
Contents

1	Basic Genetics	1
1.1	Introduction	1
1.2	Genes and Chromosomes	1
1.3	Meiosis	2
1.4	Mendel's Laws	3
1.4.1	Mendel's First Law	3
1.4.2	Mendel's Second Law	4
1.5	Linkage and Mapping	5
1.6	Interference	8
1.7	Quantitative Genetics	8
1.7.1	Population Properties of Genes	8
1.7.2	A General Quantitative Genetic Model	9
1.7.3	Genetic Models for the Backcross and F ₂ Design	10
1.7.4	Epistatic Model	12
1.7.5	Heritability and Its Estimation	12
1.7.6	Genetic Architecture	14
1.7.7	The Estimation of Gene Number	15
1.8	Molecular Genetics	16
1.9	SNP	18
1.10	Exercises	20
1.11	Note	20
1.11.1	Modeling Unequal Genetic Effects by the Gamma Function	21
1.11.2	Modeling Unequal Genetic Effects by a Geometric Series	22
2	Basic Statistics	25
2.1	Introduction	25
2.1.1	Populations and Models	25
2.1.2	Samples	27
2.2	Likelihood Estimation	27
2.3	Hypothesis Testing	31
2.3.1	The Pearson Chi-Squared Test	31

2.3.2	Likelihood Ratio Tests	33
2.3.3	Simulation-Based Approach.....	37
2.3.4	Bayesian Estimation	38
2.4	Exercises	40
3	Linkage Analysis and Map Construction	43
3.1	Introduction	43
3.2	Experimental Design	44
3.3	Mendelian Segregation	45
3.3.1	Testing Marker Segregation Patterns.....	45
3.4	Segregation Patterns in a Full-Sib Family	46
3.5	Two-Point Analysis	49
3.5.1	Double Backcross.....	50
3.5.2	Double Intercross- F_2	52
3.6	Three-Point Analysis	56
3.7	Multilocus Likelihood and Locus Ordering	58
3.8	Estimation with Many Loci	61
3.9	Mixture Likelihoods and Order Probabilities	62
3.10	Map Functions	63
3.10.1	Mather's Formula	64
3.10.2	The Morgan Map Function	65
3.10.3	The Haldane Map Function.....	65
3.10.4	The Kosambi Map Function	66
3.11	Exercises	69
3.12	Notes: Algorithms and Software for Map Construction	71
4	A General Model for Linkage Analysis in Controlled Crosses	77
4.1	Introduction	77
4.2	Fully Informative Markers: A Diplotype Model	78
4.2.1	Two-Point Analysis.....	78
4.2.2	A More General Formulation	81
4.2.3	Three-Point Analysis	82
4.2.4	A More General Formulation	86
4.3	Fully Informative Markers: A Genotype Model	88
4.3.1	Parental Diplotypes	88
4.4	Joint modeling of the Linkage, Parental Diplotype, and Gene Order ..	91
4.5	Partially Informative Markers	96
4.5.1	Joint modeling of the Linkage and Parental Diplotype.....	96
4.5.2	Joint modeling of the Linkage, Parental Diplotype, and Gene Order	98
4.6	Exercises	99
4.6.1	One dominant marker.....	100
4.6.2	One dominant marker and one F_2 codominant marker.....	100
4.7	Notes	101
4.7.1	Linkage Analysis	101

4.7.2	The Diplotype Probability	102
4.7.3	Gene Order	105
4.7.4	M-Point Analysis	106
5	Linkage Analysis with Recombinant Inbred Lines	107
5.1	Introduction	107
5.2	RILs by Selfing	107
5.2.1	Two-Point Analysis	107
5.2.2	Three-Point Analysis	111
5.3	RILs by Sibling Mating	117
5.3.1	Two-point Analysis	117
5.3.2	Three-point Analysis	118
5.4	Bias Reduction	118
5.4.1	RILs by Selfing	118
5.4.2	RILs by Sibling Mating	119
5.5	Multiway RILs	120
5.6	Exercises	120
5.7	Note	122
6	Linkage Analysis for Distorted and Misclassified Markers	123
6.1	Introduction	123
6.2	Gametic Differential Viability	123
6.2.1	One-Gene Model	123
6.2.2	Two-Gene Model	127
6.2.3	Simulation	130
6.3	Zygotic Differential Viability	132
6.3.1	One-Gene Model	132
6.3.2	Two-Gene Model	133
6.3.3	Simulation	134
6.4	Misclassification	134
6.4.1	One-gene Model	134
6.4.2	Two-Gene Model	138
6.5	Simulation	141
6.6	Exercises	142
7	Special Considerations in Linkage Analysis	145
7.1	Introduction	145
7.2	Linkage Analysis with a Complicated Pedigree	145
7.2.1	A Nuclear Family	145
7.2.2	Multipoint Estimation of Identical-By-Descent Sharing	149
7.2.3	A Complex Pedigree	150
7.3	Information Analysis of Dominant Markers	160
7.3.1	Introduction	160
7.3.2	Segregation Analysis	161
7.3.3	Linkage Analysis	165

