

Case study II

In this chapter, we present a second case study. This case study illustrates an investigation of interactions between QTL and a covariate. It also shows how we may deal with genotype data organized in linkage groups (as opposed to chromosomes). Data of that type are rare for established model organisms, but common for emerging ones without extensive genomic resources. As with the case study in Chap. 10, there are some unusual features here, but most QTL analyses will be unusual in one way or another.

We consider the data of Nichols *et al.* (2007), included in the R/qtlbook package as the data set `trout`. This is a set of doubled haploid individuals derived from a cross between the Oregon State University (OSU) and Clearwater (CW) River rainbow trout (*Oncorhynchus mykiss*) clonal lines. Doubled haploids are similar to a backcross, but with a single recombinant genome being doubled (rather than matched to a nonrecombinant genome), so that at any genomic position, individuals are homozygous for one of two possible genotypes.

Eggs from one of eight outbred females, two from Troutlodge (TL) and six from the Spokane (SP) hatchery, were irradiated to destroy maternal nuclear DNA and fertilized with sperm from a single F₁ male. The first embryonic cleavage was blocked by heat shock to restore diploidy. There are a total of 554 individuals, with between 8 and 168 individuals from each of the eight females.

The primary phenotype is time to hatch (`tth`). An additional “phenotype,” `female`, indicates the maternal cytoplasmic environment (MCE; the source of the egg).

There are data on 171 markers on 28 linkage groups. The linkage groups are named as in Nichols *et al.* (2003), though a pair of markers are assigned to linkage group “un,” as they do not connect to any of the linkage groups in Nichols *et al.* (2003). Some care is required in referring to the linkage groups in R/qtl code; in some cases we will need to refer to the linkage group names in double-quotes (e.g., “1”).

As with all QTL data analyses, we begin with an exploration of the phenotype and genotype data. In this case, further light is shed on the linkage groups, and on MCE. Next, we analyze the data for QTL without regard to interactions with MCE. We follow by considering QTL $\times$ MCE interactions, and we conclude with a discussion of the analysis.

11.1 Diagnostics

We begin by exploring the phenotype and genotype data. We must first load the R/qtl and R/qtlbook packages, and then the data set, `trout`.

```
> library(qtl)
> library(qtlbook)
> data(trout)
```

Note that the cross type is "dh", indicating doubled haploids. In R/qtl, they are treated just like a backcross, except in references to the names of the genotypes.

```
> class(trout)

[1] "dh"      "cross"
```

If one were to import such data into R using `read.cross`, it would initially be read as a backcross. One would use code something like the following.

```
> trout <- read.cross("csv", file="trout.csv",
+                    genotypes=c("C", "O"), alleles=c("C", "O"))
```

One would then need to change the cross type to "dh", as follows.

```
> class(trout)[1] <- "dh"
```

Turning back to the data, let us first consider a quick summary.

```
> summary(trout)
```

Doubled haploids	
No. individuals:	554
No. phenotypes:	2
Percent phenotyped:	100 100
No. chromosomes:	28
Autosomes:	1 2 3 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 27 29 31 un
Total markers:	171

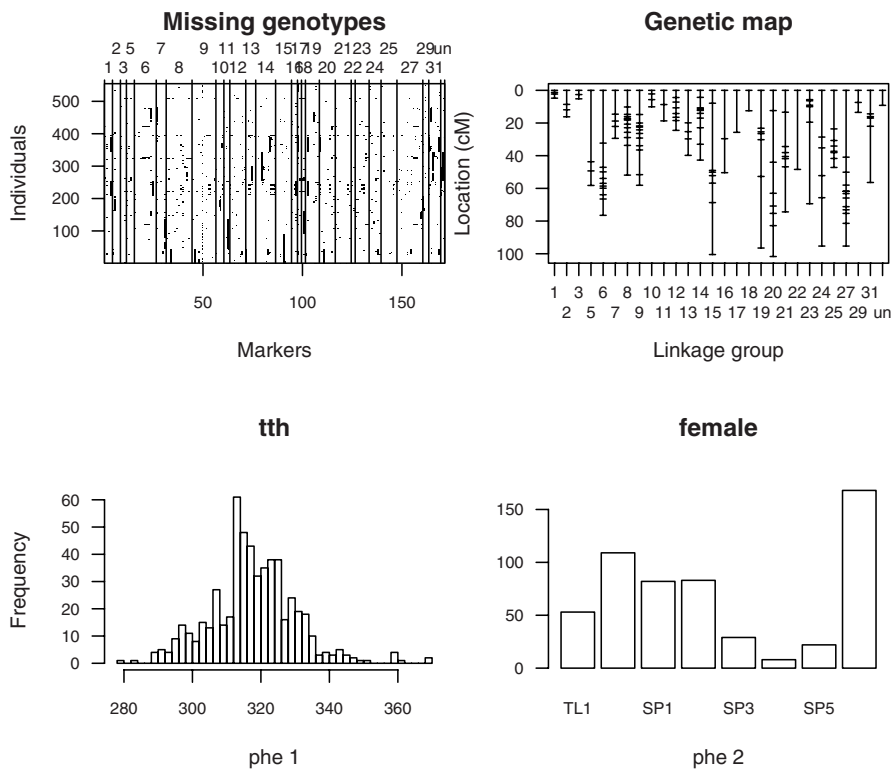


Figure 11.1. The summary plot of the `trout` data provided by the `plot.cross` function, including the pattern of missing genotype data (upper left; black pixels indicate missing data), the genetic map of the typed markers (upper right), a histogram of the `tth` phenotype (time to hatch), and a bar plot of the `female` phenotype, which indicates the maternal cytoplasmic environment (MCE; the source of the egg for each individual).

```
No. markers:      4 4 3 4 11 5 13 12 4 3 8 5 10 8 3 2 2 7
                  8 8 2 7 6 8 13 3 6 2
Percent genotyped: 95.2
Genotypes (%):    CC:50  00:50
```

This is just as described above. The C and O alleles refer to the CW and OSU clonal lines, respectively.

We make a summary plot as follows; see Fig. 11.1.

```
> plot(trout)
```

There are some very large gaps in the genetic map, but the marker genotype data are remarkably complete.

We first investigate whether there are differences in the `tth` phenotype, among the eight sources for eggs. We begin with a set of box plots.

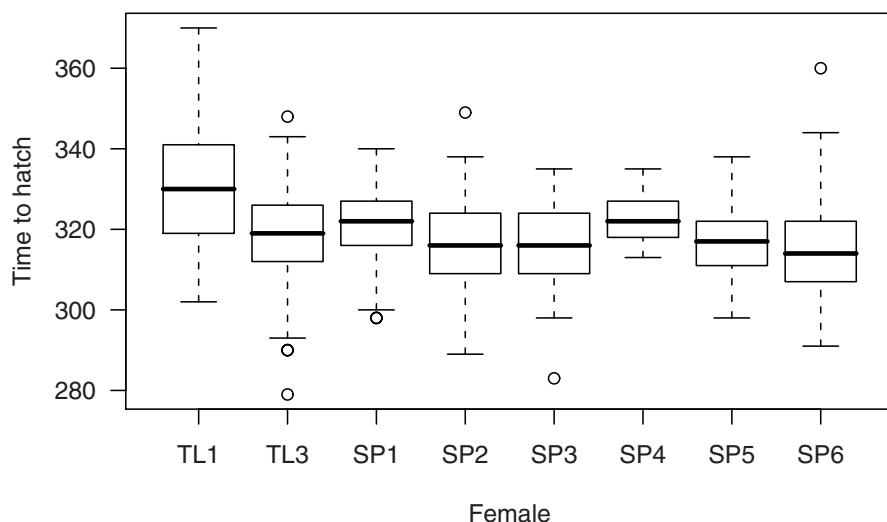


Figure 11.2. Box plots of the `tth` phenotype (time to hatch) according to the female source of the egg for the individuals, for the `trout` data.

```
> boxplot(tth ~ female, data=trout$pheno,
+         xlab="Female", ylab="Time to hatch")
```

There are clear differences among the females (see Fig. 11.2), particularly in that `tth` is larger for individuals whose egg came from the TL1 female.

An analysis of variance (ANOVA) will make clear that the observed differences cannot reasonably be ascribed to chance variation.

```
> anova(aov(tth ~ female, data=trout$pheno))
```

Analysis of Variance Table

Response: tth

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
female	7	12757	1822	13.5	3.6e-16
Residuals	546	73510	135		

Turning now to the genotype data, we first look at the segregation patterns for each marker with `geno.table`. Since there are 171 markers, this will make quite a long table, and so we focus on those markers for which the two genotypes deviate significantly from the expected 1:1 ratio. Applying a Bonferroni correction for the number of markers, we pull out the unusual ones as follows.

```
> gt <- geno.table(trout)
> gt[gt$P.value < 0.05/totmar(trout),]
```

	chr	missing	CC	00	P.value
AGCCGT8	8	11	320	223	3.145e-05
agcagc11	13	49	187	318	5.562e-09
ACCAAG16	19	15	396	143	1.185e-27

There are three markers that are behaving oddly, with marker ACCAAG16 segregating closer to 3:1 than 1:1. We do not know whether this is due to segregation distortion or genotyping error (in which case we might omit these markers), and so we will leave them in, but we should pay attention to these regions in the later results.

We next turn to the estimated recombination fractions between all pairs of markers, estimated via `est.rf` and plotted via `plot.rf`. We use `alternate.chrid=TRUE` to make the names of the many linkage groups more easily distinguished.

```
> trout <- est.rf(trout)
> plot.rf(trout, alternate.chrid=TRUE)
```

The results (Fig. 11.3) are a bit of a surprise. There are five pairs of linkage groups that are quite tightly associated (with large LOD scores, in red), but not tightly linked (estimated recombination fractions > 0.5 , in blue). Namely, the pairs of linkage groups 2 and 29, 10 and 18, 12 and 16, 14 and 20, and 27 and 31, all look to be tightly associated. We can focus on just these ten linkage groups, to make the observation more clear (see Fig. 11.4).

```
> plot.rf(trout, chr=c(2,29, 10,18, 12,16, 14,20, 27,31),
+         alternate.chrid=TRUE)
```

Let's look at a table of two-locus genotypes for a marker from a couple of these linkage groups, to figure out what is going on. We can use `find.marker` to pull out the first marker on each of linkage groups 10 and 18, and then use `geno.crosstab` to create a table of the two-locus genotypes.

```
> mar <- find.marker(trout, c(10,18), c(0,0))
> geno.crosstab(trout, mar[1], mar[2])
```

	AGCCAG12
agcagc9 - CC 00	
- 0 4 4	
CC 7 33 235	
00 15 204 52	

We see that most individuals have opposite genotypes at these two markers.

While we were surprised by these results, with a more complete understanding of the genome of this organism, they might have been anticipated. As described in Nichols *et al.* (2003), an ancestor of this species experienced an autotetraploidy event (a duplication of the genome). While diploidy has since been restored, a number of pairs of linkage groups are homeologous (meaning partially homologous), including the five pairs that we see to exhibit this odd

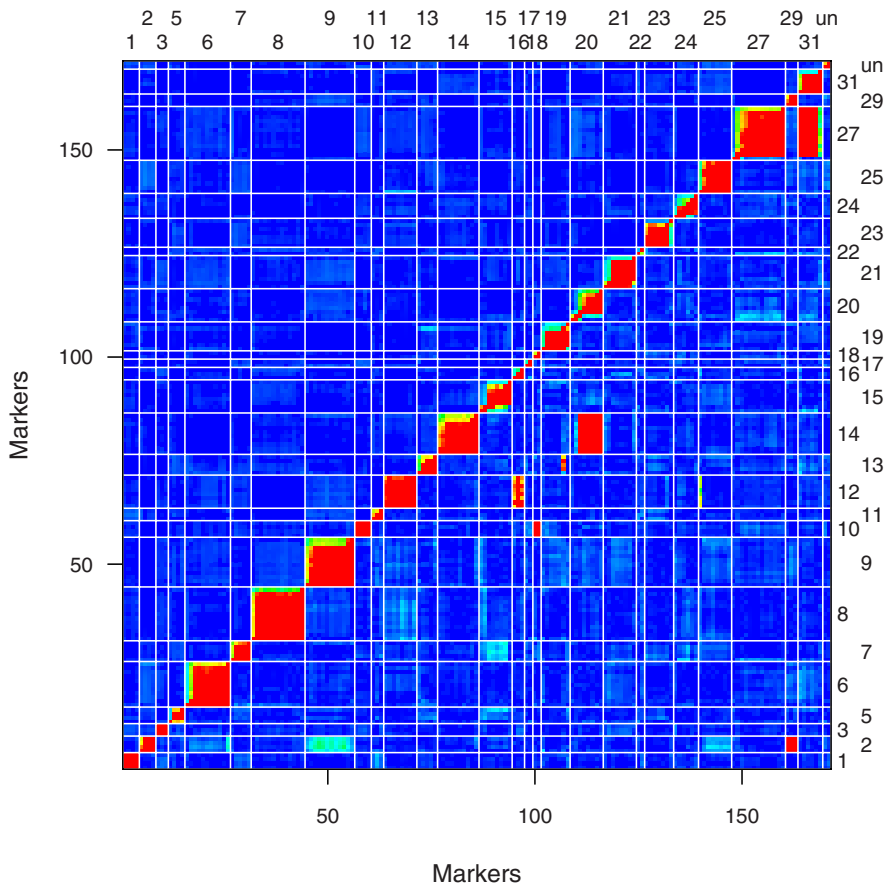


Figure 11.3. Plot of estimated recombination fractions (upper left) and LOD scores for a test of $r = 1/2$ (lower right) for all pairs of markers for the trout data. Red indicates linkage; blue indicates no linkage.

