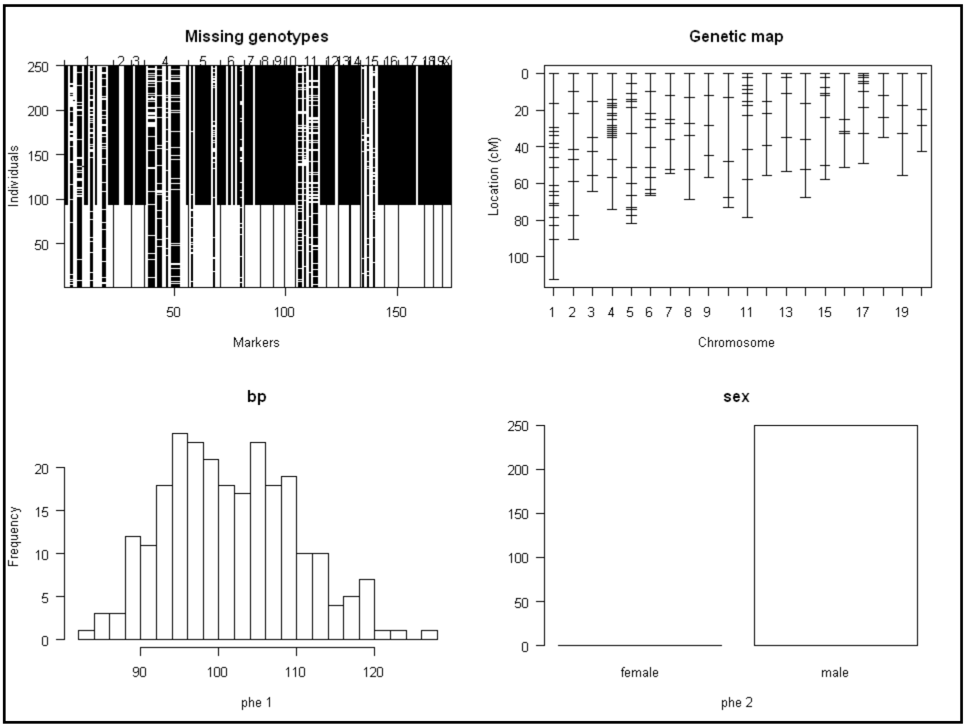
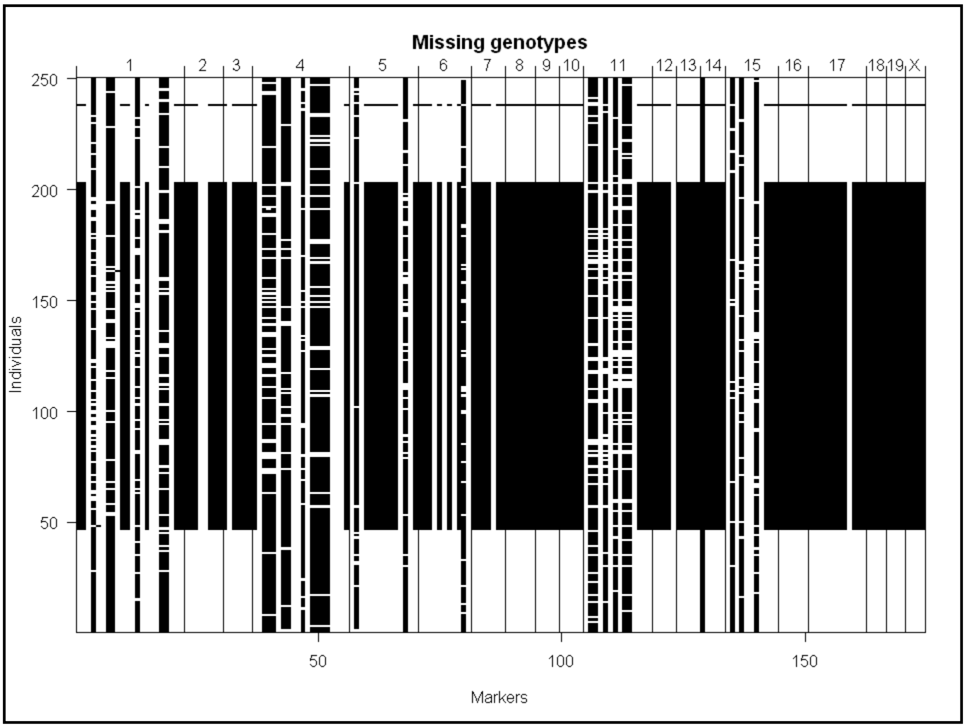
R/qtl & R/qtlbim Tutorials