7.4	Exercises	168
8	Marker Analysis of Phenotypes	171
8.1	Introduction	171
8.2	QTL Regression Model	172
8.3	Analysis at the Marker	174
8.3.1	Two-Sample t Test	174
8.3.2	Analysis of Variance	176
8.3.3	Genetic Analysis	179
8.4	Moving Away from the Marker	181
8.4.1	Likelihood	181
8.5	Power Calculation	184
8.6	Marker Interaction Analysis	188
8.6.1	ANOVA	188
8.6.2	Genetic Analysis	192
8.7	Whole-Genome Marker Analysis	195
8.8	Exercises	198
9	The Structure of QTL Mapping	203
9.1	Introduction	203
9.2	The Mixture Model	204
9.2.1	Formulation	204
9.2.2	Structure, Setting, and Estimation	206
9.3	Population Genetic Structure of the Mixture Model	207
9.3.1	Backcross/ F_2	207
9.3.2	Outbred Crosses	208
9.3.3	Recombinant Inbred Lines	208
9.3.4	Natural Populations	208
9.4	Quantitative Genetic Structure of the Mixture Model	208
9.4.1	Additive-Dominance Model	209
9.4.2	Additive-Dominance-Epistasis Model	210
9.4.3	Multiplicative-Epistatic Model	211
9.4.4	Mechanistic Model	212
9.5	Experimental Setting of the Mixture Model	213
9.6	Estimation in the Mixture Model	214
9.7	Computational Algorithms for the Mixture Model	216
9.7.1	EM Algorithm	216
9.7.2	Monte Carlo EM	216
9.7.3	Stochastic EM	217
9.7.4	An EM Algorithm/Newton-Raphson Hybrid	217
9.7.5	Some Cautions	218
9.7.6	Bayesian Methods	218
9.7.7	Estimating the Number of Components in a Mixture Model	220
9.8	Exercises	221

10	Interval Mapping with Regression Analysis	223
10.1	Introduction	223
10.2	Linear Regression Model	224
10.3	Interval Mapping in the Backcross	224
10.3.1	Conditional Probabilities	224
10.3.2	Conditional Regression Model	226
10.3.3	Estimation and Test	228
10.4	Interval Mapping in an F_2	230
10.5	Remarks	233
10.6	Exercises	234
11	Interval Mapping by Maximum Likelihood Approach	237
11.1	Introduction	237
11.2	QTL Interval Mapping in a Backcross	238
11.2.1	The Likelihood	238
11.2.2	Maximizing the Likelihood	240
11.3	Hypothesis Testing	246
11.3.1	Model for Incorporating Double Recombination	252
11.3.2	Model for Incorporating Interference	252
11.4	QTL Interval Mapping in an F_2	254
11.4.1	No Double Recombination	254
11.4.2	Independence	257
11.4.3	Interference	260
11.4.4	Testing Hypotheses	262
11.5	Factors That Affect QTL Detection	262
11.6	Procedures for QTL Mapping	263
11.6.1	The Number of QTLs	263
11.6.2	Locations of Individual QTLs	266
11.7	Exercises	267
12	Threshold and Precision Analysis	269
12.1	Introduction	269
12.2	Threshold Determination	270
12.2.1	Background	270
12.2.2	Analytical Approximations	271
12.2.3	Simulation Studies	275
12.2.4	Permutation Tests	275
12.2.5	A Quick Approach	276
12.2.6	A Score Statistic	277
12.3	Precision of Parameter Estimation	279
12.3.1	Asymptotic Variance-Covariance Matrix	279
12.3.2	Simulation Studies	282
12.4	Confidence Intervals for the QTL Location	283
12.5	Exercises	285

13 Composite QTL Mapping	287
13.1 Introduction	287
13.2 Composite Interval Mapping for a Backcross	288
13.2.1 The Likelihood	288
13.2.2 Maximizing the Likelihood	289
13.2.3 Hypothesis Testing	290
13.3 Composite Interval Mapping for an F_2	291
13.4 A Statistical Justification of Composite Interval Mapping	293
13.4.1 Conditional Marker (Co)variances	293
13.4.2 Conditional QTL Variance	295
13.4.3 Marker Selection	298
13.5 Comparisons Between Composite Interval Mapping and Interval Mapping	298
13.6 Multiple Interval Mapping	301
13.7 Exercises	302
14 QTL Mapping in Outbred Pedigrees	303
14.1 Introduction	303
14.2 A Fixed-Effect Model for a Full-Sib Family	304
14.2.1 Introduction	304
14.2.2 A Mixture Model for a Parental Diplotype	304
14.2.3 Quantitative Genetic Model	308
14.2.4 Likelihood Analysis	309
14.2.5 Fitting Marker Phenotypes	312
14.2.6 Hypothesis Tests	313
14.2.7 The Influence of Linkage Phases	316
14.3 Random-Effect Mapping Model for a Complicated Pedigree	317
14.3.1 Introduction	317
14.3.2 Statistical Model	319
14.3.3 IBD at a QTL	320
14.3.4 The Likelihood	321
14.3.5 Hypothesis Testing	323
14.4 Exercises	325
A General Statistical Results and Algorithms	331
A.1 Likelihood Asymptotics	331
A.2 General Form of the EM Algorithm	332
B R Programs	335
B.1 Chapter 2	335
B.2 Chapter 8	337
C References	343
Author Index	355
Subject Index	361