Contents

1 The foundations of evolutionary genetics **1**

1.1 Concepts and ideas in a historical context 1
1.1.1 Evolutionary thoughts before Darwin 1
1.1.2 Evolution by natural selection—Darwin 2
1.1.3 Mendel and the origin of genetics 3
1.1.4 The mutation confusion 4
1.1.5 Theoretical population genetics 4
1.1.6 The modern synthesis of evolution 6
1.1.7 Toward molecular genetics 7
1.1.8 Diverging fields 10
1.2 Contemporary evolutionary genetics: toward a new synthesis 13
1.2.1 Molecular biology meets evolutionary genetics 13
1.2.2 The genotype–phenotype map: opportunities and challenges 14
1.3 Generating molecular genetic data 15
1.3.1 Gel electrophoresis 15
1.3.2 Polymerase chain reaction 15
1.3.3 Restriction enzyme digestion 16
1.3.4 Sanger sequencing 16
1.3.5 High-throughput sequencing 18
1.4 Analyzing genetic data 21
1.4.1 Statistics and bioinformatics 22
1.4.2 Internet resources 23
Study questions 24

2 Genomes and the origin of genetic variation **25**

2.1 Genome structure 25
2.1.1 Genome organization in bacteria, archaea, and eukarya 25
2.1.2 Structure of the eukaryote genome—functional elements 29
2.1.3 Structure of the eukaryote genome—parasitic and non-functional DNA 31
2.2 Mutations 35
2.2.1 Random mutations 35
2.2.2 Point mutations 35
2.2.3 Replication slippage 36

2.3 Recombination and associated mutations 38
2.3.1 Recombination and reshuffling of alleles 38
2.3.2 Gene conversion 38
2.3.3 Unequal crossing over 38
2.3.4 Concerted evolution 42
2.4 Origin of new genes and protein functions 43
2.4.1 Gene and genome duplication 43
2.4.2 Exon shuffling 44
2.4.3 Protein moonlighting, RNA editing, and alternative splicing 45
2.5 External sources of genetic variation 46
2.5.1 Introgression 46
2.5.2 Horizontal gene transfer 47
Study questions 47

3 Changes in allele and genotype frequency 49

3.1 The Hardy–Weinberg model 49
3.1.1 Assumptions 49
3.1.2 Testing for deviations from the Hardy–Weinberg expectation 51
3.2 Non-random mating 52
3.2.1 Inbreeding and outbreeding 52
3.2.2 Assortative and disassortative mating 54
3.2.3 Inbreeding/outbreeding versus assortative/disassortative mating 55
3.2.4 F_{IS}—a coefficient for measuring within-population deviations from HWE 55
3.3 Genetic drift—evolution in finite populations 56
3.3.1 Within-population effects of genetic drift 56
3.3.2 Effects of genetic drift in a structured metapopulation 58
3.3.3 F_{ST}—the fixation index 59
3.3.4 Empirical studies of population structure 60
3.4 Other deviations from the Hardy–Weinberg model 63
3.4.1 Selection and the Hardy–Weinberg model 63
3.4.2 Mutation and the Hardy–Weinberg model 63
3.4.3 Gene flow and the Hardy–Weinberg model 64
3.5 The gene pool 65
3.5.1 Hardy–Weinberg and real populations 65
3.5.2 Effective population size N_e 65
Study questions 67

4 The theory of natural selection 69

4.1 Fitness 69
4.1.1 Defining fitness 69
4.1.2 Absolute, relative, and marginal fitness 70
4.2 One-locus model of viability selection 71
4.2.1 The model 71
4.2.2 Finding all possible equilibria using Δp 72
4.2.3 Invasion fitness analysis 72
4.2.4 Adaptive landscapes 73

4.2.5 Directional selection 74
4.2.6 Overdominance and underdominance 77
4.3 Frequency-dependent selection 79
4.3.1 Negative and positive frequency-dependent selection 79
4.3.2 Relationship between balancing, disruptive, and frequency-dependent selection 80
4.3.3 Evolutionary game theory 82
4.4 Mutation–selection balance 84
4.4.1 Approximating the mutation–selection equilibrium 84
4.4.2 Genetic load 87
4.5 Selection at other stages of the life cycle 87
4.5.1 Sexual selection 87
4.5.2 Fertility selection 88
4.5.3 Selection at the gametic level 89
Study questions 90

5 The power of natural selection 92

5.1 The limits of natural selection 92
5.1.1 Genetic constraints 92
5.1.2 Historic and ontogenetic constraints 93
5.1.3 The problem of stasis in evolution 94
5.2 Level of selection 97
5.2.1 Level of sorting, adaptation, and immortal genes 97
5.2.2 Genetic conflict and gene level selection 98
5.2.3 Kin selection 99
5.2.4 Group selection 104
5.2.5 Individual adaptations to living in groups 105
5.3 Adaptation to a heterogeneous and unpredictable world 106
5.3.1 Phenotypic plasticity and learning 106
5.3.2 The paradox of sex 108
5.3.3 Sexual reproduction as an individual adaptation 109
5.3.4 Long–term advantages of sex and recombination 111
5.4 Evolving communities 113
5.4.1 Coevolution among mutualists 113
5.4.2 Coevolution among natural enemies 114
5.4.3 Coevolution between competitors 115
Study questions 115

