Gene Mapping:
The Why and How of Multiple QTL
Brian S. Yandell

• why: strategy
  – bias with single QTL
  – advantages of multiple QTL
• how: software
  – WinQTLCart intro
  – R/qtl demo
  – R/qtlbim demo

Real knowledge is to know the extent of one’s ignorance.
Confucius (on a bench in Seattle)

outline

1. What is the goal of multiple QTL study?
2. Gene action and epistasis
3. Bayesian vs. classical QTL
4. QTL software options
cross two inbred lines
→ linkage disequilibrium
→ associations
→ linked segregating QTL
(after Gary Churchill)

1. what is the goal of QTL study?

- uncover underlying biochemistry
  - identify how networks function, break down
  - find useful candidates for (medical) intervention
  - epistasis may play key role
  - statistical goal: maximize number of correctly identified QTL
- basic science/evolution
  - how is the genome organized?
  - identify units of natural selection
  - additive effects may be most important (Wright/Fisher debate)
  - statistical goal: maximize number of correctly identified QTL
- select “elite” individuals
  - predict phenotype (breeding value) using suite of characteristics (phenotypes) translated into a few QTL
  - statistical goal: minimize prediction error
problems of single QTL approach

- wrong model: biased view
  - fool yourself: bad guess at locations, effects
  - detect ghost QTL between linked loci
  - miss epistasis completely
- low power
- bad science
  - use best tools for the job
  - maximize scarce research resources
  - leverage already big investment in experiment
advantages of multiple QTL approach

- improve statistical power, precision
  - increase number of QTL detected
  - better estimates of loci: less bias, smaller intervals
- improve inference of complex genetic architecture
  - patterns and individual elements of epistasis
  - appropriate estimates of means, variances, covariances
    - asymptotically unbiased, efficient
  - assess relative contributions of different QTL
- improve estimates of genotypic values
  - less bias (more accurate) and smaller variance (more precise)
  - mean squared error = MSE = (bias)$^2$ + variance

2. Gene Action and Epistasis

additive, dominant, recessive, general effects
of a single QTL (Gary Churchill)
additive effects of two QTL (Gary Churchill)

$$\mu_q = \mu + \beta_{q1} + \beta_{q2}$$

Epistasis (Gary Churchill)

The allelic state at one locus can mask or uncover the effects of allelic variation at another.

- W. Bateson, 1907.
epistasis in parallel pathways (GAC)

- Z keeps trait value low
- neither E₁ nor E₂ is rate limiting
- loss of function alleles are segregating from parent A at E₁ and from parent B at E₂

epistasis in a serial pathway (GAC)

- Z keeps trait value high
- neither E₁ nor E₂ is rate limiting
- loss of function alleles are segregating from parent B at E₁ and from parent A at E₂
epistatic interactions

• model space issues
  – 2-QTL interactions only?
    • or general interactions among multiple QTL?
  – partition of effects
    • Fisher-Cockerham or tree-structured or ?
• model search issues
  – epistasis between significant QTL
    • check all possible pairs when QTL included?
    • allow higher order epistasis?
  – epistasis with non-significant QTL
    • whole genome paired with each significant QTL?
    • pairs of non-significant QTL?
• see papers of Nengjun Yi (2000-7) in Genetics

limits of epistatic inference

• power to detect effects
  – epistatic model sizes grow quickly
    • $|A| = 3^n$ for general interactions
  – power tradeoff
    • depends sample size vs. model size
    • want $n / |A|$ to be fairly large (say $> 5$)
    • 3 QTL, $n = 100$ F2: $n / |A| \approx 4$ with $n = 100$
• rare genotypes may not be observed
  – $aa/BB \& AA/bb$ rare for linked loci
  – empty cells mess up balance
    • adjusted tests (type III) are wrong
  – confounds main effects & interactions
limits of multiple QTL?

