

# **User Manual for**

# **QTL.gCIMapping.GUI**

**QTL genome-wide Composite Interval Mapping GUI**

**(version 1.0)**

**Zhang Ya-Wen, Wen Yang-Jun, Wang Shi-Bo,  
Zhang Yuan-Ming (soyzzhang@mail.hzau.edu.cn)**

**Last updated on June 12, 2018**

**Disclaimer:** While extensive testing has been performed by Yuan-Ming Zhang's Lab (Statistical Genomics Lab) at Huazhong Agricultural University, the results are, in general, reliable, correct or appropriate. However, results are not guaranteed for any specific datasets. We strongly recommend that users validate the GCIM results with other software packages, such as **Windows QTL Cartographer V2.5\_011** (<https://brcwebportal.cos.ncsu.edu/qtlcart/WQTLCart.htm>) and **QTL IciMapping V4.1** (<http://www.isbreeding.net/software/?type=detail&id=18>).

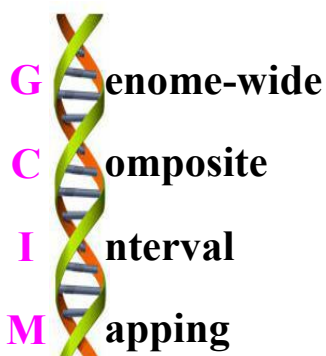
**Download website:**

<https://cran.r-project.org/web/packages/QTL.gCIMapping.GUI/index.html>

**References**

1. Wang Shi-Bo, Wen Yang-Jun, Ren Wen-Long, Ni Yuan-Li, Zhang Jin, Feng Jian-Ying, Zhang Yuan-Ming\*. Mapping small-effect and linked quantitative trait loci for complex traits in backcross or DH populations via a multi-locus GWAS methodology. *Scientific Reports* 2016, 6: 29951.
2. Wen Yang-Jun, Zhang Ya-Wen, Zhang Jin, Feng Jian-Ying, Jim M. Dunwell, Zhang Yuan-Ming\*. An efficient multi-locus mixed model framework for the detection of small and linked QTLs in F<sub>2</sub>. *Briefings in Bioinformatics*, 2018, Accepted

## Quantitative Trait Loci



**Funding**

The works were supported by the National Natural Science Foundation of China (31571268, 31701071 and U1602261), Huazhong Agricultural University Scientific & Technological Self-innovation Foundation (Program No. 2014RC020), and State Key Laboratory of Cotton Biology Open Fund (CB2017B01).

## INTRODUCTION

### 1.1 Why GCIM?

**QTL.gCIMapping.GUI** (**QTL** **G**enome-wide **C**omposite **I**nterval **M**apping **G**raphical **U**ser **I**nterface) is an R package for multi-QTL mapping of quantitative traits in bi-parental segregation populations. QTL.gCIMapping.GUI v1.0 is able to work on the popular platforms, like Windows, Linux (desktop) and MacOS.

### 1.2 Getting started

QTL.gCIMapping.GUI is a package that runs in the R software environment, which can be freely downloaded from <https://cran.r-project.org/web/packages/QTL.gCIMapping.GUI/index.html>, or request from the maintainer, Dr Yuan-Ming Zhang at the College of Plant Science and Technology, Huazhong Agricultural University ([soyzzhang@mail.hzau.edu.cn](mailto:soyzzhang@mail.hzau.edu.cn) or [soyzzhang@hotmail.com](mailto:soyzzhang@hotmail.com)).

#### 1.2.1 One-Click installation

Within R environment, the QTL.gCIMapping.GUI software can be installed directly using the below command:

```
install.packages(pkgs="QTL.gCIMapping.GUI")
```

#### 1.2.2 Step-by-step installation

##### 1.2.2.1 Install the add-on packages

**Online installation** Within R environment on the internet, the R package QTL.gCIMapping.GUI can be installed online, using the below command:

```
install.packages(pkgs=c("shiny","qtl","doParallel","foreach","iterators","openxlsx","MASS","stringr","parcor","data.table","Rcpp"))
```

**Offline installation** Users should download the below 38 packages from CRAN, github (<https://github.com/>), or google search:

```
"cmprsk","corpcor","data.table","digest","doParallel","Epi","etm","fdrtool","foreach","GeneNet","glmnet","htmltools","httpuv","iterators","jsonlite","later","longitudinal","magrittr","MASS","mime","numDeriv","openxlsx","parcor","plyr","ppals","promises","qtl","R6","Rcpp","shiny","sourcetools","stringi","stringr","testthat","utf8","xtable","zip","zoo"
```

Then, install them offline (under the R environment, select all the 38 packages and

install them offline).

