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%

posterior
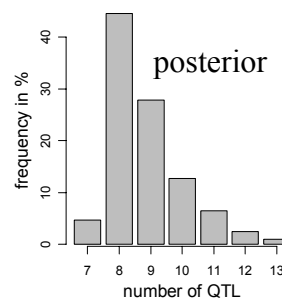
QTL	chr	loci	effect
1	1	11	–3
2	1	50	–5
3	3	62	+2
4	6	107	–3
5	6	152	+3
6	8	32	–4
7	8	54	+1
8	9	195	+2
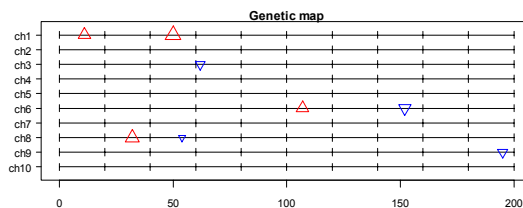
loci pattern across genome

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

Chromosome

| *m*	1	2	3	4	5	6	7	8	9	10	Count of 8000
8	**2**	**0**	**1**	**0**	**0**	**2**	**0**	**2**	**1**	**0**	3371
9	*3*	0	1	0	0	2	0	2	1	0	751
7	2	0	1	0	0	2	0	*1*	1	0	377
9	2	0	1	0	0	2	0	2	1	0	218
9	2	0	1	0	0	*3*	0	2	1	0	218
9	2	0	1	0	0	2	0	2	*2*	0	198

---

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*

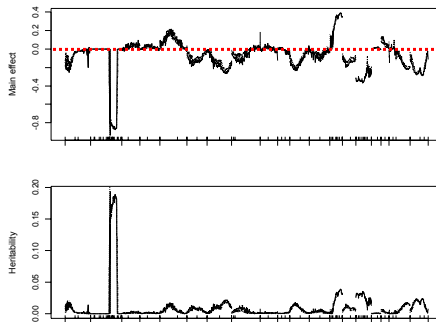# non-epistatic analysis
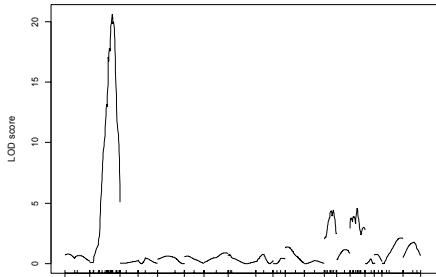
single QTL LOD profile
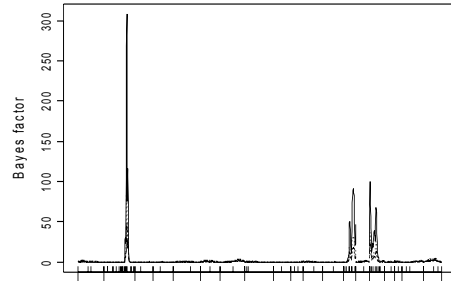
multiple QTL
Bayes factor profile

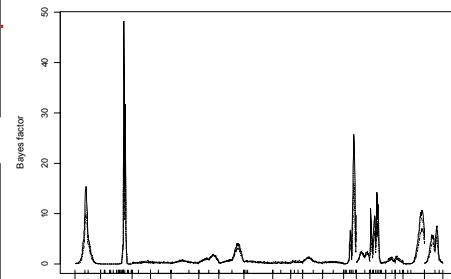# posterior profile of main effects
in epistatic analysis

main effects & heritability profile

Bayes factor profile

posterior profile of main effects in epistatic analysis

result.5b summaries with pattern ≥ 1,2,7,13,15,18,19
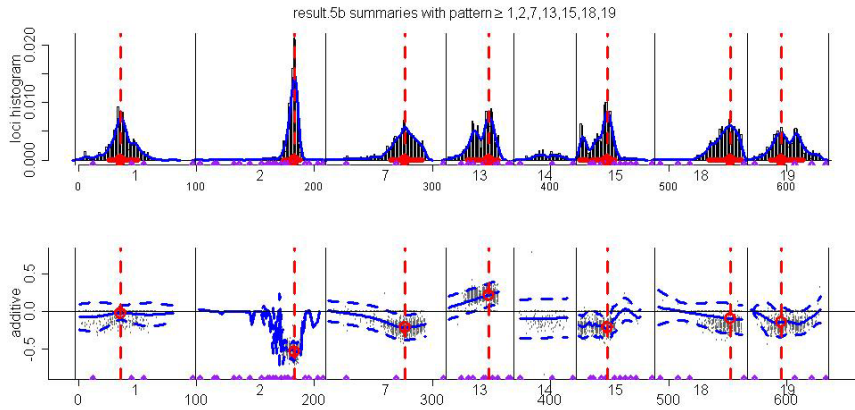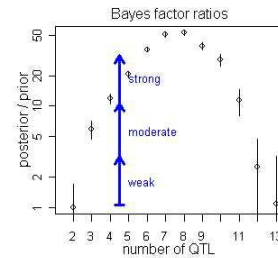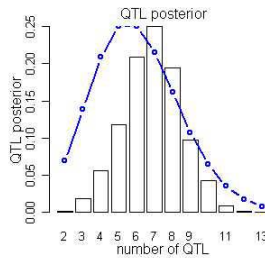

