

# *Evolutionary Quantitative Genetics*

Bruce Walsh. [jbwalsh@u.arizona.edu](mailto:jbwalsh@u.arizona.edu). University of Arizona.

*Notes from a short course taught June 2006 at University of Aarhus*

*The notes for this lecture were last corrected on 23 June 2006. Please email me any errors.*

<b>I. Basic Statistical Machinery</b>	<b>1.1</b>
Probabilities, Distributions, and Expectations	1.1
Discrete and Continuous Random Variables	1.1
Joint and Conditional Probabilities	1.2
Bayes' Theorem	1.2
Expectations of Random Variables	1.3
The Normal Distribution	1.4
Covariances	1.4
The Normal Distribution	1.5
Regressions	1.6
Properties of Least-squares Regressions	1.6
Regressions and Correlations	1.7
Maximum Likelihood	1.7
Large-sample Properties of MLEs	1.8
Likelihood Ratio Tests	1.8
Bayesian Statistics	1.9
Problems	1.10
Solutions to Problems	1.11
<b>2. Linear Algebra and Linear Models</b>	<b>2.1</b>
Elementary Matrix Algebra	2.1
Basic Matrix Notation	2.1
Partitioned Matrices	2.1
Addition and Subtraction	2.2
Multiplication	2.2
Dimensional Properties and Matrix Multiplication	2.4
Transposition	2.5
Inverses and Solutions to Systems of Equations	2.5
Expectations of Random Vectors and Matrices	2.7
Covariance Matrices of Transformed Vectors	2.8
The Multivariate Normal Distribution	2.9
Properties of the MVN	2.10
Example: The Regression of Offspring Value on Parental Value	2.11
Regression of Offspring Breeding Value on Parental Breeding Values	2.12
Overview of Linear Models	2.12
Examples of GLMs	2.13
Ordinary Least Squares	2.14
Example: GLM Solution for Regression Through the Origin	2.15

ii CONTENTS

Example: Partial (or Multiple) Regression	2.15
Polynomial Regressions and Interaction Effects	2.16
Fixed vs. Random Effects	2.17
Example: Fixed vs. Random Effects in the Sire Model	2.17
Generalized Least Squares	2.18
Example: Weighted Least Squares	2.18
Model Goodness-of-fit and Hypothesis Testing	2.19
Chi-square and F-distributions	2.19
Sums of Squares	2.20
Hypothesis Testing	2.21
Problems	2.23
Solutions to Problems	2.24

**3. Basic Concepts in Mendelian, Population and Quantitative Genetics**

	<b>3.1</b>
Overview	3.1
A Tale of Two Papers: Darwin vs. Mendel	3.1
Basic Mendelian Genetics	3.2
Mendel's View of Inheritance: Single Locus	3.2
The Genotype to Phenotype Mapping: Dominance and Epistasis	3.3
Mendel's View of Inheritance: Multiple Loci	3.3
Mendel was Wrong: Linkage	3.4
Interlude: Chromosomal Theory of Inheritance	3.5
Map Distances are Obtained from Recombination Frequencies via Mapping Functions	3.5
The Prior Probability of Linkage and Morton's Posterior Error Rate	3.7
Molecular Marker	3.7
Basic Population Genetics	3.8
Allele and Genotype Frequencies	3.8
Gamete Frequencies, Linkage, and Linkage Disequilibrium	3.9
The Effects of Population Structure	3.10
Forces that Change Allele Frequencies: Genetic Drift	3.11
Coalescence Theory	3.11
Forces that Change Allele Frequencies: Mutation	3.12
Forces that Change Allele Frequencies: Selection	3.13
Interaction of Selection and Drift	3.13
Basic Quantitative Genetics	3.14
Dichotomous (Binary) Traits	3.14
Contribution of a Locus to the Phenotypic Value of a Trait	3.15
Example: Apolipoprotein E and Alzheimer's age of onset	3.16
Example: the Booroola ( <i>B</i> ) gene	3.16
Fisher's Decomposition of the Genotypic Value	3.16
Average Effects and Additive Genetic Values	3.17
Genetic Variances	3.18
Epistasis	3.19

