



Genetics and Analysis of Quantitative Traits

Michael Lynch Bruce Walsh











Contents

CC	ONTENTS	i
PR	EFACE	xiii
I.	FOUNDATIONS OF QUANTITATIVE GENETICS	1
1.	AN OVERVIEW OF QUANTITATIVE GENETICS	3
	The Adaptationist Approach to Phenotypic Evolution	. 3
	Quantitative Genetics and Phenotypic Evolution	. 4
	Historical Background	. 7
	The Major Goals of Quantitative Genetics	13
	The nature of quantitative-trait variation	13
	The consequences of inbreeding and outcrossing	14
	The constraints on the evolutionary process	15
	The estimation of breeding values	15
	The development of predictive models for evolutionary change	16
	Mathematics in Biology	16
2.	PROPERTIES OF DISTRIBUTIONS	19
	Parameters of Univariate Distributions	19
	The Normal Distribution	26
	The truncated normal distribution	29
	Confidence Intervals	32
3.	COVARIANCE, REGRESSION, AND CORRELATION	35
	Jointly Distributed Random Variables	35
	Expectations of jointly distributed variables	36
	Covariance	36
	Useful identities for variances and covariances	38
	Regression	39
	Derivation of the least-squares linear regression	39
	Properties of least-squares regressions	41
	Correlation	43
	A Taste of Quantitative-Genetic Theory	45
	Directional selection differentials and the Robertson-Price identity	45
	The correlation between genotypic and phenotypic values	47
	Regression of offspring phenotype on parental phenotype	48

ii CONTENTS

4.	PROPERTIES OF SINGLE LOCI	51
	Allele and Genotype Frequencies	52
	The Transmission of Genetic Information	54
	The Hardy-Weinberg principle	54
	Sex-linked loci	56
	Polyploidy	57
	Age structure	60
	Testing for Hardy-Weinberg proportions	60
	Characterizing the Influence of a Locus on the Phenotype	63
	The Basis of Dominance	65
	Fisher's Decomposition of the Genotypic Value	60
	Partitioning the Genetic Variance	71
	Additive Effects, Average Excesses, and Breeding Values	74
	Extensions to Multiple Alleles and Nonrandom Mating	74
	Average excess	75
	Additive effects	76
	Additive genetic variance	
5.	SOURCES OF GENETIC VARIATION FOR	
0.	MULTHOCUS TRAITS	81
	Enistacis	82
	A Coneral Least-Squares Model for Genetic Effects	85
	Extension to haploids and polyploids	92
	Linkage	94
	Estimation of gametic phase disequilibrium	97
	Effect of Disequilibrium on the Genetic Variance	100
	The evidence	103
		107
6.	COMPONENTS OF ENVIRONMENTAL VARIATION	107
	Extension of the Linear Model to Phenotypes	108
	Special Environmental Effects	111
	Within-individual variation	112
	Developmental homeostasis and homozygosity	116
	Repeatability	121
	General Environmental Effects of Maternal Influence	123
	Genotype × Environment Interaction	127
7	RESEMBLANCE BETWEEN RELATIVES	131
	Measures of Relatedness	132
	Coefficients of identity	133
	Coefficients of coancestry and inbreeding	135
	The coefficient of fraternity	140
	The Genetic Covariance Between Relatives	141
	The Effects of Linkage and Gametic Phase Disequilibrium	146

	Linkage	146
	Gametic phase disequilibrium	150
	Assortative Mating	153
	Polyploidy	161
	Environmental Sources of Covariance Between Relatives	162
	The Heritability Concept	170
	Evolvability	175
8.	INTRODUCTION OF MATRIX ALGEBRA	
	AND LINEAR MODELS	177
	Multiple Regression	177
	An application to multivariate selection	180
	Elementary Matrix Algebra	182

Elementary Matrix Algebra	182
Basic notation	182
Partitioned matrices	183
Addition and subtraction	183
Multiplication	184
Transposition	186
Inverses and solutions to systems of equations	187
Determinants and minors	189
Computing inverses	190
Expectations of Random Vectors and Matrices	192
Covariance Matrices of Transformed Vectors	193
The Multivariate Normal Distribution	194
Properties of the MVN	195
Overview of Linear Models	198
Ordinary least squares	200
Generalized least squares	202
ANALYSIS OF LINE CROSSES	205
Expectations for Line-cross Means	206

9.