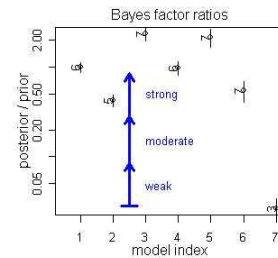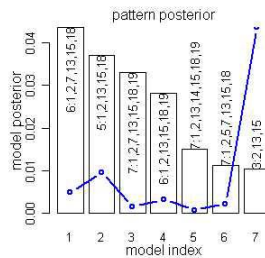QTL 2: Data   Seattle SISG: Yandell © 2008   7

---

model selection
via
Bayes factors
for
epistatic model

number of QTL

QTL pattern

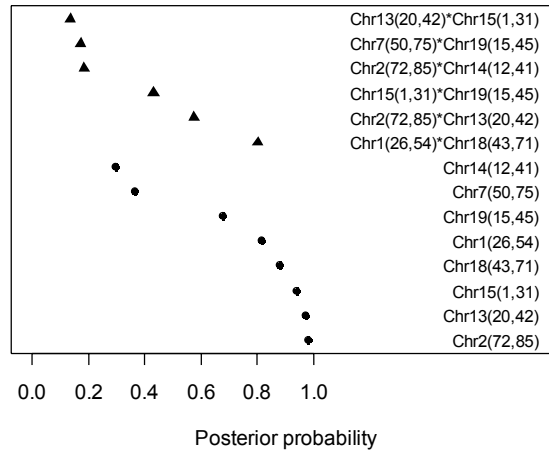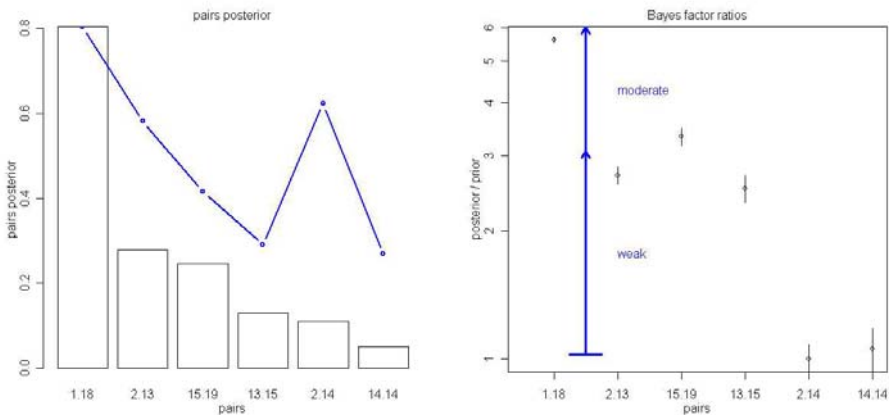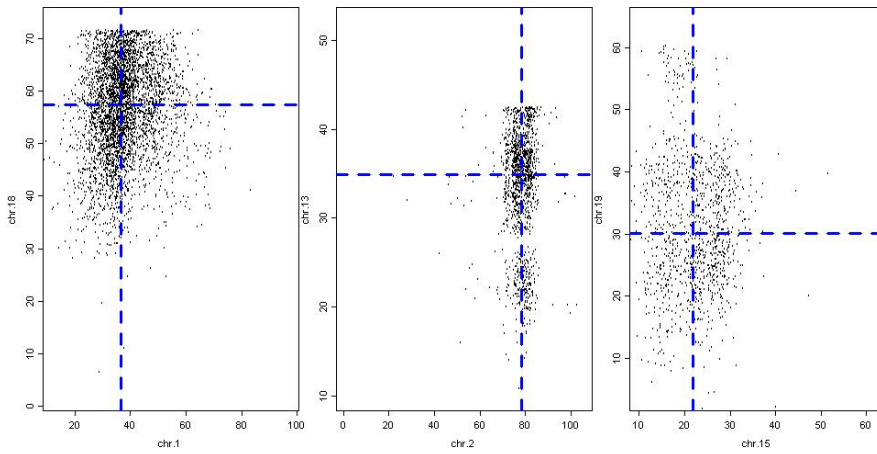QTL 2: Data   Seattle SISG: Yandell © 2008   8