- limits of statistical inference
  - power to detect QTL depends on many things
    - larger sample, higher heritability, smaller environmental variation
  - difficult to sort out effects of closely linked loci
  - “best” model balances data fit against model size

- limits of biological utility
  - marker assisted selection (Bernardo 2001 *Crop Sci*)
    - 10 QTL ok, 50 QTL are too many
    - phenotype better predictor than genotype when too many QTL
    - increasing sample size may not give multiple QTL any advantage
  - hard to select many QTL simultaneously
    - 3^n possible genotypes to choose from

QTL below detection level?

- problem of selection bias
  - QTL of modest effect only detected sometimes
  - effects overestimated when detected
  - repeat studies may fail to detect these QTL

- think of probability of detecting QTL
  - avoids sharp in/out dichotomy
  - avoid pitfalls of one “best” model
  - examine “better” models with more probable QTL

- rethink formal approach for QTL
  - directly allow uncertainty in genetic architecture
  - QTL model selection over genetic architecture
3. Bayesian vs. classical QTL study

- classical study
  - maximize over unknown effects
  - test for detection of QTL at loci
  - model selection in stepwise fashion
- Bayesian study
  - average over unknown effects
  - estimate chance of detecting QTL
  - sample all possible models
- both approaches
  - average over missing QTL genotypes
  - scan over possible loci

QTL model selection: key players

- observed measurements
  - $y$ = phenotypic trait
  - $m$ = markers & linkage map
  - $i = \text{individual index (1,…,} n\text{)}$
- missing data
  - missing marker data
  - $q$ = QT genotypes
    - alleles QQ, Qq, or qq at locus
- unknown quantities
  - $\lambda$ = QT locus (or loci)
  - $\mu$ = phenotype model parameters
  - $A$ = QTL model/genetic architecture
- $p_r(q|m,\lambda,A)$ genotype model
  - grounded by linkage map, experimental cross
  - recombination yields multinomial for $q$ given $m$
- $p_r(y,q,\mu,A)$ phenotype model
  - distribution shape (assumed normal here)
  - unknown parameters $\mu$ (could be non-parametric)

after Sen Churchill (2001)
Bayes posterior vs. maximum likelihood

- **LOD**: classical Log ODds
  - maximize likelihood over effects $\mu$
  - $R/qtl$ scanone/scantwo: method = “em”

- **LPD**: Bayesian Log Posterior Density
  - average posterior over effects $\mu$
  - $R/qtl$ scanone/scantwo: method = “imp”

\[
\text{LOD}(\lambda) = \log_{10} \{\max_{\mu} \text{pr}(\gamma \mid m, \mu, \lambda)\} + c
\]

\[
\text{LPD}(\lambda) = \log_{10} \{\int \text{pr}(\gamma \mid m, \mu, \lambda)\text{pr}(\mu)d\mu\} + C
\]

likelihood mixes over missing QTL genotypes:

\[
\text{pr}(\gamma \mid m, \mu, \lambda) = \sum_{q} \text{pr}(\gamma \mid q, \mu)\text{pr}(q \mid m, \lambda)
\]
marginal LOD or LPD

- compare two architectures at each locus
  - with \(A_2\) or without \(A_1\) another QTL at separate locus \(\lambda_2\)
    - preserve model hierarchy (e.g. drop any epistasis with QTL at \(\lambda_2\))
  - with \(A_2\) or without \(A_1\) epistasis with second locus \(\lambda_2\)
- allow for multiple QTL besides locus being scanned
  - allow for QTL at all other loci \(\lambda_1\) in architecture \(A_1\)
- use marginal LOD, LPD or other diagnostic
  - posterior, Bayes factor, heritability

\[
\begin{align*}
\text{LOD}(\hat{\lambda}_1, \hat{\lambda}_2 | A_2) &= \text{LOD}(\hat{\lambda}_1 | A_1) \\
\text{LPD}(\hat{\lambda}_1, \hat{\lambda}_2 | A_2) &= \text{LPD}(\hat{\lambda}_1 | A_1)
\end{align*}
\]
LPD: 1 QTL vs. multi-QTL
marginal contribution to LPD from QTL at $\lambda$

substitution effect: 1 QTL vs. multi-QTL
single QTL effect vs. marginal effect from QTL at $\lambda$
comparing models

