

User Manual for QTLMapper Version 1.0

A Computer Software for Mapping Quantitative Trait Loci (QTLs) with Main Effects, Epistatic Effects and QTL × Environment Interactions

Copyright 1999

by

Daolong Wang ¹, Jun Zhu ², Zhikang Li ³, Andrew H. Paterson ⁴

- 1 Biometrics Section, Anhui Agricultural University, Hefei, Anhui 230036, China
- 2 Department of Agronomy, Zhejiang University, Hangzhou 310029, China.
- 3 Plant Genetics, Breeding, and Biochemistry Department, IRRI, 1099 Manila, Philippines
- 4 Department of Soil and Crop Sciences, Texas A&M University, College Station, TX77843, USA

This software can be freely redistributed with non-commercial purpose and under the condition of no changes to the software and its related documents.

No warranty of any sort is provided for use of the software.



System Requirements

Hardware: PC 486 or higher; 16MB RAM or more; at least 2-MB free space on hard disk.

Operation System: Microsoft Windows 95/98™.

Installing QTLMapper 1.0


All related files for QTLMapper 1.0 were compressed into QTLMapr.exe, a self-extracted file.

To install the software,

- (1) Run QTLMapr.exe under *Windows Explorer*, the file manager of Windows 95/98™, then a WinZip Self-Extractor dialogue box will appear;
- (2) In the Unzip to edit box, input the name of the hard disk (e.g. C:) on which the software is going to be installed, then click on Unzip button. All the files for QTLMapper 1.0 will be automatically extracted into a directory named as “QTLMapper” on the specified hard disk (e.g. C:\QTLMapper).


Composition of QTLMapper 1.0

QTLMapper 1.0 is composed of Qtlmaper.exe, Qtlmaper.hlp, Qtlmaper.ini, and some sample files. Qtlmaper.exe, an executable file, is the core of QTLMapper 1.0. It performs all the operations necessary for QTL mapping analysis. Qtlmaper.hlp is a text file that provides on-line help. Qtlmaper.ini, also a text file, is used for initialization of settings when Qtlmaper.exe is started (see a later section for more details). If Qtlmaper.ini does not exist, Qtlmaper.exe will automatically create one that contains default settings. Sample input files are provided for demonstrating the format required by QTLMapper 1.0.

 Warning: Do not edit Qtlmaper.hlp and Qtlmaper.ini!

Starting and Exiting QTLMapper 1.0

Method of Starting: QTLMapper 1.0 can be started in three ways.

- (1) **Run from *Windows Explorer* of Windows 95/98** . In *Windows Explorer*, find the folder (e.g. C:\QTLMapper) containing the files of QTLMapper 1.0. Double click on Qtlmaper.exe.
- (2) **Run from Start menu.** Create a shortcut in Start menu of Windows 95/98™ for QTLMapper 1.0. If you do not know how to do so, please consult a computer expert. Run the shortcut from Start menu, just like running any other Windows applications.
- (3) **Run under MS-DOS prompt.** First run MS-DOS prompt from Start menu, then take the

following steps (assuming that the folder containing QTLMapper 1.0 is C:\QTLMapper):

```
C:\Windows>CD C:\QTLMapper ↵
```

```
C:\QTLMapper>QTLMapper ↵
```

Running QTLMapper 1.0 under full-screen mode will provide a better user interface. File conagent.pif is given for this purpose. But sometimes conagent.pif may cause problem. In this case, it needs to be renamed or deleted. If the software is running under full screen mode, to switch to other Windows applications, press Ctrl+Esc.

Method of Exiting: To exit QTLMapper 1.0, press Esc key one or more times, depending on the running status of the software. However, it cannot be exited in this way when it is in the process of performing analyses for QTL mapping. To terminate the process and exit the software in this case, press Ctrl + C.

Menu System of QTLMapper 1.0

QTLMapper 1.0 runs on a menu system, instead of DOS-command style. Each operation with QTLMapper 1.0 is a matter of hitting a number key (or a letter key) corresponding to a specific menu item. Many dialogue boxes are provided for guiding users through operations.

The menu system of QTLMapper 1.0 consists of a main menu with five sub-menu items (Figure 1).

The detailed usage of the menu system will be described in the following sections.

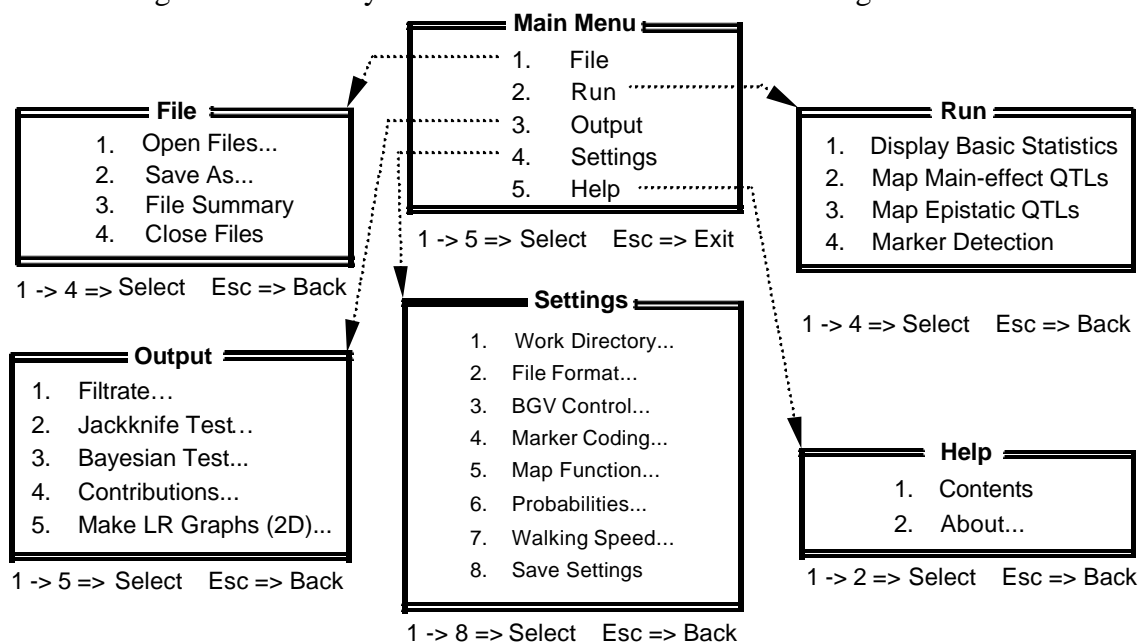


Figure 1. Menu system of QTLMapper Version 1.0

Working with *Settings* Sub-Menu

Settings are a series of conditions that control the way QTLMapper 1.0 performs analysis for QTL mapping. These include work directory, file format, BGV control, marker coding scheme, map

function, probabilities, and walking speed. Usually the first thing the user needs to do is to change the settings so that QTLMapper 1.0 runs under conditions suitable for his/her specific input files and objective of the analysis.

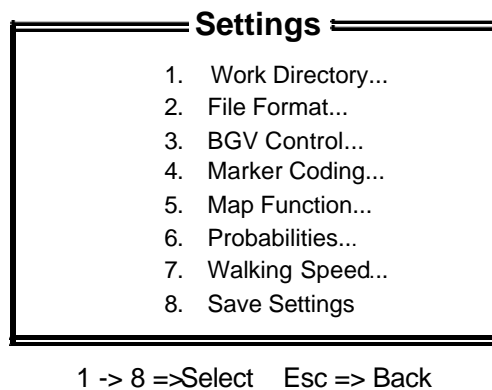


Figure 2. Settings sub-menu

To change the settings, open Settings sub-menu (Figure 2) in Main Menu by pressing number key “4”. If current status is not in Main Menu, you need to press Esc one or more times to return to Main Menu first.

Detailed descriptions for changing settings under Settings sub-menu are made below.

1. Work Directory

A **work directory** is where input files and all sorts of output files are saved. Once the user has specified a work directory, it would not be necessary to add the path name before a file name when opening the file; the program will automatically look for the specified file in the work directory.

The default work directory is where Qtlmapper.exe is started. It is suggested that a separate work directory be created to avoid confusion. To change the default setting (e.g. C:\QTLMAPER), press number “1” under Settings sub-menu, and input a new directory name in the input box (Figure 3). The character string (C:\QTLMAPER) in the square brackets is the current work directory.

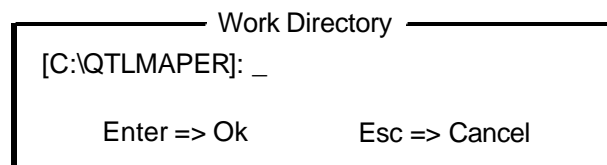


Figure 3. Input box for changing work directory

If the specified directory does not exist, a Decision box (Figure 4) will pop up to let you decide whether to create it.

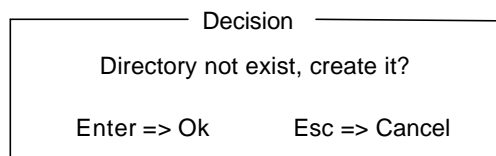


Figure 4. Decision box for creating a new work directory

If Enter key pressed, the directory will be automatically created, and the work directory is changed to this new directory; if Esc key pressed, the current work directory will remain unchanged.

2. File Format

For performing analyses with QTLMapper 1.0, two input files are required: a marker-linkage-map file (for simplification, we call it *map file*) and a data file. A map file contains information about the relative order and genetic distances of all observed markers on the chromosomes or linkage groups; a data file contains observations of the markers and the traits under study for all individuals. **File format** regulates the way of preparing these input files. QTLMapper 1.0 recognizes three formats of map and data files. These include the format for QTLMapper 1.0, the format for the transitive files (output files of Rmap.exe and Rcross.exe) of QTL Cartographer 1.12c, and the format for Mapmaker/QTL 1.1, respectively. The format for QTLMapper 1.0 will be described in detail in the section about “Preparing Input Files”.

Before opening the input files, the user needs to make the setting of file format consistent with that of his/her specific input files. Wrong setting will cause failure in opening files. To view or change current file format, press number “2” under Settings sub-menu to open a dialogue box (Figure 5). Here the current setting is the format for QTLMapper 1.0. To change the setting, further press letter “A”, “B”, or “C” (case insensitive and same hereafter).

File Format		
A.	Format for QTLMapper 1.0	Yes
B.	Format for QTL Cart. 1.10b	No
C.	Format for Mapmaker/QTL 1.1	No

A -> C => Select Esc => Back

Figure 5. Dialogue box for changing file format.
The current setting is format for QTLMapper 1.0.

It is noted that Mapmaker/QTL 1.1 does not require a separate map file for mapping QTL. So if you choose file format for Mapmaker/QTL 1.1, you will be asked to choose a format for your map file from the other two alternatives (one for QTLMapper 1.0, the other for QTL Cartographer 1.12c).

3. BGV Control

Background genetic variation (BGV) is one of the major factors affecting the precision, resolution and power in mapping QTLs. A good control of BGV is the key to obtaining high

precision, resolution and power in mapping QTLs. The capability of controlling BGV is an important feature of QTLMapper 1.0. It provides three choices for BGV control:

- (A) Controlling main and interaction effects of important markers;
- (B) Controlling main effects of important markers;
- (C) No BGV control.

The choice (A) is recommended, and has been taken as the default setting. But the user has the freedom of changing it. To view or change current setting, press number “3” under Settings sub-menu (Figure 6). The current setting in Figure 6 is Main & Interaction Markers. To change the setting, further press letter “A”, “B” or “C”.

BGV Control		
A.	Main & Interaction Markers	Yes
B.	Main-Effect Markers	No
C.	No BGV Control	No

A -> C => Select Esc => Back

Figure 6. Dialogue box for changing BGV control

4. Marker Coding

Marker data in data file need to be presented in a coding manner, or exactly, observations of markers need to be represented by a set of letters and/or numbers, each standing for one marker type. The user can define his/her coding scheme in the data file, or use the default setting shown in Marker Coding dialogue box (Figure 7). QTLMapper 1.0 also allows the user to change the default setting using this dialogue box. To open the dialogue box, press number “4” under Settings sub-menu.

Marker Coding			
A.	P1	Type	A
B.	P2	Type	B
C.	F1	Type	H
D.	F1P1	Type	C
E.	F1P2	Type	D

A -> E => Select Esc => Back

Figure 7. Dialogue box for changing default maker coding scheme

In the above dialogue box,

- “P1 Type”: Marker phenotype that is the same as that of the parent named as P1. If the backcross population is used, the recurrent parent must be named as P1.
- “P2 Type”: Marker phenotype that is the same as that of the other parent named as P2.
- “F1 Type”: Marker phenotype that is the same as that of the heterozygote.

“F1P1 Type”: Marker phenotype that is not P2 type (P1 dominant or not clear between P1 type and F1 type).

“F1P2 Type”: Marker phenotype that is not P1 type (P2 dominant or not clear between P2 type and F1 type).

The letters “**A**”, “**B**”, “**H**”, “**C**”, and “**D**” on the right column are the default codes for the five marker types. To change the setting, further press letter “**A**”, “**B**”, “**C**”, “**D**”, or “**E**”. Each marker type can be coded with only one unique letter (case sensitive) or number.