Expectations for Line-cross Means	206
Estimation of Composite Effects	213
Hypothesis testing	215
Line crosses in Nicotiana rustica	219
Additional data	221
The Genetic Interpretation of Heterosis and Outbreeding Depression	222
Variance of Line-cross Derivatives	226
Biometrical Approaches to the Estimation of Gene Number	231
The Castle-Wright estimator	233
Effect of the leading factor	238
Extensions to haploids	241
Other Biometrical Approaches to Gene Number Estimation	244
The inbred-backcross technique	244
Genotype assay	246

CONTENTS iv

10.	INBREEDING DEPRESSION	251
	The Genetic Basis of Inbreeding Depression	252
	A more general model	256
	Methodological Considerations	259
	Single-generation analysis	260
	Multigenerational analyses	262
	Ritland's method	266
	Epistasis and inbreeding depression	267
	Variance in inbreeding depression	268
	The Evidence	269
	Purging Inbreeding Depression	274
	Number of Lethal Equivalents	276
	Results from vertebrates	278
	Results from Drosophila	2/9
	Results from plants	281
	Partial Recessives vs. Overdominance	200
	The $(A+B)/A$ ratio	285
	Estimating the average degree of dominance	204
	Inferences from molecular markers	207
11	MATTERS OF SCALE	293
11.	Transformations to Achieve Normality	293
	Log normal distributions and the log transform	294
	Tests for normality	295
	Ctshilizing the Variance	300
	Klockowski's transformation	300
	Coneral variance stabilizing-transformations	301
	The Roginskii-Yablokov effect	302
	The Kluge-Kerfoot phenomenon	305
	Allometry: the Scaling Implications of Body Size	305
	Removing Interaction Effects	307
	Developmental Maps, Canalization, and Genetic Assimilation	309
	Estimating developmental maps	310
	Selection and canalization	314
	Genetic assimilation	316
т	ΟΠΑΝΤΙΤΑΤΙΛΕ ΤΒΑΙΤΙ ΟΟΙ	319
11.	QUANTITATIVE TRAIT LOCI	
12.	POLYGENES AND POLYGENIC MUTATION	321
	The Genetic Basis of Quantitative-Genetic Variation	322
	Maior genes and isoalleles	322
	The molecular nature of QTL variation	323
	The Mutational Rate of Production of Quantitative Variation	328

	Estimation from divergence experiments	330
	Bristle numbers in Drosophila	333
	Additional data	335
	The Deleterious Effects of New Mutations	340
	The Bateman-Mukai technique	341
	Results from flies, plants, and bacteria	343
	Analysis of natural populations	348
	The persistence of new mutations	351
13.	DETECTING MAJOR GENES	353
	Elementary Tests	354
	Departures from normality	354
	Tests based on sibship variances	355
	Major-gene indices (MGI)	357
	Nonparametric line-cross tests	358
	Mixture Models	359
	The distribution under a mixture model	360
	Parameter estimation	360
	Hypothesis testing	361
	Complex Segregation Analysis	364
	Likelihood functions assuming a single major gene	366
	Common-family effects	370
	Polygenic background	371
	Other extensions	373
	Ascertainment bias	374
	Estimating individual genotypes	374
	Analysis of Discrete Characters	375
	Single-locus penetrance model	376
	Major gene plus a polygenic background	377
14.	PRINCIPLES OF MARKER-BASED ANALYSIS	379
	Classical Approaches	379
	Chromosomal assays	380
	Thoday's method	381
	Genetics of Drosophila bristle number	385
	Genetics of Drosophila speciation	387
	Molecular Markers	390
	Genetic Maps	393
	Map distance vs. recombination frequencies	394
	How many markers are needed?	397
	Marker-trait Associations	398
	Selective genotyping and progeny testing	401
	Recombinant inbred lines (RILs)	401
	Bulked segregant analysis	402