### 1.2.2.2 Install QTL.gCIMapping.GUI

Open R GUI, select "Packages"—"Install package(s) from local files..." and then find the QTL.gCIMapping.GUI package which you have downloaded on your desktop. Within R environment, launch QTL.gCIMapping.GUI by command:`library(QTL.gCIMapping.GUI)` and start it by command:`QTL.gCIMapping.GUI()`

**User Manual file** Users can decompress the QTL.gCIMapping.GUI package and find the User Manual file (name: **Instruction.pdf**) in the folder of ".../QTL.gCIMapping.GUI/inst/doc".

## 2. Dataset format

**GCIM format for Dataset** The first three columns, named "**marker**", "**chr**" and "**pos**", stand for marker name, chromosome and marker position (cM) on the chromosome, respectively. Among the remaining columns, each column lists all the genotypes of one individual while the first row for each column shows the individual name. For the genotypes of each marker, the coding criteria are shown as [Table 1](#). The phenotype and covariate information are followed the marker genotypes, and each covariate or trait are listed on one row. On each row, the first column is empty followed by "**trait1**", "real trait name", and "phenotypic values for all the individuals". If multiple traits exist, more rows will be added. If covariates exist, all the information for the covariates list after the trait information. The format is seen in [Table 1](#). If there is no covariate, users should delete the last row in [Table 2](#).

**Table 1. Coding criteria for GCIM format**

Marker genotype	Code	Meaning
AA	A	Homozygous genotype (P <sub>1</sub> )
Aa	H	Heterozygous genotype (F <sub>1</sub> )
aa	B	Homozygous genotype (P <sub>2</sub> )
Not AA (Aa + aa)	C	Dominance to P <sub>2</sub>
Not aa (AA + Aa)	D	Dominance to P <sub>1</sub>
Missing	-	Missing or unclear genotype

**Table 2. The GCIM format of the dataset**

marker	chr	pos	DH6-10	DH6-101	DH6-102
RGA3(1)	1	0	B	-	B
wPt-6358	1	3.034	B	-	-
Hplc2	1	8.8291	A	A	B
wPt-9752	1	10.1452	A	-	-
abc156a	1	41.3408	A	A	B
:	:	:	:	:	:
gwm437	21	162.5218	A	B	-
gwm121	21	180.2878	A	B	-
wmc157	21	197.9196	A	B	A
*stm1actc	21	200.4216	-	-	-
	trait1	T19	75.33	105	96.33
	trait2	T191	74	105.68	97.16
	trait3	T192	75.37	104.67	95.55
	Covar1	CovarName	A	B	B

**ICIM format for Dataset** If users have the dataset files with QTL IciMapping format, these files are also available in our software. Details can be seen in the folder of “.../QTL.gCIMapping.GUI/inst/extdata”, i.e., [WheatDH\\_QTLIciMapping\\_Format.xlsx](#).

**WinQTLCart format for Dataset** If users have the dataset file with WinQTLCart format, its file is also available in our software. Details can be seen in the folder of “.../ QTL.gCIMapping.GUI/inst/extdata”, i.e., [env1-jun3\\_WinQTLCart\\_Format.mcd](#).

**The dataset fileICIMcov format** If users select ICIM format and the covariate exists in the dataset, it needs to input a covariate file. In the file, the first column indicates individual name and the second column is the covariate information ([Table 3](#)). In [Table 3](#), the covariate values are A, B and C.

**Table 3. The covariate file format**

Individual ID	Covariate
DH6-10	A
DH6-101	A
DH6-102	A
DH6-104	A
DH6-164	B
DH6-165	B
DH6-166	B
DH6-170	B
DH7-124	C
DH7-125	C

### 3. Operation process

#### 3.1 The graphical interface of QTL.gCIMapping.GUI

GCIM

Start

QTL.gCIMapping (QTL genome-wide Composite Interval Mapping)

**Coding criteria**

Genotype	Code	Meaning
AA	A	Homozygous genotype (P1)
Aa	H	Heterozygous genotype (F1)
aa	B	Homozygous genotype (P2)
AA+Aa(Not aa)	D	Dominance to P1
Aa+aa(Not AA)	C	Dominance to P2
Missing	-	Missing or unclear genotype

