

# Contents

<i>Foreword to the First Edition</i>	<i>vii</i>
<i>Foreword to the Second Edition</i>	<i>viii</i>
<i>Preface</i>	<i>xi</i>
<i>Acknowledgments</i>	<i>xv</i>
<b>1</b> <i>Molecular Genetics</i>	<b>1</b>
1.1 <i>Genetic information</i>	2
1.1.1 <i>Location of genetic information</i>	2
1.1.2 <i>Interpretation of genetic information</i>	5
1.1.3 <i>Translation of genetic information</i>	5
1.2 <i>Transmission of genetic information</i>	7
1.3 <i>Variations in genetic information</i>	10
1.3.1 <i>Individual differences in genetic information</i>	10
1.3.2 <i>Detection of variations</i>	12
1.3.3 <i>Probability for detection of variations</i>	16
1.4 <i>Problems</i>	18
<b>2</b> <i>Formal Genetics</i>	<b>21</b>
2.1 <i>Mendel and his laws</i>	22
2.2 <i>Segregation patterns</i>	23
2.2.1 <i>Autosomal dominant inheritance</i>	24
2.2.2 <i>Autosomal recessive inheritance</i>	25
2.2.3 <i>X-chromosomal dominant inheritance</i>	26

2.2.4	<i>X-chromosomal recessive inheritance</i>	27
2.2.5	<i>Y-chromosomal inheritance</i>	28
2.3	<i>Complications of Mendelian segregation</i>	28
2.3.1	<i>Variable penetrance and expression</i>	29
2.3.2	<i>Age-dependent penetrance</i>	31
2.3.3	<i>Imprinting</i>	33
2.3.4	<i>Phenotypic and genotypic heterogeneity</i>	35
2.3.5	<i>Complex diseases</i>	36
2.4	<i>Hardy–Weinberg law</i>	38
2.5	<i>Problems</i>	43
3	<i>Genetic Markers</i>	47
3.1	<i>Properties of genetic markers</i>	47
3.2	<i>Types of genetic markers</i>	52
3.2.1	<i>Short tandem repeats (STRs)</i>	52
3.2.2	<i>Single nucleotide polymorphisms (SNPs)</i>	54
3.3	<i>Genotyping methods for SNPs</i>	57
3.3.1	<i>Restriction fragment length polymorphism analysis</i>	58
3.3.2	<i>Real-time polymerase chain reaction</i>	58
3.3.3	<i>Matrix assisted laser desorption/ionization time of flight genotyping</i>	61
3.3.4	<i>Chip-based genotyping</i>	61
3.3.5	<i>Choice of genotyping method</i>	63
3.4	<i>Problems</i>	65
4	<i>Data Quality</i>	67
4.1	<i>Pedigree errors</i>	68
4.2	<i>Genotyping errors in pedigrees</i>	70
4.2.1	<i>Frequency of genotyping errors</i>	70
4.2.2	<i>Reasons for genotyping errors</i>	71
4.2.3	<i>Mendel checks</i>	72
4.2.4	<i>Checks for double recombinants</i>	74
4.3	<i>Genotyping errors and Hardy–Weinberg equilibrium (HWE)</i>	76
4.3.1	<i>Causes of deviations from HWE</i>	77
4.3.2	<i>Tests for deviation from HWE for SNPs</i>	78
4.3.3	<i>Tests for deviation from HWE for STRs</i>	81
4.3.4	<i>Measures for deviation from HWE</i>	83
4.3.5	<i>Tests for compatibility with HWE for SNPs</i>	86

4.4	<i>Quality control in high-throughput studies</i>	91
4.4.1	<i>Sample quality control</i>	94
4.4.2	<i>SNP quality control</i>	97
4.5	<i>Cluster plot checks and internal validity</i>	98
4.5.1	<i>Cluster compactness measures</i>	101
4.5.2	<i>Cluster connectedness measures</i>	101
4.5.3	<i>Cluster separation measures</i>	101
4.5.4	<i>Genotype stability measures</i>	102
4.5.5	<i>Combinations of criteria</i>	102
4.6	<i>Problems</i>	109
5	<i>Genetic Map Distances</i>	113
5.1	<i>Physical distance</i>	113
5.2	<i>Map distance</i>	114
5.2.1	<i>Distance</i>	114
5.2.2	<i>Specific map functions</i>	115
5.2.3	<i>Correspondence between physical distance and map distance</i>	116
5.2.4	<i>Multilocus feasibility</i>	117
5.3	<i>Linkage disequilibrium distance</i>	118
5.4	<i>Problems</i>	123
6	<i>Family Studies</i>	125
6.1	<i>Family history method and family study method</i>	127
6.2	<i>Familial correlations and recurrence risks</i>	129
6.2.1	<i>Familial resemblance</i>	129
6.2.2	<i>Recurrence risk ratios</i>	131
6.3	<i>Heritability</i>	134
6.3.1	<i>The simple Falconer model</i>	135
6.3.2	<i>The general Falconer model</i>	137
6.3.3	<i>Kinship coefficient and Jacquard's <math>\Delta_7</math> coefficient</i>	138
6.4	<i>Twin and adoption studies</i>	141
6.4.1	<i>Twin studies</i>	141
6.4.2	<i>Adoption studies</i>	142
6.5	<i>Critique on investigating familial resemblance</i>	143
6.6	<i>Segregation analysis</i>	144
6.7	<i>Problems</i>	154