• balance model fit against model complexity
  – want to fit data well (maximum likelihood)
  – without getting too complicated a model

<table>
<thead>
<tr>
<th>fit model</th>
<th>smaller model</th>
<th>bigger model</th>
</tr>
</thead>
<tbody>
<tr>
<td>estimate phenotype</td>
<td>miss key features</td>
<td>fits better</td>
</tr>
<tr>
<td>predict new data</td>
<td>may be biased</td>
<td>no bias</td>
</tr>
<tr>
<td>interpret model</td>
<td>easier</td>
<td>more complicated</td>
</tr>
<tr>
<td>estimate effects</td>
<td>low variance</td>
<td>high variance</td>
</tr>
</tbody>
</table>

information criteria
to balance fit against complexity

• classical information criteria
  – penalize likelihood $L$ by model size $|A|$
  – $IC = -2 \log L(A \mid y) + \text{penalty}(A)$
  – maximize over unknowns

• Bayes factors
  – marginal posteriors $pr(y \mid A)$
  – average over unknowns
4. QTL software options

- methods
  - approximate QTL by markers
  - exact multiple QTL interval mapping
- software platforms
  - MapMaker/QTL (obsolete)
  - QTLCart (statgen.ncsu.edu/qtlcart)
  - R/qtl (www.rqtl.org)
  - R/qtlbim (www.qtlbim.org)

approximate QTL methods

- marker regression
  - locus & effect confounded
  - lose power with missing data
- Haley-Knott (least squares) regression
  - correct mean, wrong variance
  - biased by pattern of missing data (Kao 2000)
- extended HK regression
  - correct mean and variance
  - minimizes bias issue (R/qtl “ehk” method)
- composite interval mapping (QTLCart)
  - use markers to approximate other QTL
  - properties depend on marker spacing, missing data
exact QTL methods

- interval mapping (Lander, Botstein 1989)
  - scan whole genome for single QTL
  - bias for linked QTL, low power
- multiple interval mapping (Kao, Zeng, Teasdale 1999)
  - sequential scan of all QTL
  - stepwise model selection
- multiple imputation (Sen, Churchill 2001)
  - fill in (impute) missing genotypes along genome
  - average over multiple imputations
- Bayesian interval mapping (Yi et al. 2005)
  - sample most likely models
  - marginal scans conditional on other QTL

QTL software platforms

- QTLCart (statgen.ncsu.edu/qtlcart)
  - includes features of original MapMaker/QTL
    - not designed for building a linkage map
  - easy to use Windows version WinQTLCart
  - based on Lander-Botstein maximum likelihood LOD
    - extended to marker cofactors (CIM) and multiple QTL (MIM)
    - epistasis, some covariates (GxE)
    - stepwise model selection using information criteria
  - some multiple trait options
  - OK graphics
- R/qtl (www.rqtl.org)
  - includes functionality of classical interval mapping
  - many useful tools to check genotype data, build linkage maps
  - excellent graphics
  - several methods for 1-QTL and 2-QTL mapping
    - epistasis, covariates (GxE)
  - tools available for multiple QTL model selection
> library(qtl)
> data(hyper)
> summary(hyper)

**Backcross**

- No. individuals: 250
- No. phenotypes: 2
- Percent phenotyped: 100 100
- No. chromosomes: 20
  - Autosomes: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
  - X chr: X
- Total markers: 174
- No. markers: 22 8 6 20 14 11 7 6 5 5 14 5 5 11 6 12 4 4 4
- Percent genotyped: 47.7
- Genotypes (%): AA:50.2 AB:49.8

> plot(hyper)
> plot.missing(hyper, reorder = TRUE)
R/qtl: find genotyping errors

> hyper <- calc.errorlod(hyper, error.prob=0.01)
> top.errorlod(hyper)
  chr  id  marker errorlod
  1   118 D1Mit14  8.372794
  2   162 D1Mit14  8.372794
  3   170 D1Mit14  8.372794
  4   159 D1Mit14  8.350341
  5   173 D1Mit14  6.165395
  6   165 D1Mit14  6.165395
  7   188 D1Mit14  6.165395
  8   184 D1Mit14  6.151606
  9   1241 D1Mit14  6.151606