“reverse linkage.” Johnson *et al.* (1987) observed this phenomenon in related species, of pseudolinkage between loci in separate linkage groups, with recombinants being more frequent than the parental types. It appears that these homeologous chromosomes are pairing at meiosis, though in a special way.

While we might leave things as they are, the tight negative association between these pairs of linkage groups would make the interpretation of QTL mapping results rather confusing, as a QTL on one linkage group could give a linkage signal on a second linkage group.

Alternatively, we could swap the genotypes in one of each of these pairs, and then merge the linkage groups. This would have the advantage of ensuring that a single QTL would show only one linkage signal, but the disadvantage that the parental origins of the alleles will not be clear. Moreover,

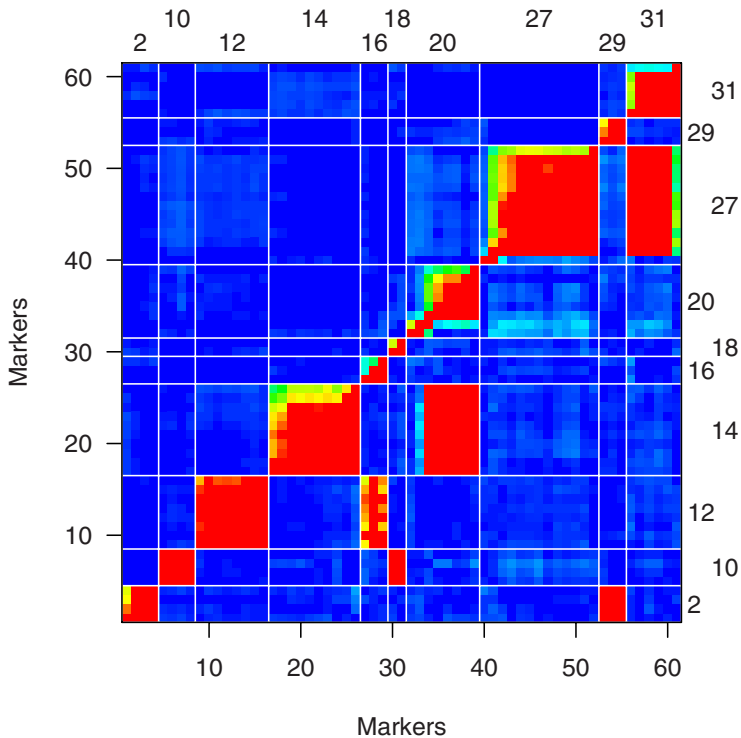


Figure 11.4. Plot of estimated recombination fractions (upper left) and LOD scores for a test of $r = 1/2$ (lower right) for all pairs of markers on selected linkage groups, for the `trout` data. Red indicates linkage; blue indicates no linkage.

our assumption, that the chromosomes are essentially joined end-to-end at meiosis, may be wrong. Nevertheless, we will follow this line of thinking. We will start by swapping the genotypes for the second of each of these pairs of linkage groups. (Note that the genotypes are coded 1 and 2, and so $3 - g$ will swap 1 for 2 and 2 for 1.)

```
> trout$geno[["29"]]$data <- 3 - trout$geno[["29"]]$data
> trout$geno[["18"]]$data <- 3 - trout$geno[["18"]]$data
> trout$geno[["16"]]$data <- 3 - trout$geno[["16"]]$data
> trout$geno[["20"]]$data <- 3 - trout$geno[["20"]]$data
> trout$geno[["31"]]$data <- 3 - trout$geno[["31"]]$data
```

The next step is to merge the pairs and then seek to establish marker order within the merged linkage groups.

We start with linkage groups 10 and 18, as they each have a small number of markers. First, we use `markernames` to obtain the names of the markers on linkage group 18 and then `movemarker` to move the markers from linkage group 18 to linkage group 10.

```
> lg18mar <- markernames(trout, 18)
> for(i in lg18mar)
+   trout <- movemarker(trout, i, 10)
```

We also change the name of the merged linkage group to "10.18."

```
> nam <- names(trout$geno)
> nam[nam=="10"] <- "10.18"
> names(trout$geno) <- nam
```

We now use `ripple` to consider all possible orders of the markers. Since there are only 6 markers (and so 360 possible marker orders), we will do a full likelihood analysis. We use the Kosambi map function and assume a 1% genotyping error rate.

```
> rip <- ripple(trout, chr="10.18", window=6, method="lik",
+               error.prob=0.01, map.function="kosambi",
+               verbose=FALSE)
```

Note that we referred to the chromosome ID in double-quotes. Chromosome identifiers are matched by name, with numbers being converted to character strings, and so one might use, for example `chr=5` in place of `chr="5"`. However, with the more complex chromosome names that we will be forming, it will be best to surround them in double-quotes, though we can actually mix numbers and character strings (and we will do so).

The summary of the `ripple` results indicates that we should switch the order of the markers. (We merged the groups at the wrong ends, and we might also switch the order of the third and fourth markers on linkage group 10.)

```
> summary(rip)
```

							LOD	chr	len
Initial	1	2	3	4	5	6	0.0	28.1	
1		3	4	2	1	5	3.7	28.7	
2		4	3	2	1	5	3.1	28.7	
3		3	4	1	2	5	2.6	28.5	
4		4	3	1	2	5	1.7	28.3	
5		3	2	4	1	5	1.7	28.4	
...	[6	additional rows]							

We use `switch.order` to switch the order of the markers to that with the maximum likelihood. The arguments `error.prob` and `map.function` are used in estimating the genetic map with the new order.

```
> trout <- switch.order(trout, "10.18", rip[2,],
+                        error.prob=0.01, map.function="kosambi")
```

The genetic map of the newly merged linkage groups is reasonably tight (with just a 15 cM gap between the two groups), which gives us some comfort that we are doing the right thing.

```
> pull.map(trout, chr="10.18")
```

```
ACGACA11 AGCAGT15 ACCAAG6 agcagc9 AGCCAG12 AGCCAG13
0.000 1.992 2.754 3.445 18.031 28.692
```

We will omit most of the details for dealing with the other four pairs of linkage groups. The techniques are similar, though with many markers on a linkage group, an iterative process is required in establishing marker order (see Sec. 3.4.2). In the following, we merge the linkage groups and switch the orders to the best that we could find.

```
> lg29mar <- markernames(trout, 29)
> for(i in lg29mar)
+   trout <- movemarker(trout, i, 2)
> lg16mar <- markernames(trout, 16)
> for(i in lg16mar)
+   trout <- movemarker(trout, i, 12)
> trout <- switch.order(trout, chr=12, c(1:8,10,11,9),
+   error.prob=0.01, map.function="kosambi")
> lg20mar <- markernames(trout, 20)
> for(i in lg20mar)
+   trout <- movemarker(trout, i, 14)
> trout <- switch.order(trout, chr=14, error.prob=0.01,
+   c(10:1,14,16,15,17,18,13:11),
+   map.function="kosambi")
> lg31mar <- markernames(trout, 31)
> for(i in lg31mar)
+   trout <- movemarker(trout, i, 27)
> trout <- switch.order(trout, chr=27, error.prob=0.01,
+   c(1:6,8,7,9:13,15:18,14,19),
+   map.function="kosambi")
```

Finally, we change the names of the linkage groups that have been merged.

```
> nam <- names(trout$geno)
> nam[nam=="2"] <- "2.29"
> nam[nam=="12"] <- "12.16"
> nam[nam=="14"] <- "14.20"
> nam[nam=="27"] <- "27.31"
> names(trout$geno) <- nam
```

There are a few remaining issues in the pairwise marker linkages: there is a marker on linkage group 19 that appears to be linked to linkage group 13, and there is a marker on linkage group 25 that appears to be linked to linkage group "12.16." Let us plot the estimated recombination fractions for those four linkage groups.

```
> plot.rf(trout, chr=c(13, 19, "12.16", 25))
```

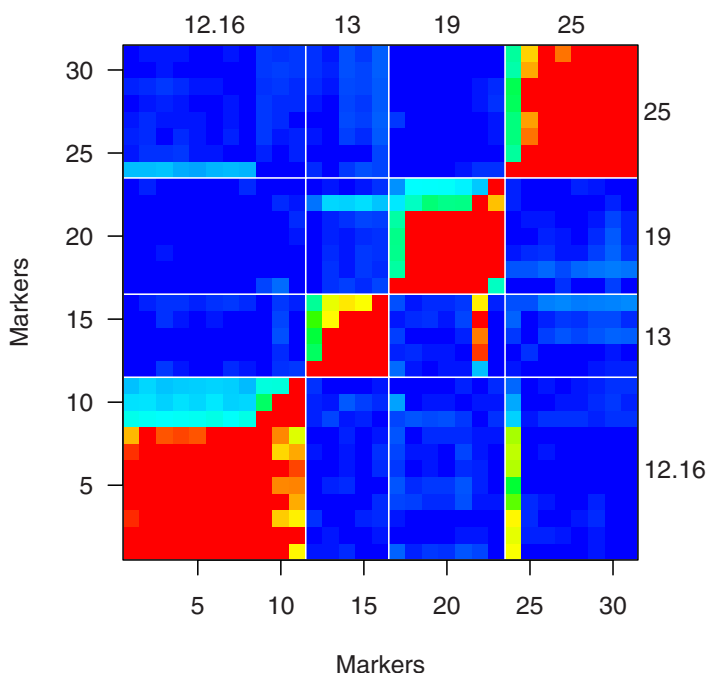


Figure 11.5. Plot of estimated recombination fractions (upper left) and LOD scores for a test of $r = 1/2$ (lower right) for all pairs of markers on linkage groups 13, 19, “12.16” and 25, for the `trout` data. Red indicates linkage; blue indicates no linkage.

The potentially problematic markers (see Fig. 11.5) are linked to another linkage group only weakly, but the large number of individuals results in reasonably large LOD scores. This may indicate that these linkage groups should also be merged, but it is likely best to keep these linkage groups separate, though we should keep in mind that a single QTL has the potential to result in LOD peaks on multiple linkage groups.

Let us reestimate the intermarker distances, using the observed data, keeping marker order fixed, and plot the estimated map against that which came with the data. We will use the Kosambi map function, as in Nichols *et al.* (2007).

```
> newmap <- est.map(trout, error.prob=0.01, verbose=FALSE,
+                   map.function="kosambi")
> plot.map(trout, newmap, alternate.chrid=TRUE)
```

As seen in Fig. 11.6, many linkage groups become slightly shorter, though overall there is good agreement. (The linkage groups that we had modified show no differences, but this is because, in switching the orders of markers, we have replaced the maps with those estimated from these data.) There are a few places (e.g., linkage group 8) where a number of markers appear to be

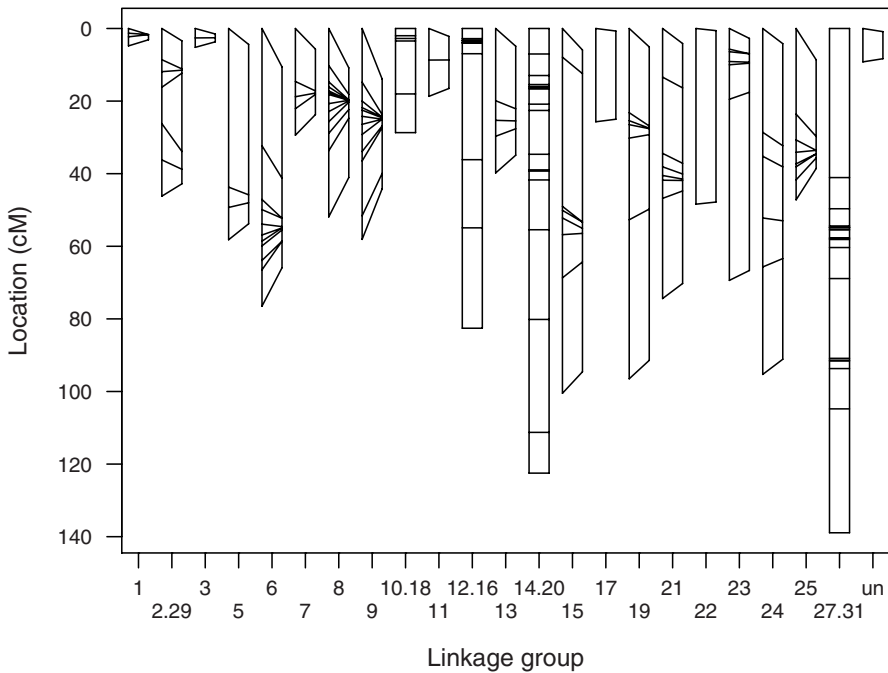


Figure 11.6. The genetic map in the `trout` data plotted against the map estimated from the genotype data. For each linkage group, the line on the left is that map provided with the data, and the line on the right is the estimated map; line segments connect the positions for each marker.

moved closer together. Overall, the estimated map is 1144 cM, while the initial map was 1328 cM (a difference of 14%).