These settings will be disabled if you define your own marker coding scheme in you data file (see later section). But it will also be effective when you use the function Save As in File sub-menu to save your current input files in a different format.

Note that the default setting for missing marker data cannot be changed, and should always be a dot “.”.

5. Map Function

The user needs to indicate the map function that was used for creating the marker linkage map, so that QTLMapper 1.0 can conduct correct transformation between genetic distances and recombination fractions. The user can specify it in the map file, or use the default setting shown in Map Function dialogue box (Figure 8). To open the dialogue box, press number “**5**” under Settings sub-menu; to change the default setting, further press letter “**A**” or “**B**”.

Map Function	
A. Haldane Function	Yes
B. Kosambi Function	No

A -> B => Select Esc => Back

Figure 8. Dialogue box for changing default map function.
The current setting is Haldane Function.

6. Probabilities

Before doing analyses using **Map Additive-effect QTLs**, **Map Epistatic QTLs** or **Marker Detection** in Run sub-menu, appropriate probabilities for significance tests need to be decided. For opening the dialogue box (Figure 9) to do the settings, press number “**6**” under Settings sub-menu; to change the default settings, further press letter “**A**”, “**B**”, or “**C**”, and input an appropriate value (between 0 and 1) in the input box that appears subsequently.

Probabilities	
A. Marker Inclusion	0.005
B. Marker Exclusion	0.005
C. Threshold Prob.	0.005

A -> B => Select Esc => Back

Figure 9. Dialogue box for changing default probabilities

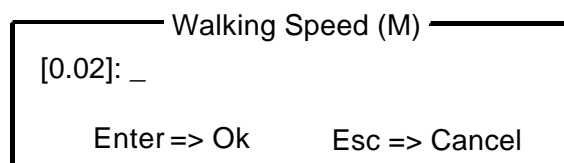
Probability for Marker Inclusion is used for selecting markers with stepwise regression or analysis of variance (ANOVA).

Probability for Marker Exclusion is used only in stepwise regression for ruling out selected marker(s) from regression model when a new marker or marker pair is selected in. Probability for Marker Inclusion cannot be larger than that for Marker Exclusion. Otherwise, it can cause an endless loop. To prevent this from happening, the program will pop up an error message when you input wrong probabilities.

Threshold Probability is used in mapping additive-effect QTLs or mapping epistatic QTLs for testing likelihood ratio (LR) at a test point in mapping additive-effect QTLs or a pair of test points in mapping epistatic QTLs. But this setting becomes effective only when you choose to save mapping results “By Threshold”.

7. Walking Speed

Walking speed means the genetic distance (in Morgan) between adjacent testing points in mapping additive-effect QTLs or epistatic QTLs. The default setting is 0.02 M. One can use a smaller value (slower speed) to perform more detailed search, or a larger value (faster speed) to get a rougher knowledge of QTLs on the genome. To change the setting, press number “7” under Settings sub-menu and input a new value in the input box (Figure 10).



Walking Speed (M)

[0.02]: _

Enter => Ok Esc => Cancel

Figure 10. Input box for changing walking speed**8. Save Settings**

This item in Settings sub-menu is designed specially for saving the settings to file Qtlmaper.ini, so that all the changes can be kept when you run QTLMapper 1.0 the next time. To save the settings, press number “8”.

If you have made changes to the settings, but forgot to save them, then QTLMapper 1.0 will also prompt you to save the settings before exiting.

Preparing Input Files

To use QTLMapper 1.0 for QTL mapping analysis, you need to get your marker linkage map, and data of markers and traits into two plain text files (a map file and a data file) in a format recognized by QTLMapper 1.0. These files are collectively called Input Files.

1. What software to use

Any text file editor can be used for preparing input files. But using MS-Excel™, Lotus 1-2-3™ or other electronic table software can make it much easier to prepare and manage your input files. Note that QTLMapper 1.0 can only open text file, so you need to use “Save As..” to obtain a text version of your input files if you use an electronic table software.

2. Formats of input files recognizable

QTLMapper 1.0 can open three formats of input files, including the format for QTLMapper 1.0 itself, the format for QTL Cartographer 1.12c (transitive files outputted by Rmap.exe and Rcross.exe), and the format for Mapmaker/QTL 1.1. For details about the formats for QTL Cartographer 1.12c and Mapmaker/QTL 1.1, please check their related documents and sample files. If your input files are already in the format for QTL Cartographer 1.12c or Mapmaker/QTL 1.1, you can open the files directly in QTLMapper 1.0 under correct setting for file format.

3. Format of input files for QTLMapper 1.0

We provide two sample files (sample.map and sample.txt) for briefly demonstrating the format of input files for QTLMapper 1.0. (You can open these files without removing any description lines.) Here, we describe the format for QTLMapper 1.0 in more details.

Format of marker-linkage-map file

This file contains information about the marker linkage map, such as the number of chromosomes, number of markers on each of the chromosomes, flanking marker distances, etc. It consists of *general description*, *map body* and some comment lines. Comment lines are not a necessary part of the map file, but they are helpful for reading the file.

General Description: This part is usually put in the front of map file. But it can be anywhere except inside the map body. A typical general description looks like:

```
_DistanceUnit    cM
_MapFunction     K
_Chromosomes     6
```

MarkerNumbers 21 20 19 15 6 18

There are a total of four possible items for general description. They can be arranged in any order. Each item in general description is a key character string (or key string) followed by certain specification(s). Each key string must be started with an underline “_”, and there should not be any white spaces within the key string. The specification(s) must be separated from the key word by at least one white space, and there should also be at least one white space between any two neighboring specifications if two or more specifications are included for the item. Key strings and specifications (if characters) are case insensitive.

“_DistanceUnit cM” specifies the unit of genetic distances used in the map file. The specification string cM stands for centi-Morgan. Another valid character M is for Morgan. If the unit is cM, this item can be omitted, and QTLMapper 1.0 will use its default distance unit (cM) in reading the map file.

“_MapFunction K” indicates the map function used in creating the marker linkage map for turning recombination fractions into genetic distances. QTLMapper 1.0 needs it for correctly transferring genetic distances into recombination fractions in QTL mapping analysis. Specification character K is for Kosambi function. Another valid character is H for Haldane function. This item can also be omitted if you use the current setting specified for map function in Settings sub-menu.

“_Chromosomes 4” is for specifying the total number of chromosomes or linkage groups involved in the map file. This is an essential item, and cannot be omitted.

“_MarkerNumbers 18 15 15 16” is for specifying the numbers of markers on all the chromosomes. The order of the numbers must be consistent with that for genetic-distance columns in the map body. The key string line here indicates 18 markers on chromosome I, 15 markers on chromosome II, 15 markers on chromosome III, and 16 markers on chromosome IV.

Map Body: This part starts from key string **MapBegin** and ends at key string **MapEnd**. Wrong typing of the key strings will cause failure in opening map file. It is convenient to place map body after general description, but it can also be placed anywhere in map file. A typical map body looks like:

<i>*MapBegin*</i>				
Marker#	Ch1	Ch2	Ch3	Ch4
1	0.00	0.00	0.00	0.00
2	9.84	11.26	7.45	9.85
3	10.22	8.69	9.10	10.93
⋮	⋮	⋮	⋮	⋮
15	8.25	9.87	10.66	10.70
16	7.47			11.30
17	12.30			
18	11.78			
<i>*MapEnd*</i>				

The character strings (Marker#, Ch1, Ch2, Ch3, Ch4) in the second row show the contents of the columns below them. QTLMapper 1.0 will skip these strings, so any string can be used, but they cannot be omitted. The Marker# column (first column) is for the order # of all markers on each chromosome; the maximum order # is equal to the number of markers on the chromosome that has the most markers among all the chromosomes. The Ch1 column (second column) to Ch4 column (last column) each represents a chromosome, and contains genetic distances between *adjacent* markers on the chromosome. The genetic distance for the first marker on each chromosome must be set to zero as the start point of the linkage map for the chromosome. The distance for the second marker is between the first and the second markers; the distance for the third marker is between the second and the third markers, and so on.

The order of Ch1 column (second column) to Ch4 (last column) must be consistent with that for the numbers following the key string `_MarkerNumbers`.

Comment Lines: Comment lines can be placed anywhere except between `*MapBegin*` and `*MapEnd*` in the map file. You can put them even between general description lines or at the end of general description lines (see the sample map file).

Format of data file

This file contains information on population type, number of genotypes sampled from the population, number of observations, observations for both markers and quantitative traits, etc. It is composed of four parts: *general description*, *marker data body*, *trait data body*, and some comment lines. Comment lines are not a necessary part of data file.

General description: This part is for specifying the basic features of the data file, and is usually put in the front of the data file. Like in the map file, each item in general description is a key character string followed by certain specification(s). Each key string must be started with an underline “_”, and no white space is allowed within it. There are eight possible items for general description. They can be arranged in any order. A typical general description for data file looks like:

<code>_Population</code>	DH
<code>_Genotypes</code>	200
<code>_Observations</code>	400
<code>_Environments</code>	yes
<code>_Replications</code>	no
<code>_TraitNumber</code>	1
<code>_TotalMarker</code>	64

`_MarkerCode` P1=1 P2=2 F1=3 F1P1=4 F1P2=5

“`_Population DH`” specifies the population type used. This is a necessary item. Specification DH (case insensitive) is for *doubled-haploid population* derived from F_1 between two inbred lines. Other three specifications are also acceptable to QTLMapper 1.0, including:

- RI *Recombinant-inbred-line population* derived from a line cross.
- BC *Backcross population* derived from crossing F_1 with one of the inbred parents.
- F2 *F_2 population* derived from selfing or sib-mating F_1 that is made from two inbred lines.

“`_Genotypes 200`” specifies the total number of individuals or genotypes sampled from the mapping population. This is also a necessary item.

“`_Observations 400`” specifies the total number of observations for each trait under study. If your experimental design does not include replications and environments, the value is the same as the total number of genotypes. In this case, this item can be omitted. However, if you include such factors as replications and/or environments, it becomes necessary.

“`_Environments yes`” specifies the status of experimental design for environments. The specification yes or y (case insensitive) indicates that one environmental factor (such as year factor, or location factor) is included in the phenotyping experiments. If your phenotyping experiment was carried out in only one environment, you can use specification no or n, or simply omit this item. It is noted that QTLMapper 1.0 cannot analyze data sets with designs for more than one environmental factor (e.g. both year factor and location factor).

“`_Replications no`” specifies the status of experimental design for replications or randomized blocks. The specification no or n (case insensitive) indicates that only one or no replication was used in the phenotyping experiment. In this case, this item can also be omitted. When replications or randomized blocks were used, however, it becomes necessary, and the specification must be yes or y.

“`_TraitNumber 1`” specifies the total number of traits whose data are included in the data file.

“`_TotalMarker 64`” specifies the total number of the markers included in the data file. This number must also be equal to the summation of the numbers for `_MarkerNumbers` in the map file.

“`_MarkerCode P1=1 P2=2 F1=3 F1P1=4 F1P2=5`” defines a marker coding scheme. There are five possible strings for the specifications. Each of the strings looks like an equation, but no white space is allowed within the string. On the left side of the equation symbol is the marker genotype/phenotype specification:

P1:	Marker phenotype that is the same as that of the parent named as P1. If BC population is used, the recurrent parent must be named as P1.
P2:	Marker phenotype that is the same as that of the other parent named as P2.
F1:	Marker phenotype that is the same as that of the heterozygote.
F1P1:	Marker phenotype that is not P2 type (P1 dominant or not clear between P1 type and F1 type).
F1P2:	Marker phenotype that is not P1 type (P2 dominant or not clear between P2 type and F1 type).

On the right side of the equation symbol (1, 2, 3, 4, or 5) is the code for the marker type. The marker code should always be a single character (a number or a letter). This definition of marker coding scheme will disable the settings for marker coding that can be viewed and changed in Settings sub-menu. It is noted that the code for missing marker data should always be a dot “.”, and cannot be defined in your data file. It is not necessary to specify codes for all possible marker types except for F₂ population. For example, if your marker data were collected from a DH population, only the specifications for P1 and P2 types are enough.

Marker data body: This part is embraced by two key strings **MarkerBegin** and **MarkerEnd** (not case sensitive). In marker data body, the input of marker data is genotype-based. That is, the marker data at all marker loci for an individual genotype must be inputted before inputting marker data for next genotype. The order of the marker data for different marker loci must be consistent with the order of markers on each chromosome determined in the map file. Below is an example for marker data body.

```
*MarkerBegin*
Geno#  Mk1  ... Mk18  Mk19  ... Mk33  Mk34  ... Mk48  Mk49  ... Mk64  ;
  1     2   ...   2     1     ...   1     1     ...   2     2     ...   1     ;
  2     2   ...   2     2     ...   2     1     ...   1     2     ...   1     ;
      ...
 200   2   ...   2     1     ...   1     1     ...   2     1     ...   2     ;
*MarkerEnd*
```

The second row “Geno# Mk1.....” includes strings for **genotype #** and **marker names**. The marker names must be arranged in the order given in the map file. In this example, Mk1, Mk2, ..., Mk18 are located on Chromosome I; Mk19, Mk20, ..., Mk33 are located on Chromosome II; Mk34, Mk35, ..., Mk48 are located on Chromosome III; Mk49, Mk50, ..., Mk64 are located on Chromosome IV. At the end of the row is a semicolon “;” with at least one white space away from the last marker name (Mk64). Like in the previous sections, the strings for genotype # and marker names should not contain any white space inside. The marker names will be read to memory and used for outputting results later.

