Whitehead Institute

Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL Version 1.1: A Tutorial and Reference Manual

Stephen E. Lincoln, Mark J. Daly, and Eric S. Lander

A Whitehead Institute for Biomedical Research Technical Report
Second Edition, January, 1993
(Beta Distribution 3B)

Please address all correspondence to:

MAPMAKER
c/o Dr. Eric Lander
Whitehead Institute
9 Cambridge Center
Cambridge, MA 02142 USA

Internet: mapmaker@genome.wi.mit.edu
Bitnet: mapm@mitwibr
FAX: 617-258-6505

Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL

Table of Contents:

MAPPING QUANTITATIVE TRAITS WITH MAPMAKER/QTL

Over the last half century, a number of methods have been developed to map genes controlling quantitatively measured phenotypes segregating in experimental populations. In general, these techniques work by finding correlations between the inheritance of particular genetic markers and variation in the phenotype for each individual in the population. MAPMAKER/QTL extends these methods to (1) provide support for "interval mapping", allowing one to fully exploit the information provided by a genetic linkage map, and (2) to calculate LOD scores for putative QTLs, providing a measure of the support for any particular hypothesis.

MAPMAKER/QTL implements these features with a user-friendly interface which is similar to that used by the multipoint genetic linkage analysis package, MAPMAKER/EXP, and actually shares data files with the new versions of that program. In fact, MAPMAKER/QTL and MAPMAKER/EXP together form a complete linkage analysis package (which we call MAPMAKER) for much of your mapping work. However, MAPMAKER is not a closed package: it is fairly easy to use other methods along with the MAPMAKER analyses.

In this section we use MAPMAKER/QTL to map putative QTLs controlling a quantitative trait varying across 333 F2 progeny of a cross between two different strains of tomato. For simplicity, we will show the results of searching for QTLs on only two chromosomes, although in practice one would normally examine an entire genome (so long as a linkage map of polymorphic markers were available). We demonstrate here many of the important MAPMAKER/QTL functions as we step through the analysis of this data set. After reading this section, you should be able to begin using MAPMAKER/QTL to analyze your own data right away. For reference, you can find a more thorough description of each of MAPMAKER/QTL's features in section III of this manual.

The data set we analyze here simply describes the genetic marker (RFLP) and quantitative trait values for each individual in the data set. We also include a previously calculated genetic linkage map, describing the map orders and distances for the genetic markers (In fact, this data set is precisely that used in the the MAPMAKER/EXP tutorial! You are free to use linkage maps calculated with other methods, however.) Users who are familiar with the material covered in the MAPMAKER/EXP tutorial will find that many of the concepts used in MAPMAKER/QTL are quite similar, and that much of the user interface is preserved between the two programs.

Throughout this example, all input to the computer is presented using ***bold italics***, while MAPMAKER/QTL output is shown in `ordinary type`. The entire process we describe here required less than 10 minutes on a Sun SPARCstation-2 computer.

Getting Data into MAPMAKER/QTL

The first step in running MAPMAKER/QTL is, oddly enough, to run MAPMAKER/EXP. In the current version, this is the only way to export data into MAPMAKER/QTL. We refer you to the Installation Guide for instructions as to how to start MAPMAKER/EXP.

In our MAPMAKER/QTL tutorial, we will use the "sample" data file which we also use in the MAPMAKER/EXP tutorial. Moreover, we will use the linkage maps which we found in that tutorial. Because we assume that you have already "prepared" this file, we now load in back into MAPMAKER/EXP with the "load data" command (as opposed to the "prepare data" command -- see the MAPMAKER/EXP tutorial for details).

Preparing data for MAPMAKER/QTL requires basic knowledge of the genome gnalysis features present in MAPMAKER/EXP 3.0. In particular you need to do the following:

1. Define each chromosome you wish to search for QTLs.
2. Assign the markers in each chromosome's map to the appropriate chromosome.
3. Set the framework of each chromosome to specify the correct order of these markers.

These issues are discussed in detail in the MAPMAKER/EXP manual, and are not discussed here. In fact, to make this process most simple, we have written a *batch file* containing the MAPMAKER/EXP commands needed to accomplish these steps (it is easy to create such a file using any standard text editor). We now execute the commands in this file, named "sample.inp", using MAPMAKER/EXP's "run" command. This batch file has been supplied with MAPMAKER, and may serve as a guide to help you write your own such files.

It is important to note that you do *not* need to use MAPMAKER/EXP's linkage calculations to prepare maps for MAPMAKER/QTL. In fact, you can make maps using any program, and specify those orders and distances in your "sequences" used to set the frameworks for MAPMAKER/QTL. See the MAPMAKER/EXP reference manual for details.

Having setup the relevant data, we know quit MAPMAKER/EXP, which will cause it to save our "sample" data file in a format which can be read by MAPMAKER/QTL.

```
************************************************************************
* Welcome to:                                                          *
*                                                                      *
*                              MAPMAKER/EXP                            *
*                             (version 3.0b)                           *
*                                                                      *
************************************************************************

1> load  sample
loading data from file 'sample.data... ok
  F2 intercross data  (333 individuals, 12 loci)... ok
loading two-point data from file 'sample.2pt'.... ok







2> run  sample.inp

        ...Running commands from input file 'sample.inp'...

3> make chromosome one two
chromosomes defined: one two

4> sequence 1
sequence #1= 1

5> anchor one
1    - anchor locus on chromosome one

                                    <--  Output  from  batch  file  continues

        ...end of input file...









15> quit
save data before quitting? [yes] yes
saving map data in file 'sample.maps'... ok

        ...goodbye...
```

Starting MAPMAKER/QTL

Like MAPMAKER/EXP, when MAPMAKER/QTL starts running you will see its start-up banner and a prompt ("1>") for the first command. Precisely how you start MAPMAKER/QTL depends on your computer and on how the program was installed: These issues are described in the Installation Guide supplied with the software.

The first step in most MAPMAKER/QTL sessions is to load in a data file for analysis. As is the convention for MAPMAKER/EXP and MAPMAKER/QTL, the data used in this example are split between two files, here named "sample.data", containing the RFLP data, and "sample.traits", containing the quantitative trait values. We now load these data into MAPMAKER/QTL using the "load data" command.

In most cases, you will also want to instruct MAPMAKER/QTL to save a transcript of the session for later reference. Here, we save the transcript in a file named "qtl.out" using the "photo" command. Note that if this file already exists, MAPMAKER/QTL will append the transcript to the end of it.

```
**********************************************************************
* Welcome to:                                                        *
*                                                                    *
*                             MAPMAKER/QTL                           *
*                            (version 1.1b)                          *
*                                                                    *
* Copyright 1987-1992, Whitehead Institute for Biomedical Research    *
**********************************************************************

Type 'help' for help.
```

```
1> load data quant
data file 'sample.data' and 'sample.traits' loaded
(333 F2 intercross progeny, 12 loci, 1 trait)
```

```
2> photo quant.out
'photo' is on. File is 'qtl.out'
```

Preliminary Examination of The Trait Data

The algorithms on which MAPMAKER/QTL is based assume that the values of the quantitative trait vary across the population following a normal distribution. Thus, the first step in most QTL mapping projects is to examine the trait data and determine how well they fit this assumption. If they do not closely follow a normal distribution, is is often possible to transform the data into a "derived trait" which more closely fits the assumption.

In this example, we wish to study a quantitative trait in our data set named "weight". Using the "trait" command, we first inform MAPMAKER/QTL to perform all further analyses on these trait data. Then, we use the "show trait" command to display a simple statistical analysis of these data. This command causes MAPMAKER/QTL to display values for the mean (μ), standard deviation (σ^2), kurtosis, skewness, and quartile ratio of the trait data, as well as the fraction of the individuals whose phenotypes fall within 0.25σ of the mean, 0.5σ, 1.0σ, 2.0σ, and 3.0σ. In addition, a rough histogram of the data is displayed.

