Ver. 1.0

November 2024

Prepared by:

Ishikawa Prefectural University

Nonoichi, Ishikawa, Japan

introduction

QTL-seq GUI

QTL-seq GUI は、異なる品種間交雑から得られた集団内で特定の表現型を示す個体に共通する遺伝子領域を、全ゲノムシーケンスデータに基づいて迅速に同定するためのツール。
本ツールに関する問題報告やご要望は、GitHub (https://github.com/sakemiki/QTL-seq_GUI_ver1) の「Issues」または「Pull Request」からご連絡ください。

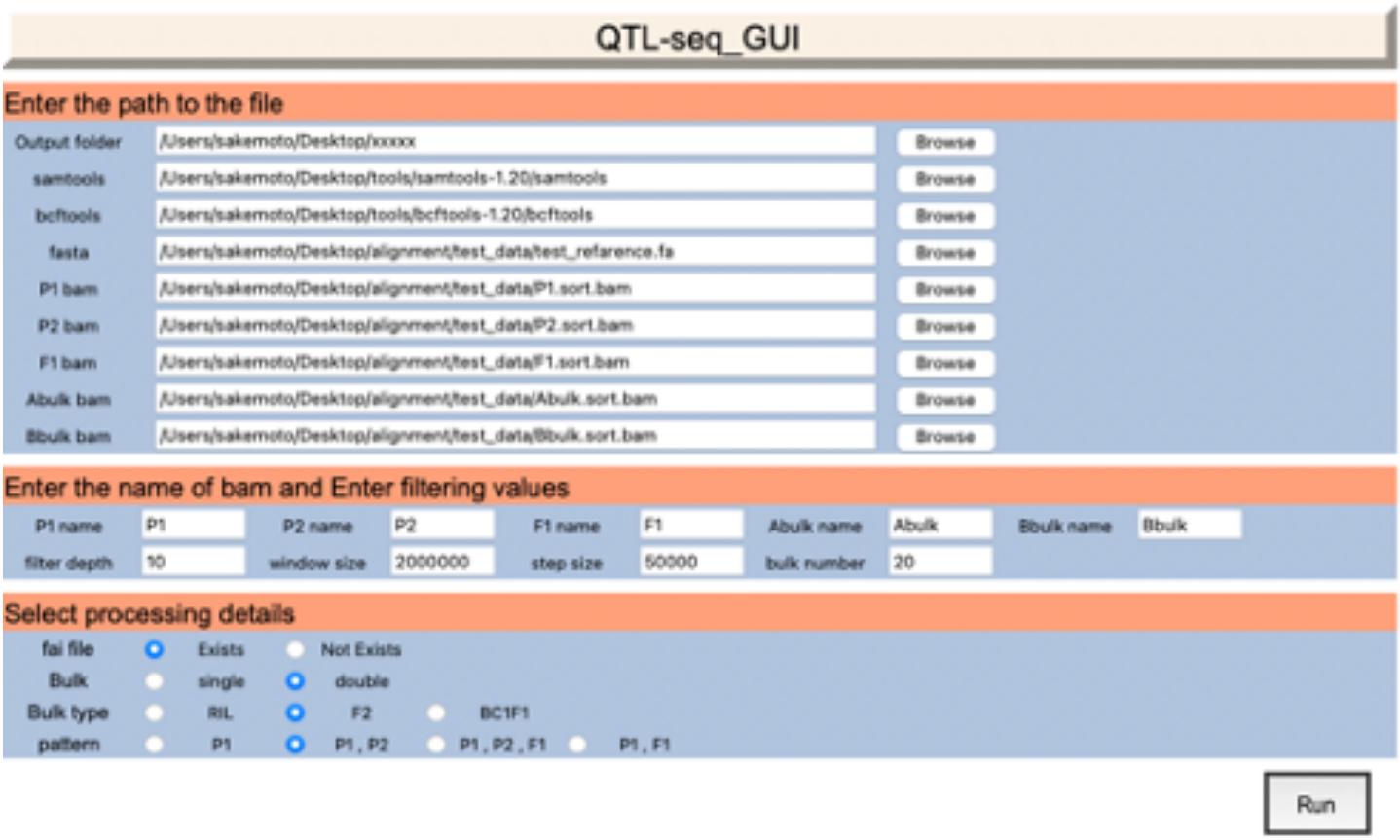
preparation

Install the following tools:
Samtools (URL : http://www.htslib.org/download/)
Bcftools (URL : http://www.htslib.org/download/)
Python3 (URL : https://www.python.org/downloads/)

Reference FASTA file
bam file

⚠️Confirm that the .bam.bai file exists in the same directory as the bam file. If not, the process will stop.