  16   1215 D1Mit267  5.822192
  17   1108 D1Mit267  5.822192
  18   1138 D1Mit267  5.822192
  19   1226 D1Mit267  5.822192
  20   1199 D1Mit267  5.815250
  21   184 D1Mit267  5.808400
> plot.geno(hyper, chr=1, ind=c(117:119,137:139,157:184))
R/qtl: 1 QTL interval mapping

```r
> hyper <- calc.genoprob(hyper, step=1, error.prob=0.01)
> out.em <- scanone(hyper)
> out.hk <- scanone(hyper, method="hk")
> summary(out.em, threshold=3)
  chr  pos  lod
  c1.loc45   1 48.3 3.52
  D4Mit164   4 29.5 8.02
> summary(out.hk, threshold=3)
  chr  pos  lod
  c1.loc45   1 48.3 3.55
  D4Mit164   4 29.5 8.09
```
black = EM
blue = HK

note bias where marker data are missing systematically

R/qtl: permutation threshold

```r
> operm.hk <- scanone(hyper, method="hk",
  n.perm=1000)
Doing permutation in batch mode ...
> summary(operm.hk, alpha=c(0.01,0.05))
LOD thresholds (1000 permutations)
  lod
1%  3.79
5%  2.78
> summary(out.hk, perms=operm.hk, alpha=0.05,
  pvalues=TRUE)
chr pos lod  pval
 1  1  48.3  3.55  0.015
 2  4  29.5  8.09  0.000
```
R/qtl: 2 QTL scan

```r
> hyper <- calc.genoprob(hyper, step=5, error.prob=0.01)
> out2.hk <- scantwo(hyper, method="hk")
--Running scanone
--Running scantwo
(1,1)
(1,2)
...
(19,19)
(19,X)
(X,X)
> summary(out2.hk, thresholds=c(6.0, 4.7, 4.4, 4.7, 2.6))

<table>
<thead>
<tr>
<th>pos1f</th>
<th>pos2f</th>
<th>lod.full</th>
<th>lod.fv1</th>
<th>lod.int</th>
<th>pos1a</th>
<th>pos2a</th>
<th>lod.add</th>
<th>lod.av1</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>:c4</td>
<td>68.3</td>
<td>30.0</td>
<td>14.13</td>
<td>6.51</td>
<td>0.225</td>
<td>68.3</td>
<td>30.0</td>
</tr>
<tr>
<td>c2</td>
<td>:c19</td>
<td>47.7</td>
<td>0.0</td>
<td>6.71</td>
<td>5.01</td>
<td>3.458</td>
<td>52.7</td>
<td>0.0</td>
</tr>
<tr>
<td>c3</td>
<td>:c3</td>
<td>37.2</td>
<td>42.2</td>
<td>6.10</td>
<td>5.08</td>
<td>0.226</td>
<td>37.2</td>
<td>42.2</td>
</tr>
<tr>
<td>c6</td>
<td>:c15</td>
<td>60.0</td>
<td>20.5</td>
<td>7.17</td>
<td>5.22</td>
<td>3.237</td>
<td>25.0</td>
<td>20.5</td>
</tr>
<tr>
<td>c9</td>
<td>:c18</td>
<td>67.0</td>
<td>37.2</td>
<td>6.31</td>
<td>4.79</td>
<td>4.083</td>
<td>67.0</td>
<td>12.2</td>
</tr>
<tr>
<td>c12</td>
<td>:c19</td>
<td>1.1</td>
<td>40.0</td>
<td>6.48</td>
<td>4.79</td>
<td>4.090</td>
<td>1.1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

> plot(out2.hk, chr=c(1,4,6,15))
```
R/qtl: ANOVA imputation at QTL

```r
> hyper <- sim.geno(hyper, step=2, n.draws=16, error.prob=0.01)
> qtl <- makeqtl(hyper, chr = c(1, 1, 4, 6, 15), pos = c(50, 76, 30, 70, 20))
> my.formula <- y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q4:Q5
> out.fitqtl <- fitqtl(hyper$pheno[,1], qtl, formula=my.formula)
> summary(out.fitqtl)
```