As you can see, the distribution of these trait data is slightly skewed toward higher values. To correct for this skewing, we may instead choose to analyze the $\log_{10}$ of the weight data. (This being consistent with the hypothesis that the QTLs somehow control the size of the plant, thus affecting its weight exponentially). To do this, we employ MAPMAKER/QTL's "make trait" command. This allows us to create a new trait with values which are a function of those for another trait existing in the data set. Here we make a new trait named "logwt" which is the log base 10 of the "weight" trait. After applying this function, MAPMAKER/QTL displays the same statistical analysis as for the "show trait" command. As you can see, the transformed trait more closely follows a normal distribution.

```
3> trait 1
The current trait is now: 1 (weight)

4> show trait

Trait 1 (weight):
-----------------------------------------------------------------------------
distribution:                      quartile |   fraction within n deviations:
mean   sigma   skewness  kurtosis  ratio    |   1/4    1/2   1     2     3
6.11   3.51    1.59      3.53      0.83     |   0.23   0.42  0.75  0.95  0.99
 -----------------------------------------------------------------------------

 -0.91   |
  0.85   |
  2.60   |*************************
  4.35   |*******************************************************
  6.11   |************************************************************
  7.86   |***************************************
  9.61   |********************
 11.37   |*************
 13.12   |********
 14.87   |****


5> make trait logwt = log(weight)

New trait number 2 (logwt) had been added to the data set.
-----------------------------------------------------------------------------
distribution:                      quartile |   fraction within n deviations:
mean   sigma   skewness  kurtosis  ratio    |   1/4    1/2   1     2     3
0.72   0.24    -0.06     -0.07     0.97     |   0.21   0.41  0.67  0.94  1.00
 -----------------------------------------------------------------------------

  0.25   |******
  0.36   |********
  0.48   |****************************
  0.60   |**********************************
  0.72   |****************************************************
  0.84   |************************************************************
  0.96   |***********************************
  1.08   |***********************
  1.20   |************
  1.32   |*****
```

As a last preliminary step, we simply print out lists of the genetic markers, linkage maps, and traits included in our data files, for our reference as we proceed with the analysis.

First, we print out the genetic markers and their assigned map orders and distances. To display these maps, we first use MAPMAKER/QTL's "sequence" command to tell the program which regions of the genome we are interested in examining. As we wish to examine all regions for which data are available (which in this example is just three chromosomes), we simply enter a sequence command which specifies "all" genetic markers. (Notice the square brackets used in the sequence command we type. These are required in most MAPMAKER/QTL sequences, as we will discuss later).

We then type the "show maps" command to view the map orders and distances for these markers. By default, MAPMAKER/QTL indicates genetic markers by number, where the numbering is assigned based on their order in the data file (<u>not</u> their order in the genetic linkage map). This feature keeps the marker numbers consistent between MAPMAKER/EXP and MAPMAKER/QTL. We could just as easily <u>always</u> have MAPMAKER/QTL display the marker names instead of their numbers by simply changing the "print names" option setting as described in the reference section of this manual.

Note that the linkage map, including both the map orders and positions of the genetic markers, was calculated using the MAPMAKER/EXP program and was inserted into the data files for MAPMAKER/QTL's use. MAPMAKER/QTL itself does not ever try to recompute map orders or distances.

To display a list of the trait data included in our file, we simply type the "list traits" command. Notice that the "logwt" trait, which we just added to the data set, is displayed along with the function we used to create it.

```
6> sequence  [all]
The sequence is now '[all]'

7> show  linkage  maps

linkage maps:
==============================================================
 1-3           4.2 cM    4.0 %
 3-2          15.0 cM   13.0 %
 2-5          11.9 cM   10.6 %
 5-7          12.2 cM   10.8 %
==============================================================
 4-11         14.8 cM   12.8 %
 11-8          6.4 cM    6.0 %
 8-12         18.9 cM   15.7 %
 12-9         24.0 cM   19.1 %
 9-6          18.1 cM   15.2 %
 6-10         28.6 cM   21.8 %
==============================================================
```

```
8> list  traits
==============================================================
TRAITS:

  1 weight
  2 logwt      = log(weight)
==============================================================
```

Scanning the Genome for Putative QTLs

Perhaps the most useful capability of MAPMAKER/QTL is to compute "QTL likelihood plots" covering the entire genome. In a manner similar to LOD score plots (e.g. likelihood vs. map distance), these graphs provide a visual representation of (1) regions in the genome which are likely to contain putative QTLs, (2) the strength of the data supporting the hypotheses that particular QTLs exist, and (3) the likely position of these putative QTLs.

The fundamental method MAPMAKER/QTL uses to generate these plots works as follows: The program iteratively "steps" along the genome, and at each point calculates a "maximum likelihood QTL map". That is, at each point, MAPMAKER/QTL asks the question, "assuming there is a QTL right here, what is the maximally likely manner in which it's inheritance affects the trait? Moreover, what is the strength of the data supporting this hypothesis?". These results are expressed as a number of real valued parameters, including:

- The effect of the QTL on the trait, expressed in terms of additive "weight" and "dominance" effects for F2 data, or a "weight" term alone for backcross data.
- The fraction of the total variation in the trait across the population explained by the QTL (this is equivalent to an R^2 value calculated by linear regression).
- The mean and standard deviation of the variation in the trait not controlled by the QTL.
- A LOD score, also called a log-likelihood, indicating the strength of the data supporting this hypothesis, and the corresponding

We will discuss the precise interpretation of these values shortly.

To perform this analysis, we first need to set MAPMAKER/QTL's "sequence", telling the program which regions of the genome we wish to search. Because we have already set the sequence to specify "all" regions, we do not need to repeat this step. Just to be sure, we can type "sequence" alone, to make sure that it is correctly set.

We also need to tell MAPMAKER/QTL which quantitative trait we are interested in studying. Here we wish to analyze the $\log_{10}$(weight) data we calculated earlier, and we simply tell the program this using the "trait" command.