Using the GUI



How to specify files and folders using the GUI in QTL-sq_GUI:

- Enter the path: the path of the file or folder can be entered directly in the input field.

- Using the Browse button: clicking on the ‘Browse’ button opens the file dialog.

Run.
Once the required options, such as ‘output folder’ and process commands, have been entered, execution can be started by clicking on the ‘Run’ button at the bottom right of the screen.

Using the GUI

The QTL-seq_GUI has the following three input fields.

1. Enter the path to the file
Enter the paths for the output folder, required tools, and the bam files to be used, or select them using the ‘ Browse ' button.

2. Enter the name of bam and Enter filtering values
Please enter the name of the bam file to be used and the numeric settings for the filter process.

3. Select processing details
Please select the appropriate option. ‘ pattern ' refers to the type of bam file used for filtering.

Run

1. Enter the path to the file :

Output Folder:

Select the folder where the results generated by QTL-seq_GUI are stored.

⚠Note that selecting a folder that has been used for output in the past may overwrite existing results. To avoid loss of previous data, select a different folder for each new run.

samtools

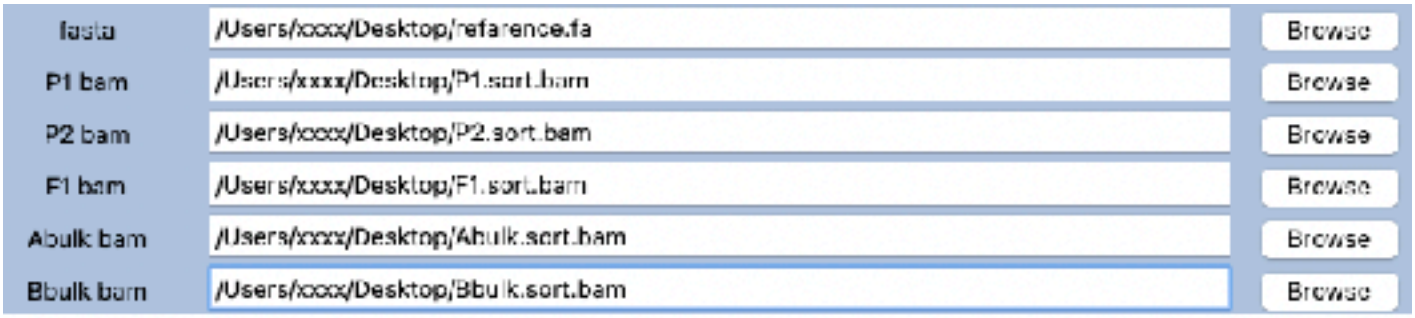
Select samtools file.

bcftools

Select bcftools file.

Run

1. Enter the path to the file：



fasta
Select fasta file used for alignment.

P1 bam , P2 bam
Select bam files of parents' lineages.

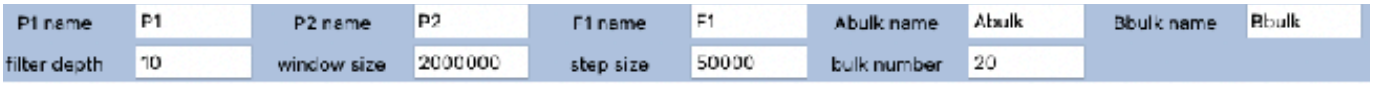
F1 bam
Select the F1 bam file for the parents' lineage.

Abulk bam , Bbulk bam
Select upper or lower bulk bam files

⚠️ By default, the path is assumed to be set correctly. If the path is not set or is incorrect, enter it correctly.

Run

2. Enter the name of bam and Enter filtering values



P1 name
Enter the name of P1.

P2 name
Enter the name of P2.

F1 name
Enter the name of F1.

