examples in detail

- simulation study (after Stephens & Fisch (1998)) 2-3
- obesity in mice ($n = 421$) 4-12
  - epistatic QTLs with no main effects
- expression phenotype (SCD1) in mice ($n = 108$) 13-22
  - multiple QTL and epistasis
- mapping two correlated phenotypes 23-35
  - Jiang & Zeng 1995 paper
  - *Brassica napus* vernalization
- gonad shape in *Drosophila* spp. (insect) ($n = 1000$) 36-42
  - multiple traits reduced by PC
  - many QTL and epistasis

simulation with 8 QTL

- simulated F2 intercross, 8 QTL
  - (Stephens, Fisch 1998)
  - $n=200$, heritability = 50%
  - detected 3 QTL
- increase to detect all 8
  - $n=500$, heritability to 97%
loci pattern across genome

- notice which chromosomes have persistent loci
- best pattern found 42% of the time

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Count of 8000</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>3371</td>
</tr>
<tr>
<td>9</td>
<td>751</td>
</tr>
<tr>
<td>7</td>
<td>377</td>
</tr>
<tr>
<td>9</td>
<td>218</td>
</tr>
<tr>
<td>9</td>
<td>218</td>
</tr>
<tr>
<td>9</td>
<td>198</td>
</tr>
</tbody>
</table>

obesity in CAST/Ei BC onto M16i

- 421 mice (Daniel Pomp)
  - (213 male, 208 female)
- 92 microsatellites on 19 chromosomes
  - 1214 cM map
- subcutaneous fat pads
  - pre-adjusted for sex and dam effects
- Yi, Yandell, Churchill, Allison, Eisen, Pomp (2005) *Genetics*
non-epistatic analysis

single QTL LOD profile

multiple QTL Bayes factor profile

posterior profile of main effects in epistatic analysis

main effects & heritability profile

Bayes factor profile
posterior profile of main effects in epistatic analysis

model selection via Bayes factors for epistatic model

number of QTL QTL pattern
posterior probability of effects

model selection for pairs
scatterplot estimates of epistatic loci

stronger epistatic effects
studying diabetes in an F2

- segregating cross of inbred lines
  - B6.ob x BTBR.ob → F1 → F2
  - selected mice with ob/ob alleles at leptin gene (chr 6)
  - measured and mapped body weight, insulin, glucose at various ages
    (Stoehr et al. 2000 Diabetes)
  - sacrificed at 14 weeks, tissues preserved
- gene expression data
  - Affymetrix microarrays on parental strains, F1
    - key tissues: adipose, liver, muscle, β-cells
    - novel discoveries of differential expression (Nadler et al. 2000 PNAS; Lan et al. 2002 in review; Ntambi et al. 2002 PNAS)
  - RT-PCR on 108 F2 mice liver tissues
    - 15 genes, selected as important in diabetes pathways
    - SCD1, PEPCK, ACO, FAS, GPAT, PPARgamma, PPARalpha, G6Pase, PDI,…

Multiple Interval Mapping (QTLCart)
SCD1: multiple QTL plus epistasis!
Bayesian model assessment: number of QTL for SCD1

Bayesian LOD and $h^2$ for SCD1
Bayesian model assessment:
chromosome QTL pattern for SCD1

trans-acting QTL for SCD1
(no epistasis yet: see Yi, Xu, Allison 2003)
2-D scan: assumes only 2 QTL!

epistasis LOD peaks

Joint LOD peaks

sub-peaks can be easily overlooked!
epistatic model fit

Cockerham epistatic effects
co-mapping multiple traits

- 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

interplay of pleiotropy & correlation

pleiotropy only

Korol et al. (2001)

QTL 2: Data
Seattle SISG: Yandell © 2009
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

pleiotropy or close linkage?

2 traits, 2 qtl/trait
pleiotropy @ 54cM
linkage @ 114,128cM
Brassica napus: 2 correlated traits

- 4-week & 8-week vernalization effect
  - log(days to flower)
- genetic cross of
  - Stellar (annual canola)
  - Major (biennial rapeseed)
- 105 F1-derived double haploid (DH) lines
  - homozygous at every locus (QQ or qq)
- 10 molecular markers (RFLPs) on LG9
  - two QTLs inferred on LG9 (now chromosome N2)
  - corroborated by Butruille (1998)
  - exploiting synteny with Arabidopsis thaliana