Satisfied that the sequence and trait are correctly set, we then type the "scan" command, instructing MAPMAKER/QTL to search for QTLs throughout the genome. By default, MAPMAKER/QTL looks for QTLs every 2.0 cM. Note that the "scan" command can generate quite a large amount of output, and on slower computers can also require a significant amount of time. For this example, we have only included output for the first three chromosomes in this data set. Scanning these three chromosomes for this trait required about 3 minutes on a Sun SPARCStation-1 computer.

```
9> sequence
The sequence is '[all]'

10> trait 2
The current trait is now: 2 (logwt)

11> scan
QTL maps for trait 2 (logwt):
Sequence: [all]
LOD threshold: 2.00  Scale: 0.25 per '*'
No fixed-QTLs.
Scanned QTL genetics are free.

POS     WEIGHT  DOM     %VAR  LOG-LIKE |
---------------------------------------| 1-3 4.2 cM
0.0    -0.033  -0.072   4.7%    3.083  | *****
2.0    -0.053  -0.048   4.5%    2.814  | ****
4.0    -0.067  -0.022   3.8%    2.474  | **
---------------------------------------| 3-2 15.0 cM
0.0    -0.068  -0.021   3.8%    2.442  | **
2.0    -0.076  -0.023   4.8%    2.873  | ****
4.0    -0.080  -0.029   5.8%    3.291  | ******
6.0    -0.081  -0.036   6.6%    3.667  | *******
8.0    -0.081  -0.041   7.2%    3.983  | ********
10.0   -0.079  -0.046   7.4%    4.228  | *********
12.0   -0.077  -0.048   7.4%    4.395  | **********
14.0   -0.075  -0.048   7.1%    4.483  | **********
---------------------------------------| 2-5 11.9 cM
0.0    -0.073  -0.048   6.8%    4.500  | ***********
2.0    -0.079  -0.048   7.7%    4.755  | ************
4.0    -0.084  -0.046   8.2%    4.912  | ************
6.0    -0.088  -0.043   8.4%    4.969  | ************
8.0    -0.088  -0.041   8.2%    4.920  | ************
10.0   -0.087  -0.038   7.7%    4.757  | ************
---------------------------------------| 5-7 12.2 cM
0.0    -0.083  -0.034   6.8%    4.501  | ***********
2.0    -0.084  -0.037   7.2%    4.427  | **********
4.0    -0.084  -0.038   7.3%    4.236  | *********
6.0    -0.081  -0.038   6.9%    3.931  | ********
8.0    -0.077  -0.036   6.2%    3.526  | *******
10.0   -0.071  -0.032   5.2%    3.046  | *****
12.0   -0.063  -0.026   4.0%    2.535  | ***
---------------------------------------|

---------------------------------------| 4-11 14.8 cM
0.0    -0.102  -0.007   9.0%    5.645  | ***************
2.0    -0.110  -0.008  10.4%    6.159  | *****************
4.0    -0.116  -0.008  11.4%    6.584  | *******************
6.0    -0.119  -0.007  12.1%    6.897  | ********************
8.0    -0.120  -0.006  12.3%    7.083  | *********************
10.0   -0.120  -0.005  12.1%    7.135  | *********************
12.0   -0.117  -0.006  11.4%    7.054  | *********************
14.0   -0.111  -0.009  10.4%    6.853  | ********************
---------------------------------------| 11-8 6.4 cM
0.0    -0.109  -0.010   9.9%    6.752  | ********************
2.0    -0.118  -0.012  11.4%    7.418  | **********************
4.0    -0.122  -0.014  12.0%    7.802  | ************************
6.0    -0.122  -0.016  11.8%    7.932  | ************************
```

MAPMAKER/QTL's "scan" command generates one line of output for each point examined in the genome, with the intervals between genetic markers the boundaries between chromosomes indicated. Note that we are using MAPMAKER/QTL's default output mode, which displays markers by their assigned numbers and map distances in Haldane centimorgans. Alternatively, we could tell MAPMAKER/QTL to use the marker's names, and/or to use recombination fractions as our units of map distance. For details of how to do this, see the discussion of the "units" and "print names" commands in the reference section of this manual. MAPMAKER Version 2 users should note that MAPMAKER/QTL defaults to using the same settings of these (and some other) options which you set while using MAPMAKER.

To declare the presence of any QTLs, one needs to select an appropriate LOD score threshold, above which the QTL map data will be considered significant. The selection of an appropriate LOD score must be based on a number of parameters, including the particular genome, type of cross, and density of linkage map used. Also, because MAPMAKER/QTL in effect performs many independent tests, selection of a LOD threshold must take into account the chance of erroneously declaring a QTL <u>anywhere</u> in the genome, rather than the error rate of each particular test. These issues are discussed in detail by Lander and Botstein (Genetics 121: 185-199). For these experiments, we have selected a conservative LOD threshold of 3.0.

If we examine the "like" (short for "log-likelihood") column in the program's output, we see that our data indicate that a number of regions of the genome are likely to contain QTLs affecting the weight trait, a hypotheses supported by log-likelihoods greatly exceeding our threshold of 3.0. We can quickly examine the output for these regions by looking at the row of asterisk (*) characters printed next to each data line. These visually indicate the shape and height of the 'log-likelihood surface'. (By default, one star is printed if the log-likelihood exceeds 2.0, with each additional star indicating an increase of 0.25 over this level. These thresholds may be changed by the user.)

For example, on chromosome one, covered by markers 1, 3, 2, 5, and 7, we see that a putative QTL has been detected on the distal end, with the most likely position (e.g. the position with the highest log-likelihood) falling in the 11.9 cM interval between markers 2 and 5. As you can see however the likelihood surface around this putative QTL is fairly flat, that is, the QTL's log-likelihoods do not change substantially with hypothesized position in the genome. (In this case, the log-likelihood falls by 1.0 only outside of a roughly 22 cM region containing the likelihood surface maximum). This result is typical of most data sets, which do not contain sufficient information to resolve QTL position more accurately.

Similarly, on the proximal end of chromosome two, we detect a putative QTL with a somewhat lower log-likelihood peak between markers 8-12. The 'confidence interval', outside of which the log-likelihood falls by 1.0, spans a 14 cM region around this point.

We may find it easier to view these LOD score plots graphically, rather than as text output. Do do this, we now use the "draw scan" command, which draws the LOD score curves for the previous scan in a PostScript file you can print out. These images are shown on the following pages.

```
---------------------------------------| 8-12 18.9 cM
0.0    -0.121  -0.016  11.7%    7.931   | ************************
2.0    -0.130  -0.014  13.6%    8.409   | **************************
4.0    -0.136  -0.011  15.1%    8.753   | ****************************
6.0    -0.140  -0.009  16.0%    8.926   | ****************************
8.0    -0.140  -0.009  16.3%    8.914   | ****************************
10.0   -0.138  -0.010  16.0%    8.723   | ***************************
12.0   -0.134  -0.013  15.2%    8.369   | **************************
14.0   -0.128  -0.016  13.9%    7.880   | ************************
16.0   -0.119  -0.020  12.2%    7.292   | **********************
18.0   -0.109  -0.022  10.3%    6.647   | *******************
---------------------------------------| 12-9 24.0 cM
0.0    -0.104  -0.022   9.5%    6.357   | ******************
2.0    -0.106  -0.022   9.8%    6.123   | *****************
4.0    -0.107  -0.021  10.0%    5.825   | ****************
6.0    -0.107  -0.020   9.9%    5.461   | **************
8.0    -0.105  -0.019   9.7%    5.032   | *************
10.0   -0.102  -0.018   9.1%    4.543   | ***********
12.0   -0.097  -0.017   8.3%    4.004   | *********
14.0   -0.090  -0.015   7.2%    3.434   | ******
16.0   -0.082  -0.013   5.9%    2.856   | ****
18.0   -0.072  -0.011   4.6%    2.301   | **
20.0   -0.062  -0.008   3.4%    1.798   |
22.0   -0.052  -0.006   2.3%    1.365   |
24.0   -0.042  -0.003   1.6%    1.010   |
---------------------------------------| 9-6 18.1 cM
0.0    -0.042  -0.003   1.5%    1.003   |
2.0    -0.045  -0.010   1.8%    1.078   |
4.0    -0.049  -0.018   2.2%    1.171   |
6.0    -0.052  -0.027   2.6%    1.275   |
8.0    -0.054  -0.034   3.0%    1.375   |
10.0   -0.055  -0.040   3.2%    1.457   |
12.0   -0.055  -0.044   3.3%    1.505   |
14.0   -0.053  -0.046   3.1%    1.515   |
16.0   -0.051  -0.045   2.8%    1.487   |
18.0   -0.048  -0.043   2.4%    1.428   |
---------------------------------------| 6-10 28.6 cM
0.0    -0.047  -0.043   2.4%    1.423   |
2.0    -0.048  -0.046   2.7%    1.378   |
4.0    -0.047  -0.048   2.9%    1.310   |
6.0    -0.046  -0.049   3.0%    1.217   |
8.0    -0.045  -0.048   2.9%    1.099   |
10.0   -0.043  -0.046   2.7%    0.960   |
12.0   -0.040  -0.041   2.4%    0.807   |
14.0   -0.036  -0.034   1.9%    0.651   |
16.0   -0.032  -0.025   1.4%    0.506   |
18.0   -0.028  -0.015   0.9%    0.383   |
20.0   -0.025  -0.006   0.6%    0.291   |
22.0   -0.022   0.002   0.4%    0.231   |
24.0   -0.020   0.009   0.4%    0.199   |
26.0   -0.018   0.015   0.3%    0.190   |
28.0   -0.017   0.020   0.3%    0.199   |
---------------------------------------|
Results have been stored as scan number 1
```

12> ***draw scan***

```
scan 1.1 saved in PostScript file 'scan1_1.ps'
```

Wiggle 1

Wiggle 2

We now discuss the interpretations of the QTL effects calculated for these putative QTLs.