- R statistical graphics & language system
- R/qtl tutorial
  - R/qtl web site: www.rqtl.org
  - Tutorial: www.rqtl.org/tutorials/rqtltour.pdf
  - R code: www.rqtl.org/tutorials/rqtltour.R
- R/qtlbim tutorial
  - R/qtlbim web site: www.qtlbim.org
  - Tutorial: www.stat.wisc.edu/~yandell/qtlbim/rqtltour.pdf
  - R code: www.stat.wisc.edu/~yandell/qtlbim/rqtltour.R

R/qtl tutorial (www.rqtl.org)

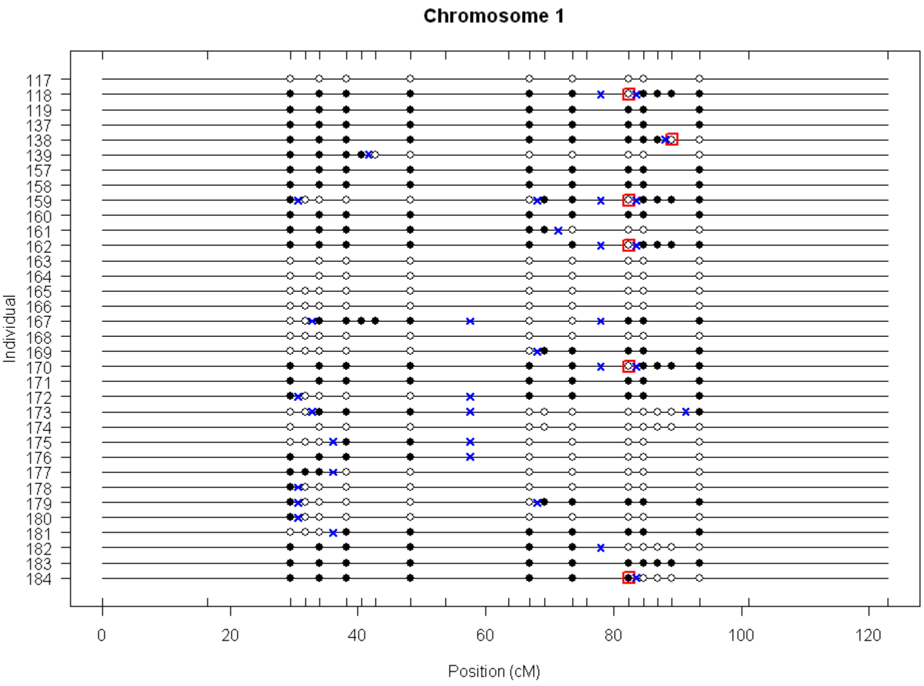
```
> library(qtl)
> data(hyper)
> summary(hyper)
    Backcross

    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 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)
```


Missing genotypes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
Individuals
250
200
150
100
50
50
100
150
Markers
Genetic map
Location (cM)
0
20
40
60
80
100
1 2 3 4 5 6 7 8 9 11 13 15 17 19
Chromosome
bp
Frequency
0
5
10
15
20
90
100
110
120
phe 1
sex
250
200
150
100
50
0
female
male
phe 2


Missing genotypes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
Individuals
250
200
150
100
50
50
100
150
Markers

