable success, and at this early stage in its history a fully satisfying definition remains elusive.

Molecular ecology as we know it today did not really exist before the mid-1980s, although its foundations go much deeper. The subject area currently encompasses a wide range of research topics including population and evolutionary genetics, behavioural ecology, microbial ecology, conservation biology, the identification and assessment of species diversity, and the release of genetically modified organisms into the environment. Molecular ecology therefore brings together aspects of many sciences including molecular biology, ecology, evolution, behavioural biology, and genetics. The emergence of molecular ecology is clearly reflected in meeting agendas and learned journals, those central bastions of scientific culture. Thus the British Ecological Society, established in 1913 and the oldest ecological society in the world, held a conference on *Genes in Ecology* at the University of East Anglia in 1991. The Spring of 1992 witnessed publication of the subject's first dedicated journal, *Molecular Ecology*. By 2001 this journal expanded to 12 issues a year covering more than 2800 pages, enjoyed among the highest impact ratings of all ecology journals, and had spawned a sister publication *Molecular Ecology Notes*. Examples of molecular ecological research now appear regularly in almost every ecological, behavioural, and conservation biology journal as well as in mainstream evolution and genetics publications. Another sign of the success of molecular ecology has been the recent development of an important sub-discipline, conservation genetics, with its own conferences and another new journal *Conservation Genetics*. There have also been books that deal with various aspects of molecular ecology, but none that cover the topic in its

broad sense at a level suitable for undergraduates or graduates new to the subject. This is the role we hope the present volume will fulfil.

This book contains a glossary in which terms that may be unfamiliar to the reader are defined, and an extensive bibliography where original papers and reviews relating to the main topics of molecular ecology can be found. There are also two appendices that outline the major practical and theoretical aspects of the subject.

There is every reason to expect that molecular ecology will continue to grow for the foreseeable future, and increasingly penetrate important areas such as quantitative genetics and adaptive variation. We may still be at too early a stage for clear definitions, but the excitement of molecular ecology is real enough.

Acknowledgements

Our special thanks to Femmie Kraaijeveld-Smit for writing the bulk of Chapter 4 (Behavioural Ecology), and to Jonathan Crowe of Oxford University Press for his unstinting support throughout the manuscript preparation. We are grateful to Ingela Dahloff, Philip Damiani, Shannon Gowans, Alec Jeffreys, Ramon Massana, Christiane Saeglitz Kornelia Smalla, and Christopher Tebbe for the donation of original pictures and photographs; and to The American Association for the Advancement of Science, Annual Reviews, John Avise, Blackwell Publishing, Cambridge University Press, Elsevier, Springer-Verlag, Harvard University Press, Kluwer Academic Publishers, Margaret Ramsey, Nature Publishing Group, Oxford University Press, The American Society for Microbiology, The British Herpetological Society, and The Royal Society for permissions to reproduce figures originally published elsewhere. Many friends and colleagues have stimulated our interest in molecular ecology, in particular Pim Arntzen, Gillian Baker, Eddie Brede, Terry Burke, Trent Garner, Susan Hitchings, and Inga Zeisset. Finally, we are indebted to reviewers of early drafts of the text (Ian Baldwin, Amanda Callaghan, Gary Carvalho, Mary Alice Coffroth, Scott Edwards, David Lambert, P. L. M. Lee, A. J. McCarthy, Michael Veith, and Peter Young) for their constructive criticisms. We found their comments invaluable and have made numerous changes, hopefully for the better, in light of them. Final responsibility for the work, including any errors that may have survived the vetting process, is of course ours alone.