As our data were collected from an F2 intercross, each individual will have one of three possible genotypes at any QTL: A/A, A/B, or B/B. MAPMAKER/QTL expresses the effect of these QTL genotypes on the trait using the classical "additive and dominance" method, with the QTL effects measured in terms of the amount that alleles derived from the B parent, individually or together, contribute to the phenotypes of individuals with an otherwise A/A background. (Note that the assignment of 'A' and 'B' alleles here follows the coding used in the raw data file, as described in the appendices.). In this method, the measured phenotype of individual number i in the data set is given by:

F2: $\text{Trait}_i = \text{Mean} + (\text{Weight} \times \text{Num}_i) + (\text{Dominance} \times \text{Het}_i) + \text{Noise}$

where:

Mean =	the mean value of the component of the trait not controlled by this QTL (in effect, the average trait value for A/A individuals).
Weight =	the additive component of the QTL B allele effect
Num_i =	the number of B alleles carried by individual number i, either 0, 1, or 2
Dominance =	the dominance component of the QTL B allele effect
Het_i =	1 if individual number i is an A/B heterozygote, and 0 otherwise
Noise =	variation in the trait not controlled by this QTL (a normal random variable)

Had this experiment been a backcross, rather than intercross, then only a weight term would apply, describing the effect of the heterozygote QTL genotype, A/B, over the homozygote A/A background. That is:

BC1: $\text{Trait}_i = \text{Mean} + (\text{Weight} \times \text{Num}_i) + \text{Noise}$

In our example, the peak of the likelihood surface on chromosome one falls between markers 2 and 5. The most likely QTL map, describing the most likely estimate of the QTL effects, is found at a point about 6 cM from marker number 2. This map indicates a hypothetical QTL with an additive effect of -0.088 and a dominance effect of -0.043 on each plant's measured trait. Thus, on average, an individual heterozygous at the QTL has a phenotype 0.131 units *lower* than a homozygote for the 'A' allele, while a typical allele 'B' homozygote has a phenotype 0.176 units lower than an 'A' homozygote.

Similarly, the putative QTL on chromosome two, at the log-likelihood maximum occurring 6.0 cM from marker 8, has an additive QTL effect of -0.140, and a a dominance effect of -0.009. Typical A/B heterozygotes have a phenotype 0.149 (= 0.140 + 0.009) units lower than typical A/A individuals, while B/B homozygotes have phenotypes 0.280 (= 0.140 $\times$ 2) units lower than the A/A homozygotes on average. Note that in both of these cases, the QTL alleles inherited from the B parents give rise to lower measured phenotypes than those from the A parent.

It is worth noting that the effects described by QTL maps in cases with no missing data (e.g. all progeny are informative and the putative QTL is located directly at a genetic marker) are literally identical to those produced using standard regression analysis (a.k.a. single factor analysis). Moreover, the "variance explained" by MAPMAKER/QTL will be the same as the R^2 value calculated by linear regression.

A DIGRESSION: We must again emphasize that the analysis performed by MAPMAKER/QTL, (i.e. fitting the model shown on page II-12 to the available data) is premised on the assumption that the phenotype can be explained by a single QTL located at the position being tested together with normally distributed noise. Obviously, the analysis can be affected by deviations from these simplifying assumptions. (These same assumptions also apply to simple regression analyses.)

One possible problem could be the presence of two linked QTLs. In this case, a "scan" analysis will typically show two nearby peaks, each near one of the QTLs. Because the assumption of a single QTL may be incorrect, the analysis for each putative QTL may be affected by the other. Consequently, the QTL positions may appear somewhat shifted, the weights incorrect, and the LOD score increased.

Two peaks do not necessarily imply two QTLs, however. Suppose that there was a single QTL lying midway between markers 1 and 3. The LOD score would be highest in the middle of interval 1-3, would drop at locus 3 (owing to the data from individuals in whom crossovers resulted in genotypes which did not predict their phenotypes well), and could rise somewhat in middle of interval 3-2. Although it may seem odd at first, that the likelihood should rise in interval 3-2, it is mathematically correct. The hypothesis that a QTL lies in interval 3-2 actually does fit the data better than the hypothesis that a QTL exists very near locus 3 (although neither hypothesis is correct) because some portion of the discrepancies between phenotype and genotype can be attributed to double crossovers in the former case but not the latter.

Alternatively, two peaks may be observed if a marker has been mistyped in a few individuals. Although there may be strong evidence for a QTL nearby, the LOD score may drop near the marker due to genotype errors again providing data which conflict with the hypothesis. If the LOD score is very different at two nearby loci (say within 5 cM), this is almost certainly the case! Whenever the LOD score falls precipitously near a marker, one should always consider the possibility of data errors, usually either mistyped loci or errors in the linkage map order.

Because the existence of two likelihood peaks can indicate many different things, whenever we see two high peaks within about 30-40 cM with a well-defined valley in between, we thus consider the above possibilities and analyze the situation carefully. To test for linked QTLs, we perform an analysis in which one putative QTL is declared at a fixed location (at one of the peaks), and a scan is performed to look for a second QTL given the presence of the first QTL (fitting a two-QTL model to the data, as described later in this tutorial and in more detail in the reference section of this manual). A second QTL is likely to be present if there is still a substantial likelihood rise when controlling for the first QTL. If instead, once we have controlled for the presence of one QTL, there is no significant evidence for a second one, then the two QTL hypothesis is considered unlikely.

Testing for data errors is more difficult. If the marker in question has flanking markers nearby, we might choose to simply re-scan the interval omitting all genotypic data for that marker. For example, to scan the interval between markers 3 and 5, omitting data for maker 2, we could use the sequence "[3|5]", as described in the reference section of this manual.

Automatically Summarizing the Results of a Genome Scan.

For the purposes of illustration, we examined the MAPMAKER/QTL "scan" output by hand, finding the maximally likely QTL positions, effects, and confidence intervals. In analyses of larger data sets however, it can be time consuming to browse the "scan" output by hand in this manner. MAPMAKER/QTL provides a facility for automatically summarizing these data, the "show peaks" command.

This command, by default, instructs MAPMAKER/QTL to search the results of the genome "scan" executed immediately beforehand, and to report various pieces of information about the maximally likely QTL maps. Again by default, the "show peaks" command ignores log-likelihood maxima which fall below a specified LOD threshold, usually 2.0, and considers neighboring peaks not separated by a log-likelihood drop of 1.0 to be local maxima of the same peak. One nice feature of this command is that it displays more information about the QTL map than the "scan" command itself does, including:

μ= the mean value of the component of the trait not controlled by this QTL
σ^2= the variance of the environmental noise in the trait not controlled by this QTL
χ^2= an alternate measure of the significance of the QTL map data, where: $\chi^2 = 2 \times (\ln 10) \times \mathrm{LOD}$, a value which is asymptotically distributed as χ^2 for large populations.