We will replace the map in the data with our newly estimated one.

```
> trout <- replace.map(trout, newmap)
```

While more might be done to investigate the genotype data, let us trust the data and the genetic map and move on to the QTL mapping.

11.2 Initial QTL analyses

We will begin with the simpler aspects of the QTL analysis. In Sec. 11.3, we will investigate the possibility of $QTL \times MCE$ interactions. Since the marker genotype data are quite complete (though, admittedly, with a few large gaps between markers), we will use Haley–Knott regression (see Sec. 4.2.2). We must first calculate the conditional QTL genotype probabilities, given the available marker data.

```
> trout <- calc.genoprob(trout, step=1, err=0.01,
+                          map.function="kosambi")
```

While we will postpone the investigation of QTL $\times$ MCE interactions to the next section, the systematic differences in the phenotype among MCE groups (see Fig. 11.2 on page 316) suggests that we should include MCE as a set of additive covariates in the QTL mapping. In doing so, we assume that the effects of any QTL are the same in the eight groups, but we allow shifts in the average phenotype between the groups.

We form a matrix of indicators, with a column for each of the MCE groups except the first one.

```
> female <- pull.pheno(trout, "female")
> lev <- levels(female)
> nlev <- length(lev)
> femcov <- matrix(0, nrow=nind(trout), ncol=nlev-1)
> colnames(femcov) <- lev[-1]
> for(i in 2:nlev)
+   femcov[female==lev[i],i-1] <- 1
```

We now perform the genome scan, including femcov as a set of additive covariates.

```
> out <- scanone(trout, method="hk", addcovar=femcov)
```

We plot the LOD curves as follows.

```
> plot(out, ylab="LOD score", alternate.chrid=TRUE)
```

The results (see Fig. 11.7) indicate overwhelming support for a QTL on linkage group 8, with maximum LOD score = 43.2.

Let us perform a quick permutation test. With Haley–Knott regression, the permutations are quite fast, and so we will use 4000 permutations. As in the next section we will want to investigate the presence of QTL $\times$ MCE interactions, which will require matched permutation tests, with and without MCE as a set of interactive covariates, we use `set.seed` to define the seed for the random number generator, so that this can be repeated with the same permutations.

```
> set.seed(523938)
> operm <- scanone(trout, method="hk", addcovar=femcov,
+                  n.perm=4000)
```

We can now pull out the results from our initial scan that meet a 10% genome-wide significance threshold.

```
> summary(out, perms=operm, alpha=0.1, pvalues=TRUE)
```

	chr	pos	lod	pval
OmyFGT12	8	9.11	43.18	0.00000

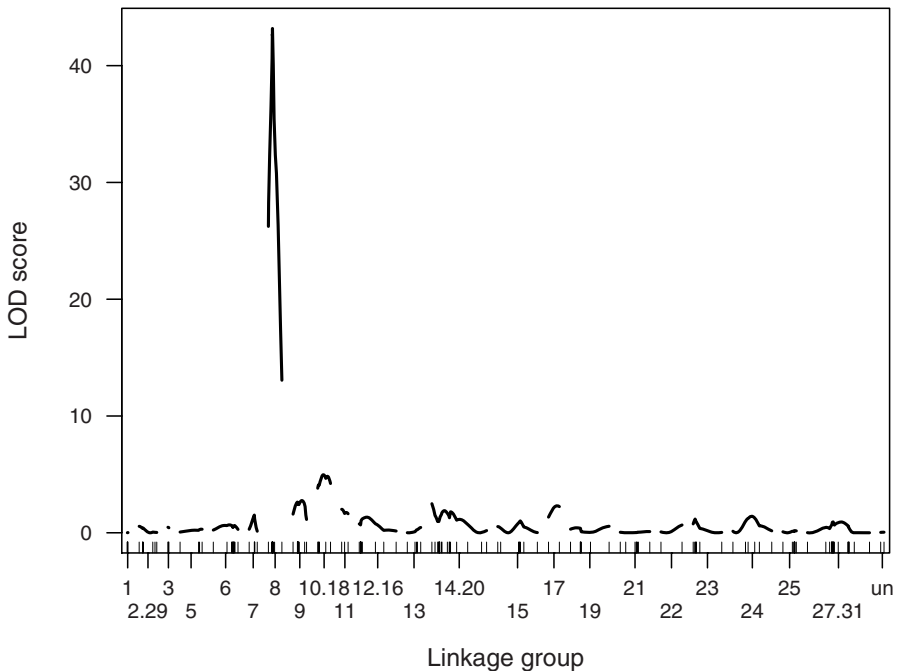


Figure 11.7. LOD curves from a genome scan by Haley–Knott regression for the trout data, with MCE groups included as additive covariates.

```
c9.loc20      9 20.00  2.76 0.03775
c10.18.loc13 10.18 13.00  4.97 0.00025
AGCAGT11     14.20  0.00  2.49 0.06950
```

In addition to the clear QTL on linkage group 8, we see evidence for QTL on linkage groups 9, 10.18, and 14.20.

Let us repeat the genome scan, controlling for the locus on linkage group 8. We use `makeqtl` to create a QTL object; as we will be performing Haley–Knott regression, we call `makeqtl` with `what="prob"` (rather than the default, `what="draws"`). We use `addqtl` to perform the scan.

```
> qtl <- makeqtl(trout, 8, 9.11, what="prob")
> out.c8 <- addqtl(trout, qtl=qtl, method="hk", covar=femcov)
```

While our permutation results do not formally apply in the present case, in which we are controlling for the locus on linkage group 8, they nevertheless provide a reasonable guide.

```
> summary(out.c8, perms=operm, alpha=0.1, pvalues=TRUE)

      chr pos  lod   pval
c9.loc21      9  21 2.59 0.05575
```

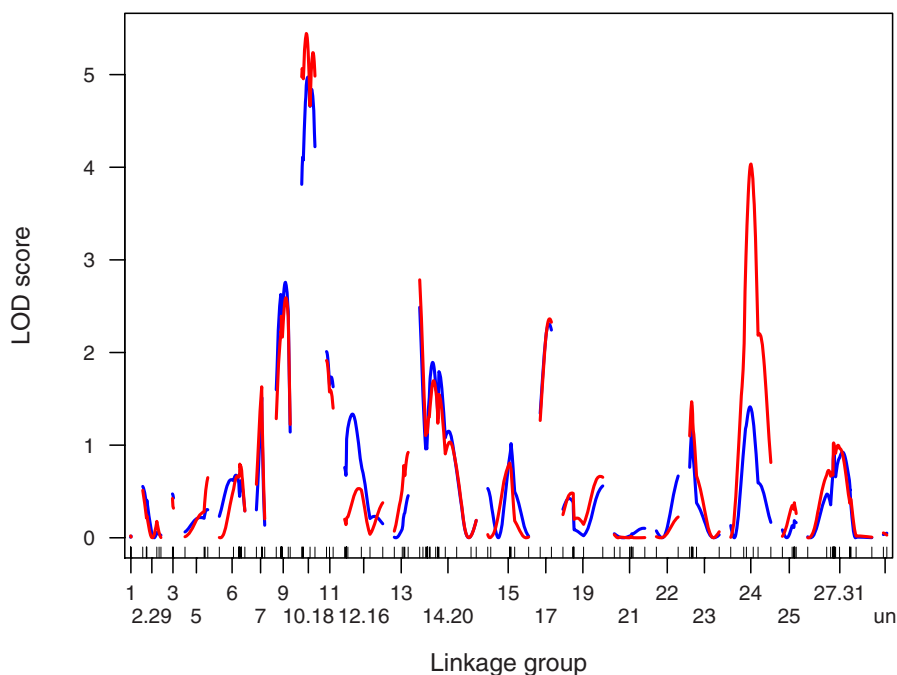


Figure 11.8. LOD curves from a genome scan by Haley-Knott regression for the trout data, with MCE groups included as additive covariates. The curves in blue are as in Fig. 11.7; those in red were calculated controlling for a QTL on linkage group 8.

```
c10.18.loc10 10.18 10 5.44 0.00025
AGCAGT11     14.20 0 2.78 0.03575
c17.loc20     17 20 2.36 0.09150
c24.loc44     24 44 4.03 0.00125
```

We see evidence for additional QTL on linkage groups 17 and 24. The three previously-identified QTL remain on the list, though their LOD scores have changed slightly.

We can plot the LOD curves, with and without controlling for the QTL on linkage group 8, as follows.

```
> plot(out, out.c8, chr = -8, col=c("blue","red"),
+       ylab="LOD score", alternate.chrid=TRUE)
```

Note the use of `chr = -8` to plot all but linkage group 8. The minus sign may also be used with character strings. For example, use of `chr="-un"` would result in a plot of all linkage groups except "un".

As seen in Fig. 11.8, controlling for the QTL on linkage group 8 leads to a great increase in the evidence for a QTL on linkage group 24, and small increases and decreases in the LOD scores on many other linkage groups.

Let us bring all of the terms together into one model, and then refine our estimates of the locations of the QTL with `refineqtl`.

```
> qtl <- makeqtl(trout, c(8, 9, "10.18", "14.20", 17, 24),
+               c(9.11, 21, 10, 0, 20, 44), what="prob")
> rqtl <- refineqtl(trout, qtl=qtl, covar=femcov,
+                  method="hk", verbose=FALSE)
```

The location of the QTL changed just slightly.

```
> options(width=64)
> rqtl
```

	name	chr	pos	n.gen
Q1	8@9.1	8	9.112	2
Q2	9@11.0	9	11.040	2
Q3	10.18@11.0	10.18	11.000	2
Q4	14.20@0.0	14.20	0.000	2
Q5	17@23.0	17	23.000	2
Q6	24@46.0	24	46.000	2

Let us fit the multiple-QTL model and perform a drop-one-QTL analysis with `fitqtl`. We do not find the two columns of p -values calculated by `fitqtl` to be particularly informative (as they fail to account for the scan across the genome), and they take up a lot of space, and so we omit them, using `pvalues=FALSE` in the `summary.fitqtl` function.

```
> summary(fitqtl(trout, qtl=rqtl, covar=femcov, method="hk"),
+         pvalues=FALSE)
```

Full model result

```
-----
Model formula: y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + TL3 + SP1 + SP2
               + SP3 + SP4 + SP5 + SP6
```

	df	SS	MS	LOD	%var
Model	13	41504	3192.6	78.92	48.11
Error	540	44762	82.9		
Total	553	86266			

Drop one QTL at a time ANOVA table:

```
-----
```

	df	Type III SS	LOD	%var	F value
8@9.1	1	21199.7196	46.6416	24.5748	255.747
9@11.0	1	911.2975	2.4245	1.0564	10.994
10.18@11.0	1	1972.8624	5.1886	2.2870	23.800
14.20@0.0	1	879.4502	2.3406	1.0195	10.609
17@23.0	1	840.3175	2.2374	0.9741	10.137

24046.0	1	1437.5384	3.8027	1.6664	17.342
TL3	1	4876.5995	12.4400	5.6530	58.830
SP1	1	2592.7989	6.7738	3.0056	31.279
SP2	1	5167.0989	13.1420	5.9897	62.334
SP3	1	4184.0042	10.7497	4.8501	50.475
SP4	1	327.2468	0.8763	0.3793	3.948
SP5	1	2534.1374	6.6247	2.9376	30.571
SP6	1	10414.5254	25.1638	12.0726	125.638

Note that our major locus on linkage group 8 is responsible for an estimated 25% of the variation in the phenotype. The conditional LOD scores for the other five QTL are reduced a bit relative to what we had seen when they were considered individually, though still controlling for the QTL on linkage group 8. The evidence for QTL on linkage groups 10.18 and 24 remains strong. The evidence for the other three loci is much weaker, but they are still interesting.

We have not yet considered the possibility of epistatic interactions among QTL, and so we now use `addint` to look for epistatic interactions among the identified QTL; `qtl.only=TRUE` indicates that QTL $\times$ covariate interactions should not be considered.

```
> addint(trout, qtl=rqtl, qtl.only=TRUE, method="hk",
+        covar=femcov, pvalues=FALSE)
```

The output was extensive and not too interesting, and so we do not display it here. There is little evidence for interactions among any of these QTL; the largest interaction LOD score was 1.1, for the QTL on linkage groups 8 and 17.

Of course, we should also perform a two-dimensional, two-QTL genome scan, which will allow us to identify additional QTL with important interactions and also potential pairs of linked QTL (particularly those linked in repulsion, with effects of opposite sign, which would generally not show up in a single-QTL genome scan).