7	<i>Model-Based Linkage Analysis</i>	155
7.1	<i>Linkage analysis between two genetic markers</i>	156
7.1.1	<i>Linkage analysis in phase-known pedigrees</i>	156
7.1.2	<i>Linkage analysis in phase-unknown pedigrees</i>	160
7.1.3	<i>Linkage analysis in pedigrees with missing genotypes</i>	161
7.2	<i>Linkage analysis between a genetic marker and a disease</i>	167
7.2.1	<i>Linkage analysis between a genetic marker and a disease in phase-known pedigrees</i>	168
7.2.2	<i>Linkage analysis between a genetic marker and a disease in general cases</i>	172
7.2.3	<i>Gain in information by genotyping additional individuals; power calculations</i>	177
7.3	<i>Significance levels in linkage analysis</i>	180
7.4	<i>Problems</i>	184
8	<i>Model-Free Linkage Analysis</i>	189
8.1	<i>The principle of similarity</i>	190
8.2	<i>Mathematical foundation of affected sib-pair analysis</i>	192
8.3	<i>Common tests for affected sib-pair analysis</i>	193
8.3.1	<i>The maximum LOD score and the triangle test</i>	194
8.3.2	<i>Score- and Wald-type 1 degree of freedom tests</i>	201
8.3.3	<i>Affected sib-pair tests using alleles shared identical by state</i>	206
8.4	<i>Properties of affected sib-pair tests</i>	206
8.5	<i>Sample size and power calculations for affected sib-pair studies</i>	207
8.5.1	<i>Functional relation between identical by descent probabilities and recurrence risk ratios</i>	207
8.5.2	<i>Sample size and power calculations for the mean test using recurrence risk ratios</i>	209
8.6	<i>Extensions to multiple marker loci</i>	212
8.7	<i>Extension to large sibships</i>	213
8.8	<i>Extension to large pedigrees</i>	214
8.9	<i>Extensions of the affected sib-pair approach</i>	216
8.9.1	<i>Covariates in affected sib-pair analyses</i>	216
8.9.2	<i>Multiple disease loci in affected sib-pair analyses</i>	216

8.9.3	<i>Estimating the position of the disease locus in affected sib-pair analyses</i>	217
8.9.4	<i>Typing unaffected relatives in sib-pair analyses</i>	217
8.10	<i>Problems</i>	218
9	<i>Quantitative Traits</i>	221
9.1	<i>Quantitative versus qualitative traits</i>	222
9.2	<i>The Haseman–Elston method</i>	223
9.2.1	<i>The expected squared phenotypic difference at the trait locus</i>	225
9.2.2	<i>The expected squared phenotypic difference at the marker locus</i>	227
9.3	<i>Extensions of the Haseman–Elston method</i>	229
9.3.1	<i>Double squared trait difference</i>	230
9.3.2	<i>Extension to large sibships</i>	230
9.3.3	<i>Haseman–Elston revisited and the new Haseman–Elston method</i>	231
9.3.4	<i>Power and sample size calculations</i>	234
9.4	<i>Variance components models</i>	237
9.4.1	<i>The univariate variance components model</i>	237
9.4.2	<i>The multivariate variance components model</i>	238
9.5	<i>Random sib-pairs, extreme probands and extreme sib-pairs</i>	240
9.6	<i>Empirical determination of p-values</i>	243
9.7	<i>Problems</i>	245
10	<i>Fundamental Concepts of Association Analyses</i>	247
10.1	<i>Introduction to association</i>	247
10.1.1	<i>Principles of association</i>	247
10.1.2	<i>Study designs for association</i>	249
10.2	<i>Linkage disequilibrium</i>	250
10.2.1	<i>Allelic linkage disequilibrium</i>	250
10.2.2	<i>Genotypic linkage disequilibrium</i>	255
10.2.3	<i>Extent of linkage disequilibrium</i>	259
10.3	<i>Problems</i>	262
11	<i>Association Analysis in Unrelated Individuals</i>	265
11.1	<i>Selection of cases and controls</i>	266