QTL with GxE or Covariates

- adjust phenotype by covariate
  - covariate(s) = environment(s) or other trait(s)
- additive covariate
  - covariate adjustment same across genotypes
  - “usual” analysis of covariance (ANCOVA)
- interacting covariate
  - address GxE
  - capture genotype-specific relationship among traits
- another way to think of multiple trait analysis
  - examine single phenotype adjusted for others
R/qtl & covariates

- additive and/or interacting covariates
- test for QTL after adjusting for covariates

```r
## Get Brassica data.
library(qtlbim)
data(Bnapus)
Bnapus <- calc.genoprob(Bnapus, step = 2, error = 0.01)

## Scatterplot of two phenotypes: 4wk & 8wk flower time.
plot(Bnapus$pheno$log10flower4, Bnapus$pheno$log10flower8)

## Unadjusted IM scans of each phenotype.
fl8 <- scanone(Bnapus, find.pheno(Bnapus, "log10flower8"))
fl4 <- scanone(Bnapus, find.pheno(Bnapus, "log10flower4"))
plot(fl4, fl8, chr = "N2", col = rep(1,2), lty = 1:2,
     main = "solid = 4wk, dashed = 8wk", lwd = 4)
```

QTL 2: Data Seattle SISG: Yandell © 2009
R/qtl & covariates

- additive and/or interacting covariates
- test for QTL after adjusting for covariates

```r
## IM scan of 8wk adjusted for 4wk.
## Adjustment independent of genotype
fl8.4 <- scanone(Bnapus, find.pheno(Bnapus, "log10flower8"),
                 addcov = Bnapus$pheno$log10flower4)

## IM scan of 8wk adjusted for 4wk.
## Adjustment changes with genotype.
fl8.4 <- scanone(Bnapus, find.pheno(Bnapus, "log10flower8"),
                 intcov = Bnapus$pheno$log10flower4)

plot(fl8, fl8.4a, fl8.4, chr = "N2",
     main = "solid = 8wk, dashed = addcov, dotted = intcov")
```
scatterplot adjusted for covariate

```r
## Set up data frame with peak markers, traits.
markers <- c("E38M50.133","ec2e5a","wg7f3a")
tmpdata <- data.frame(pull.geno(Bnapus)[,markers])
tmpdata$fl4 <- Bnapus$pheno$log10flower4
tmpdata$fl8 <- Bnapus$pheno$log10flower8

## Scatterplots grouped by marker.
library(lattice)
xyplot(fl8 ~ fl4, tmpdata, group = wg7f3a,
   col = "black", pch = 3:4, cex = 2,
   xlab = "log10(4wk flower time)",
   ylab = "log10(8wk flower time)",
   main = "marker at 47cM")
xyplot(fl8 ~ fl4, tmpdata, group = E38M50.133,
   col = "black", pch = 3:4, cex = 2,
   xlab = "log10(4wk flower time)",
   ylab = "log10(8wk flower time)",
   main = "marker at 80cM")
```
R/qtlbim and GxE

• similar idea to R/qtl
  – fixed and random additive covariates
  – GxE with fixed covariate
• multiple trait analysis tools coming soon
  – theory & code mostly in place
  – properties under study
  – expect in R/qtlbim later this year
  – Samprit Banerjee (N Yi, advisor)

reducing many phenotypes to 1

• *Drosophila mauritiana* x *D. simulans*
  – reciprocal backcrosses, ~500 per bc
• response is “shape” of reproductive piece
  – trace edge, convert to Fourier series
  – reduce dimension: first principal component
• many linked loci
  – brief comparison of CIM, MIM, BIM
PC for two correlated phenotypes

shape phenotype via PC

Figure 5.—A plot of the first two principal components of the Fourier coefficients from posterior lobe outlines. Many individuals from each of five genotypic classes are represented. Each point represents an average of scores from the left and right sides of an individual (with a few exceptions for which the score is from one side only). The percentage of variation in the Fourier coefficients accounted for by each principal component is given in parentheses. Li et al. (1996) Genetics
shape phenotype in BC study indexed by PC1

QTL 2: Data Seattle SISG: Yandell © 2009

Zeng et al. (2000)
CIM vs. MIM

composite interval mapping
(Liu et al. 1996)
narrow peaks
miss some QTL

multiple interval mapping
(Zeng et al. 2000)
triangular peaks

both conditional 1-D scans
fixing all other "QTL"
CIM, MIM and IM pairscan

multiple QTL: CIM, MIM and BIM