posterior probability of effects

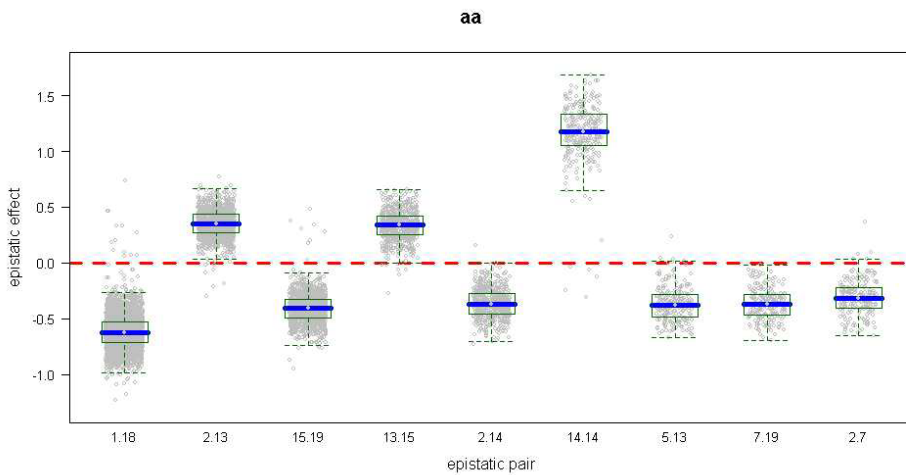# model selection for pairs

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!
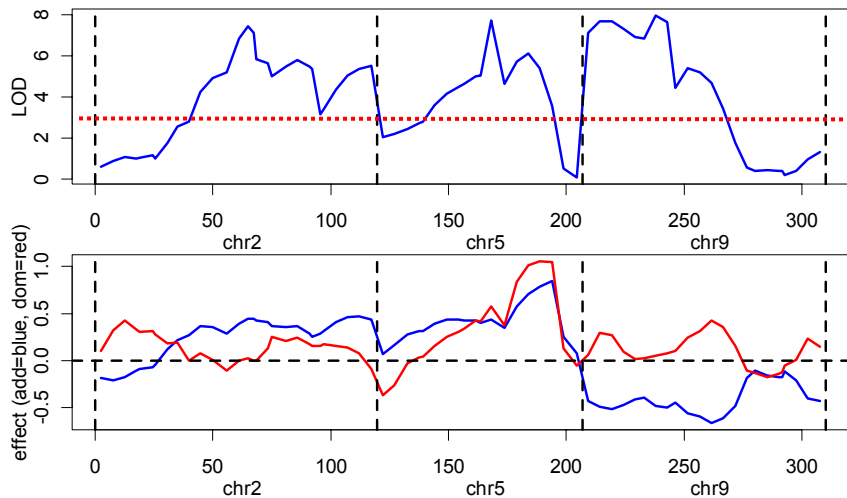
Bayesian model assessment: number of QTL for SCD1

QTL posterior

Bayes factor ratios
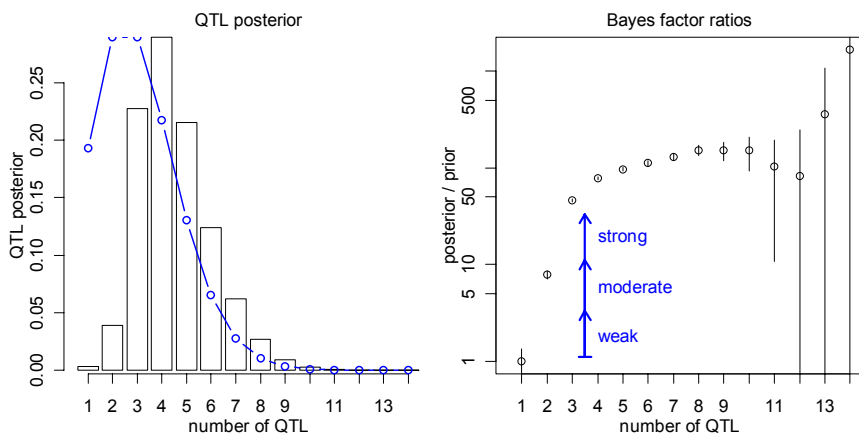
Bayesian LOD and h^2 for SCD1
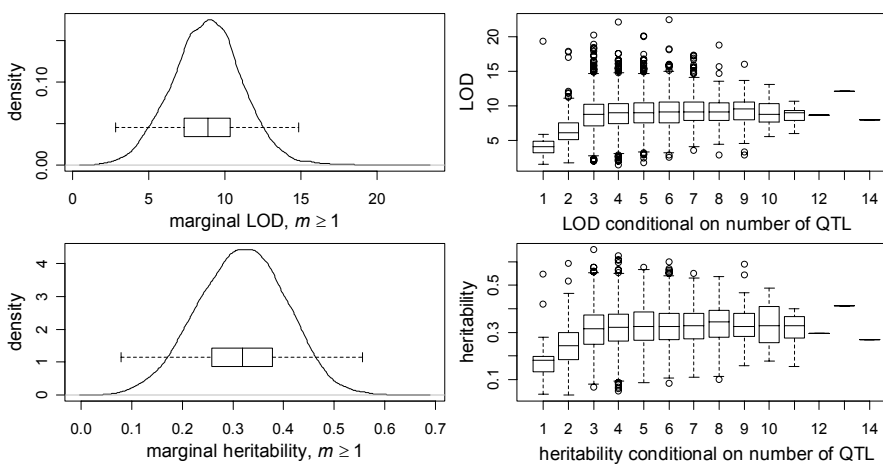
Bayesian model assessment: chromosome QTL pattern for SCD1

