9 Multiple Traits & Multiple Environments

- 2 traits: mechanics (Jiang Zeng 1995)
- close linkage or pleiotropy?
- multiple traits: efficiency beyond 2?
  - principal components on phenotypes
  - discriminant analysis (PC on residuals)
- outbred crosses
  - mixed model approach
  - linkage & linkage disequilibrium

why study multiple traits together?

- avoid reductionist approach to biology
  - address physiological/biochemical mechanisms
  - Schmalhausen (1942); Falconer (1952)
- separate close linkage from pleiotropy
  - 1 locus or 2 linked loci?
- identify epistatic interaction or canalization
  - influence of genetic background
- establish QTL x environment interactions
- decompose genetic correlation among traits
- increase power to detect QTL
why are traits correlated?

- environmental correlation
  - non-genetic, controllable by design
  - historical correlation (learned behavior)
  - physiological correlation (same body)
- pleiotropy
  - one gene, many functions
  - common biochemical pathway, splicing variants
- close linkage
  - two tightly linked genes
  - genotypes $Q$ are collinear

interplay of pleiotropy & correlation

pleiotropy only

correlation only

Korol et al. (2001)
QTL affecting multiple traits

- typical tests at a locus
  - is there a QTL here for any trait?
  - same QTL effect across traits or environments?
- power for single trait vs. multiple trait models
- Haley-Knott regression approach
  - Knott Haley (2000) simulated data
- IM & CIM approach
  - Jiang Zeng (1995); Korol et al. (1995) simulated data
  - Vieira et al. (2000) Drosophila lifespan
3 correlated traits
(Jiang Zeng 1995)

ellipses centered on genotypic value
width for nominal frequency
main axis angle environmental correlation
3 QTL, F2
27 genotypes

note signs of
genetic and
environmental
correlation

power: separate analysis vs.
multiple trait analysis

<table>
<thead>
<tr>
<th>QTL</th>
<th>Position (cM)</th>
<th>Additive effect</th>
<th>Dominance effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Trait 1</td>
<td>Trait 2</td>
</tr>
<tr>
<td></td>
<td>Parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>21.0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>84.0</td>
<td>-0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>142.0</td>
<td>-1.00</td>
<td>0.50</td>
</tr>
</tbody>
</table>

**TABLE 5**

Observed statistical power (proportion of significant replicates over all replicates) of seven methods of QTL mapping from 100 replicates of simulations

<table>
<thead>
<tr>
<th>QTL</th>
<th>J-123</th>
<th>J-12</th>
<th>J-13</th>
<th>J-23</th>
<th>S-1</th>
<th>S-2</th>
<th>S-3</th>
<th>S-125</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.80</td>
<td>0.78</td>
<td>0.64</td>
<td>0.46</td>
<td>0.04</td>
<td>0.78</td>
<td>0.46</td>
<td>0.78</td>
</tr>
<tr>
<td>2</td>
<td>0.78</td>
<td>0.77</td>
<td>0.64</td>
<td>0.44</td>
<td>0.00</td>
<td>0.41</td>
<td>0.64</td>
<td>0.78</td>
</tr>
<tr>
<td>3</td>
<td>0.80</td>
<td>0.51</td>
<td>0.84</td>
<td>0.44</td>
<td>0.00</td>
<td>0.64</td>
<td>0.79</td>
<td>0.78</td>
</tr>
</tbody>
</table>

ch. 9 © 2003 Broman, Churchill, Yandell, Zeng
history: pleiotropy vs. close linkage

- Schork et al. (1994)
- single trait analyses
  - Cheverud et al. (1997); Lebreton et al. (1998)
  - may lack power
- outbred
  - Almasy et al. (1997)
  - bivariate: unclear how to extend to multivariate
- multiple trait regression
  - easily extensible, fast
  - reasonable power, concerns about approx (Kao 2000)

pleiotropy or close linkage?

2 traits, 2 qtl/trait
pleiotropy @ 54cM
linkage @ 114,128cM
QTL x sex interaction (Vieira et al. 2000)

QTL x environment interaction (Vieira et al. 2000)
generalizing heritability to multiple traits

• multivariate normal model: \( Y_i = \text{vector of } t \text{ traits} \)
  - \( Y_i = \mu + G_i(Q) + e_i \sim \text{MVN}(0, V_E) \)
• partition trait variance-covariance matrix
  - \( V_{TOT} = V_G + V_E \)
• determinant of \( V \) generalizes variance
  - \( | V | = \text{det}(V) \)
  - \( H^2 = \frac{| V_G |}{| V_{TOT} |} \)
  - maximize \( H^2 \) by minimizing \( | V_E | \)
• use in weighted least squares (or LOD)
  - Korol et al. (1995); Knott, Haley (2000); Korol et al. (2001)