	QTL mapping by marker changes in populations under selection	404
	Marker-based Analysis Using Nearly Isogenic Lines (NILs)	405
	Marker-based introgressions	407
	Fine Mapping of Major Genes Using Population-level Disequilibrium	413
	LD mapping in expanding populations	414
	Candidate Loci	418
	The transmission / disequilibrium test	419
	Estimating effects of candidate loci	422
	Templeton and Sing's method: Using the historical	10.1
	information in haplotypes	424
	Cloning QTLs	425
	Transposon tagging	426
	Positional cloning and comparative mapping	426
15.	MAPPING AND CHARACTERIZING QTLS:	
	INBRED LINE CROSSES	431
	Foundations of Line-Cross Mapping	431
	Experimental designs	432
	Conditional probabilities of QTL genotypes	433
	Expected marker-class means	437
	Marker variance and higher-order moments	439
	Overall significance level with multiple tests	441
	QTL Detection and Estimation Using Linear Models	442
	QTL Detection and Estimation via Maximum Likelihood	445
	Likelihood maps	446
	Precision of ML estimates of QTL position	448
	ML interval mapping	450
	Approximating ML interval mapping by Haley-Knott regressions	453
	Dealing with Multiple QTLs	457
	Marker-difference regression	459
	Interval mapping with marker cofactors	463
	Detecting multiple linked QTLs using standard marker-trait regressions	467
	Sample Size Required for QTL Detection	469
	Power under selective genotyping	4/4
	Power and repeatabilty of mapping experiments	4/4
	Selected Applications	4//
	The nature of transgressive segregation	4//
	QTLs involved in reproductive isolation in <i>Mimulus</i>	478
	QTLs involved in protein regulation	478
	QTLs in the Illinois long-term selection maize lines	4/9
	QTLs involved in the differences between maize and teosinte	481
	QTLs for age-specific growth in mice	484
	Summary of OTL mapping experiments	484

vi

16.	MAPPING AND CHARACTERIZING QTLS:	
	OUTBRED POPULATIONS	491
	Measures of Informativeness	492
	Sib Analysis: Linear Models	495
	A single half-sib family	496
	Several half-sib families	498
	Power of Nested ANOVA Designs	501
	A single full-sib family	502
	Several full-sib families	504
	Sib Analysis: Maximum Likelihood	505
	Constructing likelihood functions	507
	Maximum Likelihood over General Pedigrees: Variance Components	510
	Estimating QTL position	512
	The Haseman-Elston Regression	513
	Derivation of the Haseman-Elston regression	513
	Estimating the number of marker genes ibd	516
	Power and improvements	517
	Interval mapping by a modified Haseman-Elston regression	518
	Mapping Dichotomous Characters	521
	Recurrent and relative risks of pairs of relatives	523
	Affected sib-pair tests	525
	Power of ASP tests and related issues	527
	Genomic scanning	529
	Exclusion mapping and information content mapping	530
	Affected pedigree member tests	532
III.	ESTIMATION PROCEDURES	535
17.	PARENT-OFFSPRING REGRESSION	537
	Estimation Procedures	538
	Balanced data	538
	Unequal family sizes	539
	Standardization of data from the different sexes	542
	Precision of Estimates	542
	Optimum Experimental Designs	543
	Assortative mating	547
	Estimation of Heritability in Natural Populations	548
	Linearity of the Parent-Offspring Regression	550
18.	SIB ANALYSIS	553
	Half-sib Analysis	554
	One-way analysis of variance	556
	Hypothesis testing	560

viii CONTENTS

	Sampling variance and standard errors	561
	Confidence intervals	562
	Negative estimates of heritability	563
	Optimal experimental design	564
	Unbalanced data	566
	Resampling procedures	569
	Full-sib Analysis	570
	Nested analysis of variance	573
	Hypothesis testing	574
	Sampling error	576
	Optimal design	577
19.	TWINS AND CLONES	581
	The Classical Approach	582
	Heritability estimation	584
	The Monozygotic-Twin Half-sib Method	587
	Clonal Analysis	592
20.	CROSS-CLASSIFIED DESIGNS	597
	North Carolina Design II	598
	The average degree of dominance	603
	The Cockerham-Weir model	605
	Diallels	610
	Pooled reciprocals, no self crosses	611
	Reciprocals, no self crosses	614
	Complete diallels	618
	Partial diallels	618
	Havman-Jinks analysis	619
	North Carolina Design III and the Triple Test Cross	625
	Some Closing Statistical Considerations	627
21.	CORRELATIONS BETWEEN CHARACTERS	629
	Theoretical Composition of the Genetic Covariance	630
	Estimation of the Genetic Covariance	632
	Pairwise comparison of relatives	632
	Nested analysis of variance and covariance	633
	Regression of family means	636
	Components of Phenotypic Correlation	637
	Phenotypic correlations as surrogate estimates of genetic correlations	639
	Statistical Issues	639
	Hypothesis tests	641
	Standard errors	642
	Bias due to selection	644
	Applications	648