**Full model result**

```
----------------------------------
df SS        MS      LOD     %var Pvalue(Chi2) Pvalue(F)
Model   6  5789.089 964.84822 21.54994 32.76422            0    0
Error 243 11879.847  48.88826
Total 249 17668.936
```

**Drop one QTL at a time ANOVA table:**

```
----------------------------------
df Type III SS      LOD     %var Pvalue(Chi2) Pvalue(F)
Chr1@50           1     297.149    1.341    1.682   6.078   0.01438 *
Chr1@76           1     520.664    2.329    2.947  10.650   0.00126 **
Chr4@30           1    2842.089   11.644   16.085  58.134  5.50e-13 ***
Chr6@70           2    1435.721    6.194    8.126  14.684  9.55e-07 ***
Chr15@20          2    1083.842    4.740    6.134  11.085  2.47e-05 ***
Chr6@70:Chr15@20  1     955.268    4.199    5.406 19.540  1.49e-05 ***
```

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

QTL: Why and How UW-Madison PBPG Yandell © 2007
R/qtlbim (www.qtlbim.org)

- cross-compatible with R/qtl
- model selection for genetic architecture
  - epistasis, fixed & random covariates, GxE
  - samples multiple genetic architectures
  - examines summaries over nested models
- extensive graphics

R/qtlbim: tutorial
(www.stat.wisc.edu/~yandell/qtlbim)

> data(hyper)
## Drop X chromosome (for now).
> hyper <- subset(hyper, chr=1:19)
> hyper <- qb.genoprob(hyper, step=2)
## This is the time-consuming step:
> qbHyper <- qb.mcmc(hyper, pheno.col = 1)
## Here we get stored samples.
> qb.load(hyper, qbHyper)
> summary(qbHyper)
R/qtlbim: initial summaries

> summary(qbHyper)
Bayesian model selection QTL mapping object qbHyper on cross object hyper
had 3000 iterations recorded at each 40 steps with 1200 burn-in steps.

Diagnostic summaries:

\[
\begin{array}{cccccc}
\text{nqtl} & \text{mean} & \text{envvar} & \text{varadd} & \text{varaa} & \text{var} \\
\text{Min.} & 2.000 & 97.42 & 28.07 & 5.112 & 0.000 & 5.112 \\
\text{1st Qu.} & 5.000 & 101.00 & 44.33 & 17.010 & 1.639 & 20.180 \\
\text{Mean} & 6.543 & 101.30 & 48.80 & 20.310 & 5.321 & 25.630 \\
\text{3rd Qu.} & 8.000 & 101.70 & 53.11 & 23.480 & 7.862 & 30.370 \\
\text{Max.} & 13.000 & 103.90 & 74.03 & 51.730 & 34.940 & 65.220 \\
\end{array}
\]

Percentages for number of QTL detected:

\[
\begin{array}{cccccccccccc}
2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\
2 & 3 & 9 & 14 & 21 & 19 & 17 & 10 & 4 & 1 & 0 & 0 \\
\end{array}
\]

Percentages for number of epistatic pairs detected:

\[
\begin{array}{cccccccccc}
\text{pairs} & 1 & 2 & 3 & 4 & 5 & 6 \\
29 & 31 & 23 & 11 & 5 & 1 \\
\end{array}
\]

Percentages for common epistatic pairs:

\[
\begin{array}{cccccccccccc}
6.15 & 4.15 & 6.15 & 1.7 & 15.15 & 1.4 & 1.6 & 1.5 & 1.5 & 1.15 & 1.15 & 1.15 \\
64 & 16 & 10 & 6 & 4 & 5 & 4 & 3 & 3 & 2 & 2 & 2 \\
\end{array}
\]

