The X chromosome in QTL mapping: What a pain in the @$@$!

Karl W Broman
Department of Biostatistics
Johns Hopkins University

www.biostat.jhsph.edu/~kbroman
Backcross

P₁ × P₂ → F₁

F₁ → BC
Genotype data
LOD curves

LOD score

Chromosome

log₂ liver
log₂ spleen

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Permutation test

- Genotype data
- Markers
- Phenotypes
- LOD scores
- Maximum LOD score

Diagram:
- Individuals
- Genotype data
- Markers
- Phenotypes
- LOD scores
- Maximum LOD score
Permutation results

Genome-wide maximum LOD score

![Chart showing permutation results]

Genome-wide maximum LOD score
X chr in backcross

(A x B) x A

(B x A) x A

A x (A x B)

A x (B x A)
X chr in intercross

(A x B) x (A x B)

(B x A) x (A x B)

(A x B) x (B x A)

(B x A) x (B x A)
Intercross: both dir, both sexes

♀ forward AA or AB
♀ reverse AB or BB
♂ forward AY or BY
♂ reverse AY or BY
Principles

• Sex- or cross-direction-difference in the phenotype shouldn’t lead to spurious linkage on the X chromosome

• Simple as possible

• Null nested within alternative
<table>
<thead>
<tr>
<th>Cross</th>
<th>Direction</th>
<th>Sexes</th>
<th>Contrasts</th>
<th>$H_0$</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC</td>
<td>both</td>
<td>both</td>
<td>AA:AB:AY:BY</td>
<td>$♀ : ♂$</td>
<td>2</td>
</tr>
<tr>
<td>BC</td>
<td>$♀$</td>
<td></td>
<td>AA:AB</td>
<td>grand mean</td>
<td>1</td>
</tr>
<tr>
<td>BC</td>
<td>$♂$</td>
<td></td>
<td>AY:BY</td>
<td>grand mean</td>
<td>1</td>
</tr>
<tr>
<td>$F_2$</td>
<td>both</td>
<td>both</td>
<td>AA:ABf:ABr:BB:AY:BY</td>
<td>$♀_f : ♀_r : ♂$</td>
<td>3</td>
</tr>
<tr>
<td>$F_2$</td>
<td>both</td>
<td>$♀$</td>
<td>AA:ABf:ABr:BB</td>
<td>$♀_f : ♀_r$</td>
<td>2</td>
</tr>
<tr>
<td>$F_2$</td>
<td>both</td>
<td>$♂$</td>
<td>AY:BY</td>
<td>grand mean</td>
<td>1</td>
</tr>
<tr>
<td>$F_2$</td>
<td>one</td>
<td>both</td>
<td>AA:AB:AY:BY</td>
<td>$♀ : ♂$</td>
<td>2</td>
</tr>
<tr>
<td>$F_2$</td>
<td>one</td>
<td>$♀$</td>
<td>AA:AB</td>
<td>grand mean</td>
<td>1</td>
</tr>
<tr>
<td>$F_2$</td>
<td>one</td>
<td>$♂$</td>
<td>AY:BY</td>
<td>grand mean</td>
<td>1</td>
</tr>
</tbody>
</table>
Bad Results

LOD score

2 7 8 9 16 X

log\textsubscript{2} liver
log\textsubscript{2} spleen

Chromosome

2 7 8 9 16 X
Chromosome-specific thresholds

Let $\alpha_i$ = false positive rate for chromosome $i$.

We need $1 - \alpha = \prod (1 - \alpha_i)$

For example, $\alpha_1 = \alpha$ and $\alpha_j = 0$ for $j \neq 1$

The usual method: constant LOD threshold (i.e., constant power)

My approach: $\alpha_i \propto L_i$ where $L_i = \text{length of chr } i$

Similar and more convenient: $\alpha_i = (1 - \alpha)^{L_i/L}$
Simulation results
A- and X-specific permutations

- Do separate permutations within autosomes and X chromosome.

- $M_{Ai}^* = \text{maximum LOD across autosomes in replicate } i.$
  $M_{Xi}^* = \text{maximum LOD across X chromosome in replicate } i.$

- $T_A = (1 - \alpha)^{L_A/L} \text{ quantile of } M_{Ai}^*$
  $T_X = (1 - \alpha)^{L_X/L} \text{ quantile of } M_{Xi}^*$

- Genome-scan-adjusted p-values
  $M_X = \text{observed maximum LOD on X}$
  $p_{unadj} = \text{Prop}(M_{Xi}^* \geq M_X)$
  $p_{adj} = 1 - (1 - p_{unadj})^{L/L_X}$

- Let $n_A = \text{no. permutation replicates for autosomes}$
  We want $n_X = n_A \times L_A/L_X$ to get equivalent precision.
  (This doubles the computation time.)
## LOD thresholds for the example

<table>
<thead>
<tr>
<th></th>
<th>log(_2) liver</th>
<th></th>
<th>log(_2) spleen</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>20%</td>
<td>5%</td>
<td>20%</td>
</tr>
<tr>
<td>autosomes</td>
<td>3.32</td>
<td>2.62</td>
<td>3.33</td>
<td>2.59</td>
</tr>
<tr>
<td>X chromosome</td>
<td>4.66</td>
<td>3.78</td>
<td>4.55</td>
<td>3.70</td>
</tr>
</tbody>
</table>

**Note:** 1100 permutations on autosomes  
31,075 on X chromosome
2d scans

Full: $Q_1 + Q_2 + Q_1 : Q_2$

Add: $Q_1 + Q_2$

One: $Q_1$

Null: $\emptyset$
A vs. A

\[
\begin{array}{ccc}
& AA & AB & BB \\
AA & \bullet & \bullet & \bullet \\
AB & \bullet & \bullet & \bullet \\
BB & \bullet & \bullet & \bullet \\
\end{array}
\]

full: 9 param
add: 5 param
null: 1 param

\{ int: 4 df \}
### X vs. X

<table>
<thead>
<tr>
<th></th>
<th>AA</th>
<th>ABf</th>
<th>ABr</th>
<th>BB</th>
<th>AY</th>
<th>BY</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABf</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABr</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
</tr>
<tr>
<td>AY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>BY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
</tr>
</tbody>
</table>

- **full**: 12 param
- **add**: 9 param
- **int**: 3 df
- **null**: 3 param
<table>
<thead>
<tr>
<th></th>
<th>AA</th>
<th>ABf</th>
<th>ABr</th>
<th>BB</th>
<th>AY</th>
<th>BY</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>AB</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>BB</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

- full: 18 param
- add: 8 param
- int: 10 df
- null: 3 param
Summary

• The X chromosome is a pain in the @$$!
• Must take care regarding contrasts and null hypothesis
• Need for autosome- and X-chr-specific permutations
• 2d scan requires great care; odd degrees of freedom