An Introduction to Quantitative Genetics II
Heather A Lawson
Advanced Genetics
Spring 2020

Recap...
- What is Quantitative Genetics?
- Genotypic Values and Genetic Effects
- Heritability
- Linkage and Linkage Disequilibrium

Outline
- QTL Mapping
- Genetic Architecture
  - Pleiotropy
  - Relationship to QTL
  - Epistasis
  - GxE
- Taking the next steps
  - Molecular basis

Quantitative Genetics
- The theory of the statistical relationship between genotypic variation and phenotypic variation.
  1. What is the cause of phenotypic variation in natural populations?
  2. What is the genetic architecture and molecular basis of phenotypic variation in natural populations?

Quantitative Trait Loci (QTL)
- Statistically links two types of information
  1. Phenotypic data
     - Variation in a measurable trait (e.g. height, weight)
  2. Genotypic data
     - Variation at molecular markers
- Attempt to explain genetic basis (including genetic effects) of variation in complex traits

QTL Limitations
The population used defines the genetic variation
- We cannot find loci that are not variable
- Location: the actual gene may be far away
- Some loci we find might be confounded (correlation with other traits)
- Interaction effects: if the effect of a locus differs due to interactions with other loci, environment or phenotype, it can cancel out
- Genetic background!
**General Experimental Method**

- Population
- Measure phenotype
- Extract DNA
- Genotype individuals
- Impute between Markers

**Phenotype: univariate description (one trait at a time)**

**Single locus genotypic values:** phenotypic means of the genotype classes with respect to a single trait

Example:

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>G(AA)</td>
<td>38g</td>
</tr>
<tr>
<td>G(Aa)</td>
<td>31g</td>
</tr>
<tr>
<td>G(aa)</td>
<td>18g</td>
</tr>
</tbody>
</table>

\[
\begin{pmatrix}
L_L + S_S \\
L_L - S_S \\
L_S + S_L \\
L_S - S_L
\end{pmatrix}
\]

**Linear Models**

- Try to explain a dependent variable $y$ as a linear function of a number of independent (predictor) variables
- A multiple regression is a typical linear model

\[
y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + e
\]

- Here $e$ is the residual or deviation between the true value observed and the value predicted by the linear model
- The regression coefficients are interpreted as a unit change in $x_i$ while holding all other variables constant results in a change of $\beta_i$ in $y$

**Find the equation of the straight line that would provide the best fit for the data points**

**QTL mapping: correlating univariate phenotypic variation with genotypic variation**

Analysis of variance (ANOVA):

Is the difference between genotypes greater than the variation within each genotype?
Intercrossing Leads to Accumulation of Recombination

QTL Mapping: correlating phenotypic variation with genotypic variation
QTL: region of the genome affecting quantitative traits

What happens on the phenotypic level?

Traits covary because they are affected by common genes

A gene affects many phenotypes = pleiotropy

Pleiotropy and linkage disequilibrium can account for the heritable correlation between traits
Modularity of genetic effects

Ubiquitous pleiotropy vs Modular pleiotropy

DMetS1b: A Pleotropic QTL With Context-Dependent Genetic Effects

Ubiquitous pleiotropy

Phenotypic structure can be characterized by the average relationships between pairs of traits in a population (not just the means and variances of single traits)

Variance/covariance matrix

<table>
<thead>
<tr>
<th></th>
<th>Trait1</th>
<th>Trait2</th>
<th>Trait3</th>
<th>Trait4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait1</td>
<td>Var 1</td>
<td>Cov (1, 2)</td>
<td>Cov (1, 3)</td>
<td>Cov (1, 4)</td>
</tr>
<tr>
<td>Trait2</td>
<td>Cov (2, 1)</td>
<td>Var 2</td>
<td>Cov (2, 3)</td>
<td>Cov (2, 4)</td>
</tr>
<tr>
<td>Trait3</td>
<td>Cov (3, 1)</td>
<td>Cov (3, 2)</td>
<td>Var 3</td>
<td>Cov (3, 4)</td>
</tr>
<tr>
<td>Trait4</td>
<td>Cov (4, 1)</td>
<td>Cov (4, 2)</td>
<td>Cov (4, 3)</td>
<td>Var 4</td>
</tr>
</tbody>
</table>

Phenotypes: multivariate description
It follows, we could look at the effect of a single locus on multivariate phenotype:

**Relationship QTL:**
The effect of a QTL on T2 depends on a value of T1

How can slopes vary?