pattern posterior

Bayes factor ratios
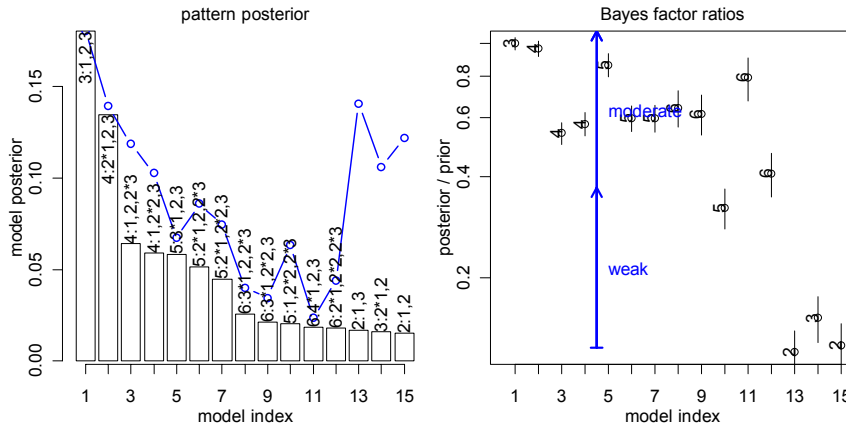
trans-acting QTL for SCD1
(no epistasis yet: see Yi, Xu, Allison 2003)

hong7scd.bim summaries with pattern ≥ ch2,ch5,ch9

dominance?
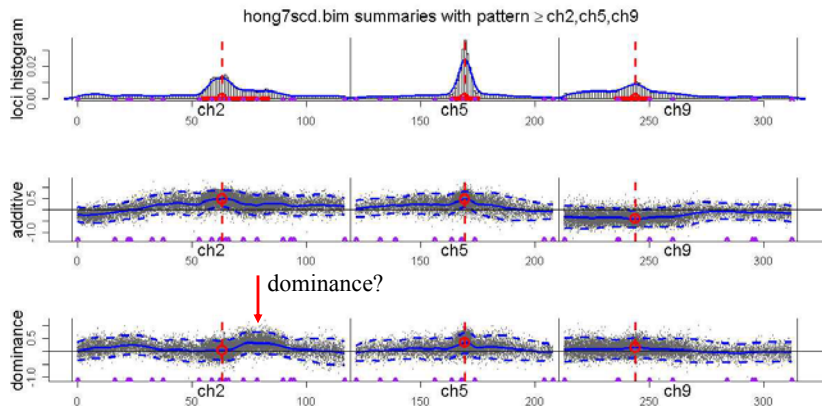
2-D scan: assumes only 2 QTL!

SCD on chr 2,5,9

epistasis
LOD
peaks

joint
LOD
peaks
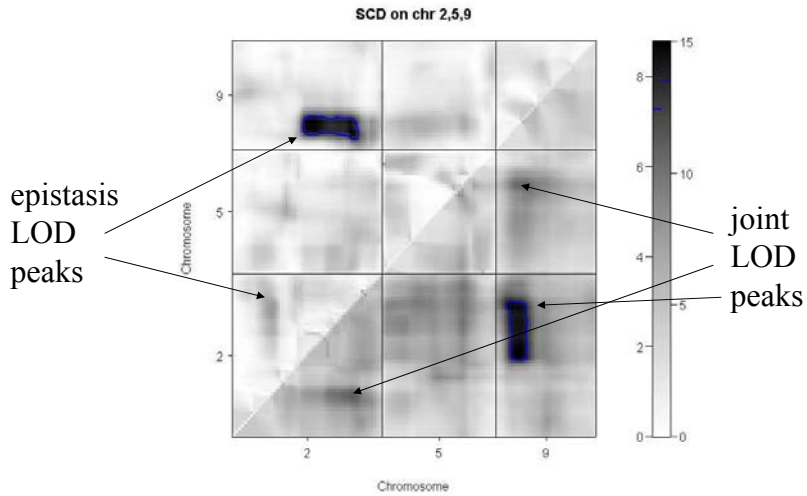
sub-peaks can be easily overlooked!
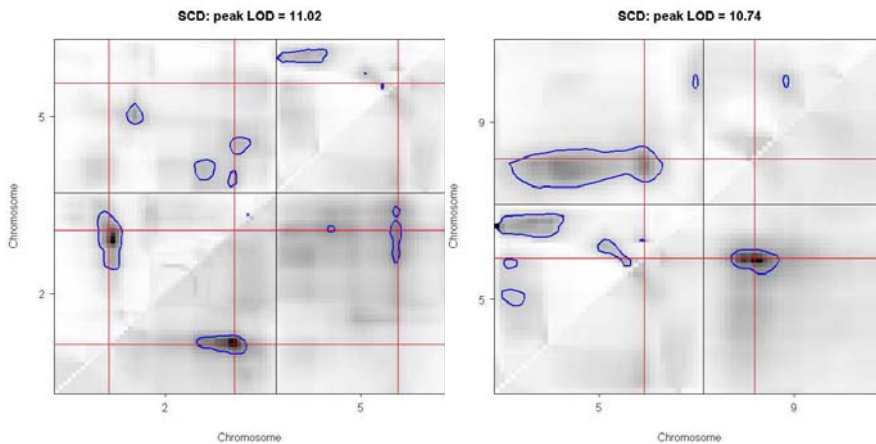
SCD: peak LOD = 11.02   **SCD: peak LOD = 10.74**