List of Figures

1.1	Male pied wagtail <i>Motacilla alba yarrellii</i> and white wagtail <i>Motacilla alba alba</i> , both in summer plumage.	4
1.2	(a) Carrion crow <i>Corvus corone corone</i> and hooded crow <i>Corvus corone cornix</i> . (b) Crow distributions in Europe.	5
1.3	Adaptive peaks and valleys. Peaks are local fitness optima for combinations of genotypes across multiple loci. Valleys represent suboptimal genotype combinations that will be selected against, inhibiting evolutionary transitions from one peak to another.	7
1.4	Chromosomes of the closely related Indian muntjac deer <i>Muntiacus muntjak</i> and Chinese muntjac deer <i>Muntiacus reevesi</i> .	10
1.5	The Oxford district showing sites where <i>Cepaea nemoralis</i> were collected.	11
1.6	The noctuid moth <i>Pseudaletia unipuncta</i> .	13
1.7	Human multilocus minisatellite DNA fingerprint.	19
1.8	Molecular phylogeny of modern and extinct elephants based on mitochondrial DNA cytochrome <i>b</i> sequences.	21
1.9	A microsatellite locus showing two allelic variants.	23
1.10	Single nucleotide polymorphisms (SNPs) in part of the <i>Drosophila melanogaster</i> alcohol dehydrogenase gene.	24
2.1	Nucleotides of (a) DNA and (b) RNA.	30
2.2	A short (3 nucleotide) oligonucleotide.	31
2.3	Diagrammatic representation of native and denatured DNA.	31
2.4	Part of the 16S ribosomal RNA molecule of <i>Escherichia coli</i> , showing stems (with base-pairing) and loops.	33
2.5	Model of a DNA replication 'fork' showing how new DNA fragments are initiated from primers.	34
2.6	The 20 amino acids from which proteins are synthesized.	36
2.7	The levels of structure in a haemoglobin protein molecule.	37
2.8	Antibody molecule with two 'heavy' and two 'light' polypeptide chains held together by disulphide bridges.	38
2.9	Arrangement of ribosomal genes in eukaryotes.	48
2.10	Typical vertebrate (human) mitochondrial DNA molecule.	50
2.11	Outline of procedure for cloning DNA fragments in plasmid vectors.	53

2.12	The main types of mutation.	56
3.1	Recent (1990s) distribution of koala bears in Australia with picture of koala bear.	63
3.2	Unrooted mtDNA haplotype network of <i>Pterodroma</i> species.	65
3.3	Changes in gene frequency, linkage disequilibrium, and embryonic mortality across a toad hybrid zone in Poland.	66
3.4	Distribution of <i>Spartina</i> (cord grass) species and hybrids around San Francisco Bay in 1997–8.	68
3.5	RAPD identification of larvae of two toad species, <i>Bufo calamita</i> and <i>Bufo bufo</i> .	70
3.6	Male and female bottle-nosed whales.	74
3.7	Estimation of coyote population size from faecal sampling.	77
4.1	(a) Dunnocks ‘cloacal pecking’ to remove sperm plugs. (b) DNA fingerprint of polyandrous matings.	84
4.2	Multiple paternity in a litter of agile <i>Antechinus</i> , as shown by microsatellite alleles.	85
4.3	Relationship between head depth and realized number of offspring in male dragon lizards.	87
4.4	Sperm fertilization success as a function of ejaculation time.	90
4.5	Male sand lizard mating success. Percentage sired relative to other males in a multiply sired clutch is shown as a function of multilocus minisatellite bandsharing.	92
4.6	Odour preference of gravid female three-spined sticklebacks for males in a flow tank.	93
4.7	Cladogram showing the phylogenetic relationships of 18 <i>Gp-9</i> alleles found in 10 <i>Solenopsis</i> species.	99
4.8	Four geographic forms of <i>Habronattus pugilis</i> and the mountain ranges where they occur.	104
4.9	Molecular phylogeny of <i>Xiphophorus</i> species showing the distribution of sword tails.	106
4.10	Map of Europe and Africa showing the breeding range of willow warblers of two subspecies (speckled <i>trochilus</i> , striped <i>acredula</i>), their contact zone (black) and wintering grounds.	107
5.1	(a) Expected relationship between heterozygosity (H) and \log_{10} of N_e for three different mutation rates; (b) Actual relationship for a range of animal species.	114
5.2	Indian rhinoceros.	116
5.3	Mean F_{is} in dog whelks as a function of increasing pooling stage for samples.	119
5.4	Metapopulation models. (a) Classical; (b) Mainland-island.	123
5.5	Isolation-by-distance effects in two natterjack toad metapopulations.	126