**Dataset example**

marker	chr	pos	DH6.10	DH6.101	DH6.102
RGA3(1)	1	0	B	-	B
wPt-6356	1	3.034	B	-	-
Hplc2	1	8.8291	A	A	B
...	...	...	...	...	...
gwm437	21	162.5216	A	B	-
gwm121	21	180.2878	A	B	-
wmc157	21	197.9196	A	B	A
trait1	T19	75.33	105	96.33	
trait2	T191	74	105.68	97.16	
Covar	CovarName	A	B	B	

**Reference**

1. Wang Shi-Bo, Wen Yang-Jun, Ren Wen-Long, Ni Yuan-Li, Zhang Jin, Feng Jian-Ying, Zhang Yuan-Ming\*. Mapping small-effect and linked quantitative trait loci for complex traits in backcross or DH populations via a multi-locus GWAS methodology. Scientific Reports 2016.6:29951.
2. Wen Yang-Jun, Zhang Ya-Wen, Zhang Jin, Feng Jian-Ying, Jim M. Dunwell, Zhang Yuan-Ming\*. An efficient multi-locus mixed model framework for the detection of small and linked QTLs in F2. Submitted

Authors: Zhang Ya-Wen, Wen Yang-Jun, Wang Shi-Bo, Zhang Yuan-Ming  
Maintainer: Zhang Yuan-Ming (soy Zhang at mail.hzau.edu.cn)  
QTL.gCIMapping version 1.0, Released April 2018

**Figure 1. The Graphical User Interface of QTL.gCIMapping.GUI**

#### 3.2 Input dataset

Users must upload the dataset files with three formats (Figs 2 to 4). If users select the QTL IciMapping format and the covariate exists in the dataset, users should upload the

covariate matrix (**Fig 5**).

QCIM

Start

## QTL.gCIMapping

Please select data format

☒ GOM
 ☐ WinQTLcart
 ☐ QTL IciMapping

Input dataset

Browse...

GOM\_Format\_DH.csv

Upload image

Show dataset:

Genotype

Parameter Settings

Figure

User manual

Dataset

Parameter Settings

Figure

Show

25

entries

Search:

marker	DH6-10	DH6-101	DH6-102	DH6-104	DH6-105	DH6-108	DH6-11	DH6-111	DH6-112	DH6-114	DH6-119	DH6-124	DH6-125	DH6-128	DH6-129
RG3A3(1)	B	-	B	A	B	B	A	A	A	-	B	B	B	A	B
wPt-6358	B	-	-	-	-	B	A	A	A	-	B	-	-	A	-
tpic2	A	A	B	A	B	B	A	A	A	B	B	B	B	A	B
wPt-9752	A	-	-	-	-	-	A	A	A	-	B	-	-	A	B
idc156a	A	A	B	A	B	B	B	B	A	B	-	B	B	A	B
RG360(2)	A	-	B	A	-	B	B	B	A	-	B	B	B	A	B
bc098	B	A	B	A	B	B	B	B	A	B	A	B	B	A	B
wmc24	B	A	B	A	B	B	B	B	A	B	A	B	B	A	B
ksuG9c	B	A	B	A	B	B	B	B	A	B	A	B	B	A	B
wPt-2436	B	-	-	-	-	B	B	B	B	-	A	B	-	B	-
wPt-4095	B	-	-	-	-	B	B	B	B	-	A	B	-	B	B
wmc129	B	A	B	A	B	B	B	B	B	A	A	B	B	B	B
cdo105	B	A	B	A	B	B	B	B	B	A	A	B	B	B	B
wPt-6074	B	-	-	-	-	B	B	B	B	-	A	B	-	B	-
GluA1	B	A	B	A	B	A	B	B	B	A	A	B	B	B	B
bcd080b	B	A	A	B	B	A	B	A	B	B	A	B	B	B	A

### Fig 2. Dataset GCIM format

### gCIMapping

Please select data format

☒ VCF

☐ WIG

☐ WIG+VCF

Input dataset

Browse: chr1.p10\_WinQTL-Cat\_Format.mcd

Load from cache

Show dataset

Genotype

Parameter Settings

Figure

Use results

Dataset	Parameter Settings	Figure
#chr10 1146697100 Mychromosome	-type interval -function l -suits chr -chromosome 3 -maxium 20 -used yes -start -Chromosome Chr-1 chr1 10.0000 chr2	