Another parameter calculated by MAPMAKER/QTL in this case is the "variance explained" by each QTL. This number is simply the percentage decrease in the estimated σ when the QTL is allowed to control a portion of the trait, as opposed to when no QTLs are assumed (and thus, when all of the variation in the population is attributed to noise). This number is equivalent to the R^2 value calculated by linear regression.

For each likelihood peak found, MAPMAKER/QTL determines the map positions of the boundaries of the 10:1 "confidence interval" surrounding the peak. These boundaries indicate the width of the peak before the likelihood drops 10-fold (a log-likelihood drop of 1.0), and indicate the range of QTL positions which can best explain the observed data.

Here, we invoke the "show peaks" command to summarize our previous calculation. As we can see, MAPMAKER/QTL simply described the same peaks that we had found by eye. However, because this tells us the values of μ and σ^2, we can simply calculate the expected phenotypic means of the various sub-populations within our data set, as well as quantify the non-genetic "noise" in the trait.

For the putative QTL on chromosome one, with a log-likelihood of 4.97, we have:

$\mu_{A/A}$= μ= 0.805	σ^2= 0.052
$\mu_{A/B}$= μ + Weight + Dominance = 0.674	σ= 0.228
$\mu_{B/B}$= $\mu + 2 \times$ Weight = 0.630	variance explained= 8.4%

For the putative QTL on chromosome two:

$\mu_{A/A}$= μ= 0.866	σ^2= 0.047
$\mu_{A/B}$= μ + Weight + Dominance = 0.717	σ= 0.217
$\mu_{B/B}$= $\mu + 2 \times$ Weight = 0.587	variance explained= 16.0%

13> ***show peaks***

```
LOD score peaks for scan 1.1 of trait 2 (logwt).
Sequence: [all]
No fixed-QTLs.
Scanned QTL genetics are unconstrained.
Peak Threshold: 2.00  Falloff: -1.00

==============================================================
QTL-Map for peak 1:
Confidence Interval:  Left Boundary= 3-2 + 4.0
                     Right Boundary= 5-7 + 10.0

INTERVAL   LENGTH  QTL-POS  GENETICS    WEIGHT  DOMINANCE
2-5        11.9    6.0      free       -0.0875 -0.0434

chi^2= 22.884 (2  D.F.)        log-likelihood= 4.97
mean= 0.805   sigma^2= 0.052   variance-explained= 8.4 %
==============================================================
QTL-Map for peak 2:
Confidence Interval:  Left Boundary= 11-8 + 2.0
                     Right Boundary= 8-12 + 16.0

INTERVAL   LENGTH  QTL-POS  GENETICS    WEIGHT  DOMINANCE
8-12       18.9    6.0      free       -0.1396 -0.0093

chi^2= 41.106 (2  D.F.)        log-likelihood= 8.93
mean= 0.866   sigma^2= 0.047   variance-explained= 16.0%
==============================================================
```

Other Methods for Displaying QTL Maps

The "show peaks" command provides a convenient method for viewing the QTL map data calculated as part of a genome scan. There are situations where we instead want to see the analyses for QTLs in other locations in the genome, to see analyses allowing multiple QTLs together, or analyses for QTLs under other particular assumptions. The remaining chapters in this tutorial will briefly cover examples of such analyses.

As a simple example, we start by calculating and displaying the QTL map for the putative QTL located 6 cM to the right of marker 8 on chromosome two. To do this, we simply type:

```
sequence [8+6]
map
```

as shown on the facing page.

Here, we briefly note some features of MAPMAKER/QTL's "sequence" command. The location (or possible locations) for one QTL may be listed inside square brackets, where we use the syntax:

sequence [*marker* + *distance*]

to specify a QTL located at a particular point on the map. The marker may be specified by name or number, and the distance may be given by recombination fraction (assumed if it is less than 0.50), or in centimorgans (assumed if it is greater than or equal to 0.50). To show multiple QTL maps for more than one position in the genome, we simply list multiple possible locations within the brackets. For example, the command:

```
sequence [8+6 2+6]
map
```

calculated the QTL maps for the two genome positions of the likelihood peaks we examined earlier. (Note that this sequence does <u>not</u> specify the situation where two QTLs are allowed to work together to control a trait, a subject we will discuss shortly).

If we want MAPMAKER/QTL to find the most likely QTL position in the interval following a marker, we can simply omit the distance specification. For example, to find the maximally likely QTL map anywhere in the interval following marker 2 (which happens to be the interval between markers 2 and 5) we type:

```
sequence [2]
map
```

as shown. This feature operates somewhat differently than the "scan" example shown earlier. In this case, only the interval immediately between two flanking genetic markers can be searched, although the resolution in QTL position is much greater (owing to the fact that this search may be performed more efficiently). In fact, this test shows us that the maximally likely position for the QTL is exactly 5.9 cM distal of marker 2. (Given the flat likelihood surface however, this degree of accuracy is not necessarily of great value).

```
13> sequence  [8+6]
The sequence is now '[8+6]'

14> map
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  WEIGHT  DOMINANCE
8-12       18.9    6.0     -0.1396 -0.0093

chi^2= 41.106 (2  D.F.)        log-likelihood= 8.93
mean= 0.866   sigma^2= 0.047   variance-explained= 16.0%
==============================================================
```

```
15> sequence  [8+6  2+8]
The sequence is now '[8+6 2+8]'

16> map
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  WEIGHT  DOMINANCE
8-12       18.9    6.0     -0.1396 -0.0093

chi^2= 41.106 (2  D.F.)        log-likelihood= 8.93
mean= 0.866   sigma^2= 0.047   variance-explained= 16.0%
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  WEIGHT  DOMINANCE
2-5        11.9    8.0     -0.0884 -0.0406

chi^2= 22.656 (2  D.F.)        log-likelihood= 4.92
mean= 0.806   sigma^2= 0.052   variance-explained= 8.2 %
==============================================================
```

```
17> sequence  [2]
The interval-list is '[2]'

18> map
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  WEIGHT  DOMINANCE
2-5        11.9    5.9     -0.0874 -0.0435

chi^2= 22.883 (2  D.F.)        log-likelihood= 4.97
mean= 0.805   sigma^2= 0.052   variance-explained= 8.4 %
==============================================================
```

One method to instruct MAPAKER/QTL to search multiple intervals for QTLs in this manner might use:

```
sequence [1-7]
map
```

which displays four QTL maps, each one showing the maximally likely QTL position in the four intervals between markers 1 and 7 (that is, the interval between 1 and 3, the interval between 3 and 2, the interval between 2 and 5, and the interval between 5 and 7). Thus, this is equivalent to typing:

```
sequence [1 3 2 5]
map
```

or:

```
sequence [1]
map
sequence [3]
map
sequence [2]
map
sequence [5]
map
```

From these four maps, we can easily pick out the QTL position with the largest log-likelihood. However, as opposed to using the "scan" command, we do not get to see the likelihood surface itself. (These data are not displayed simply because they are not fully computed along the likelihood surface). This can make it hard to discriminate multiple QTLs or observe anomalies in the results indicating problems within your data set.

```
19> sequence  [1-7]
The interval-list is '[1-7]'

20> map
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  WEIGHT  DOMINANCE
1-3        4.2     0.2     -0.0346 -0.0701

chi^2= 14.092 (2  D.F.)        log-likelihood= 3.06
mean= 0.774   sigma^2= 0.054   variance-explained= 4.7 %
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  WEIGHT  DOMINANCE
3-2        15.0    13.2    -0.0756 -0.0483

chi^2= 20.526 (2  D.F.)        log-likelihood= 4.46
mean= 0.793   sigma^2= 0.052   variance-explained= 7.2 %
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  WEIGHT  DOMINANCE
2-5        11.9    5.9     -0.0874 -0.0435

chi^2= 22.883 (2  D.F.)        log-likelihood= 4.97
mean= 0.805   sigma^2= 0.052   variance-explained= 8.4 %
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  WEIGHT  DOMINANCE
5-7        12.2    1.3     -0.0838 -0.0359

chi^2= 20.568 (2  D.F.)        log-likelihood= 4.47
mean= 0.803   sigma^2= 0.052   variance-explained= 7.1 %
==============================================================
```

Fitting QTLs to Particular Genetic Models (for F2 Intercross Data)

We have now mapped two putative QTLs which affect a quantitative trait in an F2 intercross. For each QTL, we estimated two independent parameters, the weight and dominance for the QTL, which together describe the effects of the three possible QTL genotypes on an individual's phenotype. Because this model of QTL action does not require the any fixed model of QTL-phenotype interaction, we call this a "free" model of QTL effect.