5.6 (a) Island and (b) stepping-stone models of gene flow between subpopulations.	128
5.7 Non-equilibrium (recent separation, no gene flow) and equilibrium (longstanding separation with continuous gene flow) models to explain a current level of genetic differentiation (d) between two subpopulations a and b .	129
5.8 Pairwise F_{st} —geographic distance relationships among collared lizard populations in (a) Texas; (b) Kansas (c) north-eastern Ozark; and (d) south-western Ozark.	131
5.9 Likelihood estimates of N_e for the St Lucia African buffalo population.	136
6.1 Evolutionary correlates of protein evolution.	146
6.2 (a) Normal and (b) ermine colour variants of European stoats.	147
6.3 Frequency changes of <i>Littorina Aat</i> allele 120 between 1987 and 1993 in surf and splash zones.	149
6.4 Mean numbers of developmental abnormalities as a function of mean heterozygosity in <i>Bufo bufo</i> populations.	154
6.5 Predicted male and female red deer calf first-winter survival, modelled as a function of mean d^2 .	156
6.6 Arrangement of major histocompatibility complex genes along 4000 kb of DNA in humans.	159
6.7 MHC IIB allele frequencies in salmon among families with low and high resistance to furunculosis.	160
6.8 RAPD map of the region surrounding the gene for resistance to white blister rust (R) in sugar pine.	161
7.1 (a) MtDNA phylogeny of 87 pocket gophers. (b) Schematic maximum parsimony network of the western pocket gopher clade.	167
7.2 Maximum parsimony network showing ‘star pattern’ for greenfinches.	168
7.3 (a) Schematic representation of lineage sorting of mtDNA haplotypes within a population. (b) Schematic representation of coalescence of nuclear genes, where recombination can occur.	169
7.4 Schematic diagram of maximum parsimony network showing complex pattern of relationships among mtDNA haplotypes from a single population.	171
7.5 Sample sites for the lake cisco and geographical distribution of the two major mtDNA clades (shown as frequencies within pie charts).	172
7.6 Schematic representation of three categories of phylogenetic relationship: reciprocal monophyly (a), polyphyly (b) and paraphyly (c) for haplotypic lineages of two daughter populations A and B derived from a common ancestral gene pool.	174

7.7	European phylogeographic hybrid zones (dark bars). Arrows show main postglacial colonization routes from glacial refugia.	177
7.8	European phylogeography of the grasshopper <i>Chorthippus parallelus</i> showing postglacial colonization routes from southern refugia.	178
7.9	Example of a nested clade analysis based on the alcohol dehydrogenase locus of <i>Drosophila melanogaster</i> .	180
7.10	Phylogeographic tree of bear (<i>Ursus</i>) lineages showing relationships among cave, brown, and polar bears.	185
7.11	(a) Distribution of pool frog in Europe showing sampling sites. (b) Phylogeographic tree of European pool frog populations showing percentage bootstrap support.	190
7.12	Geographical distribution of rhododendron trees.	192
7.13	Maximum parsimony network of <i>Polysiphonia harveyi rbcL</i> haplotypes.	194
7.14	Geographic distribution patterns of <i>Luciola cruciata</i> COII haplotypes in Japan, clustering into three clades.	195
8.1	Geological time scale showing principal extinction events.	200
8.2	Martha, the last passenger pigeon.	201
8.3	The Chatham Islands in relation to mainland New Zealand, showing the locations of Mangere and South-east Islands, the habitat of the endangered Chatham Island black robin.	204
8.4	Relationship between larval growth rate and average microsatellite heterozygosity in six natterjack toad populations with contrasting demographic histories.	207
8.5	Lady's slipper orchid.	212
8.6	Noah, a gaur, the first cloned endangered animal.	215
8.7	The extinct thylacine.	216
8.8	(a) Geographic distribution of seaside sparrow subspecies. (b) Dendrogram showing distinction between Gulf and Atlantic coast seaside sparrow lineages.	220
9.1	The 'universal' tree of life, based on comparisons of small subunit ribosomal RNA sequences.	226
9.2	Outlines of the major (a) carbon, (b) nitrogen, and (c) sulphur cycles in the biosphere.	228
9.3	(a) <i>Anurofeca richardsi</i> cells labelled with fluorescent polyclonal antibodies. (b) Monthly variation in numbers of <i>A. richardsi</i> cells in sediments of two ponds used for breeding by amphibians in spring.	235
9.4	Structures of (a) prokaryotic and (b) eukaryotic nuclear ribosomal RNA genes.	237
9.5	Variation in numbers of sulphate reducing bacterial cells (SRB) and amounts of SRB-cell ribosomal RNA (SRB rRNA) through a marine sediment core.	240