### Fig 3. Dataset WinQTLCart format

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20
ROA3(1)	0	-1	0	2	0	0	2	2	2	-1	0	0	0	2	0	2	2	2	0	2
WFI-6208	0	-1	-1	-1	-1	0	2	2	2	-1	0	-1	-1	2	-1	-1	2	2	0	2
WFI-2	2	2	0	2	0	0	2	2	2	0	0	0	0	2	0	2	2	2	0	2
WFI-4952	2	-1	-1	-1	-1	2	2	2	2	-1	0	-1	-1	2	0	-1	2	2	-1	2
Abx1564	2	2	0	2	0	0	0	0	2	0	-1	0	0	2	0	2	2	0	0	0
ROA3(bv2)	2	-1	0	2	-1	0	0	0	2	-1	0	0	0	2	0	2	2	0	0	0
bcd88	0	2	0	2	0	0	0	0	2	0	2	0	0	2	0	2	2	0	2	0
WFI-24	0	2	0	2	0	0	0	0	2	0	2	0	0	2	0	2	2	0	2	0
low29k	0	2	0	2	0	0	0	0	2	0	2	0	0	2	0	2	2	0	2	0
WFI-2436	0	-1	-1	-1	-1	0	0	0	0	-1	2	0	-1	0	-1	-1	2	0	2	0
WFI-4956	0	-1	-1	-1	-1	0	0	0	0	-1	2	0	-1	0	0	-1	2	0	2	0
WFI-120	0	2	0	2	0	0	0	0	0	2	2	0	0	0	0	0	2	2	0	2
cdv165	0	2	0	2	0	0	0	0	0	2	2	0	0	0	0	0	2	2	0	2
WFI-4373	0	-1	-1	-1	-1	0	0	0	0	-1	2	0	-1	0	-1	-1	-1	0	2	0
CluA1	0	2	0	2	0	2	0	0	0	2	2	0	0	0	0	2	0	0	2	0
bcd868b	0	2	2	0	0	2	0	2	0	0	2	0	0	0	2	2	0	0	2	0
WFI-2	0	2	2	0	0	2	0	2	0	-1	2	0	2	0	2	2	0	-1	2	0

Fig 4. Dataset QTLgCIMapping format

V1	V2
Individual ID	Covariate
DH6-10	A
DH6-101	A
DH6-102	A
DH6-104	A
DH6-105	A
DH6-106	A
DH6-11	A
DH6-111	A
DH6-112	A
DH6-114	A
DH6-119	A
DH6-124	A
DH6-125	A
DH6-128	A
DH6-129	A
DH6-13	A

Fig 5. Covariate input in the QTLgCIMapping dataset format

### 3.3 Parameter settings (Fig 6)

**Select population:** BC1 ( $F1 \times P1$ ), BC2 ( $F1 \times P2$ ), DH, RIL, and F2.

**Select model:** Random or Fixed model for QTL effects.

**Walk Speed for Genome-wide Scanning (cM):** Set walk speed for genome-wide scanning (centi-Morgan, cM), for example, 1 cM.

**Critical LOD score:** Critical LOD scores for significant QTL, for example, 2.5 or 3.0.

**Likelihood function:** This parameter is only for F2 population, including restricted maximum likelihood (REML) and maximum likelihood (ML).

**Completing CIM in one neighborhood:** This parameter is only for F2 population. In the



first running, please set "FALSE". If the other software detects only one QTL in a neighborhood but the current software finds two linked QTLs (one with additive effect and another with dominant effect) in the [neighborhood](#), please set "TRUE" and run again.

**Draw plot or not:** This parameter setup includes FALSE and TRUE. "FALSE" indicates no figure output, and "TRUE" indicates the output of QTL mapping curve, for example, the LOD score [or  $-\log_{10}(P\text{-value})$ ] curve against genome position.