The effect of the QTL alleles may however follow a Mendelian model, either dominant, recessive, or additive. MAPMAKER/QTL provides a number of features for estimating QTL effects under these models and determining the likely mode of inheritance of each QTL allele.

By eye, we can simply observe in our previous analyses that the QTL on chromosome two has a maximum likelihood dominance effect of nearly zero. From this, it seems that the QTL affects the phenotype in a additive fashion in this cross (meaning that each B allele an individual inherits adds a fixed amount to its phenotype). Suppose however that we wanted to test whether this QTL's effect is also consistent with a dominant/recessive model, where (for example) we'll take the A allele as dominant and the B allele as recessive.

We can test this hypothesis quantitatively. Recall that in the previous section, we simply calculated the maximum likelihood map of this QTL by typing:

```
sequence [8+6]
map
```

One way to test our hypothesis would be to recalculate the QTL map, this time forcing the QTL's B allele to exhibit strictly recessive inheritance (recall that we always express our QTL effects in terms of the B allele's contribution). In terms of the F2 trait model we discussed earlier, where:

$$\text{Trait}_i = \text{Mean} + (\text{Weight} \times \text{Num}_i) + (\text{Dominance} \times \text{Het}_i) + \text{Noise}$$

we can easily see that B allele recessive inheritance is equivalent to requiring that:

$$\text{Weight} = -\text{Dominance}$$

so that only individuals with two B alleles exhibit a genetic effect (where the effect will be $2 \times$ Weight). We can calculate the QTL map under this assumption by typing:

```
sequence [8+6:recessive]
map
```

as shown. We see that the log-likelihood of this QTL map is 5.40, as opposed the 8.93 log-likelihood of the free map. Thus, the likelihood ratio is 3388:1 favoring the free model over the recessive one [$3388=10^{(8.93-5.40)}$], and we can confidently rule out the possibility that this QTL displays recessive genetics. (Again, keep in mind that we always use "recessive" and "dominant" to refer to the behavior of the 'B' parent's alleles.)

```
21> sequence  [8+6:recessive]
The sequence is now '[8+6:recessive]'

22> map
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  GENETICS    WEIGHT  DOMINANCE
8-12       18.9    6.0      recessive  -0.0936  0.0936

chi^2= 24.853 (2  D.F.)        log-likelihood= 5.40
mean= 0.766   sigma^2= 0.050   variance-explained= 11.2%
==============================================================
```

Because the recessive QTL map is further constrained than the free QTL map, its likelihood must be less than, or at most, equal to, that of the free map (e.g. the constrained map can certainly not explain the observed data any better than the unconstrained map). In other words, we can never strictly rule out free models, we can only compare to them the constrained models of Mendelian inheritance, and declare particular constraints unlikely.

We can quickly test all of the possible modes of inheritance of our chromosome one QTL. This time we type:

```
sequence [2+6:try]
map
```

instructing MAPMAKER/QTL to calculate four QTL maps in succession: one each for free, additive, dominant, and recessive QTL genetics. In this case, the QTL map assuming recessive genetics has a log-likelihood of 1.10, as opposed the 4.97 log-likelihood of the free map. Thus, we can confidently rule out the possibility that this QTL displays recessive genetics. The additive and dominant maps however both exhibit very small log-likelihood decreases when compared to the free map (0.34 and 0.15, respectively). If, as we did with QTL position, we continue to require a 10:1 likelihood ratio to reject any hypothesis, then we can not rule out either of these hypotheses as possible explanations of the observed data.

Precisely speaking, the four genetic models MAPMAKER/QTL can test are:

Free: A/A effect = 0
$\rightarrow$ $\text{Trait}_i = \text{Mean} + (\text{Weight} \times \text{Num}_i) + (\text{Dominance} \times \text{Het}_i) + \text{Noise}$

Additive: A/A effect = 0 and B/B effect = 2 $\times$ A/B effect
$\rightarrow$ $\text{Trait}_i = \text{Mean} + (\text{Weight} \times \text{Num}_i) + (0 \times \text{Het}_i) + \text{Noise}$
$\rightarrow$ dominance = 0

Dominant A/A effect = 0 and B/B effect = A/B effect
$\rightarrow$ $\text{Trait}_i = \text{Mean} + (\text{Weight} \times \text{Num}_i) + (\text{Weight} \times \text{Het}_i) + \text{Noise}$
$\rightarrow$ dominance = weight

Recessive A/A effect = 0 and A/B effect = A/A effect
$\rightarrow$ $\text{Trait}_i = \text{Mean} + (\text{Weight} \times \text{Num}_i) + (-\text{Weight} \times \text{Het}_i) + \text{Noise}$
$\rightarrow$ dominance = -weight

It is also possible to "scan" the genome for QTLs simultaneously testing each of the genetic models at each point in the genome. For example, to scan the two intervals between markers 3 and 7 in this manner, we could type:

```
sequence [3-7:try]
scan
```