9.6	DNA with phosphodiester bond and peptide nucleic acid (PNA) with peptide bond.	241
9.7	Analysis of microbial community structure by amplification, cloning, and sequencing of 16S small subunit-rRNA genes.	243
9.8	DGGE profiles of Mediterranean Sea picoplankton using two different sets of SSU rRNA primers (A and B).	246
9.9	SSCP analysis of amplified 16S SSU rRNA from pure cultures of bacteria (lanes 2–14) and a bacterial community isolated from a maize rhizosphere (lane 15).	247
9.10	DGGE of (a) 16S SSU-rRNA sequences and (b) <i>rpoB</i> gene sequences from multiple strains of bacteria isolated from a marine rock.	248
9.11	Renaturation rate of soil bacterial DNA ('C ₀ t plot') by comparison with DNA from a pure culture of <i>Escherichia coli</i> .	250
9.12	Secondary structure of proteo-rhodopsin showing seven membrane-spanning regions.	253
10.1	Procedure for introducing a new gene (conferring resistance to the herbicide glyphosate) into a plant.	260
10.2	Detection of the fungus <i>Trichoderma virens</i> in soil samples.	263
10.3	Network of horizontal gene transfer routes.	264
10.4	Mechanisms of HGT: (a) Conjugation, (b) transduction, and (c) transformation.	265
10.5	DGGE electrophoresis showing seasonal shifts of the bacterial communities of potato plants.	269
10.6	PCR amplification of <i>cpBNYVV</i> genes in transgenic beet hybrids.	270
10.7	Survival of cells bearing test plasmid pS130 in natural freshwater.	271
10.8	Possible fates of extracellular DNA in the environment.	272
10.9	Growth of swallowtail butterfly caterpillars at various distances from <i>Bt</i> corn.	274
10.10	<i>Bt</i> toxin in leaves, aphids, and cotton leafworm larvae and faeces.	275
A1.1	Protein extraction and electrophoresis.	281
A1.2	Banding patterns of a tetrameric protein with two alleles (A and B) encoded by two loci in a diploid organism.	282
A1.3	Basis of the ELISA.	283
A1.4	PCR amplification of a DNA sequence.	285
A1.5	Basis of RFLP analysis. Arrows show restriction enzyme cutting sites on a DNA molecule, and fragment sizes are indicated in kilobase pairs (kb).	286
A1.6	Basis of AFLP analysis using restriction enzymes <i>Mse</i> I and <i>Eco</i> RI.	288

A1.7	Dideoxy DNA sequencing protocol with analysis on an acrylamide gel or on an automated sequencer.	289
A1.8	Basis of SSCP analysis.	291
A1.9	A protocol for isolating and characterising microsatellite loci.	292
A1.10	Basis of DNA microarray analysis. Reverse transcriptase copies mRNA into cDNA with one of two different fluorescent labels.	294