**Resolution of plot:** Low or High: the low or high resolution for the figure file.

**Plot format:** Users can download the picture for different file formats: \*.jpeg, \*.png, \*.tiff and \*.pdf.

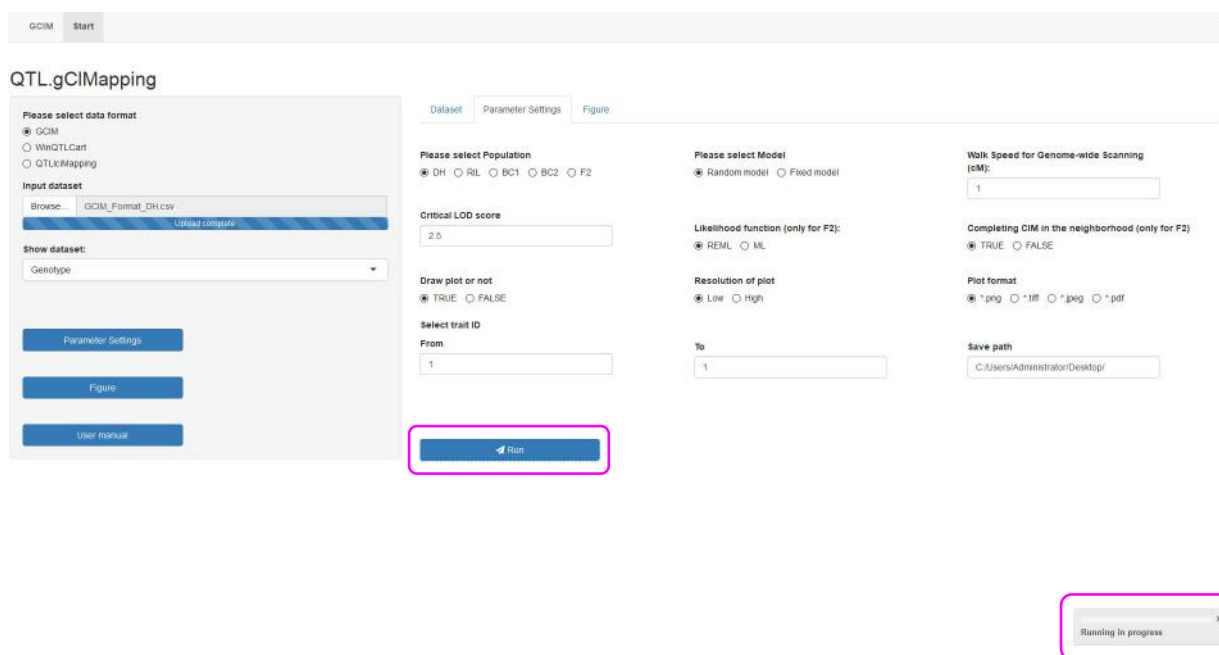
**Select trait ID:** “2:2” indicates the analyses from the second trait, and “2:4” indicates the analyses from the second to fourth traits.

**Save path:** The result will be written to the path in your computer.

The screenshot shows the QTL.gCIMapping software interface. On the left, there is a sidebar with buttons for 'Parameter Settings', 'Figure' (highlighted with a pink rectangle), and 'User Manual'. The main area is divided into three tabs: 'Dataset', 'Parameter Settings', and 'Figure'. The 'Figure' tab is currently selected. It contains several configuration options: 'Please select Population' (with radio buttons for D1, RL, BC1, BC2, F2), 'Please select Model' (with radio buttons for Random model and Fixed model), 'Walk Speed for Genome-wide Scanning (cm)' (a text input field with '1'), 'Critical LOD score' (a text input field with '2.5'), 'Likelihood function (only for F2)' (with radio buttons for REML and ML), 'Completing CIM in the neighborhood (only for F2)' (with radio buttons for TRUE and FALSE), 'Draw plot or not' (with radio buttons for TRUE and FALSE), 'Resolution of plot' (with radio buttons for Low and High), 'Plot format' (with radio buttons for \*.png, \*.tiff, \*.jpeg, and \*.pdf), 'Select trait ID' (with 'From' and 'To' text input fields, both containing '1'), and 'Save path' (a text input field containing 'C:/Users/Administrator/Desktop/'). A 'Run' button is located at the bottom center.

**Fig 6. Parameter setting in the mapping of QTL for quantitative traits**

### 3.4 Run the software



**Fig 7. Run the software QTL.gCIMapping.GUI**

### 3.5 Re-draw the plot according to your own requirement

When users finish the running, users get the resultforplot.xlsx file. With this file information, users may redraw the curve figure  $\{\text{LOD score or } -\log_{10}(P\text{-value})\}$ . With this Figure module, users may set all the figure parameters (**Fig 8**), including

**Legend and tick marks:** the size of the words in axis.

**LOD line size:** the size of the LOD line, the larger the coarse.

**Size for  $-\log_{10}(P\text{-value})$  curve:** the size of  $-\log_{10}(P\text{-value})$  curve, the larger the coarse.