Problems	3.20
Solutions to Problems	3.21
<b>4. Resemblance Between Relatives</b>	<b>4.1</b>
Phenotypic Resemblance Between Relatives	4.2
Parent-offspring regressions	4.2
Collateral relationships: ANOVAs	4.3
Causes of Phenotypic Covariance Among Relatives	4.4
The Genetic Covariance Between Relatives	4.4
Offspring and one parent	4.5
Half-sibs	4.5
Full-sibs	4.6
Environmental Causes of Relationship Between Relatives	4.7
Complex Relationships in Pedigrees: Coefficients of Coancestry	4.7
Coefficients of Coancestry and Inbreeding	4.8
The Coefficient of Fraternity	4.11
Genetic Correlations for General Relationships	4.12
Problems	4.14
Solutions to Problems	4.15
<b>5. Basic Designs for Estimation of Genetic Parameters</b>	<b>5.1</b>
Heritability	5.1
Why $h^2$ instead of $h$ ?	5.1
Heritabilities are Functions of a Population	5.1
Increasing the Heritability	5.1
Heritability and the Prediction of Breeding Values	5.1
Heritability Values and Population Divergence	5.2
Broad-Sense Heritability $H^2$ and Plant Breeding	5.2
Estimation: One-way ANOVA and the simple full-sib design	5.3
Estimating Variances and Variance Components	5.4
Estimating Heritability	5.5
Worked Example of a Full-sib Design	5.5
Estimation: The Nested Full-sib, Half-sib Analysis	5.5
Worked Example of a Nested Design	5.8
Estimation: Parent-Offspring Regressions	5.8
Natural Population Parent-Laboratory Offspring Regressions	5.10
Estimation of $\sigma_A^2$ and Breeding Values in General Pedigrees	5.10
The General Mixed Model	5.10
Estimating Fixed Effects and Predicting Random Effects	5.12
The Animal Model	5.14
ANOVA vs. REML Variance Estimates	5.15
ML Versus REML Variance Estimates	5.16
Problems	5.17
Solutions to Problems	5.18

<b>6. Inbreeding and Crossbreeding</b>	<b>6.1</b>
Changes In The Mean And Variance Under Inbreeding	6.1
Inbreeding	6.1
Inbreeding Depression in Fitness Traits	6.2
Computing the Inbreeding Depression Coefficient, $B$	6.2
Why do Traits Associated With Fitness Show Inbreeding Depression?	6.3
Minimizing the Rate of Inbreeding	6.3
Change of Variance With Inbreeding	6.3
Example: Effect of Inbreeding on the Additive Genetic Variance	6.4
Change of Variance with Inbreeding and Mutation	6.5
Hybridization And Crossbreeding	6.5
Types of Crosses	6.6
Heterosis: Change in the Mean Under Crossbreeding	6.7
Heterosis in $F_2$	6.8
Agricultural Importance of Heterosis	6.8
Hybrid Corn	6.8
Hybridization in Other Crops	6.9
Crossbreeding in Animals: General Concepts	6.10
Crossbreeding in Animals: Heterosis in <i>Bos indicus</i> $\times$ <i>taurus</i> Hybrids	6.11
Synthetics and Rotational Crossbreedings	6.11
Estimating the Amount of Heterosis in Maternal Effects	6.13
Problems	6.16
Solutions to Problems	6.17
<b>7. Genetic Drift</b>	<b>7.1</b>
Single Locus Theory Of Random Genetic Drift	7.1
The Wright-Fisher Model of Genetic Drift	7.1
Inbreeding and Drift	7.3
The Coalescent Process for Drift	7.3
The Coalescent for a Sample	7.5
Drift Generates Between-Population Variance in Replicate Lines	7.6
Effective Population Size, $N_e$	7.7
$N_e$ With Variable Population Size	7.7
$N_e$ With Unequal Sex Ratio	7.8
$N_e$ With Unequal Offspring Contribution	7.8
Mutation-Drift Equilibrium: Single Loci	7.8
Within-population Equilibrium Variance	7.8
Divergence Between Populations	7.10
Kimura's Neutral Theory of Molecular Evolution	7.11
Mutation-Drift Equilibrium: Additive Variance	7.13
Equilibrium Additive Variance within a Population	7.13
Divergence Between Populations	7.15
Diffusion Theory	7.15
Infinitesimal Means and Variances, $m(x)$ and $v(x)$	7.16
Stationary Distributions	7.16