List of Tables

1.1 Main markers used in molecular ecology.	15
2.1 The bacterial/eukaryotic nuclear genetic code, with vertebrate mtDNA variations shown in brackets.	40
2.2 Genome size variation in eukaryotes.	44
2.3 Main classes of repetitive DNA in eukaryotic genomes.	45
5.1 Pollen to seed migration ratios in various trees.	120
5.2 AMOVA of salmon microsatellite allele frequencies.	125
5.3 Effective breeding population sizes of common toads.	135
5.4 Heterozygote excess in an northern hairy-nosed wombat population.	139
6.1 Proportion of variation between <i>Pinus sylvestris</i> populations in Finland.	158

Contents

<i>Preface</i>	xiii
<i>Acknowledgements</i>	xv
<i>List of figures</i>	xvii
<i>List of tables</i>	xxiii

1 A history of molecular ecology

Introduction	1
An evolutionary perspective for molecular ecology	2
Systematics, phylogenetics, and the species concept	2
Variation within species	4
The origins of modern genetics	6
The modern synthesis	6
The neutral theory of molecular evolution	8
Behavioural ecology	8
Genetics in ecology	9
Ecological genetics	9
Genotype, phenotype, and phenotypic plasticity	12
What is a molecular marker?	13
Milestones in molecular ecology	15
Early days of molecular ecology	16
Allozyme electrophoresis	16
Restriction fragment length polymorphism	17
Minisatellite DNA fingerprinting	19
The polymerase chain reaction	20
PCR-based molecular markers	22
DNA sequencing	24
Molecular ecology today	25
SUMMARY	26

2 Molecular biology for ecologists

Introduction	27
Nucleic acids and the common origin of life	28
The structures of DNA and RNA	29

Primary structure of nucleic acids	29
Secondary structure of nucleic acids	31
DNA replication	33
Protein structure	35
Basic immunology	38
The genetic code and gene expression	39
Genome structure: an overview	41
Non-coding DNA	43
Functional (coding) DNA	47
Ribosomal DNA	47
Nuclear structural (protein-coding) genes	48
Mitochondrial DNA	49
Chloroplast DNA	51
Plasmids and genetic manipulation in molecular ecology	52
Mutation	54
Somatic mutations	54
Germline mutations	54
DNA point mutations	55
Other types of mutation	57
Evolution and the mutation rate	57
SUMMARY	58

3 Molecular identification: species, individuals, and sex

Introduction	61
The species question	62
Defining distinctiveness	62
Hybrids	65
Dealing with individuals	69
Basic identification	69
Competition between toads	69
Woodland succession	70
Identification of prey in predator guts	70
Identification of disease vectors	71
Forensic investigations	72
Sex	73
Bits of individuals	75
Molecular identification methods: an appraisal	78
Protein analysis	78
DNA analysis	78
SUMMARY	80

4 Behavioural ecology

Introduction	81
From monogamy to promiscuity	82
Animal mating systems	82
Determining unknown mating systems	84
Male reproductive success	86
Sexual dimorphism	86
Lekking	87
Female reproductive success	88
Sperm competition	88
Why is female promiscuity adaptive?	90
Mate choice and the MHC	92
Sexual conflict	94
Sex ratio biases in offspring	95
Cooperative behaviour	96
Cooperative breeding in birds	96
Social insects	97
Cheating tactics	99
Interspecific brood parasitism	99
Intraspecific brood parasitism	100
Foraging and dispersal	101
Foraging	101
Dispersal	102
Behaviourally mediated speciation	104
SUMMARY	108

5 Population genetics

Introduction	111
Genetic diversity in natural populations	112
Some basic concepts	112
Population size and genetic diversity	113
Mammalian examples of population size and genetic diversity estimates	115
Population structure	117
Assessing where subdivisions occur	117
Statistical tests for population subdivision	118
Marker selection for the study of population subdivision	119
The genetics of metapopulations	122
Metapopulations	122