It is noted that it will be better to limit the length of each marker name within eight characters.

QTLMapper 1.0 will automatically truncate the name longer than eight characters and keep it within eight characters by adding a “~” plus a number or letter in the end. This is for bettering the appearance in output files. For example, two marker names Marker1_1 and Marker1_21 will be read as Marker~1 and Marker~2, respectively, when the data file is loaded by QTLMapper 1.0. In addition, hyphen “-” should not be used within marker names, because QTLMapper 1.0 uses it to express marker intervals.

The row of marker names not only gives the names of the markers, but also designates the order of marker data arranged in each of the following marker data rows. Each of the data rows must start with an **integer number** as the genotype #, and end with a semicolon “;” (at least one white space away from the last marker value), between which is the marker data for all marker loci in the order given for the marker names. There is no requirement for the seriation of genotype #, but you must keep the genotype # unique for each genotype.

Electronic table software usually has a limit on the number of columns in spreadsheet. For example, MS-Excel cannot hold data with more than 256 columns. If you have more than 256 marker loci, you will have to input marker data in **a few consecutive rows** for each genotype with MS-Excel. QTLMapper 1.0 has no limit on the number of marker data rows for an individual genotype. But do not forget to **put a semicolon “;”** in the end of the last row. This is also the case for the marker name row.

Trait data body: This part is between two key strings *TraitBegin* and *TraitEnd* (case insensitive). The input of trait data is source-based. The *Source* includes the environment (if available), the replication (if available) and the genotype from which each set of observations was obtained for all the traits under study. The following is an example for trait data body.

```
*TraitBegin*
  Env#   Rep#   Geno#   TraitNm1   TraitNm2   TraitNm3   ;
    1     1     1       2.44       7.40       10.04      ;
    1     1     2       2.40       4.32       8.55       ;
    ⋮     ⋮     ⋮       ⋮          ⋮          ⋮          ⋮
    1     1     200     3.54       8.19       10.74      ;
    1     2     1       3.17       6.91       11.86      ;
    1     2     2       1.90       4.31       11.36      ;
    ⋮     ⋮     ⋮       ⋮          ⋮          ⋮          ⋮
    1     2     200     3.22       10.54      11.48      ;
    2     1     1       5.74       12.78      11.27      ;
    2     1     2       7.65       7.02       11.96      ;
    ⋮     ⋮     ⋮       ⋮          ⋮          ⋮          ⋮
    2     1     200     6.58       13.92      9.94       ;
    2     2     1       6.01       10.22      9.95       ;
    2     2     2       6.22       11.99      7.81       ;
    ⋮     ⋮     ⋮       ⋮          ⋮          ⋮          ⋮
    2     2     200     7.98       13.21      12.03      ;
*TraitEnd*
```

The second row includes the source strings and the names of the traits. At the end of the row is also a semicolon “;”. The number of source strings depends on the experimental design. If both environments and replications are taken, a maximum of three strings must be inputted: the first string for environment # (Env#), the second string for replication # (Rep#) and the last string for genotype # (Geno#). If only environments or replications were taken, two strings must be put because the string for genotype # must be included in any case. If there is neither environments nor replications taken in the experiment, one string for genotype # must be provided.

Following these source strings are names of the traits. These names will be read to memory for outputting results later. Like the case for marker names, trait names should not be more than eight characters. Otherwise, they will be automatically truncated, and a symbol “~” and a number or a letter are added in the ends of the names.

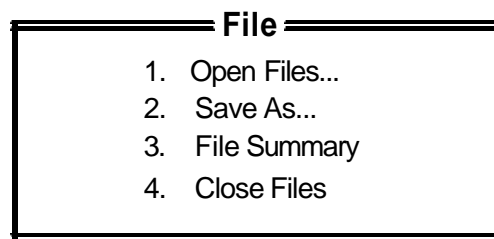
Below the row of trait names are rows of trait data. In these data rows, the first one or three integer numbers (corresponding to the source strings) are the codes of the specific environment, replication, and/or genotype. You need to specify a unique integer number for each of the environments as well as for each of the replications, but the genotype # must be consistent with that for the marker data observed on the same genotype.

Following each set of the source codes are the observations for all the traits in the order for the names given in the row of trait names. In the end of each data row must be a semicolon “;”. Trait data for each individual (from a specific source) can also be inputted in a few consecutive rows to avoid the insufficient capacity of columns provided by some electronic table software.

Comment lines can be placed anywhere in the data file except within marker data body and trait data body.

Working with *File* Sub-Menu

File sub-menu performs operations related to input files. To open File sub-menu (Figure 11), press number “1” in Main Menu. If current status is not in Main Menu, press Esc one or more times to return to Main Menu.



1 -> 4 =>Select Esc => Back

Figure 11. File sub-menu

The usage of the items in File sub-menu is described as follows in details.

1. Opening Files

Loading marker linkage map file and data file into memory is the first step to do analyses using QTLMapper 1.0. Before opening your input files, you will have to make the setting for file format consistent with that applied to your input files (see a previous section).

To open input files, press number “1” in File sub-menu. Two dialogue boxes (Figure 12) will appear in turn for you to input the file names of the map file and the data file. The input files must be already in the work directory you specified through Settings sub-menu. One can choose not to open the map file or the data file by pressing Esc to skip the corresponding dialogue box.

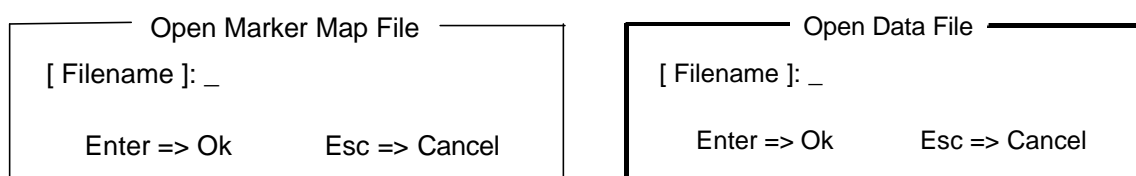


Figure 12. Input boxes for opening map file and data file

2. Save As

In File sub-menu, Save As item is designed for changing file formats between the three formats recognized by QTLMapper 1.0. One can use this function for compare mapping results from different software. To do **Save As**, press number “2” in File sub-menu, and choose one format from the three possible formats, then input filenames for the map file and/or data file through the input boxes appearing subsequently. The new map and/or data file(s) will be saved in the work directory.

3. File Summary

This item in File sub-menu is used for displaying the basic information of the opened map file and data file. For map file, file name, total number of chromosomes, total number of markers, map function used, etc. will be displayed; for data file, file name, population type, total number of genotypes, total number of observations, total number of traits, experimental design, etc. will be shown. Checking this information, the user can learn whether his/her input files are opened correctly. To display file summary information, press number “3” in File sub-menu.

4. Close Files

Close Files item in File sub-menu is designed to close current input files and free the allocated

memory. After closing current files, the user can load another set of map and data files to do QTL mapping without exiting the program.

To close files, press number “4” in File sub-menu.

Working with *Run* Sub-Menu

Run sub-menu implements all the operations related to mapping QTLs. To open Run sub-menu (Figure 13), press number “2” in Main Menu. If current status is not in Main Menu, press Esc one or more times to return to Main Menu.

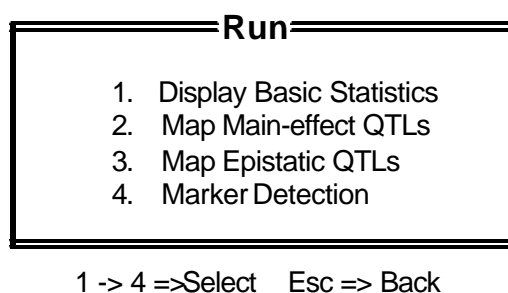


Figure 13. Run sub-menu

There are four items in Run sub-menu. Note that these items become active only after map and data files have been opened.

1. Display Basic Statistics

This item is for calculating and displaying the basic statistics of the marker data and/or trait data. The user can use this function to check the segregation of the observed marker loci and the frequency distributions of the traits under study, etc. For displaying basic statistics, press number “1” in Run sub-menu, and a dialogue box will appear for you to decide what data to display (Figure 14).

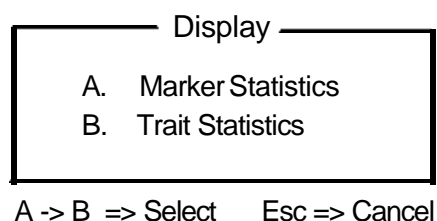


Figure 14. Dialogue box for displaying basic statistics

To display basic statistics of marker data, press letter “A” in the dialogue box. Then statistics of every four markers appear in a page. Press PgDn or PgUp to turn to next or previous page; press Esc to quit the display. The statistics include percentage of different marker phenotypes in the

sample, chi-square values and probabilities for testing marker segregation. A brief comment is given based on the probabilities (Poor for $P \leq 0.05$; Fair for $0.05 < P < 0.10$; Good for $P \geq 0.10$).

To display basic statistics of trait data, press letter “B” in the dialogue box. Then the statistics of all the traits included will be displayed page by page. Press PgDn or PgUp to turn to next or previous page; press Esc to quit the display. The statistics for data of each trait include the frequency distribution indicated with asterisks “*” (data automatically divided into ten groups), mean, standard deviation (SD), skewness, and kurtosis.

2. Map Additive-effect QTLs

This item in Run sub-menu is designed for mapping additive-effect QTLs (assuming no epistasis between QTLs). The statistical methodology used is the mixed-model approaches (Zhu, 1998; Zhu and Weir, 1998; Wang, Zhu, Li and Paterson, 1999). To map additive-effect QTLs, press number “2” in Run sub-menu. Then following operations need to be carried out step by step through various dialog boxes.

Step 1. *Setting a name for result file*

A dialogue box will appear and allow you to change the default name of the result file (we call it *original result file*) for saving the mapping results. If you do not want to change it, just press Esc. Otherwise, press Enter, and type a new file name in another dialogue box.

Step 2. *Setting mapping ranges*

After result file name has been decided, a dialogue box will pop up for you to set mapping ranges (trait range and genomic range) (Figure 15). In this step, you have a chance to choose a subset of the traits, a subset of chromosomes and certain marker intervals for detailed analysis, instead of all the traits and the entire genome.

Setting Mapping Ranges	
A. Trait Range	All
B. Genomic Range	Aut

A -> B => Select Esc => Next

Figure 15. Dialogue box for setting mapping ranges

Trait range: The default setting for trait range is “All”, namely all the traits will be analyzed. To change trait range, press letter “A” in Setting Mapping Ranges dialog box. Then an input box (Figure 16) will appear.

Trait Range Setting	
[All]: _	
Enter => Ok	Esc => Cancel

Figure 16. Input box for setting trait range

In the input box, type the order numbers (based on the order of trait names in the data file) of the traits you want to analyze, and press Enter key. At least one white space is required between the order numbers. A dash “-” symbol can be used for simplifying the input for several consecutive order numbers. For example,

1 3 5-7

is equivalent to

1 3 5 6 7

But no white spaces are allowed between number and “-”, and between “-” and number. Note that trait names are not acceptable.

Genomic range: The default setting for genomic range is “Aut” (short for *automatically*), which means that the marker intervals to be tested will be decided automatically based on selected important main-effect markers (see the following sections for details). This is a quick way for identifying important putative QTLs. To change this setting, press letter “B” in Setting Mapping Ranges dialogue box, then another dialogue box (Figure 17) will appear for you to choose from the three alternatives (A-C).

Genomic Range Setting	
A. Automatically	Yes
B. Whole Genome	No
C. Manually	No
A -> C => Select Esc => Back	

Figure 17. Dialogue box for changing genomic range

“Yes” marks the current setting (Automatically). If you want to search QTLs on entire genome, select item “B”. Pressing letter “C” will bring out an input box (Figure 18) for you to setting chromosome range.

Chromosome Setting	
[All]: _	
Enter => Ok	Esc => Cancel

Figure 18. Input box for setting chromosome range

In the input box, input the order number (based on the order of the chromosomes in map file) and press Enter key. The method of inputting chromosome order number is the same as that for inputting trait order numbers in the previous passage. After chromosome setting has been done, another dialogue box will appear for you to set marker intervals on each of the selected chromosomes for mapping analysis. You can do this in a similar way.

When you have finished the setting for mapping ranges, you will have to return the dialogue box (Figure 15) by pressing Esc. Further pressing Esc, the operation will go to next step.

Step 3. *Choosing a method for saving mapping results*

A new dialogue box will appear (Figure 19). Through this dialogue box, the user can decide how to save the mapping results.

How to Save Results	
A. By Threshold	Yes
B. For All Testing Points	No

A -> B => Select Esc => Next

Figure 19. Dialogue box for selecting a way of saving mapping results

Item “A” (saving results by threshold) is a default setting. In this way, only the results for the testing points that are likely to be putative QTLs (based on the threshold probability specified in Settings sub-menu) will be saved. This can effectively reduce the size of the result file. If you want to view the results for all the testing points, particularly to make graphs from the mapping results, you will need to select item “B”. “Yes” marks the current setting.