Abulk name
Enter the name of Abulk.

Bbulk name
Enter the name of Bbulk.

filter depth
Enter the numerical value of the depth to be used for filtering.

window size
Sliding window window size (bp): 2000000

step size
Sliding window step size (bp): 50000

bulk number
Enter the number of bulk

Run

3. Select processing details

fai file
Select “Exists” if the fai file is in the same directory as the fasta file, or “Not Exists” if not

Bulk
Select “single” if there is only one Bulk to be analyzed, or “double” if there are two Bulks

Bulk type
Select “RIL” if Bulk is RIL or “F2” if Bulk is F2

pattern
Select the appropriate one according to the bam file used for filtering

⚠️ Leave the path and name of the bam file not selected in pattern blank or as is.

Output

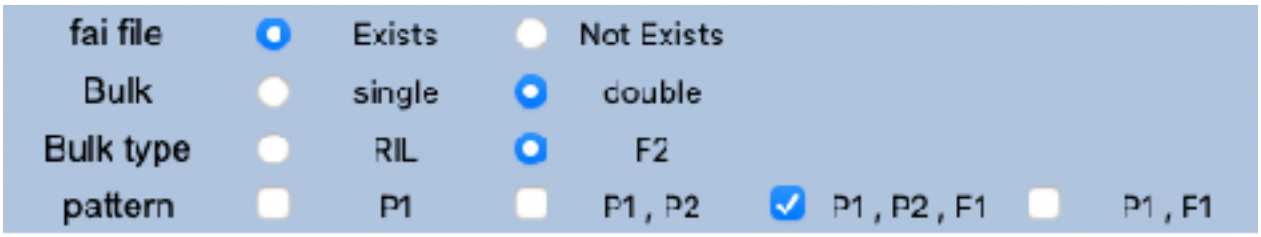
QTL-seq_GUI generates three types of folders.

```
<Output folder>
  |---bam
  |---txt
  |---result
```

The files output to each folder are described on the following pages.

bam

Example of output file when the following contents are selected

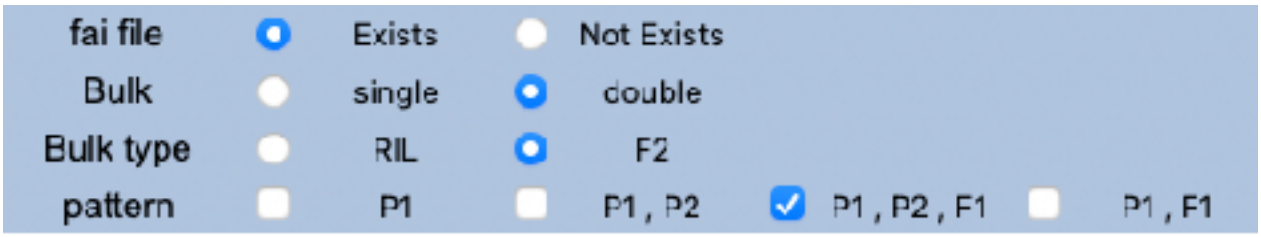



```
bam
 |---P1_chrxxx.bam
 |---P1_chrxxx.bam.bai
 |---P2_chrxxx.bam
 |---P2_chrxxx.bam.bai
 |---F1_chrxxx.bam
 |---F1_chrxxx.bam.bai
 |---Abulk_chrxxx.bam
 |---Abulk_chrxxx.bam.bai
 |---Bbulk_chrxxx.bam
 |---Bbulk_chrxxx.bam.bai      etc
```

bam file and bam bai file are output, split by chromosome

txt

Example of output file when the following contents are selected




```
txt
 |---QTLseq_input_log.txt
 |---snpindex_marge.txt
 |---average_snp_index_per_window.txt
 |---confidence_interval.txt
 |---Abulk-Bbulk_95_peak.txt
 |---Abulk-Bbulk_99_peak.txt
```

txt

QTLseq_input_log.txt

This file contains selected SNP information from the vcf file for set values ≥ depth , P1, P2 homozygous and F1 heterozygous locations. Chromosome name, position, SNP-index and depth information for each sample are listed in tab-delimited format.