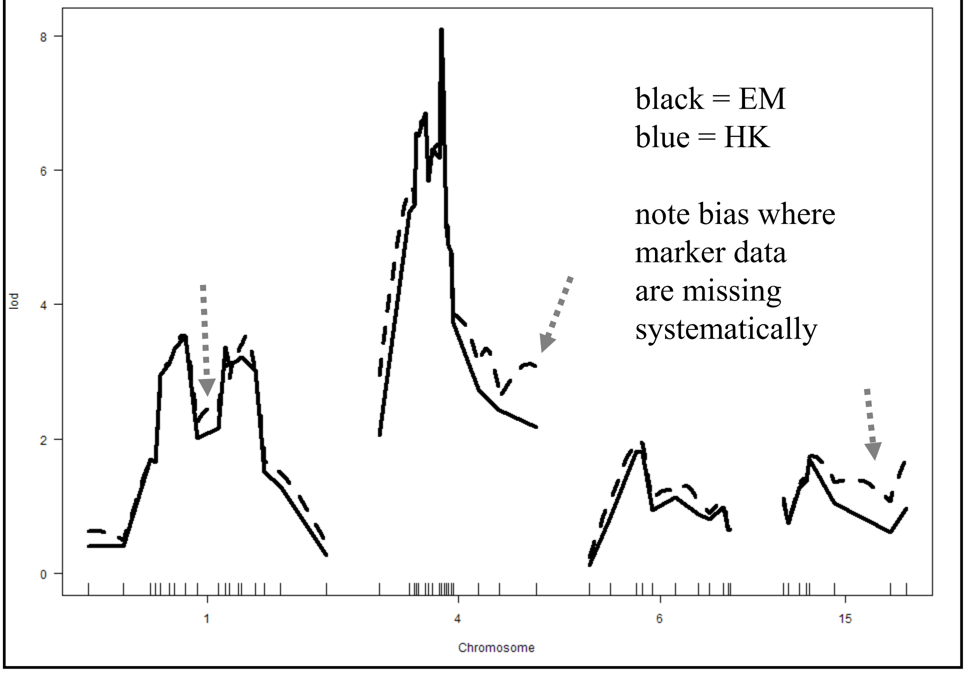
R/qtl: find genotyping errors

```
> hyper <- calc.errorlod(hyper, error.prob=0.01)
> top.errorlod(hyper)

   chr  id    marker errorlod
1    1 118   D1Mit14 8.372794
2    1 162   D1Mit14 8.372794
3    1 170   D1Mit14 8.372794
4    1 159   D1Mit14 8.350341
5    1  73   D1Mit14 6.165395
6    1  65   D1Mit14 6.165395
7    1  88   D1Mit14 6.165395
8    1 184   D1Mit14 6.151606
9    1 241   D1Mit14 6.151606
...
16   1 215  D1Mit267 5.822192
17   1 108  D1Mit267 5.822192
18   1 138  D1Mit267 5.822192
19   1 226  D1Mit267 5.822192
20   1 199  D1Mit267 5.819250
21   1  84  D1Mit267 5.808400

> plot.geno(hyper, chr=1, ind=c(117:119,137:139,157:184))
```



R/qtl: 1 QTL interval mapping

```
> 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

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

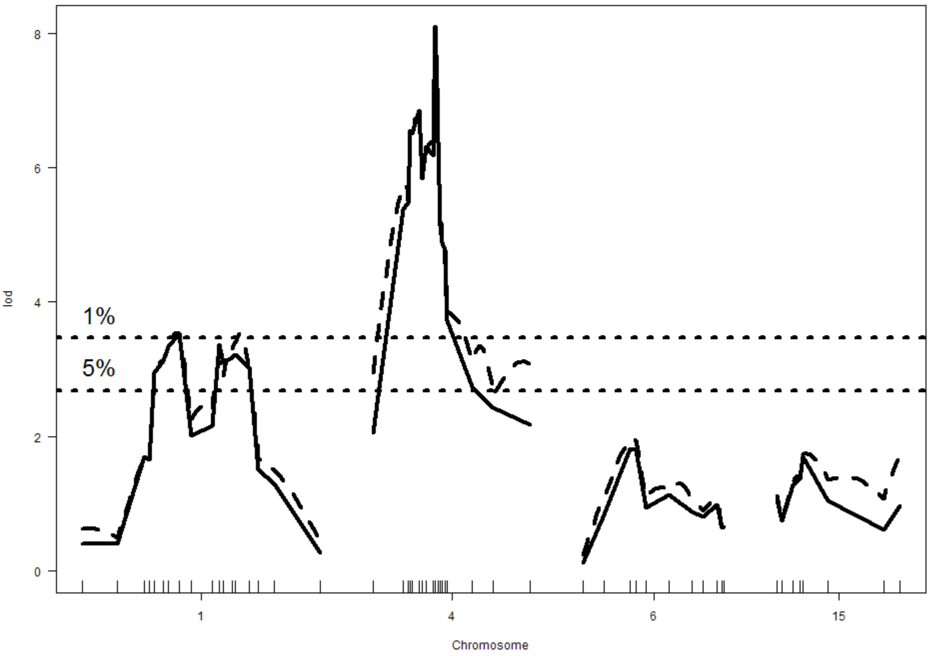


black = EM
blue = HK

note bias where
marker data
are missing
systematically

R/qtl: permutation threshold