To make sense of the results of a two-dimensional genome scan, we will want results from a permutation test, which will be quite time consuming, and so let's get that going. [The computations took a total of about 7 days of CPU time; split across 16 processors (see Sec. 8.1, page 223), it took about 10 hours in real time.]

```
> operm2 <- scantwo(trout, method="hk", addcovar=femcov,
+                   n.perm=1000)
```

Interestingly, the significance thresholds are similar to what we had seen for the `hyper` data in Sec. 8.1.

```
> summary(operm2)

tth (1000 permutations)
      full  fv1  int  add  av1  one
5%   5.43  4.42  3.99  4.40  2.64  2.71
10%  5.12  4.05  3.72  4.04  2.39  2.32
```

Let us now perform the actual two-dimensional scan with the data. We use the argument `incl.markers=TRUE` so that calculations are performed at the markers as well as on the evenly-spaced grid.

```
> out2 <- scantwo(trout, method="hk", addcovar=femcov,
+                 incl.markers=TRUE)
```

The tabular summary of the two-dimensional scan output is simplest to interpret, so we will start with that. The p -values take up a lot of space, so we won't include them.

```
> summary(out2, perms=operm2, alpha=0.05)
```

	pos1f	pos2f	lod.full	lod.fv1	lod.int	pos1a
c7 :c7	11.52	16.00	6.65	5.13	1.3915	11.52
c8 :c8	9.11	29.00	55.76	12.57	10.7777	9.11
c8 :c10.18	9.11	9.00	48.92	5.74	0.2927	9.11
c8 :c14.20	9.11	0.00	46.09	2.91	0.1275	9.11
c8 :c24	9.11	44.00	47.24	4.06	0.0232	9.11
c9 :c10.18	18.00	13.00	8.45	3.47	0.7140	9.77
c12.16:c12.16	2.00	6.95	8.10	6.76	6.1503	3.00
c13:c13	0.00	23.00	5.76	5.31	4.7070	8.00

	pos2a	lod.add	lod.av1
c7 :c7	15	5.26	3.741
c8 :c8	23	44.98	1.797
c8 :c10.18	10	48.63	5.444
c8 :c14.20	0	45.97	2.784
c8 :c24	44	47.22	4.035
c9 :c10.18	13	7.73	2.758
c12.16:c12.16	5	1.95	0.612
c13:c13	26	1.06	0.601

First note the pairs of linked QTL on linkage groups 7, 8, 12.16, and 13. The pair of QTL on linkage group 7 do not show strong evidence for an interaction, but the other three pairs show strong interaction effects. The pairs on linkage groups 7, 12.16 and 13 are relatively tightly linked, so we should be slightly skeptical. The other rows in the table indicate the multiple QTL that we had seen in our initial single-QTL scans; none show epistatic effects.

Let us consider a plot of the LOD scores for the pairs of linked QTL. We will plot the results for linkage group 8 separate from the others, as the extremely large LOD scores on linkage group 8 will make it difficult to study the others if they all are shown with the same color scale. We will focus on LOD_i , concerning epistatic interactions, and LOD_{fv1} , comparing the model with two interacting QTL to the best single-QTL model.

```
> plot(out2, lower="cond-int", chr=8)
> plot(out2, lower="cond-int", chr=c(7, "12.16", 13))
```

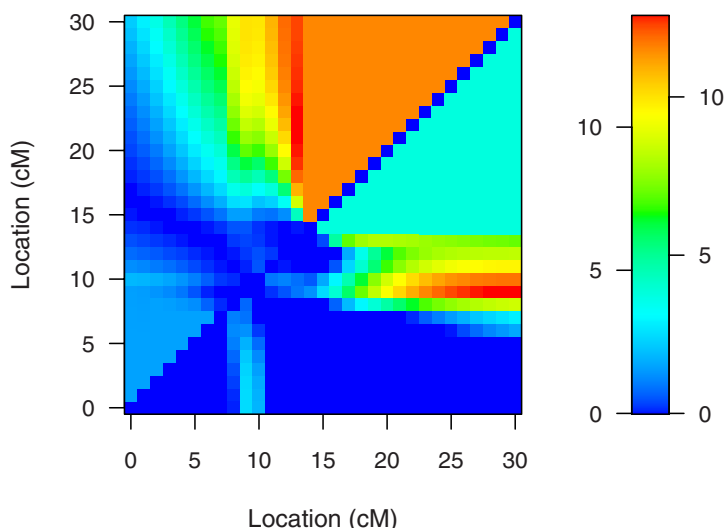


Figure 11.9. LOD scores, for linkage group 8, from a two-dimensional, two-QTL genome scan with the `trout` data. LOD_i is displayed in the upper left triangle; LOD_{fv1} is displayed in the lower right triangle. In the color scale on the right, numbers to the left and right correspond to LOD_i and LOD_{fv1} , respectively.

The LOD scores for linkage group 8 are shown in Fig. 11.9. The LOD scores for linkage groups 7, 12.16 and 13 are shown in Fig. 11.10. These figures do not provide much additional information, beyond what was obtained in the summary table above. We do get a sense of the precision of localization of the QTL, but not much else.

To alleviate our skepticism about these pairs of linked QTL, we plot the phenotypes against the two-locus genotypes at markers close to the inferred QTL positions. First, we use `find.marker` to identify the relevant markers; we also use `find.markerpos` to look at the positions of the selected markers, so that we are sure that the markers are near the QTL.

```
> mar7 <- find.marker(trout, 7, c(11.52, 16))
> find.markerpos(trout, mar7)

      chr  pos
agcagc3   7 11.52
agcatc13   7 18.05

> mar8 <- find.marker(trout, 8, c(9.11, 29))
> find.markerpos(trout, mar8)

      chr  pos
OmyFGT12  8  9.112
ACGACA8   8 30.215
```

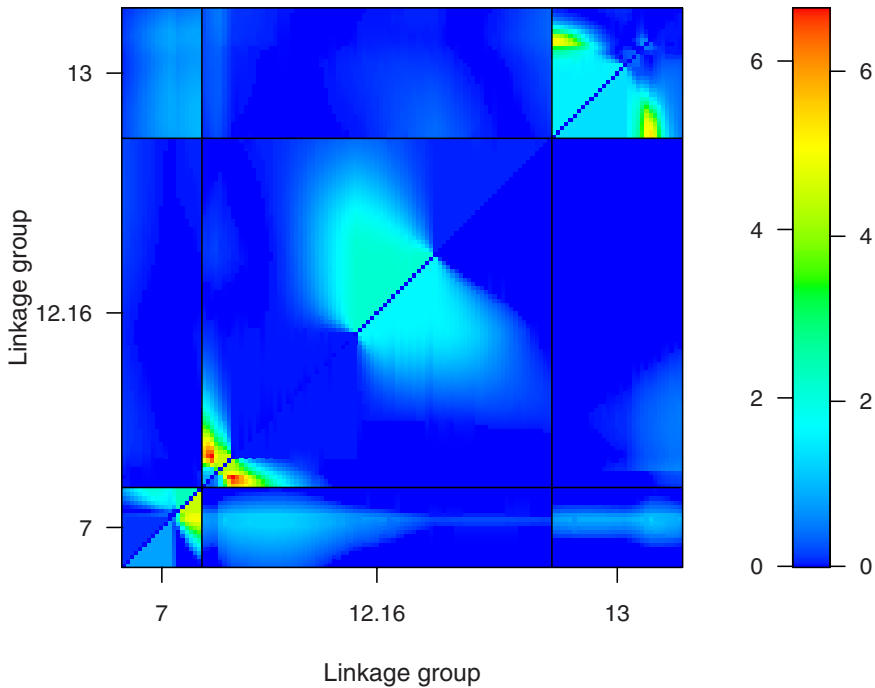


Figure 11.10. LOD scores, for selected linkage groups, from a two-dimensional, two-QTL genome scan with the `trout` data. LOD_i is displayed in the upper left triangle; LOD_{fv1} is displayed in the lower right triangle. In the color scale on the right, numbers to the left and right correspond to LOD_i and LOD_{fv1} , respectively.

```
> mar12.16 <- find.marker(trout, "12.16", c(2, 6.95))
> find.markerpos(trout, mar12.16)

      chr   pos
AGCATC6 12.16 2.794
ACGAGA5 12.16 6.952

> mar13 <- find.marker(trout, 13, c(0, 23))
> find.markerpos(trout, mar13)

      chr   pos
agcagc11 13  0.00
acgatg8  13 22.75
```

Now we call `plot.pxg` to create the plots of phenotypes against two-locus genotypes.

```
> par(mfrow=c(2,2))
> plot.pxg(trout, marker=mar7)
> plot.pxg(trout, marker=mar8)
```

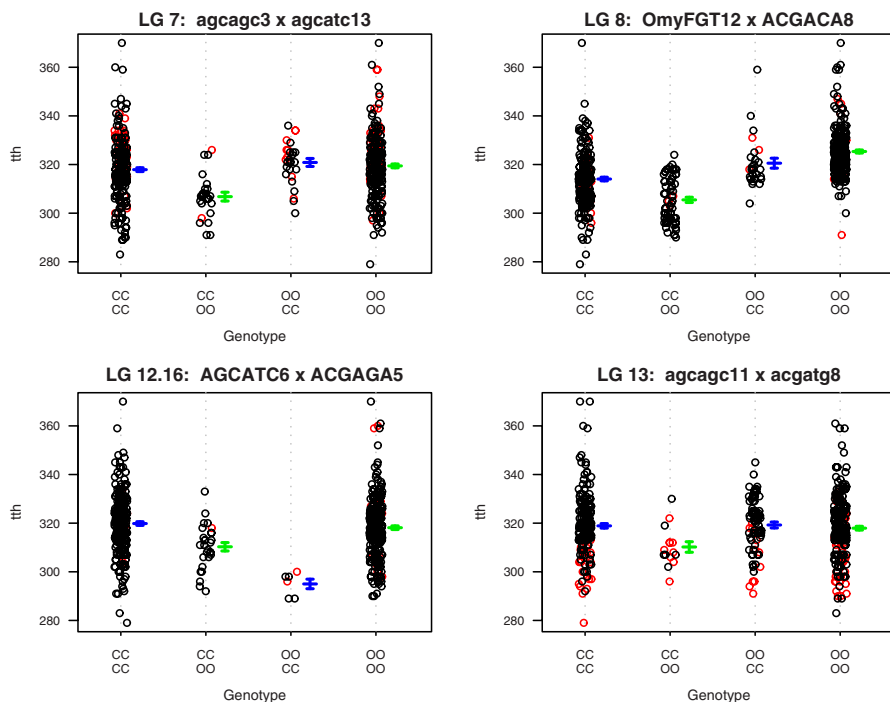


Figure 11.11. Plot of the `tth` phenotype against two-locus genotypes at four pairs of putative linked QTL, for the `trout` data. Points in red are imputed genotypes.

```
> plot.pxg(trout, marker=mar12.16)
> plot.pxg(trout, marker=mar13)
```

The plots of the phenotype against two-locus genotypes (Fig. 11.11) all look reasonable. For linkage group 12.16, the inference of a second epistatic locus depends on just a few individuals, which is worrisome, and to some extent this is also true for the loci on linkage groups 7 and 13. Nevertheless, these results are intriguing.

Let us bring all of our inferred QTL together into one model and look at the drop-one-QTL analysis from `fitqtl`. We will omit the putative loci on linkage groups 9, 14.20 and 17, which were weakly supported. We still have 10 QTL (and three pairs of interactions).

```
> qtl <- makeqtl(trout, c(7,7,8,8,"10.18","12.16","12.16",13,13,24),
+               c(11.52,16,9.11,29,11,2,6.95,0,23,46), what="prob")
> summary(fitqtl(trout, qtl=qtl, covar=femcov, method="hk",
+               formula=y~TL3+SP1+SP2+SP3+SP4+SP5+SP6+
+               Q1+Q2+Q3*Q4+Q5+Q6*Q7+Q8*Q9+Q10),
+         pvalues=FALSE)
```

Full model result

```
Model formula: y ~ TL3 + SP1 + SP2 + SP3 + SP4 + SP5 + SP6 + Q1 + Q2 +
                Q3 + Q4 + Q5 + Q6 + Q7 + Q8 + Q9 + Q10 + Q3:Q4 +
                Q6:Q7 + Q8:Q9
```

	df	SS	MS	LOD	%var
Model	20	47092	2354.6	94.97	54.59
Error	533	39175	73.5		
Total	553	86266			

Drop one QTL at a time ANOVA table:

	df	Type III SS	LOD	%var	F value
TL3	1	6096.5117	17.4002	7.0671	82.948
SP1	1	3164.2037	9.3443	3.6680	43.051
SP2	1	5339.4284	15.3714	6.1895	72.647
SP3	1	4476.5711	13.0166	5.1893	60.907
SP4	1	481.7087	1.4702	0.5584	6.554
SP5	1	1787.9517	5.3689	2.0726	24.326
SP6	1	10119.5482	27.6421	11.7306	137.684
7@11.5	1	793.7264	2.4131	0.9201	10.799
7@16.0	1	500.5065	1.5272	0.5802	6.810
8@9.1	2	15014.9633	39.0324	17.4054	102.145
8@29.0	2	3569.1533	10.4895	4.1374	24.281
10.18@11.0	1	1448.3044	4.3673	1.6789	19.705
12.16@2.0	2	938.7691	2.8488	1.0882	6.386
12.16@7.0	2	913.7320	2.7737	1.0592	6.216
13@0.0	2	971.0353	2.9456	1.1256	6.606
13@23.0	2	1066.9214	3.2325	1.2368	7.258
24@46.0	1	1476.6254	4.4511	1.7117	20.091
8@9.1:8@29.0	1	3075.7546	9.0928	3.5654	41.848
12.16@2.0:12.16@7.0	1	898.9660	2.7294	1.0421	12.231
13@0.0:13@23.0	1	881.1723	2.6760	1.0215	11.989

The support for the pairs of loci on linkage groups 7, 12.16 and 13 have dropped greatly, but there remains extremely strong support for the pair of QTL on linkage group 8, plus QTL on linkage groups 10.18 and 24.