11.2	<i>Tests, estimates, and a comparison</i>	266
11.2.1	<i>Association tests</i>	267
11.2.2	<i>Choice of a test in applications</i>	272
11.2.3	<i>Effect measures</i>	274
11.2.4	<i>Selection of the genetic model</i>	280
11.2.5	<i>Association tests for the X chromosome</i>	287
11.3	<i>Sample size calculation</i>	289
11.4	<i>Population stratification</i>	291
11.4.1	<i>Testing for population stratification</i>	293
11.4.2	<i>Structured association</i>	294
11.4.3	<i>Genomic control</i>	295
11.4.4	<i>Comparison of structured association and genomic control</i>	297
11.4.5	<i>Principal components analysis</i>	297
11.5	<i>Gene-gene and gene-environment interaction</i>	299
11.5.1	<i>Classical examples for gene-gene and gene-environment interaction</i>	299
11.5.2	<i>Coat color in the Labrador retriever</i>	301
11.5.3	<i>Concepts of interaction</i>	303
11.5.4	<i>Statistical testing of gene-environment interactions</i>	307
11.5.5	<i>Statistical testing of gene-gene interactions</i>	311
11.5.6	<i>Multifactor dimensionality reduction</i>	315
11.6	<i>Problems</i>	316
12	<i>Family-based Association Analysis</i>	319
12.1	<i>Haplotype relative risk</i>	320
12.2	<i>Transmission disequilibrium test (TDT)</i>	322
12.3	<i>Risk estimates for trio data</i>	325
12.4	<i>Sample size and power calculations for the TDT</i>	327
12.5	<i>Alternative test statistics</i>	329
12.6	<i>TDT for multiallelic markers</i>	330
12.6.1	<i>Test of single alleles</i>	330
12.6.2	<i>Global test statistics</i>	331
12.7	<i>TDT type tests for different family structures</i>	333
12.7.1	<i>TDT type tests for missing parental data</i>	334
12.7.2	<i>TDT type tests for sibship data</i>	336
12.7.3	<i>TDT type tests for extended pedigrees</i>	341

12.8	<i>Association analysis for quantitative traits</i>	344
12.9	<i>Problems</i>	346
13	<i>Haplotypes in Association Analyses</i>	349
13.1	<i>Reasons for studying haplotypes</i>	350
13.2	<i>Inference of haplotypes</i>	351
13.2.1	<i>Algorithms for haplotype assignment</i>	352
13.2.2	<i>Algorithms for estimating haplotype probabilities</i>	353
13.3	<i>Association tests using haplotypes</i>	356
13.4	<i>Haplotype blocks and tagging SNPs</i>	359
13.4.1	<i>Selection of markers by haplotypes or linkage disequilibrium</i>	360
13.4.2	<i>Evaluation of marker selection approaches</i>	363
13.5	<i>Problems</i>	364
14	<i>Genome-wide Association (GWA) Studies</i>	367
14.1	<i>Design options in GWA studies</i>	369
14.2	<i>Genotype imputation</i>	370
14.2.1	<i>Imputation algorithms</i>	370
14.2.2	<i>Quality of imputation</i>	371
14.3	<i>Statistical analysis of GWA studies</i>	372
14.4	<i>Multiple testing</i>	374
14.4.1	<i>Region-wide multiple testing adjustment by simulation</i>	375
14.4.2	<i>Genome-wide multiple testing adjustment by simulation</i>	376
14.4.3	<i>Multiple testing adjustment by effective number of tests</i>	377
14.5	<i>Analysis of accumulating GWA data</i>	378
14.5.1	<i>Multistage designs for GWA studies</i>	378
14.5.2	<i>Replication in GWA studies</i>	379
14.5.3	<i>Meta-analysis of GWA studies</i>	380
14.6	<i>Clinical impact of a GWA study</i>	383
14.6.1	<i>Evaluation of a genetic predictive test</i>	383
14.6.2	<i>Clinical validity of a single genetic marker</i>	385
14.6.3	<i>Clinical validity of multiple genetic markers</i>	386
14.7	<i>Outlook</i>	389
14.8	<i>Problems</i>	391

<i>Appendix</i>	
<i>Algorithms Used in Linkage Analyses</i>	393
A.1 <i>The Elston–Stewart algorithm</i>	394
A.1.1 <i>The fundamental ideas of the Elston–Stewart algorithm</i>	394
A.1.2 <i>The Elston–Stewart algorithm for a trait and a linked marker locus</i>	400
A.2 <i>The Lander–Green algorithm</i>	401
A.2.1 <i>The inheritance vector at a single genetic marker</i>	401
A.2.2 <i>The inheritance distribution given all genetic markers</i>	405
A.3 <i>The Cardon–Fulker algorithm</i>	412
A.4 <i>Problem</i>	414
<i>Solutions</i>	415
<i>References</i>	451
<i>Index</i>	489