Probability of Fixation	7.18
Diffusion Applications to Quantitative Characters	7.19
Brownian Motion Model of Evolution	7.19
Drift and Divergence in the Fossil Record	7.20
Ornstein-Uhlenbeck Model of Evolution	7.22
Problems	7.24
Solutions to Problems	7.25
<b>8. Tests for Molecular Signatures of Selection</b>	<b>8.1</b>
Basic Logic of Sequence-Based Selection Tests	8.1
Logic Behind Polymorphism-Based Tests	8.1
Recombination and Polymorphism	8.2
Logic Behind Divergence Tests	8.3
Logic Behind Joint Polymorphism and Divergence Tests	8.3
Tests Based Strictly on Within-Population Variation	8.4
The Infinite Alleles Model: Ewen's Sampling Formula	8.5
The Infinite Sites Model	8.5
Tajma's $D$ Test	8.6
Fu and Li's $D^*$ and $F^*$ tests	8.7
Depaulis and Veuille's $K$ and $H$ tests	8.7
Fu's $W$ and $F_S$ Tests	8.8
Fay and Wu's Test	8.9
Genome-Wide Polymorphism Tests	8.9
The Ghost of Lewontin-Krakauer: Genome Wide $F_{ST}$ -based Scans	8.9
The Linkage Disequilibrium Decay (LDD) Test	8.10
Joint Polymorphism and Divergence Tests	8.10
McDonald-Kreitman Test	8.10
Hudson-Kreitman-Aquade (HKA) Test	8.11
Tests Based on Between-Population (Species) Divergence	8.12
Parsimony-Based Ancestral Reconstruction Tests	8.13
Likelihood-Based Codon Tests	8.13
Bayesian Estimator of Sites Under Positive Selection	8.14
<b>9. Short-term Selection Response</b>	<b>9.1</b>
Changes in the Mean: the Breeder's Equation and its Extensions	9.1
The Selection Differential $S$ and Response $R$	9.1
The Selection Intensity $i$	9.1
The Breeders' Equation: Translating $S$ into $R$	9.2
The Generalized Breeders' Equation: Accuracy	9.3
The Generalized Breeders' Equation: Generation Intervals	9.4
Truncation Selection	9.5
Selection Intensities and Differentials Under Truncation Selection	9.5
Correcting the Selection Intensity for Finite Samples	9.6
Selection on Threshold Traits	9.8
Permanent Versus Transient Response	9.11

Response with Epistasis	9.11
Maternal Effects: Response Under Falconer's Model	9.12
Gene Frequency Changes Under Selection	9.14
Short-term Changes in the Variance	9.15
Bulmer's Equation for the Change in Variance	9.15
Change in Variance Under Truncation Selection	9.17
Problems	9.20
Solutions to Problems	9.21
<b>10. Analysis of Short-term Selection Experiments</b>	<b>10.1</b>
Variance in Response	10.1
Realized Heritabilities	10.2
Estimators for Several Generations of Selection	10.3
Weighted Least-Squares Estimates of Realized Heritability	10.4
Standard Errors for Realized Heritability Estimates	10.5
Experimental Evaluation of the Breeders' Equation	10.7
Asymmetric Selection Response	10.8
Control Populations and Experimental Designs	10.9
Basic Theory of Control Populations	10.10
Divergent Selection Designs	10.11
Variance in Response	10.11
Control Populations and Variance in Response	10.12
Optimal Experimental Design	10.12
Nicholas' Criterion	10.14
Mixed-model Estimation	10.14
The Relationship Matrix Accounts for Drift and Disequilibrium	10.17
Model Validation	10.18
Problems	10.20
Solutions to Problems	10.21
<b>11. Long-Term Response and Selection Limits</b>	<b>11.1</b>
Idealized Long-Term Response In A Large Population	11.1
Deterministic Single-Locus Theory	11.2
An Overview Of Long-Term Selection Experiments	11.6
Estimating Selection Limits and Half-Lives	11.6
General Features of Long-Term Selection Experiments	11.7
Increases In Variances And Accelerated Responses	11.8
Conflicts Between Natural And Artificial Selection	11.9
Characterizing The Nature Of Selection Limits	11.11
Long-Term Response In Finite Population	11.12
Fixation Probabilities of Favorable QTL Alleles	11.12
Limits Under Drift and Selection	11.13
Robertson's Theory of Selection Limits	11.14
Tests of Robertson's Theory	11.15
Weber's Selection Experiment on <i>Drosophila</i> Flight Speed	11.17