Let us drop the loci on linkage groups 12.16 and 13, refine the QTL positions, and perform the drop-one analysis again.

```
> qtl2 <- dropfromqtl(qtl, 6:9)
> qtl2 <- refineqtl(trout, qtl=qtl2, covar=femcov, method="hk",
+                   formula=y~TL3+SP1+SP2+SP3+SP4+SP5+SP6+
+                   Q1+Q2+Q3*Q4+Q5+Q6, verbose=FALSE)
> summary(fitqtl(trout, qtl=qtl2, covar=femcov, method="hk",
+                formula=y~TL3+SP1+SP2+SP3+SP4+SP5+SP6+
+                Q1+Q2+Q3*Q4+Q5+Q6),
+         pvalues=FALSE)
```

Full model result

Model formula: $y \sim \text{TL3} + \text{SP1} + \text{SP2} + \text{SP3} + \text{SP4} + \text{SP5} + \text{SP6} + \text{Q1} + \text{Q2} + \text{Q3} + \text{Q4} + \text{Q5} + \text{Q6} + \text{Q3:Q4}$

	df	SS	MS	LOD	%var
Model	14	44943	3210.19	88.54	52.1
Error	539	41323	76.67		
Total	553	86266			

Drop one QTL at a time ANOVA table:

	df	Type III SS	LOD	%var	F value
TL3	1	6194.3529	16.8028	7.1805	80.796
SP1	1	2919.7332	8.2130	3.3846	38.083
SP2	1	6663.2666	17.9841	7.7241	86.912
SP3	1	4902.4536	13.4868	5.6829	63.945
SP4	1	464.3894	1.3444	0.5383	6.057
SP5	1	1890.9159	5.3825	2.1920	24.664
SP6	1	9968.8767	25.9981	11.5560	130.028
7@11.5	1	1220.1786	3.5007	1.4144	15.915
7@17.0	1	700.8668	2.0232	0.8124	9.142
8@9.1	2	17976.8740	43.4503	20.8389	117.240
8@30.2	2	4266.6131	11.8206	4.9459	27.826
10.18@10.0	1	1686.1471	4.8112	1.9546	21.993
24@47.0	1	1521.7448	4.3504	1.7640	19.849
8@9.1:8@30.2	1	3790.5184	10.5577	4.3940	49.441

We now have good support for the proximal locus on linkage group 7, but not for the distal one. Let's drop the distal locus, refine the QTL positions, and repeat the drop-one-QTL analysis.

```
> qtl3 <- dropfromqtl(qtl2, 2)
> qtl3 <- refineqtl(trout, qtl=qtl3, covar=femcov, method="hk",
+                   formula=y~TL3+SP1+SP2+SP3+SP4+SP5+SP6+
+                   Q1+Q2*Q3+Q4+Q5, verbose=FALSE)
> summary(fitqtl(trout, qtl=qtl3, covar=femcov, method="hk",
+               formula=y~TL3+SP1+SP2+SP3+SP4+SP5+SP6+
+               Q1+Q2*Q3+Q4+Q5),
+         pvalues=FALSE)
```

Full model result

Model formula: $y \sim \text{TL3} + \text{SP1} + \text{SP2} + \text{SP3} + \text{SP4} + \text{SP5} + \text{SP6} + \text{Q1} + \text{Q2} + \text{Q3} + \text{Q4} + \text{Q5} + \text{Q2:Q3}$

	df	SS	MS	LOD	%var
Model	13	44278	3405.99	86.62	51.33

```
Error 540 41988    77.76
Total 553 86266
```

Drop one QTL at a time ANOVA table:

	df	Type III SS	LOD	%var	F value
TL3	1	5995.4909	16.0567	6.9500	77.107
SP1	1	2854.5748	7.9126	3.3090	36.712
SP2	1	6543.2056	17.4221	7.5849	84.151
SP3	1	5059.5513	13.6870	5.8651	65.070
SP4	1	443.2697	1.2633	0.5138	5.701
SP5	1	1687.4758	4.7401	1.9561	21.702
SP6	1	9898.5955	25.4645	11.4745	127.303
7@11.5	1	865.3537	2.4541	1.0031	11.129
8@9.1	2	17614.6554	42.1427	20.4190	113.269
8@28.0	2	4822.5266	13.0794	5.5903	31.011
10.18@9.0	1	1787.9916	5.0167	2.0726	22.995
24@46.0	1	1482.9148	4.1754	1.7190	19.071
8@9.1:8@28.0	1	4179.5638	11.4156	4.8450	53.752

The support for remaining locus on linkage group 7 is no longer strong, so let's omit it and rerun `refineqtl` and `fitqtl`.

```
> qtl4 <- dropfromqtl(qtl3, 1)
> qtl4 <- refineqtl(trout, qtl=qtl4, covar=femcov, method="hk",
+                   formula=y~TL3+SP1+SP2+SP3+SP4+SP5+SP6+
+                   Q1*Q2+Q3+Q4, verbose=FALSE)
> summary(fitqtl(trout, qtl=qtl4, covar=femcov, method="hk",
+               formula=y~TL3+SP1+SP2+SP3+SP4+SP5+SP6+
+               Q1*Q2+Q3+Q4),
+         pvalues=FALSE)
```

Full model result

```
Model formula: y ~ TL3 + SP1 + SP2 + SP3 + SP4 + SP5 + SP6 + Q1
+ Q2 + Q3 + Q4 + Q1:Q2
```

	df	SS	MS	LOD	%var
Model	12	43416	3618.0	84.18	50.33
Error	541	42851	79.2		
Total	553	86266			

Drop one QTL at a time ANOVA table:

	df	Type III SS	LOD	%var	F value
TL3	1	5867.4051	15.4379	6.8015	74.078

SP1	1	2687.4846	7.3178	3.1153	33.930
SP2	1	6234.3561	16.3407	7.2269	78.710
SP3	1	4748.5963	12.6430	5.5046	59.952
SP4	1	442.3387	1.2355	0.5128	5.585
SP5	1	1518.2935	4.1887	1.7600	19.169
SP6	1	9591.8949	24.3002	11.1190	121.100
8@9.1	2	17486.1464	41.1692	20.2700	110.384
8@28.0	2	4623.4707	12.3264	5.3595	29.186
10.18@10.0	1	2027.7848	5.5623	2.3506	25.601
24@46.0	1	1312.6225	3.6298	1.5216	16.572
8@9.1:8@28.0	1	3972.6388	10.6658	4.6051	50.156

We have good support for the remaining four QTL, including the interaction between the two loci on linkage group 8.

Let us complete this initial analysis (ignoring the possibility of QTL $\times$ MCE interactions) with an application of the automated model selection approach accomplished with `stepwiseqtl`. We first calculate the penalties for our model comparison criterion, using the results of the permutation test with a two-dimensional, two-QTL genome scan.

```
> print(pen <- calc.penalties(operm2))

main heavy light
2.711 3.990 1.711
```

Let us first consider strictly additive models, enforced with the argument `additive.only=TRUE`. Note that in this case, only the penalty on main effects is used.

```
> stepout1 <- stepwiseqtl(trout, covar=femcov, penalties=pen,
+                          method="hk", additive.only=TRUE,
+                          verbose=FALSE)
```

The chosen model includes loci on linkage groups 8, 9, 10.18 and 24.

```
> stepout1

      name   chr   pos n.gen
Q1    8@9.1    8  9.112    2
Q2    9@11    9 11.040    2
Q3 10.18@10 10.18 10.000    2
Q4   24@45   24 45.000    2
```

```
Formula: y ~ TL3 + SP1 + SP2 + SP3 + SP4 + SP5 + SP6 + Q1 + Q2
          + Q3 + Q4
pLOD: 44.51
```

We use `fitqtl` to perform a drop-one-QTL analysis, to look at the support for the individual terms in the model.

```
> summary(fitqtl(trout, qtl=stepout1, covar=femcov,
+               method="hk"), pvalues=FALSE)
```

Full model result

```
-----
Model formula: y ~ Q1 + Q2 + Q3 + Q4 + TL3 + SP1 + SP2 + SP3 +
                SP4 + SP5 + SP6
```

	df	SS	MS	LOD	%var
Model	11	39866	3624.20	74.6	46.21
Error	542	46400	85.61		
Total	553	86266			

Drop one QTL at a time ANOVA table:

```
-----
      df Type III SS          LOD      %var F value
8@9.1    1  21858.0351    46.4352    25.3379 255.325
9@11     1   1071.7613     2.7471     1.2424  12.519
10.18@10 1   2213.2359     5.6055     2.5656  25.853
24@45    1   1624.4151     4.1395     1.8830  18.975
TL3      1   5275.8899    12.9553     6.1158  61.628
SP1      1   2312.3039     5.8504     2.6804  27.010
SP2      1   4676.5682    11.5520     5.4211  54.627
SP3      1   3697.8501     9.2244     4.2866  43.195
SP4      1    264.4609     0.6837     0.3066   3.089
SP5      1   2328.1130     5.8895     2.6988  27.195
SP6      1   9658.9287    22.7492    11.1967 112.827
```

Note that the locus on linkage group 9 just barely enters the model, as its conditional LOD score (that is, the $\log_{10}$ likelihood ratio comparing the four-QTL model to the model with the locus on linkage group 9 omitted) is 2.75 and the penalty on main effects was 2.71.

Let us repeat the stepwise analysis, allowing for epistatic interactions (and using the combination of heavy and light penalties on interactions). We use the default, of forward selection to a model with 10 QTL, followed by backward deletion.

```
> stepout2 <- stepwiseqtl(trout, covar=femcov, penalties=pen,
+                          method="hk", verbose=FALSE)
```

Allowing for interactions, we choose a model that includes the pair of interacting loci on linkage group 8, plus loci on linkage groups 10.18 and 24. This is identical to the model that we had chose through our exploratory search, above.

```
> stepout2
```

	name	chr	pos	n.gen
Q1	8@9.1	8	9.112	2
Q2	8@28.0	8	28.000	2
Q3	10.18@10.0	10.18	10.000	2
Q4	24@46.0	24	46.000	2

Formula: $y \sim \text{TL3} + \text{SP1} + \text{SP2} + \text{SP3} + \text{SP4} + \text{SP5} + \text{SP6} + \text{Q1} + \text{Q2} + \text{Q3} + \text{Q4} + \text{Q1:Q2}$

pLOD: 52.37

Finally, let us study the estimated effects under this model. We use `fitqtl` with `get.ests=TRUE` (to obtain the estimated effects) and `dropone=FALSE` (to skip the drop-one-QTL analysis).

```
> summary(fitqtl(trout, qtl=stepout2, covar=femcov,
+               method="hk", dropone=FALSE, get.ests=TRUE,
+               formula=y~TL3+SP1+SP2+SP3+SP4+SP5+SP6+
+               Q1*Q2+Q3+Q4))
```

Full model result

Model formula: $y \sim \text{TL3} + \text{SP1} + \text{SP2} + \text{SP3} + \text{SP4} + \text{SP5} + \text{SP6} + \text{Q1} + \text{Q2} + \text{Q3} + \text{Q4} + \text{Q1:Q2}$

	df	SS	MS	LOD	%var	Pvalue(Chi2)	Pvalue(F)
Model	12	43416	3618.0	84.18	50.33	0	0
Error	541	42851	79.2				
Total	553	86266					

Estimated effects:

	est	SE	t
Intercept	327.2417	1.2985	252.018
TL3	-12.8621	1.4944	-8.607
SP1	-9.1620	1.5729	-5.825
SP2	-14.1634	1.5964	-8.872
SP3	-16.0743	2.0760	-7.743
SP4	-8.0083	3.3888	-2.363
SP5	-10.1892	2.3273	-4.378
SP6	-15.5640	1.4143	-11.005
8@9.1	11.6583	1.2913	9.028
8@28.0	-0.3157	1.3856	-0.228
10.18@10.0	4.2497	0.8399	5.060
24@46.0	3.3667	0.8270	4.071
8@9.1:8@28.0	19.4213	2.7423	7.082

Most striking, of course, are the loci on linkage group 8. The distal locus on linkage group 8 has essentially no marginal effect, but it has a large influence on the effect of the proximal locus. Our estimated QTL model explains a substantial fraction of the phenotypic variance (50%).