After this is done, pressing Esc, the operation will move to next step.

Step 4. *Deciding a hypothesis for LR test*

This step happens only when your trait data include an environmental factor. If no environmental factor was involved in your phenotyping experiment, the program will skip to the next step. In addition, QTLMapper 1.0 cannot analyze F₂ data set from multiple environments.

In this step, a dialog box will appear for you to choose a hypothesis from three alternatives (Figure 20).

LR Test for	
A. Genetic Main Effects & GE Interactions	Yes
B. Genetic Main Effects Only	No
C. GE Interactions Only	No

-19-

A -> C => Select Esc => Next

Figure 20. Dialog box for deciding a hypothesis for LR test

The default hypothesis is for testing “Genetic Main Effects & GE Interactions”. For DH or RI population, it is equivalent to $H_0: a_i = 0$ and $\mathbf{s}_{A_i E}^2 = 0$, where a_i is the additive effect of the individual testing point i , and $\mathbf{s}_{A_i E}^2$ is the variance from the main effect (a_i) \times environment interaction. Item “B” is for testing $H_0: a_i = 0$ only, while item “C” is for testing $H_0: \mathbf{s}_{A_i E}^2 = 0$ only. For BC population, just change a and A to s and S in the expressions for the hypotheses. S or s stands for the difference between the additive effect (a) and the dominance effect (d) for the testing point, $s = a - d$.

The user can decide one in this dialog box by pressing letter “A”, “B”, or “C”, then press Esc for next step.

Step 5. Deciding whether to continue the process for QTL mapping

A decision box (Figure 21) will appear and give you a chance to quit the mapping process and return to Run sub-menu.

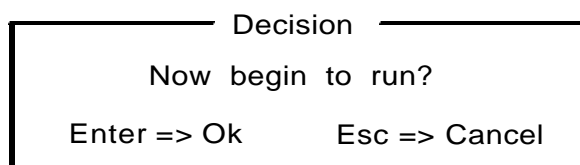


Figure 21. Decision box for continuing or quitting QTL mapping

Step 6. Deciding whether to redo marker selection

The mixed model approach uses important markers (selected by ANOVA or stepwise regression) for BGV control. The selected main-effect markers will be saved in a file with extension of “.one”, while the selected pair-wise interaction markers are saved in another file with extension of “.two”. If you set BGV control (by Settings sub-menu) as controlling “Main-effect Markers” or “Main & Interaction Markers”, or you have used “Aut” to decide your genomic range to be mapped, then a decision box (Figure 22) will appear if the .one file already exists. You can decide to redo the main-effect marker selection (here only by stepwise regression) or keep the original. If the .one file does not exist in the work directory, the main-effect marker selection will start automatically.

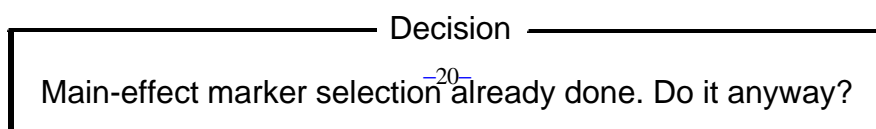


Figure 22. Decision box for main-effect marker selection

If you set BGV control as controlling “Main & Interaction Markers”, a similar dialog box (not shown here) will appear subsequently if the two file already exists. You can decide to redo the interaction marker selection (only by stepwise regression) or keep the original. If there is no such .two file in the work directory, the subroutine for selecting interaction markers will be started automatically, and a new .two file will be created.

After all the above has been done, QTLMapper 1.0 will begin the process of mapping additive-effect QTLs, and display the progress of the analysis and some other information, such as result file name and path, trait being analyzed, BGV control method, and hypothesis for LR test, etc. (Figure 23).

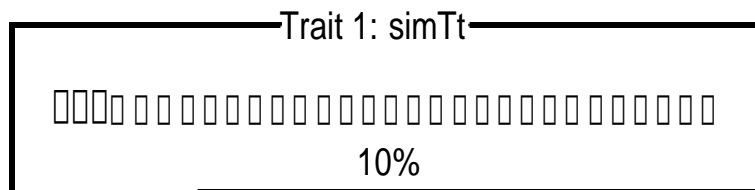


Figure 23. Progress of mapping additive-effect QTLs for first trait named simTt

Step 7. Deciding whether to filtrate putative main-effect/epistatic QTLs

QTLMapper 1.0 provides the user with a function for filtrating putative additive-effect QTLs or epistatic QTLs from the original mapping results obtained through the above steps. The filtration subroutine will search for the peak points of LR values along each chromosome and compare them with the threshold value determined by the threshold probability setting. If the LR value at a peak is above the threshold, the testing point with this LR value will be kept and saved in a file with the same stem name as the *original result file* and extension of .flq (for additive-effect QTLs) or .fle (for epistatic QTLs).

A decision box (Figure 24) will appear after the process of mapping additive-effect QTLs has been completed. The user can do the filtration at this step, or use Output sub-menu to do it later.

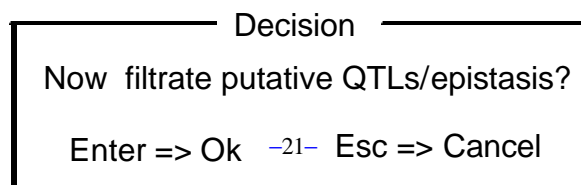


Figure 24. Decision box for filtrating putative QTLs/epistatis

Step 8. Deciding whether to conduct Bayesian test for the putative main-effect/epistatic QTLs

If the data set includes an environmental factor, a dialog box (Figure 25) will appear at this step, otherwise this step will be automatically skipped.

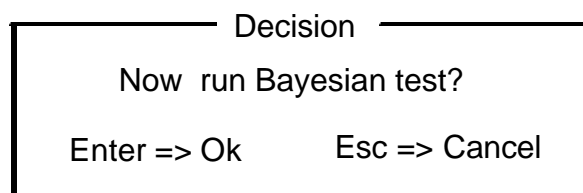


Figure 25 Decision box for running Bayesian test

Bayesian test is intended to test both fixed genetic main-effects (QTL additive effects and/or epistatic effects) and random QTL \times environment (QE) interaction effects (see later section for details) for the putative QTLs filtrated in the previous step. You can do Bayesian test at this step. If you do not take Bayesian test at this step, you will not have the predicted values for the QE interaction effects, since the QE interaction effects are also predicted in this Bayesian test step.

The results from Bayesian test will be saved in a file with the same stem as that for the *original result file* and extension of .byq (for additive-effect QTLs) or .bye (for epistatic QTLs) in the work directory.

Step 9. Deciding whether to calculating relative contributions for putative main-effect/epistatic QTLs

Relative contributions of QTLs and QE interaction effects to the phenotypic variation are important criteria by which the importance of a QTL or QE interaction effect can be judged. In this step, the user can estimate the relative contributions for all the putative QTLs and/or QE interactions based on the filtration results (if Bayesian test has not been done) or Bayesian test results (if Bayesian test has been done) in the previous steps. A decision box (Figure 26) will appear for you to decide.

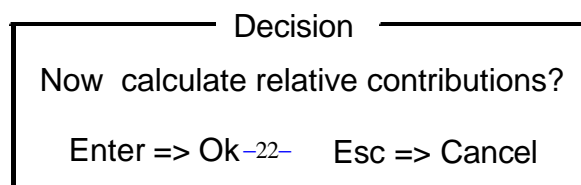


Figure 26. Decision box for calculating relative contributions

The results for the relative contributions will be saved in a file with the same stem as that for the *original result file* and extension of .ctq (for additive-effect QTLs) or .cte (for epistatic QTLs) in the work directory.

3. Map Epistatic QTLs

This item in Run sub-menu is designed for mapping both additive effect and digenic epistatic QTLs. The statistical methodology used is also the mixed-model approach proposed (Zhu, 1998; Zhu and Weir, 1998; Wang, Zhu, Li and Paterson, 1998). To map epistatic QTLs, press number “3” in Run sub-menu. The subsequent operations are quite similar to those for mapping additive-effect QTLs, with the the following differences.

Difference 1: In step 1 (Deciding a name for result file), the default file name here is the stem name of the data file plus an extension of .eps, instead of .qtl.

Difference 2: In step 4 (Deciding a hypothesis for LR test), if you select hypothesis “C” (testing GE Interactions Only), another dialog box (Figure 27) will appear, in which you can test separately:

Item “B”: $A_i \times E$ (or $S_i \times E$ for BC population) interactions ($\mathbf{s}_{A_i E}^2 = 0$ or $\mathbf{s}_{S_i E}^2 = 0$);

Item “C”: $A_j \times E$ (or $S_j \times E$ for BC population) interactions ($H_0: \mathbf{s}_{A_j E}^2 = 0$ or $\mathbf{s}_{S_j E}^2 = 0$);

Item “D”: $AA_{ij} \times E$ (or $SS_{ij} \times E$ for BC population) interactions ($H_0: \mathbf{s}_{AA_{ij} E}^2 = 0$ or $\mathbf{s}_{SS_{ij} E}^2 = 0$).

You can also test “General GE Interactions” ($H_0: \mathbf{s}_{A_i E}^2 = \mathbf{s}_{A_j E}^2 = \mathbf{s}_{AA_{ij} E}^2 = 0$) by selecting item “A”.

GE Interaction Test for	
A. General GE Interactions	Yes
B. Ai x E Interactions Only	No
C. Aj x E Interactions Only	No
D. AAij x E Interactions Only	No

A -> D => Select Esc => Next

Figure 27. Dialog box for deciding a hypothesis on QE interactions

Difference 3: Mapping epistatic QTLs with F₂ data is not yet available in QTLMapper 1.0 whether the data set is collected from one environment or from multiple environments.

4. Marker Detection

This item in Run sub-menu (Figure 13) is designed for selecting important main-effect and/or interaction markers that can be used in two ways: controlling BGV and automatically determining genomic ranges in mapping main-effect/epistatic QTLs.

To do marker detection, press number “4” in Run sub-menu. A dialogue box (Figure 28) will appear to prompt you to select a method from the two alternatives: stepwise regression and analysis of variance.

Method Choice	
A. Stepwise Regression	Yes
B. Analysis of Variance	No

A -> B => Select Esc => Next

Figure 28. Dialog box for choosing a method of marker detection

After the method selected, press Esc for next step. Then another dialog box (Figure 29) will appear.

Detection for	
A. Main-Effect Markers	Yes
B. Main & Interaction Markers	No

A -> B => Select Esc => Next

Figure 29. Dialog box for detecting important markers

In this dialog box, you can decide to detect main-effect markers (item “A”) only or both main and interaction markers (item “B”). After you have decided, press Esc for next step.

A decision box (Figure 30) will appear and let you decide to continue or quit the process of marker detection.

Decision	
Now begin to run?	
Enter => Ok	Esc => Cancel

Figure 30. Decision box for continuing or quit marker detection

Press Enter key to continue, and a progress bar similar to Figure 23 and some information will also be displayed. Press Esc key to quit the process.

After the marker detection, one (detecting main-effect markers only) or two files (detecting both main and interaction markers) will be created. The user cannot change the file names. They will always be the stem name of the data file plus an extension of .one (for main-effect markers only) or .two (for interaction markers only). For example, if you have a data file named rice.txt, then the result file for detected main-effect markers will be rice.one, and the result file for detected interaction markers will be rice.two.

Working with *Output* Sub-Menu

Output sub-menu is designed for processing the original result file from mapping QTLs with additive/epistatic effect so that the manual work for presentation with the *original result file* can be largely reduced. In addition, Output sub-menu can also be used for obtaining the results of hypothesis test using some special methods. To open Output sub-menu (Figure 31), press number “3” in Main Menu.

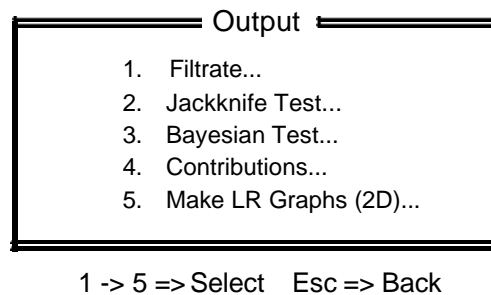


Figure 31. Output sub-menu

There are five items in Output sub-menu. Detailed usage of these items in Output sub-menu is given below.

1. Filtrate

This function is designed for filtrating putative additive-effect QTLs or epistatic QTLs from the *original result file*. The filtration subroutine will search for the peak points of LR values along each chromosome and compare them with the threshold value determined by the threshold probability setting. If the LR value at a peak is above the threshold, the testing point with this LR value will be kept and saved in a file with the same stem name as the *original result file* and extension of .flq (for additive-effect QTLs) or .fle (for epistatic QTLs). This can largely reduce the manual work in searching putative additive-effect QTLs/epistatic QTLs from the *original result file*.

To filtrate putative additive-effect QTLs or epistatic QTLs, press number “1” in Output sub-menu. A dialog box (Figure 32) will appear and let you decide what to filtrate.