Regression
\[ T2 = a + \beta T1 + e \]

Slope = \( \beta \) = \( \frac{\text{cov}(T1, T2)}{\text{var}(T1)} \)

Why important? Example: relationship between traits affects shape, physiology, growth, etc.

Some intertrait relationships are quite constant. . .

. . . others seem inaccessible. . .
Genetic variation in covariance is variation in pleiotropy.

Loss of pleiotropy
In this genotype, the traits are independent.

This is the kind of genetic variation that is needed to select on modularity of traits, i.e., to evolve modularity.

Epistasis: Variation in phenotype due to interaction between genetic loci.

What’s happening at the phenotypic level:
Two locus genotypic values: phenotypic means of the 9 genotype classes with respect to a single trait.

Example:
- G(AABB) = 40g
- G(AABb) = 38g
- G(AAbb) = 36g
- G(AaBB) = 35g
- G(AaBb) = 30g
- G(Aabb) = 28g
- G(aaBB) = 20g
- G(aaBb) = 18g
- G(aabb) = 16g

Epistasis and Genotypic Values

Basic Model of Quantitative Genetics

\[ P_{ij} = \mu + G_i + E_j \]

Phenotypic Value = the value observed when a trait is measured on an individual
Mean = the mean phenotype for the entire population
Genotypic Value = the average phenotype of those carrying the specified genotype
Environmental Deviation = the deviation of the observed phenotype in an individual from the genotypic value

Genotype by Environment Interaction

- Genotypes respond differently across a range of environments
  \[ P = \mu + G + E + GE \]
  \[ V(P) = \mu^2 + V(G) + V(E) + V(GE) + 2 \times COV(GE) \]

Two different sources for GxE Interactions

1) Heterogeneity of genotypic variances across environments (aka change in scale)

2) Lack of perfect correlation among additive genotypic values across environments (aka change in rank)
1) Heterogeneity of genotypic variances across environments (aka change in scale)

\[ V(GE) = \frac{(V(A)_{e1} - V(A)_{e2})^2}{2} + V(A)_{e1} \cdot V(A)_{e2} \cdot (1 - r_e) \]

2) Lack of perfect correlation among additive genotypic values across environments (aka change in rank)

\[ V(GE) = \frac{(V(A)_{e1} - V(A)_{e2})^2}{2} + V(A)_{e1} \cdot V(A)_{e2} \cdot (1 - r_e) \]

No GxE Interaction

Mean Phenotype

Environment 1  Environment 2

LL  LL

SS  SS

GxE Interaction due to change of scale

Mean Phenotype

Environment 1  Environment 2

LL  LL

SS  SS

GxE Interaction due to change in ranking

Mean Phenotype

Environment 1  Environment 2

LL  SS

SS  LL
QTL Lessons

- The bulk of genetic variation for quantitative traits is due to many loci of small effect sizes
- Many QTL for complex traits do not map to obvious genes
  - Novel associations!
- Many QTL are context dependent
  - Gene X Environment
  - Pleiotropy is pervasive
  - There are no “genes for” specific traits
- Traits in “modules” covary

Whole Genome Sequences Can Inform Quantitative Trait Gene Identification

Identifying Variants in Apoa2 To Follow-up

QTG to Candidate Quantitative Trait Nucleotide

- SNPs in non-HD regions
- SNP in exon of Pprs2
- Involved in bone formation
- P→R
- Predicted to be functionally damaging
**Modified Quantitative Hybrid Complementation Test**

- LG
- SM
- KO
- WT

- LG
- LG
- KO
- WT
- WT
- KO
- SM
- SM

**Genetic Engineering of Candidate Gene(s)**

- Microinjection of DNA into zygotes
- TALEN, CRISPR
- Injection of embryos with recombinant virus
- Transfection of embryonic stem cells with cloned DNA

**Recap**

- QTL Mapping
- Genetic Architecture
  - Pleiotropy
  - Epistasis
  - GxE
  - Relationship QTL
- Taking the next steps
  - Molecular basis