	Genetic basis of population differentiation	648 650
	Evolutionary allometry	653
	Evolution of life-history characters	655
22.	$\textbf{GENOTYPE} \times \textbf{ENVIRONMENT INTERACTION}$	657
	Genetic Correlation Across Two Environments	660
	Estimation procedures	663
	Iwo-way Analysis of Variance	666
	Relationship to Falconer's correlation across environments	671
	Further Characterization of Interaction Effects	672
	Joint-regression analysis	672
	lesting for Cross-over Interaction	678
	Concepts of Stability and Plasticity	680
	Additional issues	682
	The Quantitative Genetics of Genotype × Environment Interaction	683
23.	MATERNAL EFFECTS	687
	Components of Variance and Covariance	689
	Cytoplasmic transmission	693
	Postpollination reproductive traits in plants	695
	Cross-fostering experiments	696
	Body weight in mice	700
	Eisen's Approach	703
	Bondari's experiment	703
	Falconer's Approach	706
	Extension to Other Types of Relatives	711
24.	SEX LINKAGE AND SEXUAL DIMORPHISM	715
	Sex-linked Loci and Dosage Compensation	715
	Sex-modified Expression of an Autosomal Locus	718
	Gametic imprinting	718
	Extension to Multiple Loci and the Covariance Between Relatives	719
	Variation for Sexual Dimorphism	724
25.	THRESHOLD CHARACTERS	727
	Heritability on the Underlying Scale	730
	Multiple Thresholds	736
	Genetic Correlation Among Threshold Traits	739
	Heritability on the Observed Scale	741
26.	ESTIMATION OF BREEDING VALUES	745
	The General Mixed Model	746
	Estimating Fixed Factors and Predicting Random Effects	748

x CONTENTS

	Estimability of fixed factors	753
	Standard errors	754
	Models for the Estimation of Breeding Values	755
	The animal model	755
	The gametic model	758
	The reduced animal model	759
	Simple Rules for Computing A and A ⁻¹	762
	Allowing for mutation when computing A	766
	Ioint Estimation of Several Vectors of Random Effects	767
	BLUP estimates of dominance values	767
	Repeated records	769
	Maternal effects	773
	Multiple traits	774
	I	
27.	VARIANCE-COMPONENT ESTIMATION WITH	
	COMPLEX PEDIGREES	779
	ML versus REML Estimates of Variance Components	780
	A simple example of ML versus REML	781
	ML Estimates of Variance Components in the General Mixed Model	784
	Standard errors of ML estimates	788
	Restricted Maximum Likelihood	789
	Multivariate analysis	792
	ML/REML estimation in populations under selection	792
	Solving ML/REML Equations	793
	Derivative-based methods	794
	EM methods	797
	Additional approaches	799
	A Molecular-marker Based Method for Inferring Variance Components	800
	I more and a second sec	

IV. APPENDICES

805

A1.	EXPECTATIONS, VARIANCES, AND COVARIANCES	
	OF COMPOUND VARIABLES	807
	The Delta Method	807
	Expectations of complex variables	808
	Variances of complex variables	810
	Covariances of complex variables	813
	Variance of Variances and Covariances	813
	Expectations and Variances of Products	817
	Expectations and Variances of Ratios	818
	Sampling variance of regression and correlations coefficients	818
	Sampling variance of a coefficient of variation	819

A2.	PATH ANALYSIS	823
	Univariate Analysis	823
	Bivariate Analysis	826
	Applications	826
	Phenotypic correlation between parents and offspring	827
	Correlations between characters	829
	Growth analysis	831
A3.	FURTHER TOPICS IN MATRIX ALGEBRA AND	0.75
	LINEAR MODELS	835
	Generalized Inverses and Solutions to Singular Systems of Equations	835
	Generalized inverses	836
	Consistency and solutions to consistent systems	836
	Estimability of fixed factors	839
	The Square Root of a Matrix	841
	Derivation of the GLS Estimators	842
	Quadratic Forms and Sums of Squares	843
	Moments of quadratic forms	843
	The sample variance as a quadratic form	844
	Sums of squares expressed as a quadratic form	846
	Testing Hypotheses About Linear Models	848
	Equivalent Linear Models	849
	Derivatives of Vectors and Matrices	851
A4.	MAXIMUM LIKELIHOOD ESTIMATION AND	
	LIKELIHOOD-RATIO TESTS	853
	Likelihood, Support, and Score Functions	853
	Large-sample properties of MLEs	854
	The Fisher information matrix	855
	Likelihood-ratio tests	857
	The G-test	859
	Likelihood-ratio tests for the general linear model	860
	Iterative Methods for Solving ML Equations	861
	Newton-Raphson methods	861
	Expectation-maximization methods	863
	EM for mixture model likelihoods	863
	EM modifications for QTL mapping	865
Δ5	COMPUTING THE POWER OF STATISTICAL TESTS	860
10.	Power of Nermally Distributed Test Statistics	870
	One sided tests	870
	Two sided tests	070
	Applications: Depend offension recording	072
	Applications: Farent-onspring regressions	0/4