epistatic model fit
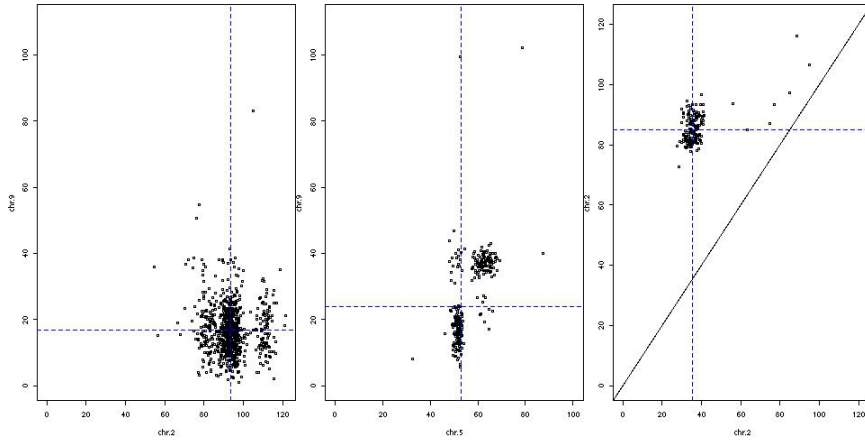
Cockerham epistatic effects
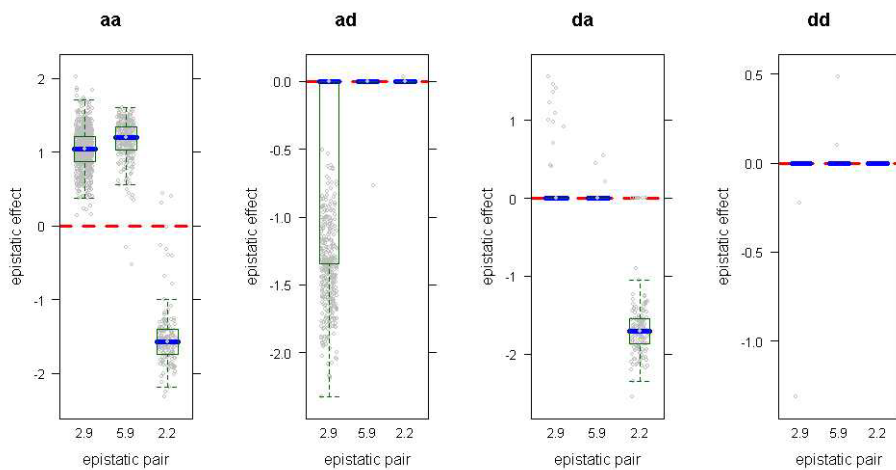
co-mapping multiple traits

- 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
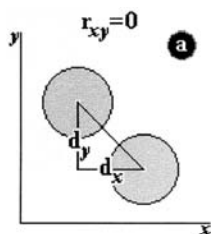
pleiotropy only   correlation only   both
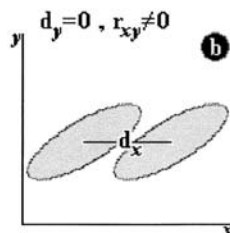
Korol et al. (2001)

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

ρ_P = 0.06, ρ_G = 0.68, ρ_E = -0.2

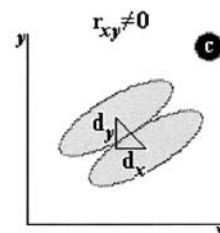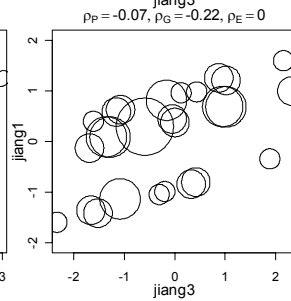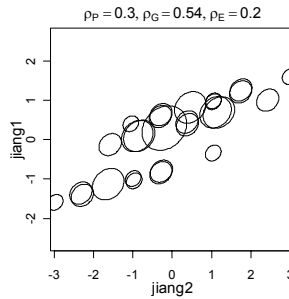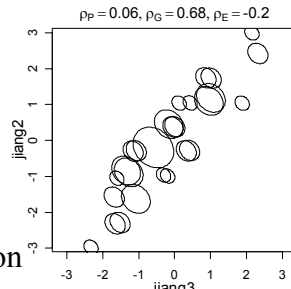ρ_P = 0.3, ρ_G = 0.54, ρ_E = 0.2

ρ_P = -0.07, ρ_G = -0.22, ρ_E = 0

pleiotropy or close linkage?