Filtrate for	
A. Putative Additive-effect QTLs	Yes
B. Putative Epistatic QTLs	No

A -> B => Select Esc =>

Figure 32. Dialog box for filtrating putative QTLs

Whatever you select in the dialog box, an input box (Figure 33) will appear subsequently. This input box let you change the default threshold probability (0.005) used for the filtration. But it cannot change the default threshold probability specified in Settings sub-menu.

Filtration Probability	
[0.004999]: _	
Enter => Ok	Esc => Cancel

Figure 33. Input box for changing the filtration probability

If you do not want to change the default, just press Enter key. Otherwise, you need to input a value between 0 and 1.

Another input box (Figure 34) will appear after you press Enter in the previous input box. You need to input the *original result file* (.qtl or .eps file) in this input box.

Open Result File	
[Filename]: _	
Enter => Ok	Esc => Cancel

Figure 34. Input box for inputting original result file name

The *original result file* must already exist in the work directory that is specified in Settings sub-menu. The resulted file with .flq or .fle extension will also be saved in the work directory.

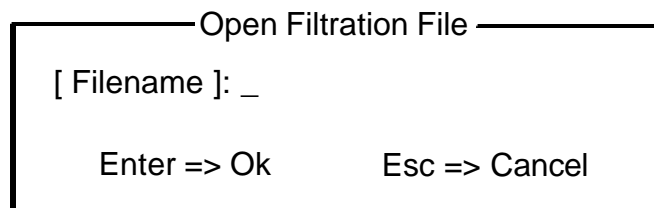
2. Jackknife Test

In the mixed model approach (Zhu,1998; Zhu and Weir, 1998; Wang, Zhu, Li, and Paterson, 1998), the environmental factor, if involved, is taken as a random factor. Then the QE interactions should also be random though the genetic main effects (QTL additive effects and epistatic effects) are fixed in the mixed model. In this case, the significance test for the predicted QE interaction

effects by best linear unbiased prediction (BLUP) has very low power using the ordinary z -test (based on the normal distribution).

Jackknife test is a re-sampling technique that has been widely used for parameter estimation and significance test. Here we use it for parameter estimation and significance test of QTL genetic main effects and QE interaction effects by cutting one genotype each round, when an environmental factor is involved in the data set. If the data set does not include an environmental factor, you are not suggested to run Jackknife test because this operation will take a long time.

To run Jackknife test, press number “2” in Output sub-menu. An input box (Figure 35) will appear. You need to input the name of the file created from filtrating additive-effect QTLs or epistatic QTLs. The Jackknife test will be based on the positions of the putative QTLs presented in this file.



Open Filtration File

[Filename]: _

Enter => Ok Esc => Cancel

Figure 35. Input box for opening filtration file

After the filtration file has been opened, you will be asked to input the map file and data file if they are not opened. The method of opening the map and data files is the same as that described in Working with File Sub-Menu. After that, the process of Jackknife test will begin, and the progress and some information will be displayed.

After the process, a result file with the same stem as the *original result file* and an extension of .jq (for additive-effect QTLs) or .jpe (for epistatic QTLs) will be created.

3. Bayesian Test

Bayesian test is designed for the same purpose as that of Jackknife test. But Bayesian test is much quicker than Jackknife test. The results from the two methods are rather similar.

To run Bayesian test, press number “3” in Output sub-menu. The method of following operations is the same as for Jackknife test.

After the process of Bayesian test, a result file with the same stem as the original result file and an extension of .byq (for additive-effect QTLs) or .bye (for epistatic QTLs) will be created.

4. Contributions

This item in Output sub-menu is designed for obtaining the relative contributions of QTLs and/or QE interactions. Relative contributions of QTLs to the phenotypic variation are important

criteria of judging a specific additive-effect QTL, epistatic QTLs, or QE interaction. To obtain relative contributions, press number “4” in Output sub-menu. Then a dialog box (Figure 36) will appear, and ask you to select a file type on which the calculation of relative contribution will be based.

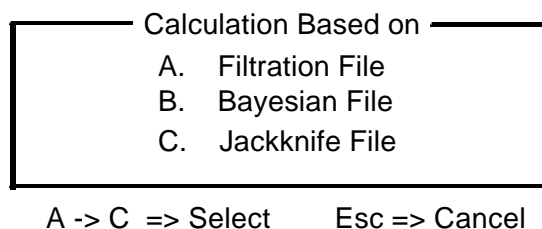


Figure 36. Dialog box for selecting a file type for calculating relative contributions

Type “A” (Filtration File) is the file with .flq or .fle from filtrating putative additive-effect or putative QTLs; type “B” (Bayesian File) is the file with .byq or .bye extension from Bayesian test; type “C” (Jackknife File) is the file with .jqk or .jke extension from Jackknife test.

After you have selected a file type, an input box will appear for you to input the file name of that type. Subsequently, you will be asked to open the map file and data file through input boxes. The operations are similar to those for Jackknife test.

After the process of calculating relative contributions, a result file with the same stem as the original result file and an extension of .ctq (for putative QTLs) or .cte (for putative epistatic QTLs) will be created.

5. Make LR Graphs (2D)

This item in Output sub-menu is designed to create some files containing two-dimensional coordinates: testing points on each chromosome as x-axis; the likelihood ratio (LR) value as y-axis. It is based only on the *original result file* from mapping additive-effect QTLs. These graph files can be used by **Wgnuplot**, a free Windows application (Copyright (C) 1986 - 1993 Thomas Williams, Colin Kelley), for showing and printing LR graphs. This can make the presentation of your mapping results more attractive.

To make LR graphs, press number “5” in Output sub-menu. An input box (Figure 37) will appear for you to input the *original result file* from mapping additive-effect QTLs.

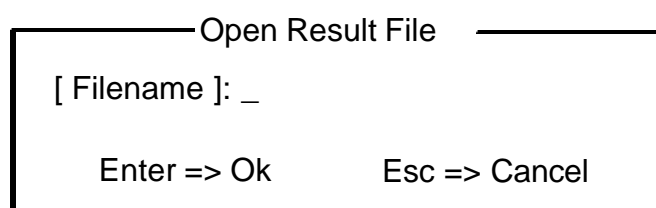


Figure 37. Input box for opening original result file

The graph files will be created in a moment. They include:

A plot file with an extension of .plt: containing commands for plotting 2D LR graphs by Wgnuplot;


A series of coordinate files like *ticj.00k*: containing the coordinates of testing positions and LR values. Each chromosome for each trait has one or more such files. For example, *t1c1.001* contains the coordinates for chromosome #1 of trait #1 (the numbers are determined in map and data files). If a relatively large gap was found between two testing segments on a chromosome in the *original result file*, there will be two separate files to express the coordinates on the same chromosome for the same trait; if there are more such gaps, then more files will be created. The extension .001 indicates that this file is the first one for the trait and the chromosome.

A series of threshold files like *chromi.sig*: These files contain coordinates for horizontal threshold lines ($y = \text{threshold}$). Each chromosome will have one such file. For example, *chrom1.sig* contains coordinates for drawing a threshold line on chromosome #1. The threshold is determined by the threshold probability specified in Settings sub-menu.

For the method of plotting 2D graphs by **Wgnuplot**, please see the next section.

Understanding Files Created by QTLMapper 1.0

Several kinds of result files will be created from the analysis for mapping QTLs with QTLMapper 1.0, as described above. To make inferences about the putative QTLs for the traits under study, the user needs to understand the contents of these result files. In general, every result file consists of two portions: *description of conditions* on which the result is obtained, and the *result body*. In these files, there is usually a word “**End**” that ends the files. In this section, the contents of all types of result files will be described in detail.

 Warning! Do not edit the part of *description of conditions*. Some information given in this part will usually be read in as conditions for further treatments by Run or Output sub-menu. If this part in a result file has been modified, particularly with BGV control method, Threshold probability, etc. changed, inappropriate results could be created in the files derived from it.

1. Files from marker detection (*.one* and *.two* files)

As we mentioned earlier, controlling background genetic variation (BGV) is an important feature of QTLMapper 1.0. The selected main-effect and/or interaction markers are used in the mixed model for BGV control. Whether you use stepwise regression or ANOVA, two types of

result files will be created. The file with .one extension contains the results for selected main-effect markers, and the file with .two extension has the results for selected interaction markers. The contents of the two types of files are described in this section.

File with .one extension: List 1 gives an example of .one file obtained for a data set without environmental factor in the phenotyping experiment.

The *description of the conditions* is placed in the front of the file. It includes: the names and paths of data file and map file, experimental design for environments and replications (yes = included; no = not included), brief contents of the file, method of the analysis used; and threshold probabilities for the marker selection. The date and time when the file was created are also provided in this part.

The *result body* includes the results of selected main-effect markers for all the traits that are inputted in the data file. For each trait, the output of the results include: trait name following a string (e.g. Trait1:, Trait2:, Trait3:, etc.) indicating its order in the data file, total number of selected markers, and general R² (multiple determination coefficient of all the selected markers). Note that there will be no general R² if you used ANOVA method. Then there is a two-way table showing results for the selected main-effect markers.

List 1. An example of .one file for a single environment: dh.one

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\dh.txt
// Marker map file name: D:\QTLMAPER\dh.map
// Environments:       no
// Replications:       no
// Contents: selected main-effect markers
// Analysis method: Stepwise Regression
// Inclusion probability: 0.010000
// Exclusion probability: 0.010000

# Date: 1999-02-01      Time: 15:46:56

Trait1: ju28ph

Total selected markers: 6   General R^2: 0.4989

  Marker#      Name      F value      Prob.      R^2
  -----
    171      Mk171      16.5499      0.0002      0.0873
     54      Mk54      20.9835      0.0000      0.1107
     18      Mk18      12.0779      0.0013      0.0637
     67      Mk67      21.4424      0.0000      0.1131
     32      Mk32      11.1240      0.0019      0.0587
    124      Mk124       8.6798      0.0054      0.0458
.....
End
```

The **Marker#** column shows the order numbers of selected markers. The order numbers are decided according to the order of the markers arranged in the data file. The order number for the first marker is 1, the second is 2, ...

The **Name** column shows the names of the selected markers. Note that if the name of a selected marker is longer than eight characters in the data file, the truncated version of the name will be given in this .one file.

The **F value** column provides the F values for testing partial regression coefficients of the selected main-effect markers.

The **Prob.** column gives the probabilities for the F values.

The **R²** column shows the partial determination coefficients of the selected main-effect markers, each calculated with all the other selected markers in the regression model. There will be no such column if you use ANOVA for the marker detection.

If the data set includes an environmental factor, the presentation of the results for selected main-effect markers is slightly different from the above in the *result body*. List 2 is a .one file obtained for a data set with two environments.

List 2. An example of .one file for multiple environments: simu.one

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:        yes
// Replications:         no
// Contents: selected main-effect markers
// Analysis method: Stepwise Regression
// Inclusion probability: 0.005000
// Exclusion probability: 0.005000

# Date: 1999-03-12      Time: 12:13:07

Trait1: simTt

T1/LE1:  1      // Environment code

Total selected markers: 3   General R^2: 0.4039

  Marker#      Name      F value      Prob.      R^2
    42         MK42      44.6769      0.0000      0.1359
    32         MK32      56.7282      0.0000      0.1725
    25         MK25      29.5473      0.0000      0.0899

T1/LE2:  2      // Environment code

Total selected markers: 4   General R^2:0.5242

  Marker#      Name      F value      Prob.      R^2
    33         MK33      78.7107      0.0000      0.1921
    42         MK42      98.2873      0.0000      0.2398
```

27	MK27	20.9713	0.0000	0.0512
2	MK2	8.4631	0.0040	0.0207

.....

End

For each trait, there is a subset of results for each environment. A string, such as T1/LE1, T1/LE2, indicates the trait order number and environment order number. For instance, T1/LE1 stands for the first trait and the first environment level. The integer number that follows the string is the code of the environment level as a subscript for the environmental level in the data file.

File with .two extension: List 3 is an example of .two file created for a data set from a single environment.