```
> 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
```

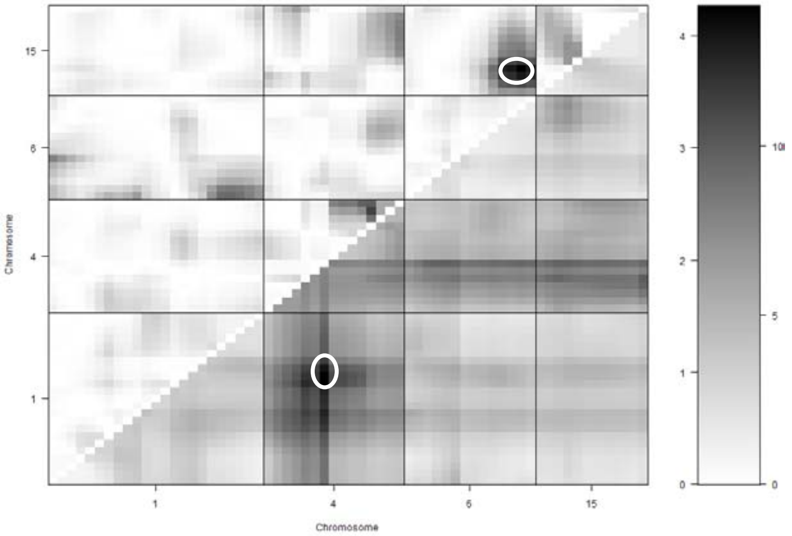


R/qtl: 2 QTL scan

```
> 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))

        pos1f pos2f lod.full lod.fv1 lod.int     pos1a pos2a lod.add lod.av1
c1 :c4   68.3  30.0    14.13    6.51   0.225      68.3  30.0   13.90   6.288
c2 :c19  47.7   0.0     6.71    5.01   3.458      52.7   0.0    3.25   1.552
c3 :c3   37.2  42.2     6.10    5.08   0.226      37.2  42.2    5.87   4.853
c6 :c15  60.0  20.5     7.17    5.22   3.237      25.0  20.5    3.93   1.984
c9 :c18  67.0  37.2     6.31    4.79   4.083      67.0  12.2    2.23   0.708
c12:c19   1.1  40.0     6.48    4.79   4.090       1.1   0.0    2.39   0.697

> plot(out2.hk, chr=c(1,4,6,15))
```

---

upper triangle/left scale: epistasis LOD
lower triangle/right scale: 2-QTL LOD



R/qtl: ANOVA imputation at QTL

```
> 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.col = 1, qtl, formula = my.formula)
> summary(out.fitqtl)

Full model result
----------------------------------
Model formula is:  y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q4:Q5

       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 F value 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
```

selected R/qtl publications

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

- www.rqtl.org
- tutorials and code at web site
  - www.rqtl.org/tutorials
- Broman et al. (2003 *Bioinformatics*)
  - R/qtl introduction
- Broman (2001 *Lab Animal*)
  - nice overview of QTL issues