2 traits, 2 qtl/trait
pleiotropy @ 54cM
linkage @ 114,128cM
Jiang Zeng (1995)
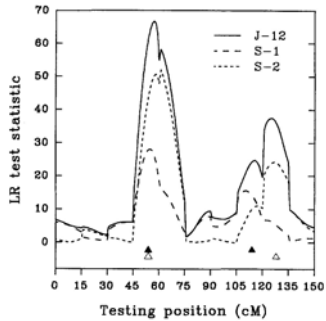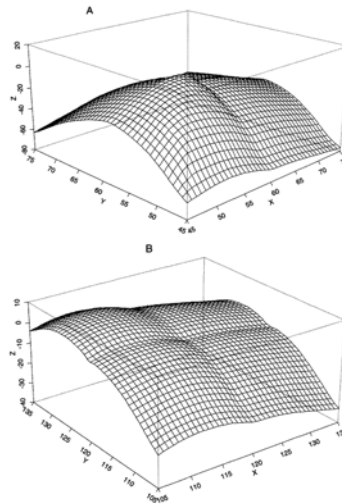
FIGURE 2.—Two-dimensional log-likelihood surfaces (expressed as deviations from the maximum of the log-likelihoods on the diagonal) for the test of pleiotropy vs. close linkage are presented for two regions: the region between 45 and 75 cM of Figure 1(A) and the region between 105 and 135 cM (B). X is the testing position for a QTL affecting trait 1 and Y is the testing position for a QTL affecting trait 2. On the diagonal of X-Y plane, two QTL are located in the same position and statistically are treated as one pleiotropic QTL. Z is the likelihood ratio test statistic scaled to zero at the maximum point of the diagonal.

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

```
## 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)
```
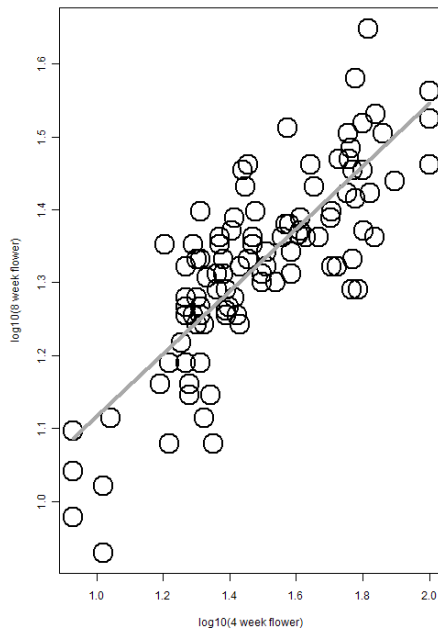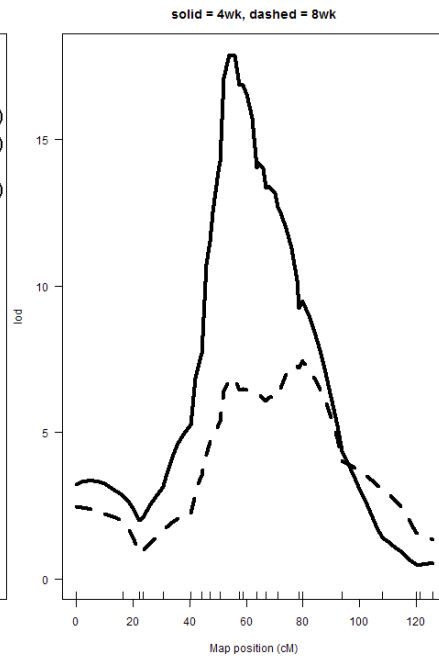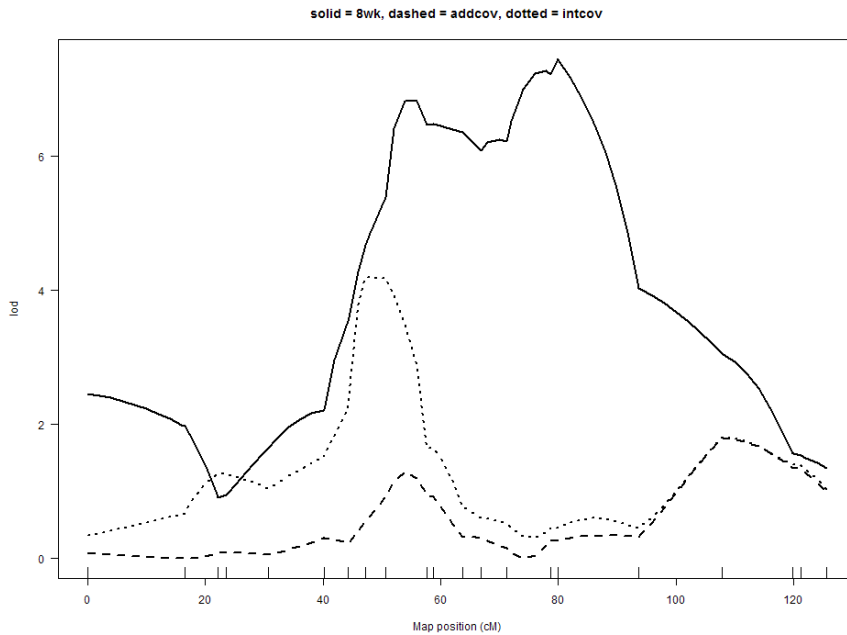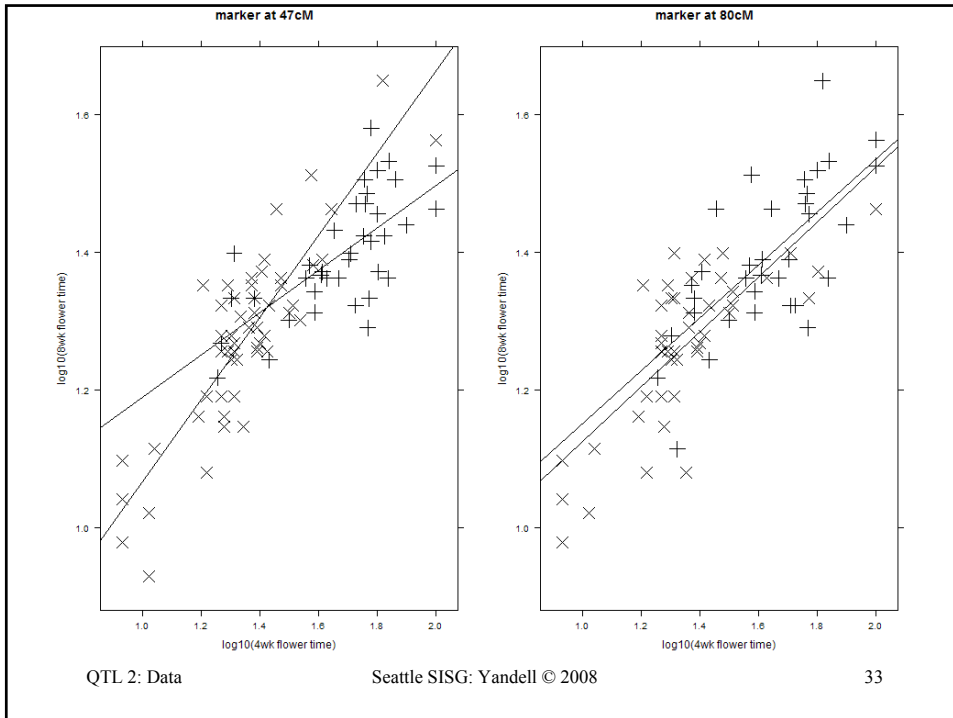
R/qtl & covariates