List 3. An example of .two file for a single environment: dh.two

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\dh.txt
// Marker map file name: D:\QTLMAPER\dh.map
// Environments:       no
// Replications:       no
// Contents: selected interaction markers
// Analysis method: Stepwise Regression
// Inclusion probability: 0.010000
// Exclusion probability: 0.010000

# Date: 1999-02-01      Time: 15:47:05

Trait1: ju28ph

Total selected marker pairs: 4   General R^2: 0.2187

  MarkerA      Name      MarkerB      Name      F value      Prob.      R^2
  -----
    51         Mk51         74         Mk74        20.1118      0.0000      0.0624
    10         Mk10         14         Mk14        35.9031      0.0000      0.1114
    96         Mk96        152        Mk152        17.4605      0.0001      0.0542
    32         Mk32         79         Mk79        10.0271      0.0022      0.0311
  -----
End
```

The *description of conditions* is quite similar to that for .one file. So we are not going to talk about it here.

In the *result body*, the trait name, total number of selected pairs of interaction markers, and general R^2 (if using stepwise regression) are given for each trait. Note that general R^2 is the multiple determination coefficient of all the selected pairs of interaction markers with control over all the selected main-effect markers in the .one file obtained for the same data set. Then there is also

a two-way table showing the results for the selected pairs of interaction markers.

The **MarkerA** and **MarkerB** columns are the order numbers of each pair of interaction markers. The order numbers are also decided according the order of the markers arranged in the data file.

The two **Name** columns give the marker names corresponding to the **MarkerA** order number and **MarkerB** order number.

The **F value** and **Prob.** columns gives the F-test results for the selected interaction markers, which is similar to that in .one file.

The **R²** column, if using stepwise regression method, presents the partial determination coefficients for the selected interaction markers with the control over all the selected main-effect markers (in the corresponding .one file) and all the other selected pairs of interaction markers. If ANOVA method was used, there will be no such column.

If the data set includes an environmental factor (several environmental levels), the presentation of the results for selected interaction markers is slightly different from the above in the *result body*. List 4 is a .two file obtained for a data set from two environments.

The subset of the data for each of the environments will be analyzed separately, and the results for selected interaction markers are presented separately. A string like T1/LE1, is used for indicating the trait and the environment level. The integer number that follows the string is the code of the environment level as appeared for the subscript in the data file.

List 4. An example of .two file for multiple environments: simu.two

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:        yes
// Replications:        no
// Contents: selected interaction markers
// Analysis method: Stepwise Regression
// Inclusion probability: 0.005000
// Exclusion probability: 0.005000
```

```
# Date: 1999-03-16      Time: 21:07:34
```

```
Trait1: simTt
```

```
T1/LE1: 1      // Environment code
```

```
Total selected marker pairs: 7  General R2: 0.4763
```

MarkerA	Name	MarkerB	Name	F value	Prob.	R ²
2	MK2	25	MK25	189.3376	0.0000	0.1200
32	MK32	42	MK42	153.9681	0.0000	0.0976
2	MK2	32	MK32	94.2345	0.0000	0.0597
25	MK25	42	MK42	88.5875	0.0000	0.0562
24	MK24	32	MK32	59.4700	0.0000	0.0377
3	MK3	42	MK42	29.2651	0.0000	0.0186

```

1           MK1           11           MK11           15.0407           0.0001           0.0095
T1/LE2:  2           // Environment code

Total selected marker pairs: 4 General R^2: 0.2356

MarkerA      Name      MarkerB      Name      F value      Prob.      R^2
25           MK25           33           MK33           62.3905           0.0000           0.0785
32           MK32           42           MK42           61.481            0.0000           0.0773
2            MK2            24           MK24           45.0538           0.0000           0.0567
23           MK23           41           MK41           14.7214           0.0002           0.0185
.....
End

```

2. Original result files (.qtl and .eps files)

So-called *original result files*, as defined previously, are created from mapping QTLs with additive effects and epistatic effects. The default extensions are .qtl for additive-effect QTLs and .eps for epistatic QTLs. But the user can change them at Step 1 (Deciding a name for result file) in performing mapping analysis. Likewise, these two types of result files are also composed of two parts: *description of conditions* used for mapping analysis and the *result body*. The following describes the details about the contents of these files.

File with .qtl extension: List 5 is an example of .qtl file for a data set from a single environment.

List 5. An example of .qtl file for a single environment: dh.qtl

```

// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\dh.txt
// Marker map file name: D:\QTLMAPER\dh.map
// Environments:       no
// Replications:       no
// Contents: putative additive-effect QTLs
// BGV control method: A (control marker main & interaction effects)
// Results saved for: points significant at 0.005000

# Date: 1999-02-03      Time: 12:28:36

Trait 1: ju28ph

Chrom      Interval      Site1(M)      Site2(M)      LR      LOD      A      Prob
1           Mk17-Mk18      2.167         0.000         13.06    2.83    -1.9994  0.0004
1           Mk17-Mk18      2.187         0.020         14.80    3.21    -2.1550  0.0001
1           Mk17-Mk18      2.207         0.040         16.16    3.51    -2.2455  0.0001
1           Mk17-Mk18      2.227         0.060         16.99    3.69    -2.2630  0.0000
1           Mk17-Mk18      2.247         0.080         17.27    3.75    -2.2134  0.0000
2           Mk31-Mk32      1.599         0.000         10.16    2.21    1.7219   0.0016
.....
End

```

The part of *description of conditions* is put in the front of the file. It includes: the names and paths of the map file and the data file, experimental design for environments and replications (yes = included; no = not included), brief contents of the file, BGV control method (decided through Settings sub-menu), and method for saving the results that is decided in Step 3 (Deciding how to save mapping results) of the mapping analysis.

The layout of the result body is also trait-based like in the .one or .two files. For each trait, a string (e.g. Trait 1:) is given to indicate the trait order number according to the arrangement of the trait data in the data file. Followed is the trait name (if more than eight characters in the data file, it will be shown here as a truncated name). Then there is a large two-way table showing the mapping results for each testing point using the method for saving mapping results. If the method is By Threshold, then only the testing points shown significant LR values are saved, otherwise the results for all testing points are saved.

The **Chrom** column gives the order numbers (integer numbers) of the chromosomes on which each testing point is set. The order numbers are decided the arrangement of chromosome columns for the genetic distances in the map file.

The **Interval** column presents the names of the marker intervals (each composed of two flanking markers) on which a testing point is set. A hyphen “-” is used between the two flanking markers. Here the truncation method will apply if the marker names exceed eight characters.

The **Site1(M)** column is the genetic distance (in Morgan) of the testing points from the left end of the chromosomes on which the testing points are set. The **Site2(M)** column is the genetic distance (in Morgan) of the testing points from the left marker on the interval on which the testing points are set.

The **LR** column gives the likelihood ratio values for the testing points. The hypothesis is $H_0: a_i = 0$ for DH or RI populations, $s_i = 0$ for BC population, $a_i = d_i = 0$ for F_2 population. Here, a_i , d_i , and s_i are the additive effect, dominance effect, and the difference between a_i and d_i , respectively. The **LOD** column is the LOD scores transferred from the LR values under $LOD = 0.217 LR$.

The **A** column shows the additive genetic effects estimated at the testing points. The title **A** is used for DH, RIL or F_2 populations. If a BC population is used, it will become **S**. **S** is the difference between the additive effects and the dominance effects at the testing points; the additive effect and dominance effect are not separable in a BC population. It is noted that the signs of the additive effects (**A**) follow that of the parent called P1 in marker coding. Positive **A** values imply that the P1 parent takes positive value for the additive effect, while the P2 parent takes the negative. Negative **A** values mean that the P1 parent takes the negative value for the additive effect, while the P2

parent takes the positive.

The **Prob** column is the two-tail probabilities of student t values for testing the additive effects. This is used for testing the significance of the estimated additive effects.

If you are using a F_2 population, then two more columns will appear following the above: a **D** column and another **Prob** column. The **D** column gives the dominance effects estimated at the testing points, with the **Prob** column is for the significance of the dominance effects.

The layout of the mapping results in the .qtl file for the data set from multiple environments is basically the same. The only difference is that one more item will be included in the *description of conditions* about the hypothesis used for LR test (List 6). In addition, the **LR** and **LOD** values are obtained based on the hypothesis, when multiple environments are included.

List 6. An example of .qtl file for multiple environments: simu.qtl

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:       yes
// Replications:       no
// Contents: putative additive-effect QTLs
// BGV control method: A (control marker main & interaction effects)
// Results saved for: points significant at 0.005000
// LR test for hypothesis on: both genetic effects and GE interactions
```

```
# Date: 1999-03-12    Time: 12:13:07
```

```
Trait1: simTt
```

Chrom	Interval	Site1(M)	Site2(M)	LR	LOD	A	Prob
1	MK1-MK2	0.000	0.000	13.89	3.01	0.6931	0.0051
1	MK1-MK2	0.020	0.020	15.44	3.35	0.7434	0.0034
1	MK1-MK2	0.040	0.040	16.74	3.63	0.7798	0.0024
1	MK1-MK2	0.060	0.060	17.66	3.83	0.7991	0.0019
1	MK1-MK2	0.080	0.080	18.17	3.94	0.8003	0.0017
1	MK2-MK3	0.098	0.000	18.26	3.96	0.7869	0.0017
1	MK2-MK3	0.118	0.020	17.71	3.84	0.7914	0.0019
1	MK2-MK3	0.138	0.040	16.78	3.64	0.7803	0.0025

```
.....
End
```

File with eps extension: List 7 is an example of eps file for a data set from a single environment.

The part of *description of conditions* is quite similar to that for .qtl files. The only difference is the brief contents. For .qtl file, it is putative additive-effect QTLs, while for .eps file, it is putative epistatic QTLs.

The **result body** is also a trait-based two-way table. Trait names and strings for trait order numbers are also provided in the same way as in .qtl files. But in .eps files, more columns will be given to show the results for epistatic QTLs.

The **Ch-Ini** and **Ch-Inj** columns show pairs of marker intervals (expressed in order numbers) on which the pair-wise testing points are set. For example, 1-9 and 1-13 expressed a pair of marker intervals. The 1-9 interval is the **9th interval** (counted from the left end) on the **1st chromosome**; 1-13 interval is the **13th interval** on the **1st chromosome**. The order numbers of chromosomes are decided in the map file.

The **Int. Namei** and **Int. Namej** columns give the corresponding names of flanking markers for **Ch-Ini** and **Ch-Inj**.

The **Sitei(M)** and **Sitej(M)** columns give the genetic distances (in Morgan) of pair-wise testing points from the left markers of their own intervals. For example, Sitei = 0.00 M means that the testing point *i* is set exactly on the left marker of the interval.

The **LR** column shows the likelihood ratios for all the testing points. The hypothesis tested is: $H_0: a_i = a_j = aa_{ij} = 0$ for DH or RI populations; $H_0: s_i = s_j = ss_{ij} = 0$ for BC population. Here a_i , a_j , s_i , and s_j are the same as defined for .qtl file, while aa_{ij} and ss_{ij} are digenic additive \times additive epistasis and digenic interaction effects between s_i and s_j . The **LOD** column provides the equivalent LOD score values under $LOD = 0.217 LR$.

The **Ai**, **Aj** and **AAij** (for DH or RI populations) columns list the estimates of additive effect for testing point *i*, additive effect for testing point *j*, and additive by additive effects between the two testing points, respectively. The signs of **Ai** and **Aj** values are determined in the same way as that for .qtl files (see previous section). A positive **AAij** value implies that the two-locus genotypes being the same as those in P1 parent or P2 parent take the positive effects, while the two-locus genotypes of recombination between the P1 parent and P2 parent take the negative effects. The case of negative **AAij** values is just the opposite.

If you are using a BC population, the three columns will be named as **Si**, **Sj**, and **SSij** for giving the estimates of s_i , s_j and ss_{ij} of the two testing points, respectively.

The **Probi**, **Probj** and **Probij** columns provide the *t*-test probabilities for the corresponding estimates of the three genetic effects.

List 8 presents an example of .eps file for a data set from multiple environments. The layout of the mapping results for multiple environments will be basically the same as the above. The difference is that there is one more item in description of conditions in the .eps file for multiple environments: the hypothesis used for the LR test. The **LR** or **LOD** values are obtained for this hypothesis.

List 7. An example of .eps file for a single environment: dh.eps

```
// Data file name:      D:\QTLMAPER\dh.txt
// Marker map file name: D:\QTLMAPER\dh.map
// Environments:       no
// Replications:       no
// Contents: putative epistatic QTLs
// BGV control method: A (control marker main & interaction effects)
// Results saved for: points significant at 0.005000
```

```
# Date: 1999-01-29      Time: 21:20:34
```

Trait 1: ju28ph

Ch-Ini	Int.Namei	Sitei(M)	Ch-Inj	Int. Namej	Sitej(M)	LR	LOD	Ai	Probi	Aj
1-9	Mk9-Mk10	0.00	1-13	Mk13-Mk14	0.00	20.78	4.51	0.218	0.802	-0.225
1-9	Mk9-Mk10	0.00	1-13	Mk13-Mk14	0.02	22.03	4.78	0.254	0.768	-0.082
1-9	Mk9-Mk10	0.00	1-13	Mk13-Mk14	0.04	22.63	4.91	0.289	0.736	0.055
1-9	Mk9-Mk10	0.00	1-13	Mk13-Mk14	0.06	22.39	4.86	0.316	0.713	0.158
1-9	Mk9-Mk10	0.00	1-13	Mk13-Mk14	0.08	21.34	4.63	0.329	0.705	0.218
1-9	Mk9-Mk10	0.02	1-13	Mk13-Mk14	0.00	19.81	4.30	0.053	0.955	-0.050
1-9	Mk9-Mk10	0.02	1-13	Mk13-Mk14	0.02	21.12	4.58	0.127	0.892	0.150
1-9	Mk9-Mk10	0.02	1-13	Mk13-Mk14	0.04	21.82	4.74	0.207	0.825	0.329

.....

End

List 8. An example of .eps for multiple environments: simu.eps

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:        yes
// Replications:        no
// Contents: putative epistatic QTLs
// BGV control method: A (control marker main & interaction effects)
// Results saved for: points significant at 0.005000
// LR test for hypothesis on: both genetic effects and GE interactions

# Date: 1999-03-12      Time: 12:15:39
```

Trait 1: simTt

Ch-Ini	Int. Namei	Sitei(M)	Ch-Inj	Int. Namej	Sitej(M)	LR	LOD	Ai	Probi	Aj	Probj	AAij	Probij
1-1	MK1-MK2	0.00	1-10	MK10-MK11	0.00	18.94	4.11	0.689	0.0051	0.045	0.8588	0.2333	0.705
1-1	MK1-MK2	0.00	1-10	MK10-MK11	0.02	21.10	4.58	0.690	0.0045	0.0432	0.8597	0.2697	0.6966
1-1	MK1-MK2	0.00	1-10	MK10-MK11	0.04	23.50	5.10	0.690	0.0040	0.0414	0.8435	0.3022	0.6881
1-1	MK1-MK2	0.00	1-10	MK10-MK11	0.06	24.81	5.38	0.689	0.0036	0.0396	0.8605	0.3263	0.6799
1-1	MK1-MK2	0.00	1-10	MK10-MK11	0.08	25.47	5.53	0.688	0.0035	0.0383	0.8888	0.3393	0.6716
1-1	MK1-MK2	0.02	1-10	MK10-MK11	0.00	20.87	4.53	0.745	0.0033	0.0268	0.9156	0.2346	0.7134
1-1	MK1-MK2	0.02	1-10	MK10-MK11	0.02	23.04	5.00	0.744	0.0030	0.0248	0.9192	0.2705	0.7052
1-1	MK1-MK2	0.02	1-10	MK10-MK11	0.04	25.43	5.52	0.744	0.0027	0.0234	0.9123	0.3023	0.6970

.....
End

3. Filtration files (*.flq* and *.fle* files)

These files are created for simplifying the manual work of picking out putative additive-effect QTLs or epistatic QTLs from the original result files (*.qtl* and *.eps*). The results for the testing points with the local maximum LR values along chromosomes are filtrated and saved in these files. It is strongly suggested that the user check the filtrated results because some filtrated testing points may not be putative QTLs.