> plot(qb.diag(qbHyper, items = c("herit", "envvar")))
R/qtlbim: 1-D (not 1-QTL!) scan

```r
> one <- qb.scanone(qbHyper, chr = c(1,4,6,15), type = "LPD")
> summary(one)
LPD of bp for main,epistasis,sum

<table>
<thead>
<tr>
<th>n.qtl</th>
<th>pos</th>
<th>m.pos</th>
<th>e.pos</th>
<th>m.epi</th>
<th>epistasis</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>1.331</td>
<td>64.5</td>
<td>67.8</td>
<td>6.10</td>
<td>0.442</td>
<td>6.27</td>
</tr>
<tr>
<td>c4</td>
<td>1.377</td>
<td>29.5</td>
<td>29.5</td>
<td>11.49</td>
<td>0.375</td>
<td>11.61</td>
</tr>
<tr>
<td>c6</td>
<td>0.838</td>
<td>59.0</td>
<td>59.0</td>
<td>3.99</td>
<td>6.265</td>
<td>9.60</td>
</tr>
<tr>
<td>c15</td>
<td>0.961</td>
<td>17.5</td>
<td>17.5</td>
<td>1.30</td>
<td>6.325</td>
<td>7.28</td>
</tr>
</tbody>
</table>

> plot(one)
> plot(out.em, chr=c(1,4,6,15), add = TRUE, col = "red", lty = 2)
```

hyper data: scanone

```
LPD of bp for main+epistasis+sum
```

```r
c1 1.331 64.5 67.8 6.10 0.442 6.27
```

```r
c4 1.377 29.5 29.5 11.49 0.375 11.61
```

```r
c6 0.838 59.0 59.0 3.99 6.265 9.60
```

```r
c15 0.961 17.5 17.5 1.30 6.325 7.28
```
R/qtlbim: automated QTL selection

```r
> hpd <- qb.hpdone(qbHyper, profile = "2logBF")
> summary(hpd)

<table>
<thead>
<tr>
<th>chr</th>
<th>n.qtl</th>
<th>pos</th>
<th>lo.50%</th>
<th>hi.50%</th>
<th>2logBF</th>
<th>A</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.829</td>
<td>64.5</td>
<td>72.1</td>
<td>6.692</td>
<td>103.611</td>
<td>99.090</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3.228</td>
<td>29.5</td>
<td>31.7</td>
<td>11.169</td>
<td>104.584</td>
<td>98.020</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>1.033</td>
<td>59.0</td>
<td>66.7</td>
<td>6.054</td>
<td>99.637</td>
<td>102.965</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>0.159</td>
<td>17.5</td>
<td>17.5</td>
<td>5.837</td>
<td>101.972</td>
<td>100.702</td>
</tr>
</tbody>
</table>

> plot(hpd)
```

2log(BF) scan with 50% HPD region
R/qtlbim: Bayes Factor evaluations

```r
> tmp <- qb.BayesFactor(qbHyper)
> summary(tmp)
$pattern
 posterior prior  bf  bfse
7:2*1,2*15,2*4,6  0.00500 3.17e-07 220.00 56.700
6:1,2*15,2*4,6  0.01400 1.02e-06 192.00 29.400
7:1,2*15,2*4,5,6  0.00600 4.49e-07 186.00 43.800
7:1,2*15,2*2*4,6  0.00433 5.39e-07 112.00 31.000
5:1,15,2*4,6   0.00867 5.81e-06  20.80  4.060
5:1,15,4,2*6   0.00733 5.22e-06  19.60  4.170
4:1,15,4,6   0.03770 2.71e-05  19.40  1.790

$chrom
 posterior prior  bf  bfse
4  0.2100 0.0595 15.00 0.529
15 0.1470 0.0464 13.40 0.589
6  0.1280 0.0534 10.10 0.483
1  0.2030 0.0901  9.55 0.345
> plot(tmp)
```

hyper: number of QTL posterior, prior, Bayes factors
R/qtlbim: 2-D (not 2-QTL) scans

```r
> two <- qb.scantwo(qbHyper, chr = c(6,15), type = "2logBF")
> plot(two)
> plot(two, chr = 6, slice = 15, show.locus = FALSE)
> plot(two, chr = 15, slice = 6, show.locus = FALSE)
> two <- qb.scantwo(qbHyper, chr = c(6,15), type = "LPD")
> plot(two, chr = 6, slice = 15, show.locus = FALSE)
> plot(two, chr = 15, slice = 6, show.locus = FALSE)
```