11.3 QTL $\times$ covariate interactions

We now turn our attention to the search for potential QTL $\times$ MCE interactions in these data: loci that show varying effects across the eight MCE groups (defined by the female source of the eggs). As described in Sec. 7.2, there are three ways that we might go about identifying such interactions. First, we could focus on the QTL identified in Sec. 11.2, which showed clear marginal effects, and test for QTL $\times$ MCE interaction at those positions. Second, we may look for loci for which the combined effect of the QTL and its possible interactions with MCE are clear (after adjustment for the genome scan), and again test for the QTL $\times$ MCE interactions at those positions, with no further adjustment for multiple testing. Finally, we may look for positions for which the LOD score for the QTL $\times$ MCE interaction is large, adjusting for the genome scan.

We start with a genome scan including MCE as an interactive covariate.

```
> outi <- scanone(trout, method="hk", addcovar=femcov,
+               intcovar=femcov)
```

The results are LOD scores for a test of the full model (including the QTL $\times$ MCE interaction) to the null model of no QTL, and so concern eight degrees of freedom (the effect of the QTL in each of the eight MCE groups). A large LOD score indicates that the QTL has an effect in at least one of the eight MCE groups.

It is best to compare these results side-by-side with the results, obtained in the previous section, in which MCE was included as an additive covariate but not an interactive covariate. For convenience, we combine the LOD scores together into one object, along with the difference, which concerns the test of QTL $\times$ MCE interaction.

```
> outi <- c(out, outi, outi-out, labels=c("a", "f", "i"))
```

We may then plot the three sets of LOD curves as follows.

```
> plot(outi, lod=1:3, ylab="LOD score", alternate.chrid=TRUE)
```

As seen in Fig. 11.12, we continue to have extremely strong evidence for the locus on linkage group eight, but note that there is little evidence for a QTL $\times$ MCE interaction at this locus.

To make sense of the statistical significance of the results, we perform a permutation test. It is critical that the permutations with MCE as an interactive covariate be precisely matched to those with MCE as a strictly additive

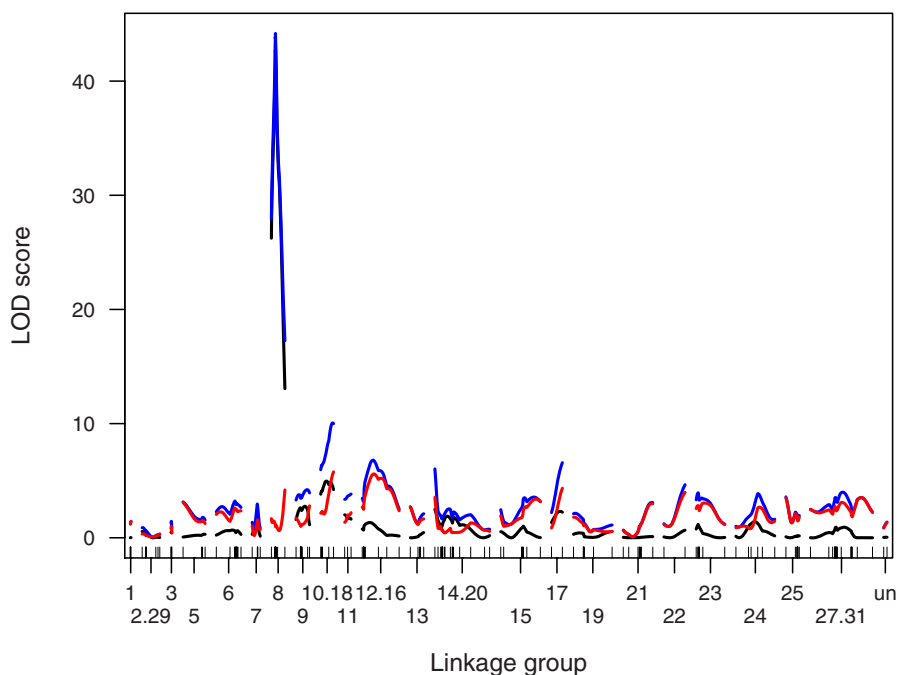


Figure 11.12. LOD curves from a genome scan by Haley-Knott regression for the trout data, with MCE groups included as additive covariates (in black), with MCE groups included as interactive covariates (in blue) and for the test of $QTL \times MCE$ interaction (in red).

covariate, so that the differences (which concern the test of $QTL \times MCE$ interaction) have meaning. And so we use `set.seed` to set the seed for the random number generator to be the same as was used in our permutations in Sec. 11.2.

```
> set.seed(523938)
> opermi <- scanone(trout, method="hk", addcovar=femcov,
+                   intcovar=femcov, n.perm=4000)
```

We again combine the results together into one object.

```
> opermi <- cbind(operm, opermi, opermi-operm,
+                 labels=c("a", "f", "i"))
```

With eight MCE groups and so seven degrees of freedom associated with the $QTL \times MCE$ interaction, the significance thresholds for the LOD scores with MCE as an interactive covariate are quite large.

```
> summary(opermi)
LOD thresholds (4000 permutations)
```

```

      lod.a lod.f lod.i
5%    2.66  6.93  5.36
10%   2.32  6.38  4.81

```

If we were to test for QTL $\times$ MCE pointwise (that is, not adjusting for the genome scan), we could make use of the approximation that $\text{LOD} \times (2 \ln 10)$ follows a $\chi^2(\text{df} = 7)$ distribution under the null hypothesis (of no interaction). The pointwise 5% significance threshold is then

```

> qchisq(0.95, 7) / (2 * log(10))

[1] 3.055

```

The advantage of pasting the three sets of LOD curves together is that we can get a combined summary. If we use `format="allpheno"` in the call to `summary.scanone`, we get a row in the summary table for each position for which any one of the three LOD scores exceeds its chosen threshold.

```

> summary(outi, perms=opermi, alpha=0.05, pvalues=TRUE,
+         format="allpheno")

```

	chr	pos	lod.a	pval	lod.f	pval	lod.i
OmyFGT12	8	9.11	43.18	0.00000	44.17	0.0000	0.988
c9.loc20	9	20.00	2.76	0.03775	4.09	0.8153	1.333
c10.18.loc13	10.18	13.00	4.97	0.00025	7.68	0.0177	2.705
c10.18.loc27	10.18	27.00	4.47	0.00050	10.06	0.0000	5.592
AGCCAG13	10.18	28.69	4.22	0.00125	10.00	0.0000	5.775
c12.16.loc26	12.16	26.00	1.17	0.78350	6.76	0.0638	5.593


```

      pval
OmyFGT12  0.9978
c9.loc20   0.9910
c10.18.loc13 0.7232
c10.18.loc27 0.0357
AGCCAG13   0.0283
c12.16.loc26 0.0357

```

With MCE as a strictly additive covariate (column `lod.a`), we see significant evidence for QTL on linkage groups 8, 9, and 10.18. The loci on linkage groups 8 and 9 show no evidence for QTL $\times$ MCE interaction ($\text{LOD}_i = \text{LOD}_f - \text{LOD}_a$ is small). However, the locus on linkage group 10.18 shows reasonably good evidence for an interaction. When allowing for QTL $\times$ MCE interaction, the QTL on linkage group 10.18 shifts a bit (from 13 cM to 27 cM), and the evidence for interaction becomes clear. In the analysis allowing QTL $\times$ MCE interaction, a locus on linkage group 12.16 nearly reaches significance, and shows a strong interaction effect.

Recall our three strategies for identifying QTL $\times$ MCE interactions. First, we could look at loci with clear marginal effects (adjusting for the genome scan), and test the interaction at these positions, pointwise. With this strategy,

we identify loci on linkage groups 8, 9, and 10.18, but only the locus on linkage group 10.18 would show a QTL $\times$ MCE interaction. Second, we could look at significant loci in the scan allowing for QTL $\times$ MCE interaction, and again test for the interaction at these positions, pointwise. If we are strict in this approach, we identify just the loci on linkage groups 8 and 10.18, and again only the locus on linkage group 10.18 would show a significant QTL $\times$ MCE interaction. Finally, we could focus on the interaction LOD score alone, adjusting for the genome scan. This reveals the QTL $\times$ MCE interactions for loci on linkage groups 10.18 and 12.16.

In the above analysis, we consider just a single locus at a time. But our analysis in Sec. 11.2 revealed two interacting loci on linkage group 8 with very large effects, and so it would be best to repeat our scan for possible QTL $\times$ MCE interaction, accounting for these large-effect loci. This may be done with `addqtl`.

We first call `makeqtl` to create a QTL object containing our two QTL on linkage group 8.

```
> qtl <- makeqtl(trout, c("8", "8"), c(9.1, 28), what="prob")
```

We then call `addqtl` twice. First, we scan for a third QTL, with the MCE groups as strictly additive covariates. Second we scan for a third QTL, allowing the MCE groups to interact with the QTL being scanned but not with the first two QTL.

The model formulas are a bit cumbersome to create, as we must refer to the seven covariates by name. One can avoid some typing (and reduce the chance of errors) by using `paste` to create a character string representation of the model formula. We could then use `as.formula` to convert it to a formula, though actually `addqtl` and related functions accept the character representation, and so conversion with `as.formula` is not needed.

```
> addform <- paste("y~Q1*Q2+Q3+",
+                 paste(colnames(femcov), collapse="+"),
+                 sep="")
> addform
```

```
[1] "y~Q1*Q2+Q3+TL3+SP1+SP2+SP3+SP4+SP5+SP6"
```

```
> intform <- paste("y~Q1*Q2+Q3+",
+                 paste("Q3", colnames(femcov),
+                 sep="*", collapse="+"),
+                 sep="")
> intform
```

```
[1] "y~Q1*Q2+Q3+Q3*TL3+Q3*SP1+Q3*SP2+Q3*SP3+Q3*SP4+Q3*SP5+Q3*SP6"
```

Now we are ready to perform the scans with `addqtl`.

```
> out.aq <- addqtl(trout, qtl=qtl, method="hk", covar=femcov,
+                 formula=addform)
> out.iq <- addqtl(trout, qtl=qtl, method="hk", covar=femcov,
+                 formula=intform)
```

We again paste the three sets of LOD scores into one object.

```
> outi.aq <- c(out.aq, outi.aq, outi.aq - out.aq,
+             labels=c("a", "f", "i"))
```

We use `summary.scanone` to pull out the interesting loci. We will assess significance using the results of our permutation tests not conditioning on linkage group 8, even though they are not strictly valid here.

```
> summary(outi.aq, perms=opermi, alpha=0.05, pvalues=TRUE,
+         format="allpheno")
```

	chr	pos	lod.a	pval	lod.f	pval	lod.i	pval
c9.loc30	9	30.0	1.12	0.81800	7.32	0.0302	6.196	0.0155
AGCCAG15	9	30.2	1.07	0.85925	7.32	0.0302	6.251	0.0145
c10.18.loc9	10.18	9.0	5.57	0.00025	6.43	0.0922	0.866	0.9988
AGCCAG13	10.18	28.7	4.81	0.00025	10.56	0.0000	5.747	0.0305
c17.loc18	17	18.0	2.91	0.02675	5.26	0.3380	2.349	0.8413
c24.loc44	24	44.0	3.66	0.00450	5.45	0.2755	1.788	0.9573

The locus on linkage group 10.18 remains, and shows a clear QTL $\times$ MCE interaction. The locus on linkage group 12.16 has disappeared. Additional loci on linkage groups 17 and 24 are seen, but neither shows evidence for a QTL $\times$ MCE interaction. Most interesting is the locus on linkage group 9, which no longer shows a marginal effect, but does show a clear QTL $\times$ MCE interaction.

A plot of the LOD curves for selected linkage groups may be useful.

```
> plot(outi.aq, lod=1:3, ylab="LOD score", alternate.chrid=TRUE,
+      chr=c(8, 9, "10.18", "12.16", 17, 24))
```

The LOD curves in Fig. 11.13 are useful in giving a sense of the precision of localization of the QTL.

Of course, we should bring all of the QTL together into one model, as this gives the best assessment of the support for the individual loci. However, the drop-one-QTL analysis with `fitqtl` can be extremely cumbersome in the case that we have a multilevel factor as an interactive covariate: each term is dropped one at a time, and we really want to see what happens when the set of terms are omitted together.

We can, however, perform our own drop-one-QTL analysis, by repeatedly calling `fitqtl` with multiple model formulas. The formulas are long and cumbersome, but with careful use of `paste`, we can create them without too much difficulty. We illustrate the process here, though it is not for the faint-of-heart.