This sequence instructs MAPMAKER/QTL to simply perform the appropriate scan four times, once each for the four genetic models. These results are shown on the following pages.

```
23> sequence  [2+6:try]
The sequence is now '[2+6:try]'

24> map

==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  GENETICS    WEIGHT  DOMINANCE
2-5        11.9    6.0      free       -0.0875 -0.0434

chi^2= 22.884 (2  D.F.)        log-likelihood= 4.97
mean= 0.805   sigma^2= 0.052   variance-explained= 8.4 %
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  GENETICS    WEIGHT  DOMINANCE
2-5        11.9    6.0      dominant   -0.0695 -0.0695

chi^2= 22.192 (2  D.F.)        log-likelihood= 4.82
mean= 0.805   sigma^2= 0.052   variance-explained= 8.2 %
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  GENETICS    WEIGHT  DOMINANCE
2-5        11.9    6.0      recessive  -0.0574  0.0574

chi^2= 5.048  (2  D.F.)        log-likelihood= 1.10
mean= 0.733   sigma^2= 0.055   variance-explained= 2.2 %
==============================================================
QTL map for trait 2 (logwt):

INTERVAL   LENGTH  QTL-POS  GENETICS    WEIGHT  DOMINANCE
2-5        11.9    6.0      additive   -0.1043  0.0000

chi^2= 21.285 (2  D.F.)        log-likelihood= 4.62
mean= 0.795   sigma^2= 0.052   variance-explained= 7.9 %
==============================================================
```

```
25> sequence  [3-7:try]
The sequence is now '[3-7:try]'

26> scan
QTL maps for trait 2 (logwt):
Sequence: [3-7:try]
LOD threshold: 2.00  Scale: 0.25 per '*'
```

```
Scanned QTL genetics are free.

POS     WEIGHT  DOM     %VAR  LOG-LIKE |
---------------------------------------| 3-2 15.0 cM
0.0     -0.068  -0.021   3.8%    2.440 | **
2.0     -0.076  -0.023   4.8%    2.871 | ****
4.0     -0.080  -0.029   5.8%    3.288 | ******
6.0     -0.081  -0.036   6.6%    3.664 | *******
8.0     -0.081  -0.041   7.2%    3.981 | ********
10.0    -0.079  -0.045   7.4%    4.225 | *********
12.0    -0.077  -0.048   7.4%    4.392 | **********
14.0    -0.075  -0.048   7.1%    4.481 | **********
---------------------------------------| 2-5 11.9 cM
0.0     -0.073  -0.047   6.8%    4.497 | **********
2.0     -0.079  -0.048   7.7%    4.752 | ************
4.0     -0.084  -0.046   8.2%    4.910 | ************
6.0     -0.088  -0.043   8.4%    4.966 | ************
8.0     -0.088  -0.041   8.2%    4.917 | ************
10.0    -0.087  -0.037   7.7%    4.755 | ************
---------------------------------------| 5-7 12.2 cM
0.0     -0.083  -0.034   6.8%    4.498 | **********
2.0     -0.084  -0.037   7.2%    4.424 | **********
4.0     -0.084  -0.038   7.3%    4.233 | *********
6.0     -0.081  -0.038   6.9%    3.928 | ********
8.0     -0.077  -0.036   6.2%    3.522 | *******
10.0    -0.071  -0.032   5.2%    3.043 | *****
12.0    -0.063  -0.026   4.0%    2.532 | ***
---------------------------------------|

Scanned QTL genetics are constrained to be: dominant

POS     WEIGHT  DOM     %VAR  LOG-LIKE |
---------------------------------------| 3-2 15.0 cM
0.0     -0.045  -0.045   3.6%    2.380 | **
2.0     -0.051  -0.051   4.7%    2.796 | ****
4.0     -0.057  -0.057   5.7%    3.210 | *****
6.0     -0.061  -0.061   6.5%    3.591 | *******
8.0     -0.064  -0.064   7.1%    3.915 | ********
10.0    -0.065  -0.065   7.4%    4.166 | *********
12.0    -0.065  -0.065   7.3%    4.337 | **********
14.0    -0.063  -0.063   7.0%    4.425 | **********
---------------------------------------| 2-5 11.9 cM
0.0     -0.062  -0.062   6.8%    4.440 | **********
2.0     -0.066  -0.066   7.6%    4.673 | ***********
4.0     -0.069  -0.069   8.0%    4.797 | ************
6.0     -0.069  -0.069   8.2%    4.816 | ************
8.0     -0.069  -0.069   8.0%    4.729 | ***********
10.0    -0.067  -0.067   7.4%    4.531 | ***********
---------------------------------------| 5-7 12.2 cM
0.0     -0.063  -0.063   6.4%    4.235 | *********
2.0     -0.065  -0.065   6.9%    4.211 | *********
4.0     -0.065  -0.065   7.0%    4.057 | *********
6.0     -0.064  -0.064   6.7%    3.779 | ********
8.0     -0.060  -0.060   6.1%    3.392 | ******
10.0    -0.055  -0.055   5.0%    2.922 | ****
12.0    -0.048  -0.048   3.8%    2.415 | **
---------------------------------------|
```

```
Scanned QTL genetics are constrained to be: recessive

POS     WEIGHT  DOM     %VAR  LOG-LIKE |
---------------------------------------| 3-2 15.0 cM
0.0     -0.049  0.049   0.5%    0.256  |
2.0     -0.063  0.063   1.0%    0.423  |
4.0     -0.067  0.067   1.4%    0.555  |
6.0     -0.065  0.065   1.6%    0.639  |
8.0     -0.061  0.061   1.6%    0.681  |
10.0    -0.056  0.056   1.4%    0.695  |
12.0    -0.050  0.050   1.3%    0.688  |
14.0    -0.045  0.045   1.1%    0.666  |
---------------------------------------| 2-5 11.9 cM
0.0     -0.043  0.043   1.0%    0.652  |
2.0     -0.050  0.050   1.5%    0.825  |
4.0     -0.055  0.055   1.9%    0.981  |
6.0     -0.057  0.057   2.2%    1.097  |
8.0     -0.057  0.057   2.2%    1.157  |
10.0    -0.054  0.054   2.1%    1.153  |
---------------------------------------| 5-7 12.2 cM
0.0     -0.050  0.050   1.8%    1.095  |
2.0     -0.052  0.052   1.9%    1.078  |
4.0     -0.053  0.053   2.0%    1.029  |
6.0     -0.052  0.052   1.9%    0.945  |
8.0     -0.050  0.050   1.7%    0.830  |
10.0    -0.046  0.046   1.3%    0.690  |
12.0    -0.039  0.039   0.9%    0.538  |
---------------------------------------|

Scanned QTL genetics are constrained to be: additive

POS     WEIGHT  DOM     %VAR  LOG-LIKE |
---------------------------------------| 3-2 15.0 cM
0.0     -0.083  0.000   3.8%    2.405  | **
2.0     -0.092  0.000   4.9%    2.828  | ****
4.0     -0.099  0.000   5.8%    3.218  | *****
6.0     -0.103  0.000   6.5%    3.542  | *******
8.0     -0.104  0.000   6.9%    3.784  | ********
10.0    -0.103  0.000   7.0%    3.941  | ********
12.0    -0.100  0.000   6.8%    4.018  | *********
14.0    -0.095  0.000   6.4%    4.029  | *********
---------------------------------------| 2-5 11.9 cM
0.0     -0.093  0.000   6.2%    4.015  | *********
2.0     -0.099  0.000   7.0%    4.310  | **********
4.0     -0.103  0.000   7.7%    4.519  | ***********
6.0     -0.104  0.000   7.9%    4.620  | ***********
8.0     -0.103  0.000   7.8%    4.602  | ***********
10.0    -0.100  0.000   7.3%    4.459  | **********
---------------------------------------| 5-7 12.2 cM
0.0     -0.093  0.000   6.5%    4.221  | *********
2.0     -0.096  0.000   6.8%    4.133  | *********
4.0     -0.097  0.000   6.8%    3.946  | ********
6.0     -0.095  0.000   6.5%    3.661  | *******
8.0     -0.090  0.000   5.8%    3.290  | ******
10.0    -0.083  0.000   4.9%    2.856  | ****
12.0    -0.074  0.000   3.8%    2.394  | **
---------------------------------------|
Results have been stored as scan numbers 3.1-3.4.
```

The output from this command can be examined by eye to compare log-likelihoods, as we did previously. More conveniently however, we can have MAPMAKER/QTL automatically summarize these results using the "show trys" command. This command (by default) analyzes the data produced by the immediately previous "scan", which must have been generated from a sequence using the ":try" notation. The four individual scans are summarized at each data point, with MAPMAKER/QTL displaying the log-likelihood and other characteristics of the free maps at each point in the genome, as well as the log-likelihood change (which of course is a negative number, representing a likelihood decrease) for the additive, dominant, and recessive maps. In this output, an asterisk is displayed for each point where the free map's log-likelihood exceeds our default threshold of 3.0, making it easy to spot likelihood peaks.

As another convenient way of viewing these results, we also wil graphically draw the results as a Postscript graphic image. This graphic image is included as the last page of this tutorial (flip ahead to see it).