Variance In Response	11.18
Response From Mutational Input	11.19
Contributions from New Mutation	11.19
Mutational Response Under the Infinitesimal Model	11.19
Problems	11.23
Solutions to Problems	11.24
<b>12. Individual Fitness and Measures of Univariate Selection</b>	<b>12.1</b>
Episodes Of Selection	12.1
Fitness Components	12.1
Assigning Fitness Components	12.2
Variance In Individual Fitness	12.3
Some Caveats in Using the Opportunity for Selection	12.6
Describing Phenotypic Selection: Introductory Remarks	12.6
Fitness Surfaces	12.7
Describing Phenotypic Selection: Changes In Phenotypic Moments	12.8
The Price-Robertson Identity	12.8
Directional Selection: Differentials ( $S$ ) and Gradients ( $\beta$ )	12.9
Quadratic Selection: Differentials ( $C$ ) and Gradients ( $\gamma$ )	12.10
$\beta$ and $\gamma$ Measure the Geometry of the Fitness Surface	12.11
$\beta$ and $\gamma$ Describe the Selection Dynamics	12.12
Describing Phenotypic Selection: Individual Fitness Surfaces	12.12
Linear and Quadratic Approximations of $W(z)$	12.13
Schluter's Cubic-Spline Estimate	12.15
Complications from Unmeasured Correlated Variables	12.15
Strenght of Selection in Natural Populations	12.16
The Importance of Experimental Manipulation	12.17
Problems	12.18
Solutions to Problems	12.19
<b>13. Genetic Correlations and Multivariate Selection Response</b>	<b>13.1</b>
Phenotypic And Genetic Correlations	13.1
Estimating the Genetic Correlation	13.2
Half sibs	13.2
Parent-Offspring	13.2
Correlated Response to Selection	13.3
Correlated Selection Differentials	13.3
Indirect Selection	13.4
General Multivariate Selection Response	13.4
The Directional Selection Gradient	13.5
Derviatiion of the Multivariate Breeders' Equation	13.5
Realized Selection Gradients	13.7
The Effects of Genetic Correlations: Direct and Correlated Responses	13.7
Realized Genetic Correlations	13.8
Evolutionary Constraints Imposed by Genetic Correlations	13.8

A Short Diversion: The Geometry of a Matrix	13.9
Comparing Vectors: Lengths and Angles	13.9
Matrices Describe Vector Transformations	13.11
Eigenvalues and Eigenvectors	13.11
Quantifying Multivariate Constraints to Response	13.13
Is There Genetic Variation in the Multivariate Direction of Selection?	13.13
Schluter's Genetic Line of Least Resistance, $\mathbf{g}_{\max}$	13.14
Blow's Matrix Subspace Projection	13.15
General Theory of Selection on a Linear Index	13.17
The Smith-Hazel Index	13.18
MAS – Marker Assisted Selection	13.18
Indirect Selection on Marker Score: Applications to Sex-limited Traits	13.21
Problems	13.23
Solutions to Problems	13.24
<b>14. Measuring Multivariate Selection</b>	<b>14.1</b>
Selection On Multivariate Phenotypes: Differentials And Gradients	14.1
Changes in the Mean Vector: The Directional Selection Differential $\mathbf{S}$	14.1
The Directional Selection Gradient $\beta$	14.2
Directional Gradients, Fitness Surface Geometry and Selection Response	14.3
Changes in the Covariance Matrix: The Quadratic Selection Differential, $\mathbf{C}$	14.4
The Quadratic Selection Gradient $\gamma$	14.5
Quadratic Gradients, Fitness Surface Geometry and Selection Response	14.7
Summary	14.9
Multidimensional Quadratic Fitness Regressions	14.9
Estimation, Hypothesis Testing and Confidence Intervals	14.9
Geometric Aspects	14.11
A Brief Digression: Orthonormal and Diagonalized Matrices	14.11
Canonical Transformation of $\gamma$	14.13
Strength of Selection: $\gamma_{ii}$ versus $\lambda$	14.15
Canonical Subspaces of $\gamma$	14.15
Unmeasured Characters and Other Biological Caveats	14.16
The Bias due to Environmental Correlations Between Fitness and Characters	14.16
Problems	14.19
Solutions to Problems	14.20
<b>15. Phenotypic Evolution Models</b>	<b>15.1</b>
Univariate Gaussian Fitness Functions	15.1
Response Under Normalizing Selection	15.1
Application: Charlesworth's Model of the Cost of Phenotypic Selection	15.2
Application: Drift vs. Selection in the Fossil Record	15.4
Multivariate Gaussian Fitness Functions	15.5
Phenotypes and Breeding Values Remain MVN After Selection	15.5