The constitutions of filtration files are the same for data sets with or without an environmental factor.

File with *.flq* extension: List 9 is an example of *.flq* file for additive-effect QTLs.

List 9. An example of *.flq* file: *simu.flq*

```
// Result file created by QTLMapper V1.0
// Contents: filtrated putative additive-effect QTLs
// Original result file: D:\QTLMAPER\simu.qtl
// BGV control method: A (control main & interaction markers)
// Threshold probability: 0.0050
// LR Threshold value: 7.8798

# Date: 1999-03-09      Time: 10:39:05

Trait 1: simTt

Chrom      Interval      Site1(M)      Site2(M)      LR      LOD      A      Prob
1          MK2-MK3         0.098         0.000         18.26   3.96     0.7869 0.0017
2          MK25-MK26      0.594         0.000         150.78  32.72    -0.9942 0.6711
2          MK32-MK33      1.262         0.040         247.09  53.62    3.6980 0.0000
.....

End
```

In this file, the *description of conditions* includes: brief contents of the file, name of the original result file, BGV control method, threshold probability used in the filtration, and corresponding LR critical value.

The *result body* is just a subset of that in the original *.qtl* file. In a *.flq* file, only testing points with local maximum LR values are retained. For detailed description, please see the previous section about the “*File with .qtl extension*”.

File with fle extension: List 10 provides an example of .fle file. The items included in the part of *description of conditions* are exactly the same as in .flq file. The *result body* is a subset of that in the original .eps file, containing the testing points with local maximum LR values. For detailed contents of the result body, please refer to the previous section on “*File with .eps extension*”.

4. Jackknife files (*.jkq* and *.jke* files)

These files are created from *Jackknife test* for genetic main effects and QE interaction effects when *multiple environments* are involved. These files provide both estimates of genetic main effects and predicted values of QE interaction effects only for the testing points presented as putative QTLs in filtration files.

File with jkq extension: This type of file is created from Jackknife test for additive-effect QTLs. List 11 is an example of .jkq file for multiple environments.

In the *description of conditions*, the following items are included:

Names of the map file and data file

Experimental designs for environments and replications

Brief contents of the .jkq file

Name of the filtration file on which the .jkq file is based

BGV control method used for mapping analysis for creating original result file

Threshold probability used for filtration of putative QTLs in the filtration file

Like in the previous sections, the layout of the *result body* is also trait-based. For each trait involved, trait order number (in such string as Trait 1:) and trait name are given first, followed by a two-way table that shows the results from Jackknife test for all the testing points listed in the filtration file.

The **Chrom** column gives the order number of chromosomes on which the putative QTLs are located.

The **Interval** column is the names of flanking marker intervals in which the putative QTLs are sited.

The **Site(M)** column gives the genetic distances (in Morgan) of the putative QTLs from the left marker on the interval.

The **Ai** column provides the Jackknife estimates for additive effects of the putative QTLs (DH or RIL populations). With a BC population, the title of the column will be **Si**, the difference of the additive effect and dominance effect for the putative QTL. The signs of **Ai** values are determined in the same way as that in .qtl files. The **Prob** column after this column is the probabilities for testing the estimates of additive effects.

If multiple environments are involved, more columns that give the results for QE interaction effects will appear after the above columns. In the example, the **AEix** columns (for DH or RI populations) shows the predicted values of interaction effects between putative QTL i and x -th environment. The order number of a specific environment is decided in the appearing order of the subscript for the environment in the data file. If a BC population is used, the columns will be titled **SEix**. Immediately following each of the **AEix** or **SEix** is a **Prob** column that gives the probability for testing the predicted values.

File with .jke extension: This type of file is created from Jackknife test for epistatic QTLs. List 12 is an example of .jke file for multiple environments.

The items included in *description of conditions* are the same as those in .jkq files.

In the *result body* of .jke file, the part represented by “.....” is the columns for **Ai**, **Probi**, **Aj**, **Probj**, **AAij**, and **Probij** for each trait. The meanings of these columns are the same as those for .eps files. In the presence of multiple environments, more columns for the results of QE interaction effects will be outputted after the “.....” part. In the example, **AEix** columns and **Prob** columns immediately after them are the predicted interaction effects between QTL i involved in the epistasis and x -th environment. The **Prob** columns provide significance tests for these effects.

List 10. An example of .fle file: simu.fle

```
// Result file created by QTLMapper V1.0
// Contents: filtrated putative epistatic QTLs
// Original result file: D:\QTLMAPER\simu.eps
// BGV control method: A (control main & interaction markers)
// Threshold probability: 0.0050
// LR Threshold value: 7.8798
```

```
# Date: 1999-03-09      Time: 17:13:35
```

Trait 1: simTt

Ch-Ini	Int. Namei	Sitei(M)	Ch-Inj	Int. Namej	Sitej(M)	LR	LOD	Ai	Probi	Aj	Proj	AAij	Probij
1-2	MK2-MK3	0.00	2-6	MK24-MK25	0.08	276.66	60.04	0.718	0.002	-1.314	0.560	1.092	0.699
1-2	MK2-MK3	0.00	2-14	MK32-MK33	0.02	303.28	65.81	0.740	0.002	3.643	0.000	1.149	0.447
1-2	MK2-MK3	0.02	3-8	MK41-MK42	0.10	234.64	50.92	0.828	0.001	-0.052	0.987	-0.374	0.694
2-6	MK24-MK25	0.06	3-8	MK41-MK42	0.10	352.99	76.60	-1.282	0.558	0.055	0.987	-1.600	0.166
2-7	MK25-MK26	0.00	2-14	MK32-MK33	0.02	375.86	81.56	-1.057	0.631	3.752	0.000	-0.810	0.538
2-14	MK32-MK33	0.02	3-9	MK42-MK43	0.02	433.92	94.16	3.628	0.000	-0.017	0.996	-0.670	0.806

.....
End

List 11. An example of .jkq file: simu.jkq

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:        yes
// Replications:         no
// Contents: Jackknife test results for additive-effect QTLs
// Jackknife based on: D:\QTLMAPER\simu.flq
// BGV control method: A (control marker main & interaction effects)
// Threshold probability: 0.005000
```

```
# Date: 1999-03-10      Time: 15:30:40
```

```
Trait 1: simTt
```

Chrom	Interval	Site(M)	Ai	Prob	AEi1	Prob	AEi2	Prob
1	MK2-MK3	0.00	0.791	0.0000	0.006	0.9454	-0.006	0.9449
2	MK25-MK26	0.00	-0.988	0.0000	-2.245	0.0000	2.231	0.0000
2	MK32-MK33	0.04	3.687	0.0000	0.393	0.2527	-0.393	0.2534

```
.....
End
```

List 12. An example of .jke file: simu.jke

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:       yes
// Replications:       no
// Contents: Jackknife test results for epistatic QTLs
// Jackknife based on: D:\QTLMAPER\simu.fle
// BGV control method: A (control marker main & interaction effects)
// Threshold probability: 0.005000
```

```
# Date: 1999-03-10      Time: 21:44:53
```

Trait 1: simTt

Ch-Ini	Int. Namei	Sitei(M)	Ch-Inj	Int. Namej	Sitej(M)	AEi1	Prob	...	AEj1	Prob	...	AAEij1	Prob	...
1-2	MK2-MK3	0.00	2-6	MK24-MK25	0.08	0.148	0.022	...	-2.349	0	...	2.769	0.000	...
1-2	MK2-MK3	0.00	2-14	MK32-MK33	0.02	0.214	0.000	...	0.387	0.175	...	1.480	0.000	...
1-2	MK2-MK3	0.02	3-8	MK41-MK42	0.10	0.000	1.000	...	3.184	0.000	...	-0.950	0.000	...
2-6	MK24-MK25	0.06	3-8	MK41-MK42	0.10	-2.105	0.000	...	3.297	0.000	...	-1.172	0.000	...
2-7	MK25-MK26	0.00	2-14	MK32-MK33	0.02	-2.320	0.000	...	0.326	0.287	...	1.332	0.010	...
2-14	MK32-MK33	0.02	3-9	MK42-MK43	0.02	0.426	0.129	...	3.198	0.000	...	-2.764	0.000	...

.....
End

Similarly, the **AE_jx** columns are the predicted interaction effects between the other QTL *j* involved in the epistasis and *x*-th environment. The **Prob** columns immediately after them provide significance tests for these effects. The **AAE_{ij}x** columns are the predicted interaction effects between the epistasis (between QTLs *i* and *j*) and *x*-th environment. The corresponding **Prob** columns are used for significance tests of these effects. If the population type is BC, the letter “A” in the titles of all these columns will be changed to “S”, as mentioned in previous sections.

5. Bayesian files (.byq and .bye files)

These files are created from *Bayesian test* for genetic main effects and QE interaction effects usually in the presence of multiple environments. The compositions of these files are quite similar to those of Jackknife files. Only the brief contents indicated in the *description of conditions* is different. Here it is “Contents: Bayesian test results for .?”, while in Jackknife files it is “Contents: Jackknife test results for .?”. List 13 and List 14 are examples of .byq and .bye files, respectively.

List 13. An example of .byq file: simu.byq

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:       yes
// Replications:       no
// Contents: Bayesian test results for additive-effect QTLs
// Bayesian test based on: D:\QTLMAPER\simu.qtl
// BGV control method: A (control marker main & interaction effects)
// Threshold probability: 0.005000
```

```
# Date: 1999-03-08      Time: 10:41:54
```

```
Trait 1: simTt
```

Chrm	Interval	Site(M)	Ai	Prob	AEi1	Prob	AEi2	Prob
1	MK2-MK3	0.00	0.787	0.000	0.001	0.994	-0.001	0.994
2	MK25-MK26	0.00	-0.994	0.000	-2.326	0.000	2.326	0.000
2	MK32-MK33	0.04	3.698	0.000	0.105	0.549	-0.105	0.549

```
.....
End
```

List 14. An example of .bye file: simu.bye

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:       yes
// Replications:       no
// Contents: Bayesian test results for epistatic QTLs
// Jackknife based on: D:\QTLMAPER\simu.fle
// BGV control method: A (control marker main & interaction effects)
// Threshold probability: 0.005000
```

```
# Date: 1999-03-10      Time: 21:44:53
```

Trait 1: simTt

Ch-Ini	Int. Namei	Sitei(M)	Ch-Inj	Int. Namej	Sitej(M)	AEi1	Prob	...	AEj1	Prob	...	AAEij1	Prob	...
1-2	MK2-MK3	0.00	2-6	MK24-MK25	0.08	0.034	0.835	...	-2.230	0.000	...	2.808	0.000	...
1-2	MK2-MK3	0.00	2-14	MK32-MK33	0.02	0.043	0.788	...	0.247	0.263	...	1.491	0.000	...
1-2	MK2-MK3	0.02	3-8	MK41-MK42	0.10	0.000	1.000	...	3.121	0.000	...	-0.916	0.001	...
2-6	MK24-MK25	0.06	3-8	MK41-MK42	0.10	-2.161	0.000	...	3.191	0.000	...	-1.112	0.000	...
2-7	MK25-MK26	0.00	2-14	MK32-MK33	0.02	-2.169	0.000	...	0.143	0.442	...	1.264	0.000	...
2-14	MK32-MK33	0.02	3-9	MK42-MK43	0.02	0.252	0.263	...	3.171	0.000	...	-2.724	0.000	...