Testing whether metapopulations exist	123
Distinguishing different types of metapopulations	125
Gene flow and migration rates	127
Genetic estimation of gene flow and migration rates	127
Isolation by distance	129
Maximum likelihood methods for estimating migration rates	131
Identification of immigrants	132
Effective population size	133
Genetic estimators of effective population size	133
Determination of effective population size using allele frequency variation between generations	134
Alternative methods for N_e determination	136
Population bottlenecks	137
The significance of population bottlenecks	137
Genetic tests for population bottlenecks	138
Molecular markers for population genetics: an appraisal	140
The basis for choice	140
Fundamental tests	141
SUMMARY	142

6 Molecular and adaptive variation

Introduction	145
Neutral markers that are not really neutral	148
Allozymes	148
Nuclear DNA markers	150
Mitochondrial DNA	152
Heterozygosity and fitness	152
Background	152
Allozyme studies	153
DNA markers	155
Molecular approaches to understanding adaptive variation	157
Comparisons of neutral and adaptive variation	157
Variation at specific loci	158
Gene mapping	160
Quantitative traits and adaptive variation	161
Genomics and the study of adaptive variation	163
SUMMARY	164

7 Phylogeography

Introduction	165
Molecular markers in phylogeography	166
Early phylogeography	166
MtDNA as the standard phylogeographic tool	166
Alternatives to mtDNA	168
Genealogies and the coalescent process	169
Genetic variation in space	170
Geographic patterns in single populations	170
Vicariance and dispersal	173
Divergence between populations: Drift versus gene flow	174
Gene flow between species	175
Genetic consequences of the Pleistocene Ice Ages	176
Phylogeography and coevolution	179
Nested clade analysis	179
Genetic variation in time	181
Geological events and molecular divergence rates	181
Measuring lineage divergence in real time	183
Applied phylogeography	185
Taxonomic decisions	185
Determining a species natural range	187
Finding the source populations of introduced species	191
Phylogeography and adaptive traits	194
SUMMARY	196

8 Conservation genetics

Introduction	199
The state of the biosphere	200
Molecular genetics in conservation biology	202
Genetic diversity as a conservation issue	203
Inbreeding and genetic load	205
The basis of inbreeding effects	205
Inbreeding depression and genetic load in the wild	206
Purging of genetic load	208
Outbreeding depression	209
Genetic restoration	210
Desperate measures	211
Plant conservation	212

Animal conservation	214
Wildlife forensics	217
Genetics in conservation biology—a wider role	217
Systematics and conservation genetics	218
Phylogeography and conservation genetics	219
Globally important areas of genetic diversity	221
Molecular markers in conservation genetics	222
SUMMARY	222

9 Microbial ecology

Introduction	225
Outstanding issues in microbial ecology	227
The role of microbial communities in nature	227
Problems for molecular microbiology	230
Immunological approaches to microbial ecology	232
Immunological methods	232
Use of antibody identification in microbial ecology	233
Ribosomal genes and microbial ecology	236
Structure and usefulness of ribosomal genes	236
Identification: the use of sequence-specific oligonucleotide probes in microbial ecology	238
New developments, problems, and alternative approaches with probes	240
Ribosomal gene sequencing and community studies	242
Genetic profiling of microbial communities	244
Alternatives to ribosomal genes in profiling approaches	247
Genome analysis and microbial ecology	250
Sequence complexity and microbial diversity	250
Microarrays and microbial ecology	251
Other genomic approaches in microbial ecology	252
Whole genome separation and viral diversity	253
Overview of microbial molecular ecology	254
Molecular markers for microbial ecology: an appraisal	255
Protein methods	255
Nucleic acid methods	255
SUMMARY	256

10 Molecular ecology and genetically modified organisms

Introduction	259
Environmental risks from GMOs	261