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

```
## 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")
```

---

solid = 8wk, dashed = addcov, dotted = intcov

scatterplot adjusted for covariate

```
## 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, type = c("p","r"),
  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, type = c("p","r"),
  xlab = "log10(4wk flower time)",
  ylab = "log10(8wk flower time)",
  main = "marker at 80cM")
```
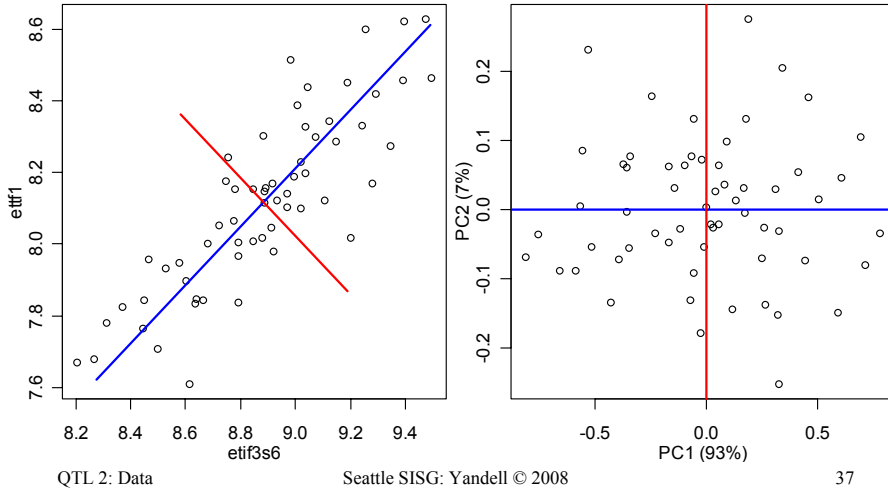
R/qtlbim and GxE

- 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 2.—The effect of harmonic number on the accuracy of reconstruction of a posterior lobe outline by elliptical Fourier analysis.