2-D plot of 2logBF: chr 6 & 15
R/qtlbim: slice of epistasis

```r
> slice = qb.slicetwo(qbHyper, c(6,15), c(59,19.5))
> summary(slice)

2logBF of bp for epistasis

<table>
<thead>
<tr>
<th>n.qtl</th>
<th>pos</th>
<th>m.pos</th>
<th>e.pos</th>
<th>epistasis</th>
<th>slice</th>
</tr>
</thead>
<tbody>
<tr>
<td>c6</td>
<td>0.838</td>
<td>59.0</td>
<td>59.0</td>
<td>66.7</td>
<td>15.8</td>
</tr>
<tr>
<td>c15</td>
<td>0.961</td>
<td>17.5</td>
<td>17.5</td>
<td>17.5</td>
<td>60.6</td>
</tr>
</tbody>
</table>

cellmean of bp for AA, HA, AH, HH

<table>
<thead>
<tr>
<th>n.qtl</th>
<th>pos</th>
<th>m.pos</th>
<th>AA</th>
<th>HA</th>
<th>AH</th>
<th>HH</th>
<th>slice</th>
</tr>
</thead>
<tbody>
<tr>
<td>c6</td>
<td>0.838</td>
<td>59.0</td>
<td>97.4</td>
<td>105</td>
<td>102</td>
<td>100.8</td>
<td>18.1</td>
</tr>
<tr>
<td>c15</td>
<td>0.961</td>
<td>17.5</td>
<td>99.8</td>
<td>103</td>
<td>104</td>
<td>98.5</td>
<td>60.6</td>
</tr>
</tbody>
</table>

estimate of bp for epistasis

<table>
<thead>
<tr>
<th>n.qtl</th>
<th>pos</th>
<th>m.pos</th>
<th>e.pos</th>
<th>epistasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>c6</td>
<td>0.838</td>
<td>59.0</td>
<td>66.7</td>
<td>-7.86</td>
</tr>
<tr>
<td>c15</td>
<td>0.961</td>
<td>17.5</td>
<td>17.5</td>
<td>-8.72</td>
</tr>
</tbody>
</table>
> plot(slice, figs = c("effects", "cellmean", "effectplot"))
```
selected publications

www.stat.wisc.edu/~yandell/statgen

- Broman et al. (2003 *Bioinformatics*)
  - R/qtl introduction
- Broman (2001 *Lab Animal*)
  - nice overview of QTL issues
  - overview/comparison of QTL methods
- Yandell et al. (2007 *Bioinformatics*)
  - R/qtlbim introduction
- Yi et al. (2005 *Genetics*)
  - methodology of R/qtlbim
many thanks

U AL Birmingham
Nengjun Yi
Tapan Mehta
Samprit Banerjee
Daniel Shriner
Ram Venkataraman
David Allison

Jackson Labs
Gary Churchill
Hao Wu
Hyuna Yang
Randy von Smith

USDA Hatch, NIH/NIDDK (Attie), NIH/R01s (Yi, Broman)

Alan Attie
Jonathan Stoehr
Hong Lan
Susie Clee
Jessica Byers
Mark Gray-Keller
Tom Osborn
David Butruille
Marcio Ferrera
Josh Udahl
Pablo Quijada

UW-Madison Stats
Yandell lab
Jaya Satagopan
Fei Zou
Patrick Gaffney
Chunfang Jin
Elias Chaibub
W Whipple Neely
Jee Young Moon
Michael Newton
Christina Kendziorski
Daniel Gianola
Liang Li
Daniel Sorensen

QTL: Why and How UW-Madison PBPG Yandell © 2007