QTL via Principal Components

• PC or SVD decomposition of multiple traits
  - \( Y = n \times t \text{ matrix}: Y = UDU^T \)
  - \( U = \text{orthonormal transformation} \)
  - \( D = \text{diagonal matrix with eigen-values} \)
• transform problem to principal components
  - \( Y^* = FY \text{ has uncorrelated PC traits} \) (\( F = D^{1/2}U^T \))
  - \( Y^* = \mu^* + G^*(Q) + e^* \)
• interval map each PC separately
  - \( Y^*_{1i} = \mu^*_{1i} + G^*_{1i}(Q_i) + e^*_{1i} \)
  - may only need to map a few PCs
QTL via Principal Components

• example: Drosophila reproduction
  – Liu et al. (1996); Zeng et al. (2000); ch. 7
• other refs
  – Weller et al. (1996); Mangin et al. (1998);
    Olson et al. (1999); Mahler et al. (2002)
• problems
  – PC may have no relation to genetics!
  – residuals from QTL correlated across PCs
  – PC is descriptive summary, not interpretive

interval-dependent PC

• want to reduce dimensionality while focusing on
  QTL differences
• interval-dependent PC on residuals
  – reduces dimensionality by identifying patterns in
    residuals not explained by QTL
  – not quite discriminant analysis: does not aim to best
    discriminate among QTL genotypes
• pleiotropy highlighted by PCs
  – find strongest correlation
  – interval map using transformed data
• Allison et al. (1998); Korol et al. (2001)
interval-dependent PC details

- flanking marker based PC
  - remove effect of flanking markers (Haley-Knott)
  - PC decomposition of residuals
  - transform original data to new PC axes
  - map transformed data on interval
  - problem: bias, variance for Haley-Knott
- improvement through iteration
  - PC decomposition of residuals of transformed data
  - repeat until estimates stabilize

LOD depends on number of traits

- higher, steeper LOD
  - more power
  - provided traits are correlated
- simulated data (BC)
  \[ n = 200 \]
  \[ t = 2, 5, 8, 10 \]
  \[ h^2 = .072, .173, .254, .294 \]

Korol et al. (2001)
power of multiple trait analysis
estimated effect of 1 QTL

multiple traits for outbred crosses

- Almasy et al. (1997); Lund et al. (2003)
- recombination model blends linkage and linkage disequilibrium (recall ch. 8)
- limited implementation
  - single QTL
  - proof of principle
outbred mixed model: multiple traits

• phenotypes = design + QTLs + polygenes + env
  – \( Y = \mu + G_i + g + e \)
  – \( Y_{ij} = \mu_j + G_j(Q_i) + g_{ij} + e_{ij}, \quad i = 1, \ldots, n, j = 1, \ldots, t \)

• QTL effects: fixed or random

• random polygenic effects
  – within trait correlation depends on relationship \( A \)
  – between trait polygenic correlations?

• random environmental errors
  – correlation among traits unexplained by polygenes
    \[ g_{ij} \sim MVN(0, W_{jj}A), \quad \text{or} \quad \text{cov}(g_{1j}, g_{2j}) = W_{jj}A_{12} \]
    \[ g_{i*} \sim MVN(0, A_{ii}W), \quad \text{or} \quad \text{cov}(g_{i1}, g_{i2}) = W_{12}A_{ii} \]
    \[ e_{ij} \sim MVN(0, V), \quad \text{or} \quad \text{cov}(e_{i1}, e_{i2}) = V_{12} \]

how: software

• WinQTL/QTL Cartographer: Jzmapqtl
  – Jiang Zeng (1995); statgen.ncsu.edu/qtlcart/

• MultiQTL
  – Korol et al. (2001); www.multiqtl.com

• QTL Express
  – Knott Haley (2000); qtl.cap.ed.ac.uk
Mapping Microarray Data

- overview, wish lists
  - Jansen, Nap (2001); Cheung, Spielman (2002); Doerge (2002); Bochner (2003)
- single gene expression as trait (single QTL)
  - Dumas et al. (2000)
- microarray scan via 1 QTL interval mapping
  - Brem et al. (2002); Schadt et al. (2003)
  - found cis and trans acting genes
- multivariate and multiple QTL approach
  - Lan et al. (2003)

central dogma via microarrays (Bochner 2003)
mutant (green) vs. parental (red) E. coli (Bochner 2003)

idea of mapping microarrays (Jansen Nap 2001)
goal: unravel biochemical pathways (Jansen Nap 2001)

(a) DNA Marker

(b) Map positions

(c) One putative pathway

yeast linkage mapping: tetrad analysis (Brem et al. 2002)
QTL mapping of gene expression (Brem et al. 2002)

Schadt et al. (2003): mapping gene expression in mouse genome