.....
End

6. Contribution files (*.ctq* and *.cte* files)

These files are created for outputting the relative contributions of detected putative additive / epistasis effect to the phenotypic variation as well as the relative contributions of QE interaction effects when multiple environments are available. The calculation of relative contributions is different from that for R-squares in multiple regression analysis. A relative contribution here is calculated as the proportion of variance caused by a specific genetic source in the total phenotypic variance. So actually it could be taken as *heritability* contributed by that genetic source. That is why we use H^2 to stand for relative contributions.

The calculation of relative contributions is based on the results of putative QTLs / epistatic QTLs given in .flx, .jlx, or .byx files (x is q or e).

File with .ctq extension: List 15 is an example of ctq file that is created for multiple environments.

List 15. An example of .ctq file: simu.ctq

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:       yes
// Replications:       no
// Contents: relative contributions ( $H^2$ ) for putative additive-effect QTLs/epistatic QTLs
// Calculations based on: D:\QTLMAPER\simu.byq
// BGV control method: A (control marker main & interaction effects)
```

Date: 1999-03-11 Time: 16:47:34

Trait 1: simTt

Chrom	Interval	Site(M)	$H^2(A_i)$	$H^2(AE_i)$
1	MK2-MK3	0.00	0.0130	0.0000
2	MK25-MK26	0.00	0.0208	0.2277
2	MK32-MK33	0.04	0.2877	0.0000

General contributions:
 Additive(A): $H^2(A)=0.3015$; QE interactions: $H^2(AE)=0.2277$

.....

End

The *description of conditions* includes the following items:

Names of the map file and data file

Experimental design for environments and replications

Brief content of the file

Name of the file on which the calculation is based. Three types of files are possible: filtration file, Jackknife file, or Bayesian file.

BGV control method used for creating the original result file.

The **result body** is also trait-based. Trait order number and trait name are given first. Then there is a two-way table showing the relative contributions of the putative additive-effect QTLs, and the contributions of QE interactions if multiple environments are involved.

The **Chrom** column shows the order number of the chromosomes on which the putative QTLs are located.

The **Interval** column gives the names of the flanking markers constructing the intervals in which the putative QTL are detected.

The **Site(M)** column provides the genetic distances (in Morgan) of the putative QTLs from the left markers on the intervals.

The **H²(Ai)** column is the relative contributions from the additive effects of the putative QTLs. If the mapping population is BC, the title will be **H²(Si)**, and the relative contributions are from the differences between the additive effects and dominance effects for the putative QTLs.

The **H²(AEi)** column is the relative contributions from the interaction effects between the additive effect at putative QTL *i* and the environments. With a BC population, it will be **H²(SEi)**. It is noted that there will be no such column if no environments are involved in the experiments.

At the bottom of each two-way table are the **General contributions**. **H²(A)** is the general contribution from the additive effects of all the putative QTLs involved. With BC population it will be **H²(S)**. **H²(AE)** is the general contributions from all the additive by environment interaction effects if multiple environments are involved.

File with .cte extension: List 16 is an example of .cte file that is created for multiple environments.

The items included in the part of *description of conditions* are the same as those in .ctq files.

List 16. An example of .cte file: simu.cte

```
// Result file created by QTLMapper V 1.0
// Data file name:      D:\QTLMAPER\simu.txt
// Marker map file name: D:\QTLMAPER\simu.map
// Environments:        yes
// Replications:        no
// Contents: relative contributions (H^2) for putative additive-effect QTLs/epistatic QTLs
// Calculations based on: D:\QTLMAPER\simu.by
// BGV control method: A (control marker main & interaction effects)
```

```
# Date: 1999-03-11      Time: 16:54:05
```

Trait 1: simTt

Ch-Ini	Int. Namei	Sitei(M)	Ch-Inj	Int. Namej	Sitej(M)	H^2(Ai)	H^2(Aj)	H^2(AAij)	H^2(AEi)	H^2(AEj)	H^2(AAEij)
1-2	MK2-MK3	0.00	2-6	MK24-MK25	0.08	0.0026	0.0087	0.0060	0.0000	0.0499	0.0791
1-2	MK2-MK3	0.00	2-14	MK32-MK33	0.02	0.0027	0.0666	0.0066	0.0000	0.0003	0.0223
1-2	MK2-MK3	0.02	3-8	MK41-MK42	0.10	0.0034	0.0000	0.0000	0.0000	0.0978	0.0084
2-6	MK24-MK25	0.06	3-8	MK41-MK42	0.10	0.0082	0.0000	0.0128	0.0469	0.1022	0.0124
2-7	MK25-MK26	0.00	2-14	MK32-MK33	0.02	0.0056	0.0706	0.0033	0.0472	0.0001	0.0160
2-14	MK32-MK33	0.02	3-9	MK42-MK43	0.02	0.0661	0.0000	0.0023	0.0000	0.1009	0.0745

General contributions:

```
Additive(A): H^2(A)=0.0852; Epistasis: H^2(AA)=0.0302
QE Interactions: H^2(AE)=0.5758; H^2(AAE)=0.2205
```

.....

End

In the *result body*, the layout of the results for the relative contributions is similar to that for .ctq files.

The **Ch-Ini** and **Chi-Inj** columns provide the order numbers for the chromosomes and intervals on which the pair of involved putative QTLs are located. The **Int. Namei** and **Int. Namej** columns give the flanking marker names of the two involved intervals.

The **Sitei(M)** and **Sitej(M)** columns show the genetic distances of the two putative QTLs from the left markers on their own intervals.

The **H²(Ai)** and **H²(Aj)** columns give the relative contributions of the two putative QTLs *i* and *j* involved in the epistasis. The **H²(AAij)** column is the contributions of the epistasis. For a BC mapping population, the titles will be **H²(Si)**, **H²(Sj)**, and **H²(SSij)**, respectively.

For the case of multiple environments, more columns will be available for the contributions of QE interactions. In this example, the **H²(AEi)** and **H²(AEj)** columns provides the contributions of additive by environment interaction effects for the two putative QTLs, while the **H²(AAEij)** is the contributions of epistasis by environment interactions. For a BC mapping population, the titles will be **H²(SEi)**, **H²(SEj)**, and **H²(SSEij)**, respectively. It is noted that there will be no those columns if no environmental factor is involved.

At the bottom of each two-way table is **General contributions**. The value for **H²(A)** is the general contribution from all the additive effects of the identified putative QTLs. The **H²(AA)** value is the general contribution from all the epistatic effects of the identified putative QTLs. **H²(AE)** and **H²(AAE)** provide the general contributions from all the additive by environment interactions and all the epistasis by environment interactions, respectively if an environmental factor is involved. For a BC population, likewise, the letter “A” in the above expressions will be changed to “S”.

7. LR Graph (2D) files (.plt, .sig, .001 files, etc.)

These files are created from the original .qtl file. They are only used for displaying or printing two-dimensional graphs of LR_s against testing positions (in Morgan) for the results from mapping additive-effect QTLs. The user usually does not need to check these files for details. In the section of Working with Output Sub-Menu, we have briefly described the contents of these files, so we here will talk about how to use these files to display your main-effect-QTL mapping results.

You need to install the **Wgnuplot** application, a free software for Windows, in your computer before you view the LR graphs. Here we copy the statement of the copyright shown in the help of the software:

Copyright (C) 1986 - 1993 Thomas Williams, Colin Kelley

Permission to use, copy, and distribute this software and its documentation for any purpose with or without fee is hereby granted, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation.

Permission to modify the software is granted, but not the right to distribute the modified code. Modifications are to be distributed as patches to released version.

This software is provided "as is" without express or implied warranty.

AUTHORS

Original Software: Thomas Williams, Colin Kelley.
 Gnuplot 2.0 additions: Russell Lang, Dave Kotz, John Campbell.
 Gnuplot 3.0 additions: Gershon Elber and many others.

We have packed this software with QTLMapper 1.0 in the compressed file **qtlmapr.exe**. After you unzip the files in **qtlmapr.exe**, there will be a directory named “Wgnuplot” as a sub-directory of “QTLMapper”. All the necessary files for **Wgnuplot** are saved in this directory.

Viewing LR graphs: In Windows Explorer, open the folder “Wgnuplot” in folder “QTLMapper”. Double-click on **wgnuplot.exe** to open a **Wgnuplot** window (Figure 39). Press Open button or select Open item in File menu to bring up a Load dialog box. Select the drive and the work directory in which your graph files are saved.

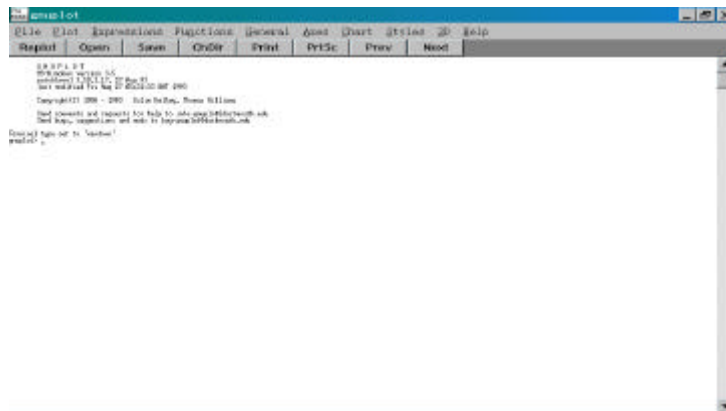


Figure 39. Windows interface of **Wgnuplot**

Find the file with .plt extension, and double-click on it. Then a LR graph will appear in another window for the first trait and first chromosome. To show next graph, press Enter or click on the OK button in a “gnuplot pause” dialog box. The following graph is presented as an example.

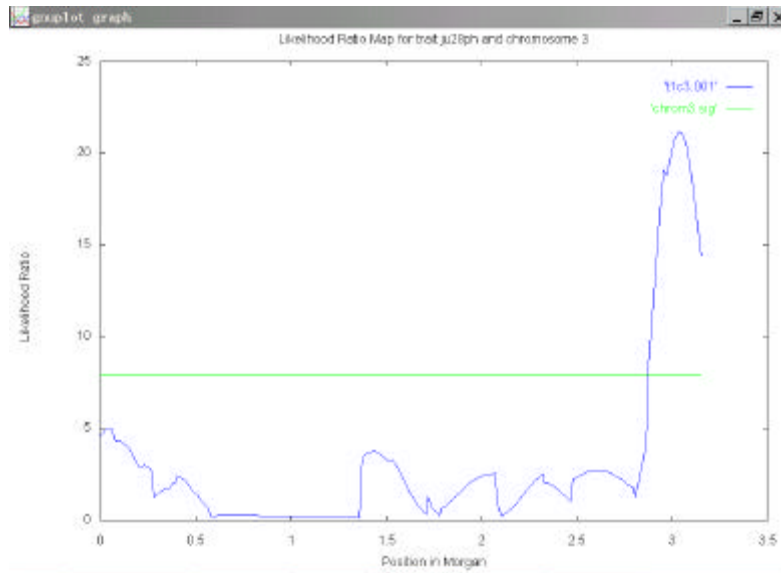


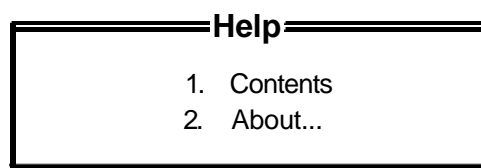
Figure 40. An example of LR graph

Each LR graph has a title like “Likelihood Ratio Map for trait **ju28ph** and chromosome **3**”. The string **ju28ph** is the name of the trait, while the number **3** is the order of the chromosome. The x-axis is the positions (in Morgan) of the testing points from the left end of the chromosome. The y-axis is LR values for each testing point. The horizontal line is the threshold line whose height is decided by the threshold probability specified in Settings sub-menu. The curve line is the LR line showing the relation between the testing points and the LR values. The peak points above the threshold line could be taken as putative additive-effect QTLs.

Getting On-line Help

The sub-menu Help in Main Menu is designed to provide on-line help. The user can get on-line help by pressing number “5” in Main Menu and then “1” in Help sub-menu (Figure 38). The program will load a text file named **Qtlmapper.hlp** and display it. By pressing PgDn, PgUp, Home, or End keys, one can view next, previous, first, or last pages. To view a specific page, press letter “P” and then page number. To quit help, press Esc key.

The About...item in the sub-menu is for showing the About information of the software.



1 -> 2 =>Select Esc => Back

Figure 41. Help sub-menu

Error Messages and Troubleshooting

If you have any questions with our software, please contact Jun Zhu, Department of Agronomy, Huajiachi Campus, Zhejiang University, Hangzhou, Zhejiang 310029, China. Email: jzhu@zjau.edu.cn.

References

- Zhu J. (1998). Mixed model approaches of mapping genes for complex quantitative traits. In: Wang LZ and Dai JR (eds) Proceedings of Genetics and Crop Breeding of China, Chinese Agricultural Science and Technology Publication House, Beijing, pp19-20
- Zhu J. and B. S. Weir. (1998). Mixed model approaches for genetic analysis of quantitative traits. In: Chen LS, Ruan SG, and Zhu J (eds) Advanced Topics in Biomathematics: Proceedings of International Conference on Mathematical Biology, World Scientific Publishing Co., Singapore, pp321~330
- Wang, D. L., J. Zhu, Z. K. Li, and A. H. Paterson. (1999). Mapping QTLs with epistatic effects and QTL \times environment interactions by mixed linear model approaches. Theor. Appl. Genet., in printing.