- ■ *simulans*
- ▲ *mauritiana*
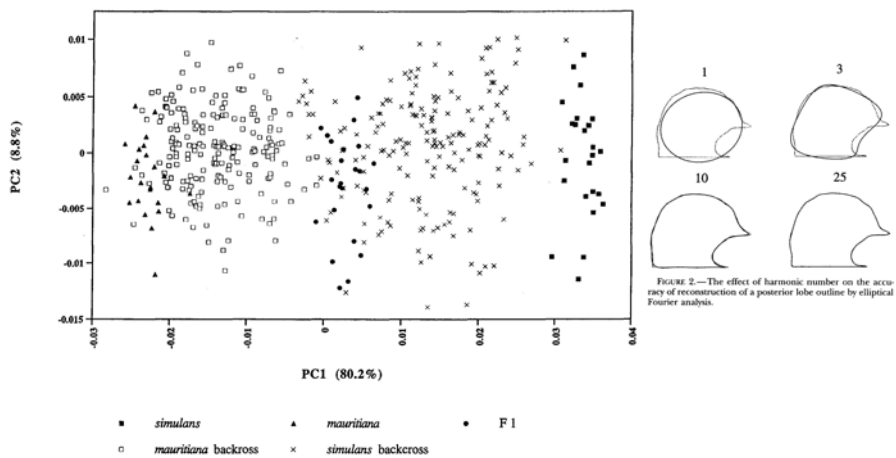- ● F 1
- □ *mauritiana* backross
- ✕ *simulans* backcross

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.   Liu et al. (1996) *Genetics*
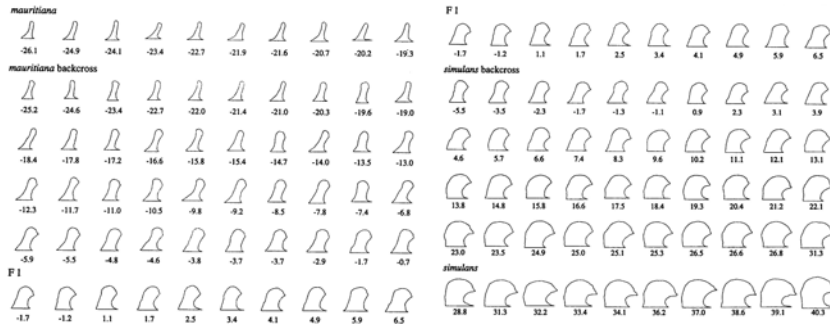
shape phenotype in BC study indexed by PC1

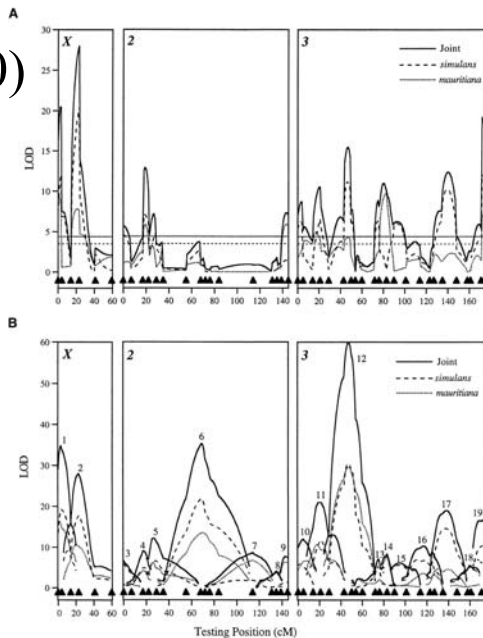Liu et al. (1996) *Genetics*

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
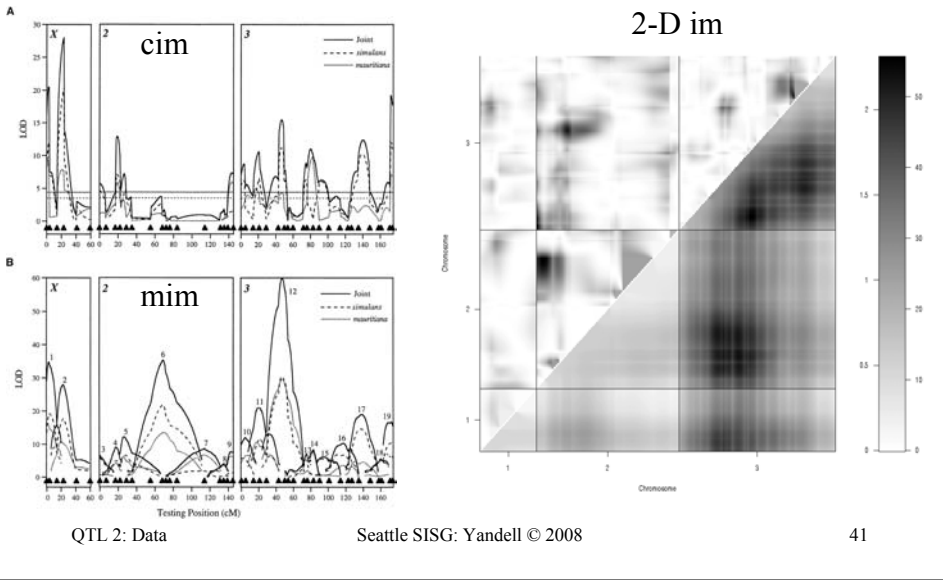   fixing all other "QTL"

CIM, MIM and IM pairscan

cim

mim

2-D im

multiple QTL: CIM, MIM and BIM

cim

mim

mauritiana bc

simulans bc