The Selection Differential and Gradient	15.7
The Within-Generation Change in $G$	15.7
Changes in $G$ Under the Infinitesimal Model	15.8
Long-term Response: Balance Between Directional and Stabilizing Selection	15.10
Long-term Response: The Infinitesimal Model with Drift and Mutation	15.10
Derivatives Of Vectors And Vector-Valued Functions	15.11
The Hessian Matrix, Local Maxima/Minima, and Multidimensional Taylor Series	15.13
Problems	15.15
Solutions to Problems	15.16
<b>16. Major Genes, Polygenes, and QTLs</b>	<b>16.1</b>
Major and Minor Genes	16.1
Major Genes and Isoalleles	16.1
Polygenic Mutation and the Mutational Variance, $\sigma_m^2$	16.2
Simple Tests for Detecting Major Genes	16.2
Mixture Models	16.3
Complex Segregation Analysis	16.4
Likelihood Functions Assuming a Single Major Gene	16.4
Common-family Effects	16.7
Genetic Maps and Candidate Genes	16.7
Map Distances vs. Recombination Frequencies	16.8
Linkage Disequilibrium Mapping	16.9
Fine-mapping Major Genes Using LD	16.9
Example of LD Fine Mapping: Diastrophic Dysplasia	16.9
The Transmission/Disequilibrium Test, TDT	16.10
Mapping Type 1 Diabetes	16.11
Linkage vs. Association	16.12
Dense SNP Association Mapping	16.12
Genomic Control	16.13
Structured Association Analysis	16.14
Regression Approaches	16.14
Problems	16.16
Solutions to Problems	16.17
<b>17. QTL Mapping</b>	<b>17.1</b>
Mapping Using Inbred Line Crosses	17.1
Experimental Designs	17.1
Conditional Probabilities of QTL Genotypes	17.1
Example: Conditional Probabilities for an $F_2$	17.2
Expected Marker Means	17.3
Linear Models for QTL Detection	17.4
Maximum Likelihood Methods for QTL Mapping and Detection	17.5

Likelihood Maps	17.5
Precision of ML Estimates of QTL Position	17.6
Interval Mapping with Marker Cofactors	17.7
Power and Repeatability: The Beavis Effect	17.8
Mapping In Outbred Populations	17.9
QTL Mapping Using Sib Families	17.10
General Pedigree Methods	17.12
Haseman-Elston Regressions	17.13
Affected Sib Pair Methods	17.14
Problems	17.16
Solutions to Problems	17.17

**18. Quantitative Analysis of Regulatory Variation 18.1**

Gene Regulation is a Complex Trait	18.1
QTLs Involved in Protein Regulation	18.1
Microarrays	18.2
A Brief Overview Of The Technology	18.2
Analysis of Microarray Data	18.3
Microarray Analysis Is Best Regarded As An EDA Approach	18.4
Problems (and Pitfalls) of Gene Discovery via Microarray Analysis	18.5
General Patterns of Transcriptional Variation	18.6
Gene Expression Levels are Typically Highly Heritable	18.6
Correlations Between Rates of Regulatory (Transcriptional) and Sequence Divergence	18.7
Correlations Between Regulatory Divergence and Expression Level/Pattern	18.8
Does Divergence in Expression Follow a Neutral Model?	18.8
Analysis of Pathways	18.9
Two-hybrid screen: Construction Protein-protein interaction maps	18.10
Flux and Pathways	18.10
Kascer-Burns Sensitivity Analysis	18.11
Regulatory Networks and Graph Theory	18.12
Erdos-Renyi Random Graphs and Random Boolean Networks	18.13
Graphs: Small Worlds, Scale-Free, and Power Laws	18.15