27> ***show trys***

Test genetics results for trait 2 (logwt).
Sequence: [3-7:try]
No fixed-QTLs.
Scan numbers: 3.1-3.4 Threshold: 2.00

GENETICS:			FREE		DOMINANT		RECESSIVE		ADDITIVE		
POS	WEIGHT	DOM	%VAR	LOG LIKE	%VAR	LIKE DIFF	%VAR	LIKE DIFF	%VAR	LIKE DIFF	
interval= 3-2 length= 15.0 cM											
0.0	-0.068	-0.021	3.8	2.44	3.6	-0.06	0.5	-2.18	3.8	-0.04	*
2.0	-0.076	-0.023	4.8	2.87	4.7	-0.08	1.0	-2.45	4.9	-0.04	*
4.0	-0.080	-0.029	5.8	3.29	5.7	-0.08	1.4	-2.73	5.8	-0.07	*
6.0	-0.081	-0.036	6.6	3.66	6.5	-0.07	1.6	-3.03	6.5	-0.12	*
8.0	-0.081	-0.041	7.2	3.98	7.1	-0.07	1.6	-3.30	6.9	-0.20	*
10.0	-0.079	-0.045	7.4	4.23	7.4	-0.06	1.4	-3.53	7.0	-0.28	*
12.0	-0.077	-0.048	7.4	4.39	7.3	-0.05	1.3	-3.70	6.8	-0.37	*
14.0	-0.075	-0.048	7.1	4.48	7.0	-0.06	1.1	-3.81	6.4	-0.45	*
interval= 2-5 length= 11.9 cM											
0.0	-0.073	-0.047	6.8	4.50	6.8	-0.06	1.0	-3.85	6.2	-0.48	*
2.0	-0.079	-0.048	7.7	4.75	7.6	-0.08	1.5	-3.93	7.0	-0.44	*
4.0	-0.084	-0.046	8.2	4.91	8.0	-0.11	1.9	-3.93	7.7	-0.39	*
6.0	-0.088	-0.043	8.4	4.97	8.2	-0.15	2.2	-3.87	7.9	-0.35	*
8.0	-0.088	-0.041	8.2	4.92	8.0	-0.19	2.2	-3.76	7.8	-0.32	*
10.0	-0.087	-0.037	7.7	4.75	7.4	-0.22	2.1	-3.60	7.3	-0.30	*
interval= 5-7 length= 12.2 cM											
0.0	-0.083	-0.034	6.8	4.50	6.4	-0.26	1.8	-3.40	6.5	-0.28	*
2.0	-0.084	-0.037	7.2	4.42	6.9	-0.21	1.9	-3.35	6.8	-0.29	*
4.0	-0.084	-0.038	7.3	4.23	7.0	-0.18	2.0	-3.20	6.8	-0.29	*
6.0	-0.081	-0.038	6.9	3.93	6.7	-0.15	1.9	-2.98	6.5	-0.27	*
8.0	-0.077	-0.036	6.2	3.52	6.1	-0.13	1.7	-2.69	5.8	-0.23	*
10.0	-0.071	-0.032	5.2	3.04	5.0	-0.12	1.3	-2.35	4.9	-0.19	*
12.0	-0.063	-0.026	4.0	2.53	3.8	-0.12	0.9	-1.99	3.8	-0.14	*

28> ***draw scan 3***
scan 3.x saved in PostScript file 'scan3_x.ps'

Fitting Multiple QTLs Simultaneously

In each of the previous analyses, we have examined QTL maps where one QTL at a time is allowed to explain the genetic component of the variation in our trait. In many cases however, multiple QTL's act together to affect the trait. Moreover, examining multiple QTL's simultaneously can extend the sensitivity of QTL mapping, simply by reducing the unexplained noise that must be accounted for. Note that MAPMAKER/QTL is currently only able to consider cases where independent QTLs act additively with each other.

Here, we perform a simple test, where we see how well the QTL map allowing both QTLs to explain the trait together fits the observed data. In other words, we will find the most likely parameters for the two-QTL model:

$$\begin{aligned}\text{Trait}_i = \ & \text{Mean} + (\text{Weight}_1 \times \text{Num}_{1i}) + (\text{Dominance}_1 \times \text{Het}_{1i}) + \\ & (\text{Weight}_2 \times \text{Num}_{2i}) + (\text{Dominance}_2 \times \text{Het}_{2i}) + \text{Noise}\end{aligned}$$

To limit the amount of searching that needs to be done, we will declare the position of these QTLs to be fixed precisely at the peaks we found in our previous genome scan. (There is no requirement to do this: you could instead let MAPAMAKER/QTL search for the most likely QTL positions). We will not however constrain the genetic behavior of either QTL (e.g. both will have free genetics), although we easily could.

To perform this test, we need only to enter a new sequence which specifies both of these qtls, and then again use the map command:

```
sequence [2+6][8+6]
map
```

Notice that we have placed the two QTL positions in separate brackets to indicate that they should be considered in one model together. Compare this to typing:

```
sequence [2+6 8+6]
map
```

which specifies two models each with one QTL!

In this test, the log-likelihood of the two QTL map is substantially higher than the log-likelihoods of the individual QTL's maps, and slightly exceeds the sum of their individual log-likelihoods (16.15, versus 8.93 + 4.97 = 13.9). This is expected if the two QTLs in fact act additively with each other and explain independent portions of the variation in the trait, assumptions of the model shown above. Cases where the likelihood of the map allowing multiple QTLs is much lower than the sum of the QTL's independent likelihoods may indicate epistasis, partially spurious results, non-normal trait data, errors within the data set, or other conditions. Note that QTL maps which allow multiple QTLs have more degrees of freedom than QTL maps allowing single QTLs, and the appropriate log-likelihood threshold for multiple QTL maps must be increased.

Because the two QTL map allows each locus to control its fraction of the variance while at the same time we estimate the effect of the other, the amount of unexplained noise in the model is reduced. Thus, we might consider these estimates of the QTL effects to be slightly more accurate than our previous ones.

```
28> sequence  [2+6][8+6]
The sequence is now '[2+6][8+6]'

29> map
==============================================================
QTL map for trait 2 (logwt):

INTERVALS  LENGTH  QTL-POS  WEIGHT  DOMINANCE
2-5        11.9    6.0     -0.1083 -0.0301
8-12       18.9    6.0     -0.1496 -0.0151

chi^2= 74.385 (4  D.F.)        log-likelihood= 16.15
mean= 0.971   sigma^2= 0.041   variance-explained= 27.0%
==============================================================
```

With more sophisticated features of the "sequence", "scan", and "map" commands, MAPMAKER/QTL provides a number of other methods for performing tests such as these. We unfortunately could not attempt to cover them all in this simple tutorial, however. By way of introduction, you can have the program:

- Perform a search similar to a two dimensional "scan", allowing pairs of QTLs at pairwise interval combinations, hoping that this might further refine the data indicating the genes' positions. Unfortunately, this is a very time consuming process, only feasible on faster computers.

- Pick one very likely QTL and "fix" it in place. Then, you can re-scan the genome, searching for other QTLs which, together with the first, provide likely genetic explanations of the data. In this case, the increased sensitivity from the reduced noise may help you detect new QTL's (some of which you would not have otherwise), and avoid some spurious QTLs. This method may also be used to determine whether multiple neighboring likelihood surface peaks are caused by one or many QTLs.

- Examine the effects of QTLs in various pairwise combinations, hoping to see signs of pairwise (epistatic) effects.

- Save the results of many "scan" commands for later examination in detail.

- Use non-genetic variables, including trait values, as components of models in QTL map calculations. (Examples of this may be found in the reference section).

For further information on these and the other capabilities of MAPMAKER/QTL, consult the reference section of this manual.

Following these tests, we make a note of the results in our output transcript file using MAPMAKER/QTL's "comment" command.

With our goal of mapping QTLs affecting our weight trait on these three chromosomes finished, we now exit MAPMAKER/QTL using the "quit" command. We then printed out the transcript of this session (the file "sample.out"), from which this tutorial was generated.

```
30> comment We mapped and characterized two QTLs.
```

```
31> quit
Do you really want to quit? [no] yes
Now saving sample.qtls...
Now saving sample.traits...

Goodbye...
```

Graphic