An Introduction to Quantitative Genetics II
Heather A. Lawson
Advanced Genetics
Spring 2017

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

Outline
• QTL Mapping
• Genetic Architecture
  • Pleiotropy
  • Relationship 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 at Genomic Markers and 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>Phenotypic Value</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>

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
LL + SS \\
LL - SS \\
LS + SL - 2SL \\
\end{bmatrix}
\]

\[
y = X\beta + e
\]

- \(y\) = vector of observed dependent values
- \(X\) = Design matrix: observations of the variables in the assumed linear model
- \(\beta\) = vector of unknown parameters to estimate
- \(e\) = vector of residuals (deviation from model fit)
- \(e = y - X\beta\)
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

$$y = \mu + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + e$$

• 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 $X$ while holding all other variables constant results in a change of $\beta$ in $y$

---

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

- Linear regression equation
  $$Y = c + \beta_1 * X + \text{error}$$

- Multivariate case
  $$Y = c + \beta_1 * X + \beta_2 * Z + \text{error}$$
  (univariate case)
  (multivariate case)

---

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?
Currently, there is no textual content to convert into a natural text representation.
Intercrossing Leads to Accumulation of Recombination
(=> recombination == increased power to localize QTL)

Experimental Population
Advanced Intercross (AIL)

QTL Mapping: correlating phenotypic variation with genotypic variation

LOD Scores (log 10)
QTL Position

Significance
Support Region
QTL: region of the genome affecting quantitative traits

F₂ Experimental Population

F₁₀ Experimental Population
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 Pleiotropic QTL With Context-Dependent Genetic Effects

Log (L/B) against Chromosome 1 (Mb)

47 genes including: Rgs5, Rgs6, Hsd17b7, Atp2a2, Kcnj11, Fil, Usp1, Dredd, Nr5a1, Krap16

Additive Effects in Full Population for Cholesterol Level

Log (L/B) against Chromosome 1 (Mb)
Phenotypes: multivariate description

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)

<table>
<thead>
<tr>
<th>Variance/covariance matrix</th>
<th>Trait1</th>
<th>Trait2</th>
<th>Trait3</th>
<th>Trait4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait1</td>
<td>Var 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trait2</td>
<td>Cov (1,2)</td>
<td>Var 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trait3</td>
<td>Cov (2,1)</td>
<td>Cov (2,3)</td>
<td>Var 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trait4</td>
<td>Cov (1,4)</td>
<td>Cov (3,4)</td>
<td>Cov (4,3)</td>
<td>Var 4</td>
<td></td>
</tr>
</tbody>
</table>
It follows, we could look at the effect of a single locus on multivariate phenotype:

\[ \text{AA} \]
\[ \text{Aa} \]
\[ \text{aa} \]

\[ \text{Trait 1} \]
\[ \text{Trait 2} \]
\[ \text{Trait 3} \]
\[ \text{Trait 4} \]

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

\[ \text{Regression} \]
\[ T2 = \alpha + \beta \cdot T1 + e \]

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

How can slopes vary?
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 the 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:

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Phenotypic Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AABB</td>
<td>40g</td>
</tr>
<tr>
<td>AABb</td>
<td>38g</td>
</tr>
<tr>
<td>AAbb</td>
<td>36g</td>
</tr>
<tr>
<td>AaBB</td>
<td>35g</td>
</tr>
<tr>
<td>AaBb</td>
<td>30g</td>
</tr>
<tr>
<td>Aabb</td>
<td>28g</td>
</tr>
<tr>
<td>aaBB</td>
<td>20g</td>
</tr>
<tr>
<td>aaBb</td>
<td>18g</td>
</tr>
<tr>
<td>aabb</td>
<td>16g</td>
</tr>
</tbody>
</table>

Epistasis and Genotypic Values
Basic Model of Quantitative Genetics

Phenotypic Value = the value observed when a trait is measured on an individual

\[ P = G + E \]

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

\[ P = G + E + GE \]

\[ V(P) = V(G) + V(E) + V(GE) + 2 \times \text{COV}(GE) \]

Two different sources for GxE Interactions

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

2) Lack of perfect correlation among breeding values across environments (aka change in rank)
V(GE) can be partitioned into these two sources:

\[ V(GE) = \frac{(V(A)_{x1} - V(A)_{x2})^2}{2} + V(A)_{x1} \cdot V(A)_{x2} \cdot (1 - r) \]

\[ V(A)_i = \text{additive variance in environment } i \]

\[ r_A = \text{additive genetic correlation across environments} \]

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

2) Lack of perfect correlation among breeding values across environments (aka change in rank)
GxE Interaction due to change of scale
GxE Interaction due to change in ranking

Mean Phenotype

<table>
<thead>
<tr>
<th></th>
<th>Environment 1</th>
<th>Environment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL</td>
<td></td>
<td>SS</td>
</tr>
<tr>
<td>SS</td>
<td>LL</td>
<td></td>
</tr>
</tbody>
</table>

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 IBD regions
• SNP in exon of Ptprz1
  • Involved in bone formation
  • P → H
  • Predicted to be functionally damaging

Modified Quantitative Hybrid Complementation Test

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