We first create a QTL object with all of the putative QTL; we will use `addtoqtl` to add additional terms to the object we had created, containing just the two loci on linkage group 8.

```
> qtl <- addtoqtl(trout, qtl, c(9, "10.18", 17, 24),
+                 c(30, 28.7, 18, 44))
```

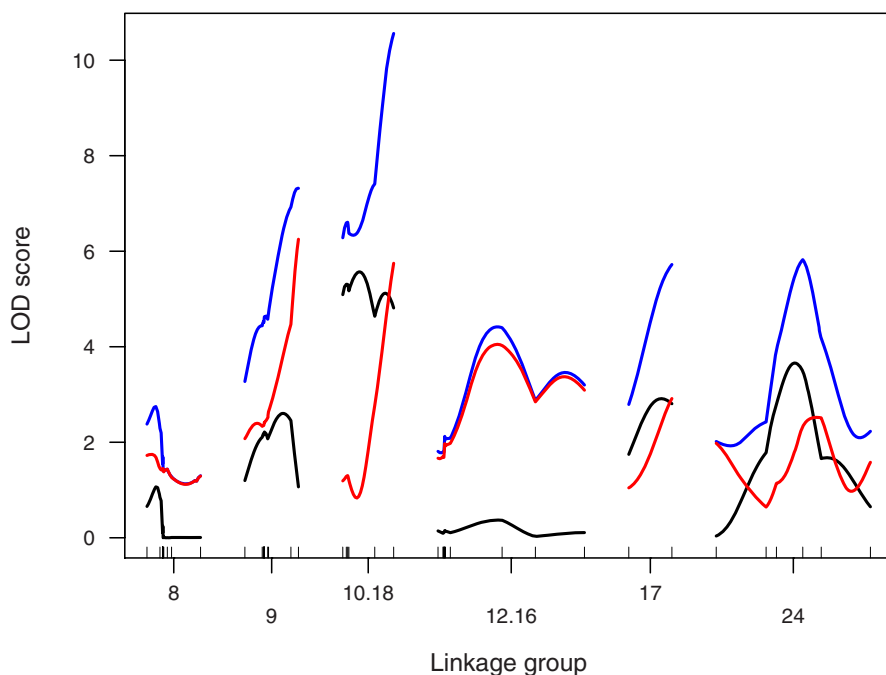


Figure 11.13. LOD curves for selected linkage groups from a genome scan by Haley–Knott regression for the `trout` data, controlling for two interacting loci on linkage group 8, with MCE groups included as additive covariates (in black), with MCE groups included as interactive covariates (in blue) and for the test of $\text{QTL} \times \text{MCE}$ interaction (in red).

Now we create our set of model formulas. We start with the full model, containing all of the QTL plus the $\text{QTL} \times \text{MCE}$ interactions for the loci on linkage groups 9 and 10.18.

```
> fullform <- paste("y~Q1*Q2+Q3+Q4+Q5+Q6",
+                   paste(colnames(femcov), collapse="+"),
+                   paste("Q3", colnames(femcov), sep=":",
+                           collapse="+"),
+                   paste("Q4", colnames(femcov), sep=":",
+                           collapse="+"), sep=":",
+                   collapse="+"), sep="+")
```

We will assume that evidence for the loci on linkage group 8 is so strong that we don't need to consider models that lack them, but we will want to fit the set of models with each of the other QTL missing.

```
> form.m9 <- paste("y~Q1*Q2+Q4+Q5+Q6",
+                   paste(colnames(femcov), collapse="+"),
+                   paste("Q4", colnames(femcov), sep=":",
+                           collapse="+"), sep=":",
+                   collapse="+"), sep="+")
```

```

+               collapse="+"), sep="+")
> form.m1018 <- paste("y~Q1*Q2+Q3+Q5+Q6",
+                    paste(colnames(femcov), collapse="+"),
+                    paste("Q3", colnames(femcov), sep=":",
+                          collapse="+"), sep="+")
> form.m17 <- paste("y~Q1*Q2+Q3+Q4+Q6",
+                  paste(colnames(femcov), collapse="+"),
+                  paste("Q3", colnames(femcov), sep=":",
+                        collapse="+"),
+                  paste("Q4", colnames(femcov), sep=":",
+                        collapse="+"), sep="+")
> form.m24 <- paste("y~Q1*Q2+Q3+Q4+Q5",
+                  paste(colnames(femcov), collapse="+"),
+                  paste("Q3", colnames(femcov), sep=":",
+                        collapse="+"),
+                  paste("Q4", colnames(femcov), sep=":",
+                        collapse="+"), sep="+")

```

We also want formulas with just the QTL $\times$ MCE interactions for the loci on linkage groups 9 and 10.18 omitted.

```

> form.m9int <- paste("y~Q1*Q2+Q3+Q4+Q5+Q6",
+                    paste(colnames(femcov), collapse="+"),
+                    paste("Q4", colnames(femcov), sep=":",
+                          collapse="+"), sep="+")
> form.m1018int <- paste("y~Q1*Q2+Q3+Q4+Q5+Q6",
+                       paste(colnames(femcov), collapse="+"),
+                       paste("Q3", colnames(femcov), sep=":",
+                             collapse="+"), sep="+")

```

With our model formulas in hand, we can calculate the LOD scores for each of these seven models. Let us start with the full model. It would be best to first use `refineqtl` to get improved estimates of the locations of the QTL, in the context of this model.

```

> qtl <- refineqtl(trout, qtl=qtl, formula=fullform,
+                method="hk", covar=femcov, verbose=FALSE)

```

We now use `fitqtl` to fit the model.

```

> full <- fitqtl(trout, qtl=qtl, formula=fullform, method="hk",
+               covar=femcov, dropone=FALSE)

```

In the summary of the result, we can see the LOD score comparing the full model to the null model (with none of the QTL or covariates).

```

> summary(full, pvalues=FALSE)

```

Full model result

```
Model formula: y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + TL3 + SP1 + SP2
               + SP3 + SP4 + SP5 + SP6 + Q1:Q2 + Q3:TL3 +
               Q3:SP1 + Q3:SP2 + Q3:SP3 + Q3:SP4 + Q3:SP5 +
               Q3:SP6 + Q4:TL3 + Q4:SP1 + Q4:SP2 + Q4:SP3 +
               Q4:SP4 + Q4:SP5 + Q4:SP6
```

	df	SS	MS	LOD	%var
Model	28	48552	1734.00	99.54	56.28
Error	525	37714	71.84		
Total	553	86266			

To pull out just the LOD score for the fit of the model (relative to the null model with no QTL), note that the output of `fitqtl` is a list, and one of the components, named "lod", is simply this LOD score.

```
> print(fulllod <- full$lod)
```

```
[1] 99.54
```

We may now use `fitqtl` to fit our other six models.

```
> m9 <- fitqtl(trout, qtl=qtl, formula=form.m9, method="hk",
+             covar=femcov, dropone=FALSE)
> m1018 <- fitqtl(trout, qtl=qtl, formula=form.m1018,
+               method="hk", covar=femcov, dropone=FALSE)
> m17 <- fitqtl(trout, qtl=qtl, formula=form.m17, method="hk",
+              covar=femcov, dropone=FALSE)
> m24 <- fitqtl(trout, qtl=qtl, formula=form.m24, method="hk",
+              covar=femcov, dropone=FALSE)
> m9int <- fitqtl(trout, qtl=qtl, formula=form.m9int,
+               method="hk", covar=femcov, dropone=FALSE)
> m1018int <- fitqtl(trout, qtl=qtl, formula=form.m1018int,
+                  method="hk", covar=femcov, dropone=FALSE)
```

We are interested in the differences between the LOD score for the full model and the LOD scores for models with individual terms omitted. Let us start with the loci on linkage groups 17 and 24, for which we did not include QTL $\times$ MCE interactions.

```
> fulllod - m17$lod
```

```
[1] 2.854
```

```
> fulllod - m24$lod
```

```
[1] 2.749
```

The LOD score for each is above the main effect penalty (2.71).

For the loci on linkage groups 9 and 10.18, we first consider the difference between the full model and the models with both the QTL and the QTL $\times$ MCE interaction terms omitted.

```
> fulllod - m9$lod
```

```
[1] 7.933
```

```
> fulllod - m1018$lod
```

```
[1] 11.32
```

These are both well above the threshold from the permutation test including QTL $\times$ MCE interactions. We turn to the interaction LOD scores, comparing the full model to the models with just the QTL $\times$ MCE interaction terms omitted.

```
> fulllod - m9int$lod
```

```
[1] 6.867
```

```
> fulllod - m1018int$lod
```

```
[1] 6.046
```

Both are quite large, indicating that the loci on linkage groups 9 and 10.18 show clear QTL $\times$ MCE interactions.

In the LOD scores above, the loci other than the one under test are kept fixed at their estimated positions under the full model. This is what is done in the drop-one-QTL analysis in `fitqtl`, but it gives a somewhat rosy view of the support for the individual loci. Ideally, for each submodel, the locations for the remaining QTL would be reestimated, and each comparison would be between the full model with QTL in their best positions under the full model, to a submodel with QTL in their best positions under that submodel.

This is simple enough to accomplish in our “by hand” comparisons of the individual models. We simply run `refineqtl` for each submodel, followed by `fitqtl`.

```
> qtl.m9 <- refineqtl(trout, qtl=qtl, formula=form.m9,
+                     method="hk", covar=femcov,
+                     verbose=FALSE)
> qtl.m1018 <- refineqtl(trout, qtl=qtl, formula=form.m1018,
+                        method="hk", covar=femcov,
+                        verbose=FALSE)
> qtl.m17 <- refineqtl(trout, qtl=qtl, formula=form.m17,
+                      method="hk", covar=femcov,
+                      verbose=FALSE)
> qtl.m24 <- refineqtl(trout, qtl=qtl, formula=form.m24,
+                      method="hk", covar=femcov,
+                      verbose=FALSE)
> qtl.m9int <- refineqtl(trout, qtl=qtl, formula=form.m9int,
+                        method="hk", covar=femcov,
+                        verbose=FALSE)
```

```
> qtl.m1018int <- refineqtl(trout, qtl=qtl, method="hk",
+                           formula=form.m1018int, covar=femcov,
+                           verbose=FALSE)
```

We now call `fitqtl` for each of these reduced models, with the QTL in the positions estimated under the corresponding model.

```
> m9r <- fitqtl(trout, qtl=qtl.m9, formula=form.m9,
+               method="hk", covar=femcov, dropone=FALSE)
> m1018r <- fitqtl(trout, qtl=qtl.m1018, formula=form.m1018,
+                 method="hk", covar=femcov, dropone=FALSE)
> m17r <- fitqtl(trout, qtl=qtl.m17, formula=form.m17,
+               method="hk", covar=femcov, dropone=FALSE)
> m24r <- fitqtl(trout, qtl=qtl.m24, formula=form.m24,
+               method="hk", covar=femcov, dropone=FALSE)
> m9intr <- fitqtl(trout, qtl=qtl.m9int, formula=form.m9int,
+                 method="hk", covar=femcov, dropone=FALSE)
> m1018intr <- fitqtl(trout, qtl=qtl.m1018int, method="hk",
+                    formula=form.m1018int, covar=femcov,
+                    dropone=FALSE)
```

We recalculate the conditional LOD scores for each locus, first for the loci on linkage groups 17 and 24.

```
> fulllod - m17r$lod
```

```
[1] 2.849
```

```
> fulllod - m24r$lod
```

```
[1] 2.736
```

The LOD scores are slightly smaller, but both are still above main effect penalty (2.71).

Now we consider the loci on linkage groups 9 and 10.18.

```
> fulllod - m9r$lod
```

```
[1] 7.921
```

```
> fulllod - m1018r$lod
```

```
[1] 11.01
```

```
> fulllod - m9intr$lod
```

```
[1] 5.168
```

```
> fulllod - m1018intr$lod
```

```
[1] 5.798
```

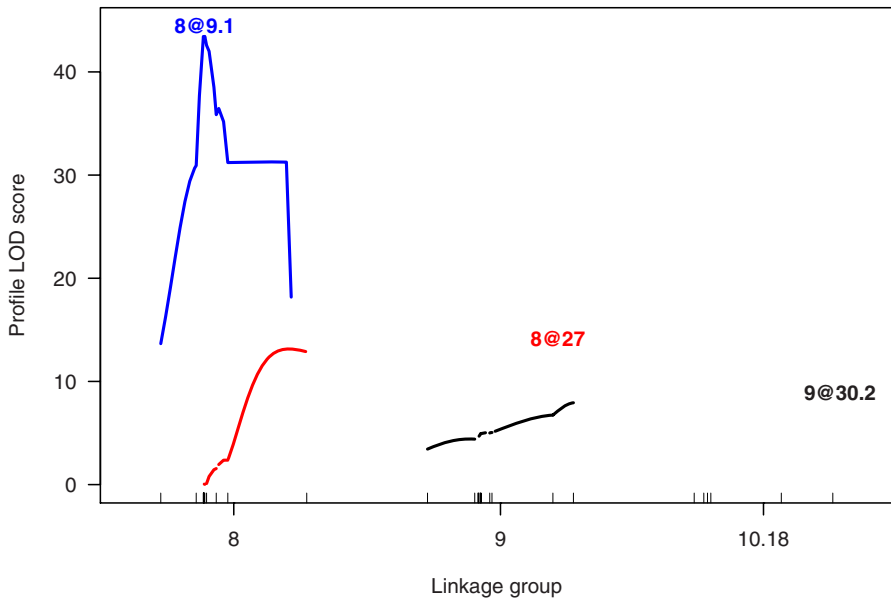


Figure 11.14. Profile LOD curves for a six-QTL model, including two epistatic QTL on linkage group 8 and QTL $\times$ MCE interactions for the QTL on linkage groups 9 and 10.18, for the `trout` data.