**Margin space:** the space between the figure and the margin of the paper.

**Critical LOD score:** The critical LOD score for significant QTL.

Before saving this Figure, please set the related parameters: **width** and **height** [with the unit of pixel (px)], **word resolution** [with the unit of 1/72 inch, being pixels per inch (ppi)], and **figure resolution** [with the unit of pixels per inch (ppi)]. Users may set the colors for the LOD line color and  $-\log_{10}(P\text{-value})$  curve, with a drop-down option. Use Download plot button to choose a path and to save the Figure, with four frequently used image formats: \*.png, \*.tiff, \*.jpeg and \*.pdf (**Fig 9**).

QTL.gCIMapping

Please select data format

- ☒ GCM
- ☐ WinQTL Cart
- ☐ QTLxMapping

Input dataset

Browse... GCM\_Format\_F2.csv Upload dataset

Show dataset: Genotype

Parameter Settings

Figure

User manual

Dataset: Parameter Settings: Figure

Genome-wide composite interval mapping (GCIM) figure

☒ Parameter Settings ☐ Draw plot

Select resolution of plot

☒ General resolution ☐ High resolution ☐ Set by yourself

Width (px): 1500 Height (px): 600 Word resolution (1/72 inch, ppi): 12 Figure resolution (ppi): 72

Legend and tick marks: 1.0 LOD line size: 1.0 Size for -log10(P) curve: 0.5 Margin space: 1.5

Space between tick marks and axis: 1.0 Times for max(-log10(P)): 1.5 Critical LOD score: 2.5

LOD line color: red -log10(P) curve color: gray50 -log10(P) curve color2 (only for F2): green

Figure 8. Parameter settings

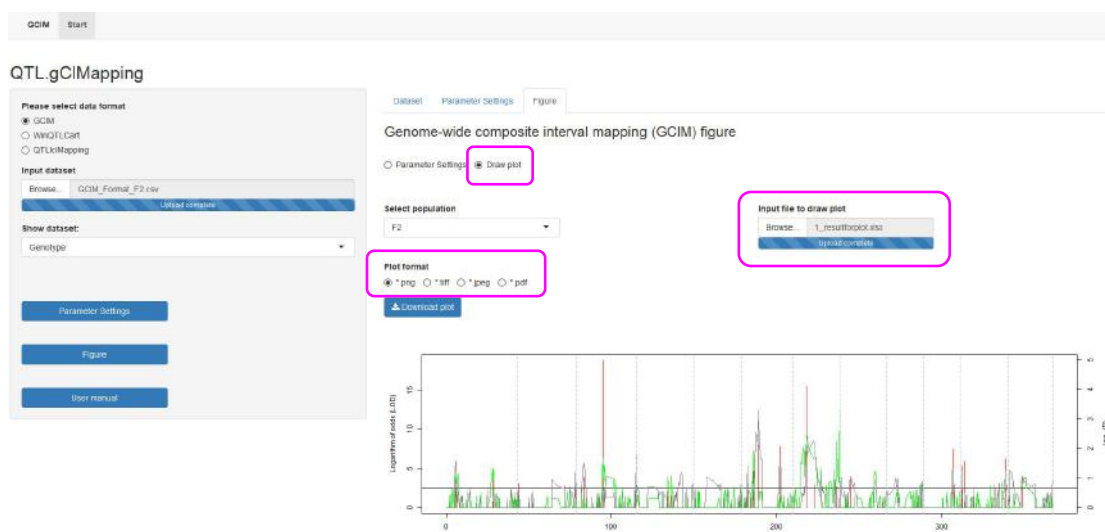


Fig 9. How to draw the QTL mapping figure

#### 4. Result

For BC1, BC2, DH and RIL populations, the **Results** file has ten columns, as shown below.

**Trait:** The trait name analyzed.

**Chr:** Chromosome, represented by an integer number.

**Position (cM):** The QTL position (cM) on the chromosome.

**Additive Effect:** Additive effect for significant QTL.

**LOD:** LOD score for significant QTL.

**Left\_Marker:** Left flanking marker name for significant QTL.

**Right\_Marker:** Right flanking marker name for significant QTL.

**Var\_Genet:** Genetic variance for each significant QTL.

**r<sup>2</sup> (%)**: Proportion of phenotypic variance explained by single QTL.

**Var\_Error**: residual variance for the full model.

**Var\_Phen (total)**: Phenotypic variance in the analyzed population.

For F<sub>2</sub> population, the **Results** file has eleven columns. Trait, Chr, Position (cM), Left\_Marker, Right\_Marker, Var\_Genet, LOD, r<sup>2</sup> (%), Var\_Error and Var\_phen are same as those in the above populations. The different columns are as follows.

**Effect.a** and **Effect.d**: Additive and dominant effects for significant QTL.