6 Multilocus evolution 117

6.1 The two-locus model 117
6.1.1 Linkage equilibrium and disequilibrium 117
6.1.2 Estimating haplotype frequencies from genotype data 123
6.1.3 Factors that cause linkage disequilibrium 123
6.2 Multilocus selection 128
6.2.1 Additive and non-additive fitness 128
6.2.2 Epistasis and the evolutionary process 130

6.3 Quantitative genetics in the genomic era 132
6.3.1 The architecture of the genotype–phenotype map 132
6.3.2 Sources of phenotypic variance and the concept of heritability 132
6.3.3 Evolvability 135
6.3.4 Quantitative trait locus (QTL) analysis 136
6.3.5 Genome wide association study (GWAS) 141
6.3.6 The enigma of missing heritability 142
Study questions 143

7 Inferring evolutionary processes from DNA sequence data 145

7.1 The neutral theory of molecular evolution 145
7.1.1 The neutral theory of molecular evolution 145
7.1.2 Mutation, genetic drift, and the molecular clock 146
7.1.3 The generation and population size problem 149
7.1.4 The value of the neutral theory as a null model 149
7.2 Describing and interpreting variation in DNA sequences 150
7.2.1 DNA sequence versus allelic data 150
7.2.2 Descriptive statistics of DNA sequence variation 152
7.2.3 Coalescent theory and the expected pattern of sequence variation 153
7.2.4 Tajima's *D* test 155
7.2.5 Natural selection and the sign of Tajima's *D* 156
7.2.6 Demographic changes and the sign of Tajima's *D* 157
7.2.7 Descriptive statistics for sequence divergence between populations 159
7.3 Neutrality tests 159
7.3.1 The HKA test (Hudson–Kreitman–Aguadé test) 159
7.3.2 The McDonald–Kreitman (MK) test and its extensions 160
7.4 Genome scans 163
7.4.1 The logic of genome scans 163
7.4.2 Sliding window genome scans 164
7.4.3 Haplotype tests and selective sweeps 167
7.4.4 Limits to genome scans 170
Study questions 172

8 Genetics and genomics of speciation 173

8.1 Species concepts 173
8.1.1 What are species? 173
8.1.2 Typological species definitions 174
8.1.3 Species concepts that rely on biological properties 175
8.1.4 Species concepts that seek to incorporate the historical dimension of speciation 178
8.2 Evolution of barriers to gene exchange 181
8.2.1 Classifying barriers to gene exchange 181
8.2.2 Ecological incompatibilities 182
8.2.3 Assortative mating 183
8.2.4 Sexual incompatibility 187

8.2.5 Developmental incompatibility 190
8.2.6 Haldane's rule and the larger X/Z effect 193
8.3 Modes of speciation 194
8.3.1 Population structure and allelic models of divergence and speciation 194
8.3.2 Non-allopatric speciation 195
8.3.3 Genomic "islands of speciation" and the divergence hitchhiking hypothesis 199
8.3.4 Allopatric and non-allopatric divergence and the study of hybrid zones 202
8.3.5 The fate of hybridizing populations 203
8.3.6 Reinforcement and coupling of barriers to gene exchange 205
8.4 The role of hybridization in speciation 208
8.4.1 Polyploid hybrid speciation 208
8.4.2 Homoploid hybrid speciation 209
8.4.3 Hybridization and adaptive radiation 211
Study questions 212

9 Reconstructing the past 214

9.1 Phylogenetics and phylogenomics 214
9.1.1 The origins of tree thinking 214
9.1.2 Trees, phylogenies, and evolutionary relationships 216
9.1.3 The basis of phylogenetic inference 218
9.1.4 Alignment and models of sequence evolution 220
9.1.5 Building trees using distance and character data 222
9.1.6 Maximum likelihood tree estimation 224
9.1.7 Bayesian phylogenetic inference 225
9.1.8 Gene tree and species tree discordance 227
9.1.9 Phylogenetics at the genomic scale 229
9.2 Phylogeography—genes in time and space 233
9.2.1 Phylogeography and its link to modern population genetics 233
9.2.2 Phylogeography and biogeography 234
9.2.3 Detecting spatial population structure with genomic data 236
9.2.4 Testing demographic models and evolutionary scenarios 239
9.2.5 Ancient DNA—a direct insight to the past 242
9.2.6 Human origins: insights from genomic data 244
9.2.7 Using the four-population test to infer gene flow and introgression 246
Study questions 248

10 Sequencing the genome and beyond 250

10.1 Assembling and building on the genome 250
10.1.1 Reference genomes and genome assembly 250
10.1.2 Genome annotation, gene expression, and gene interactions 254
10.1.3 Proteomics—bridging the gap between gene expression and the phenotype 256
10.1.4 DNA methylation, chromatin binding, and epigenomics 257
10.1.5 The 3D structure of the genome 258

10.2 Manipulating the genome 259
10.2.1 From candidate genes to an understanding of function 259
10.2.2 CRISPR/Cas9 and the future of genetic manipulation 262
10.2.3 Prospects for genome editing in evolutionary genetics 264
10.3 The future of evolutionary genetics? 266
Study questions 267

Glossary 269
References 283
Index 305
Species index 309