The biggest change is in the interaction LOD score for the locus on linkage group 9, which dropped from 6.87 to 5.17. But the evidence for both of these loci, and for their QTL $\times$ MCE interactions, remains clear.

Let us turn to interval estimates for the locations of the inferred QTL. We are particularly interested in the QTL on linkage group 10.18 (which is really the pasting together of linkage groups 10 and 18): is the QTL in the linkage group 10 part or the linkage group 18 part, or can we not tell?

First, let us plot the LOD profiles (see Sec. 9.3.2). Since we had used `refineqtl` on the object `qtl`, concerning the full model, we may immediately use `plotLodProfile`.

```
> plotLodProfile(qtl, col=c("blue", "red", rep("black", 4)),
+               ylab="Profile LOD score")
```

The LOD profiles appear in Fig. 11.14. Recall from Sec. 9.3.2 that in these curves, the location of one QTL is allowed to vary while other QTL are fixed at their best estimates. The LOD scores are for the comparison between the full model and the reduced model with the QTL of interest (and any of its interactions) omitted. For example, in the curve for the QTL on linkage group 10.18, that locus is allowed to vary while the locations of all other QTL are kept fixed, and the LOD scores compare the full model, with the locus on

linkage group 10.18 in varying position, to the reduced model in which this QTL plus its $\text{QTL} \times \text{MCE}$ interactions are omitted.

The LOD profiles for the proximal locus on linkage group 8 looks a bit odd, but note that for linked QTL (and particularly for QTL linked as tightly as these two), the profile LOD curves give a poor representation of our uncertainty in the locations of the QTL. It would be best to allow the locations of the two QTL to vary jointly. We will return to that in a moment.

We were particularly interested in the QTL on linkage group 10.18, and of whether its location could be defined to linkage group 10 or linkage group 18, or whether this was uncertain. We may use `lodint` and `bayesint` to obtain approximate confidence intervals for the location of the QTL. The intervals should be considered with caution, as they fail to capture the uncertainty in the locations of the other QTL in the model, and also the performance of these intervals, in the context of multiple-QTL models, is not well understood. The 1.5-LOD support interval and 95% Bayesian credible interval are calculated as follows.

```
> lodint(qtl, qtl.index=4)

      chr   pos   lod
29  10.18 24.00  9.498
34  10.18 28.69 11.319
34  10.18 28.69 11.319

> bayesint(qtl, qtl.index=4)

      chr   pos   lod
31  10.18 26.00 10.38
34  10.18 28.69 11.32
34  10.18 28.69 11.32
```

The intervals are remarkably small: the LOD support interval spans 4.7 cM and the Bayesian interval spans 2.7 cM. It is hard to believe that the location of the QTL could be so precisely defined.

Now, consider the genetic map of the markers on linkage group 10.18.

```
> pull.map(trout, "10.18")

ACGACA11 AGCAGT15  ACCAAG6  agcagc9  AGCCAG12  AGCCAG13
0.000    1.992    2.753    3.444    18.031    28.692
```

The first four markers were on linkage group 10, and the last two markers were on linkage group 18. Thus it appears that the QTL on the merged linkage group 10.18 is within the linkage group 18 part.

Returning to the case of the linked loci on linkage group 8, if we wished to better define the locations of these QTL in the context of our six-QTL model, we should perform a two-dimensional scan on linkage group 8, keeping the locations of the other four QTL at their best estimates. This can be

accomplished with `addpair`. We first use `dropfromqtl` to drop the two QTL on linkage group 8 from our QTL object.

```
> qtl.m8 <- dropfromqtl(qtl, 1:2)
```

We need to create a model formula including our QTL $\times$ MCE interactions, and with the two additional QTL (to be scanned) allowed to interact. Note that, having dropped the QTL on linkage group 8, the numeric indices of the other four QTL have changed.

```
> theformula <- paste("y~Q1+Q2+Q3+Q4+Q5*Q6",
+                      paste(colnames(femcov), collapse="+"),
+                      paste("Q1", colnames(femcov), sep=":",
+                      collapse="+"),
+                      paste("Q2", colnames(femcov), sep=":",
+                      collapse="+"), sep=":",
+                      collapse="+"), sep="+")
```

Now we are ready to run `addpair`.

```
> out.ap <- addpair(trout, chr="8", qtl=qtl.m8, covar=femcov,
+                  formula=theformula, method="hk",
+                  incl.markers=TRUE, verbose=FALSE)
```

The results are the two-dimensional equivalent of the LOD profiles in Fig. 11.14. At each pair of positions for the two QTL, we compare the full model to the reduced model with these two QTL (and their interaction) omitted.

We may plot the two-dimensional LOD profile as follows. The argument `contour` is used to display an approximate 1.5-LOD support region. Note that the performance of such two-dimensional regions is not well understood.

```
> plot(out.ap, contours=1.5)
```

As seen in Fig. 11.15, the 1.5-LOD support regions indicates that the location of the proximal QTL on linkage group 8 has been impressively well defined, to an interval spanning just over 1 cM. The location of the distal QTL is less well defined; the support region spans about 8 cM.

Finally, let us study the effects of the loci. We are particularly interested in the QTL on linkage groups 9 and 10.18, which exhibit QTL $\times$ MCE interactions. We will use `fitqtl` to get the estimated effects in the context of our six-QTL model.

```
> summary(fitqtl(trout, qtl=qtl, formula=fullform, covar=femcov,
+               method="hk", dropone=FALSE, get.ests=TRUE))
```

Full model result

```
-----
Model formula: y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + TL3 + SP1 + SP2
+ SP3 + SP4 + SP5 + SP6 + Q1:Q2 + Q3:TL3 +
+ Q3:SP1 + Q3:SP2 + Q3:SP3 + Q3:SP4 + Q3:SP5 +
```

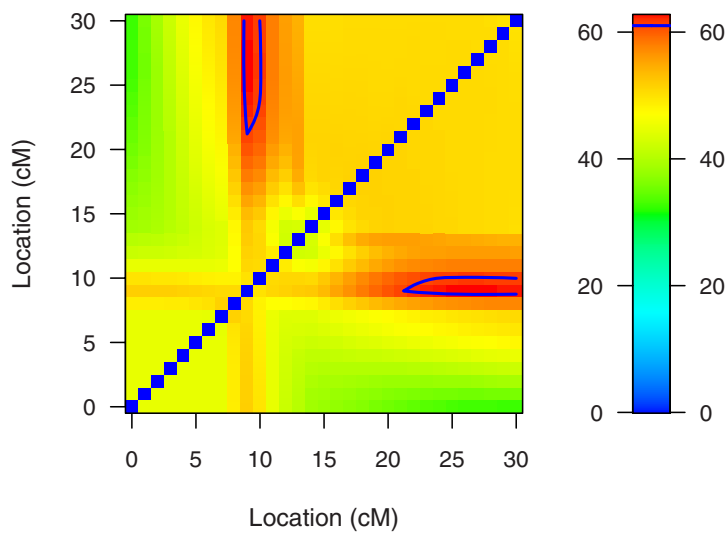


Figure 11.15. Two-dimensional profile LOD surface for the pair of interacting QTL on linkage group 8, in the context of a six-QTL model, including QTL $\times$ MCE interactions for the QTL on linkage groups 9 and 10.18 and additional loci on linkage groups 17 and 24, for the trout data.

Q3:SP6 + Q4:TL3 + Q4:SP1 + Q4:SP2 + Q4:SP3 +
Q4:SP4 + Q4:SP5 + Q4:SP6

	df	SS	MS	LOD	%var	Pvalue(Chi2)	Pvalue(F)
Model	28	48552	1734.00	99.54	56.28	0	0
Error	525	37714	71.84				
Total	553	86266					

Estimated effects:

	est	SE	t
Intercept	326.4836	1.3124	248.767
TL3	-12.1128	1.4839	-8.163
SP1	-8.5237	1.5711	-5.425
SP2	-14.0621	1.5777	-8.913
SP3	-15.2948	2.0203	-7.570
SP4	-6.0049	3.8896	-1.544
SP5	-11.1024	2.4090	-4.609
SP6	-15.2999	1.4161	-10.805
8@9.1	12.2976	1.3077	9.404
8@27.0	-1.1588	1.4225	-0.815
9@30.2	9.3167	2.5252	3.690

10.18@28.7	15.3082	3.0431	5.031
17@23.0	2.8183	0.7939	3.550
24@48.0	2.7082	0.7775	3.483
8@9.1:8@27.0	19.5517	2.8142	6.948
9@30.2:TL3	-12.9477	3.0492	-4.246
9@30.2:SP1	-3.0747	3.1703	-0.970
9@30.2:SP2	-5.8341	3.1489	-1.853
9@30.2:SP3	-12.1483	4.1183	-2.950
9@30.2:SP4	-3.3755	7.7862	-0.434
9@30.2:SP5	-14.3642	4.6790	-3.070
9@30.2:SP6	-8.3860	2.8493	-2.943
10.18@28.7:TL3	-7.5077	3.5224	-2.131
10.18@28.7:SP1	-12.7976	3.6071	-3.548
10.18@28.7:SP2	-11.4783	3.5959	-3.192
10.18@28.7:SP3	-10.2645	4.6146	-2.224
10.18@28.7:SP4	-11.6208	7.9593	-1.460
10.18@28.7:SP5	-13.8230	5.1415	-2.689
10.18@28.7:SP6	-15.0835	3.3221	-4.540

Consider the locus on linkage group 10.18, and recall that, for the linkage group 18 part of this merged “linkage group,” we had swapped the two genotypes, and so the effects have the opposite sign than might be expected. The estimated coefficient in the row labeled 10.18@28.7 is the effect of the linkage group 10.18 locus in the first MCE group (TL1). That is, for the individuals whose egg came from the TL1 female, the difference in the average phenotype between the two genotypes is 15.3 ± 3.0 . The other rows that begin 10.18@28.7 are for the differences in the QTL effect for the indicated MCE group and the QTL effect for the TL1 group. Relative to the TL1 group, the QTL on linkage group 10.18 has smaller effect in all other MCE groups; in some cases (e.g., SP6) the QTL appears to have no effect. Similar observations apply to the locus on linkage group 9.

The linked epistatic QTL on linkage group 8 are especially interesting. To make sense of the estimated effects, note that the marginal effects of the QTL are based on a coding of the two QTL genotypes as $\pm 1/2$, and the interaction term is the product of the two main effect terms. Thus the effect of the proximal QTL among individuals with the CC genotype at the distal locus is estimated to be $12.3 - 19.6/2 = 2.5$, while its effect among individuals with the OO genotype at the distal locus is estimated to be $12.3 + 19.6/2 = 22.1$. That is, the proximal locus on linkage group 8 has a large effect, but only among individuals with genotype OO at the distal locus.

11.4 Discussion

Our principal aim in this case study was to illustrate the exploration of possible QTL $\times$ covariate interactions. As with the first case study, we have focused

on the analysis techniques rather than the biological interpretation of the results; for the latter, see the original article describing these data (Nichols *et al.*, 2007).

With these data, there was a single covariate to consider, maternal cytoplasmic environment (MCE): the source of the individuals' eggs in these doubled haploids. Nevertheless, the best approach for evaluating $\text{QTL} \times \text{covariate}$ interactions, with adjustment for the multiple hypothesis tests, is still not perfectly clear. We favor the use of a scan allowing for the possibility of $\text{QTL} \times \text{covariate}$ interactions, followed by pointwise tests of the significance of the interactions. Had we been presented with a set of covariates for which $\text{QTL} \times \text{covariate}$ interactions were to be explored, the multiple testing issue would be even more onerous.

Automated multiple-QTL analyses in R/qtl do not yet allow for the possibility of $\text{QTL} \times \text{covariate}$ interactions, and so we relied on a more exploratory approach based on single- and two-dimensional genome scans. Extension of the model comparison criteria of Sec. 9.1.4, to allow $\text{QTL} \times \text{covariate}$ interactions, deserves exploration. Likely one could use a significance threshold for the interaction LOD score as a penalty on a $\text{QTL} \times \text{covariate}$ interaction. If multiple covariates were to be considered, such a significance threshold should account for the search among the covariates as well as the scan across the genome.

Our treatment of the linked pairs of homeologous linkage groups in these data (swapping the genotypes on one linkage group and merging the two together) is unorthodox, and is not beyond criticism. An advantage of our approach is that we avoid the possibility of a single QTL giving linkage signals on multiple linkage groups. However, the gaps between the parts of the merged linkage groups may contain no DNA, and the swapping of genotypes on one part makes the interpretation of QTL effects more difficult. Clearly, we need a better understanding of the underlying mechanism giving rise to these pseudolinkages.