

The Foundations of Population Genetics

Daniel M. Weinreich

**The MIT Press
Cambridge, Massachusetts
London, England**

Contents

Preface xi

1	Deterministic Single-Locus Population Genetics	1
1.1	Natural Selection	3
1.1.1	An Exponentially Growing Population	3
1.1.2	Darwin's Model of Natural Selection	7
1.1.3	The Time for a Selective Fixation	12
1.1.4	Fisher's Fundamental Theorem of Natural Selection	15
1.1.5	Natural Selection in Populations with Nonoverlapping Generations	18
1.2	Mutation	21
1.2.1	The Evolution of Allele Frequencies under Mutation Alone	21
1.2.2	Mutation/Selection Equilibrium	23
1.2.3	The Error Catastrophe	25
1.2.4	Beneficial Mutations	26
1.2.5	Soft Selection and Hard Selection	26
1.3	One-Locus Sexual Reproduction	28
1.3.1	The Population Genetics of Diploid Reproduction	30
1.3.2	Natural Selection in Diploids	35
1.4	Population Structure, Migration, and Nonrandom Mating	39
1.4.1	Models of Population Structure and Migration	40
1.4.2	Migration/Selection Equilibrium	41
1.4.3	Allele Frequency Clines	42
1.4.4	Nonrandom Mating in Diploids	43
1.4.4.1	Assortative and disassortative mating	44
1.4.4.2	Inbreeding and the probability of identity by descent	45
1.4.4.3	Population subdivision and the Wahlund effect	47
1.4.4.4	Nonrandom mating within a subdivided population	50
1.5	Chapter Summary	50
2	Stochastic Single-Locus Population Genetics	51
2.1	Stochasticity While Rare, and Establishment	53
2.1.1	The Establishment Problem	54
2.1.2	Establishment of a New Beneficial Allele	59
2.2	Stochasticity at All Frequencies and Random Genetic Drift	62
2.2.1	The Chapman–Kolmogorov Equation	63
2.2.2	The Diffusion Approximations of the Chapman–Kolmogorov Equation	67

2.2.2.1	Probability mass moves through time	68
2.2.2.2	Restricting movement of probability mass to adjacent stochastic states	69
2.2.2.3	Partitioning movement of probability mass into biased and unbiased components	69
2.2.2.4	Moving to continuous time and frequency	70
2.2.3	Representing Biology in the Chapman–Kolmogorov and Diffusion Equations	73
2.2.3.1	The Wright–Fisher model of random genetic drift	73
2.2.3.2	Incorporating biased processes into the Wright–Fisher model	76
2.2.3.3	The Moran model of random genetic drift	77
2.2.4	Using the Backward Diffusion Approximation to Study Fixation Events	79
2.2.4.1	The probability of fixation for a selectively neutral allele	79
2.2.4.2	The probability of fixation for a selected allele	82
2.2.4.3	Time to fixation for a new allele	87
2.2.5	Using the Forward Diffusion Approximation to Study Internal Equilibrium Allele Frequencies	88
2.3	Random Genetic Drift and Heterozygosity	91
2.3.1	The Rate of Decay in Heterozygosity under the Wright–Fisher Model	91
2.3.2	The Infinite Alleles Model and Heterozygosity at Mutation/Drift Equilibrium under the Wright–Fisher Model	94
2.3.3	Heterozygosity at Migration/Drift Equilibrium under the Wright–Fisher Model	97
2.4	Coalescent Theory	97
2.4.1	Modeling Genealogies under the Wright–Fisher Model	98
2.4.2	The Infinite Sites Model and the Site Frequency Spectrum	102
2.4.3	Population Subdivision and Incomplete Lineage Sorting	106
2.5	The Effective Size of a Population	108
2.6	Chapter Summary	112
3	Multilocus Population Genetics	115
3.1	Deterministic Multilocus Theory	117
3.1.1	Pairwise Linkage Disequilibrium and Genetic Recombination	117
3.1.2	The Two-Locus Wahlund Effect	121
3.1.3	Pairwise Epistasis, Natural Selection, and Recombination	123
3.1.4	Multilocus Mutation/Selection Equilibrium	128
3.1.5	The Multilocus Error Catastrophe	133
3.1.6	Sequence Space and Fitness Landscapes	136
3.1.7	Multilocus Adaptation on Fitness Landscapes under the Strong Selection/Weak Mutation Assumption	140
3.1.8	Stochastic Tunneling	144
3.2	Stochastic Multilocus Theory	147
3.2.1	Multilocus Coalescent Theory	148
3.2.1.1	The ancestral recombination graph	148
3.2.1.2	Recombination/drift equilibrium	150
3.2.2	Selective Sweeps and Genetic Draft	151
3.2.2.1	Recurrent selective sweeps without recombination	151
3.2.2.2	Coalescent time immediately after a selective sweep with recombination	154
3.2.2.3	Recurrent selective sweeps with recombination	157

3.2.3	Background Selection	159
3.2.4	The Hill–Robertson Effect	160
3.2.4.1	The Fisher–Muller effect	162
3.2.4.2	Muller’s ratchet	166
3.3	Modifier Theory	168
3.3.1	The Evolution of Sex	170
3.3.2	The Evolution of Mutation Rate	173
3.3.3	Other Modifiers	173
3.4	Chapter Summary	174
4	The Data	175
4.1	Measuring Genetic Parameters	177
4.1.1	Mutation Rate	178
4.1.2	Recombination Rate	180
4.2	Measuring Effective Population Size	181
4.2.1	Direct Observation of Reproductive Variance	182
4.2.2	Equilibrium Effective Population Size Estimates	182
4.2.3	Nonequilibrium Mutation-Based Effective Population Size Estimates in Asexuals	185
4.2.3.1	The distribution of pairwise Hamming distances	185
4.2.3.2	Skyline plots	188
4.2.4	Nonequilibrium Mutation-Based Effective Population Size Estimates in Sexuals	190
4.2.4.1	The site frequency spectrum	190
4.2.4.2	Tajima’s D statistic	191
4.2.4.3	The pairwise sequential Markovian coalescent	193
4.2.5	Nonequilibrium Recombination-Based Effective Population Size Estimates	195
4.3	Describing Population Structure, Migration, and Admixture	198
4.3.1	Detecting Population Structure	198
4.3.2	Estimating Population Structure with F_{ST}	200
4.3.3	Estimating Migration with the Joint Site Frequency Spectrum	201
4.3.4	Estimating Admixture from Haplotype Structure	203
4.4	Describing Natural Selection	205
4.4.1	Quantifying Natural Selection with the Molecular Clock	205
4.4.2	The McDonald/Kreitman Test	209
4.4.3	Scanning Whole Genome Sequence Data for Natural Selection	211
4.4.3.1	Detecting natural selection from heterozygosity	211
4.4.3.2	Detecting natural selection from the site frequency spectrum	212
4.4.3.3	Detecting natural selection from deme-specific frequency differences	213
4.4.3.4	Detecting natural selection from haplotype structure	213
4.4.4	Hard and Soft Selective Sweeps	214
4.5	Chapter Summary	217
	References	219
	Index	231