①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬
chr1	3500	0	25	1	25	0.556	18	0.333	21	0.483	29	-0.15
chr1	4500	0	19	1	19	0.471	17	0.571	14	0.429	21	0.142
chr2	7500	1	22	0	22	0.52	25	0.368	19	0.565	23	-0.197
chr2	8500	0	19	1	19	0.5	16	0.333	15	0.529	17	-0.196
chr3	10500	0	14	1	14	0.526	19	0.48	25	0.235	17	0.245
chr3	11500	0	17	1	17	0.35	20	0.25	12	0.452	31	-0.202
chr4	12500	0	22	1	22	0.5	18	0.556	18	0.471	17	0.085
chr4	15500	0	17	1	17	0.652	23	0.5	22	0.4	20	0.1

① ; chr
② ; position
③ ; P1 snp-index
④ ; P1 depth
⑤ ; P2 snp-index
⑥ ; P2 depth
⑦ ; F1 snp-index
⑧ ; F1 depth
⑨ ; Abulk snp-index
⑩ ; Abulk depth
⑪ ; Bbulk snp-index
⑫ ; Bbulk depth
⑬ ; delta snp-index

txt

snpindex_marge.txt

This file contains selected SNP information from the vcf file for set values ≥ depth , P1, P2 homozygous and F1 heterozygous locations. Chromosome name, position, SNP-index and depth information for each sample are listed in tab-delimited format.

①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬
chr1	3500	0	25	1	25	0.556	18	0.333	21	0.483	29	-0.15
chr1	4500	0	19	1	19	0.471	17	0.571	14	0.429	21	0.142
chr2	7500	1	22	0	22	0.52	25	0.368	19	0.565	23	-0.197
chr2	8500	0	19	1	19	0.5	16	0.333	15	0.529	17	-0.196
chr3	10500	0	14	1	14	0.526	19	0.48	25	0.235	17	0.245
chr3	11500	0	17	1	17	0.35	20	0.25	12	0.452	31	-0.202
chr4	12500	0	22	1	22	0.5	18	0.556	18	0.471	17	0.085
chr4	15500	0	17	1	17	0.652	23	0.5	22	0.4	20	0.1

① ; chr
② ; position
③ ; P1 snp-index
④ ; P1 depth
⑤ ; P2 snp-index
⑥ ; P2 depth
⑦ ; F1 snp-index
⑧ ; F1 depth
⑨ ; Abulk snp-index
⑩ ; Abulk depth
⑪ ; Bbulk snp-index
⑫ ; Bbulk depth
⑬ ; delta snp-index

txt

average_snp_index_per_window.txt

This file contains information from the snp-index text to the average value of the snp-index for each bulk and delta within the window size. Information on the chromosome name, middle position, start position, end position, and average snp-index for each sample is tab-delimited.

①	②	③	④	⑤	⑥	⑦
chr1	1000000	0	2000000	0.53	0.43	0.096
chr1	1050000	50000	2050000	0.58	0.45	0.125
chr2	1000000	0	2000000	0.77	0.36	0.416
chr2	1050000	50000	2050000	0.76	0.32	0.444
chr3	1000000	0	2000000	0.5	0.43	0.071
chr3	1050000	50000	2050000	0.53	0.39	0.136

① ; chr
② ; middle position
③ ; start position
④ ; end position
⑤ ; Abulk snp-index
⑥ ; Bbulk snp-index
⑦ ; delta snp-index

txt

confidence_interval.txt

This file contains the confidence intervals for the depths. It contains tab-separated values for the chromosome name, intermediate position, start position, end position, and the 99% and 95% confidence intervals for the depths calculated using the simulation file.

①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫
chr1	1000000	0	2000000	0.15	0.86	0.225	0.78	-0.511	0.52	-0.386	0.396
chr1	1050000	50000	2050000	0.15	0.85	0.226	0.78	-0.507	0.51	-0.379	0.392
chr2	1000000	0	2000000	0.15	0.86	0.228	0.78	-0.51	0.51	-0.389	0.4
chr2	1050000	50000	2050000	0.15	0.87	0.227	0.79	-0.516	0.52	-0.394	0.41
chr3	1000000	0	2000000	0.15	0.86	0.221	0.79	-0.512	0.52	-0.391	0.402
chr3	1050000	50000	2050000	0.14	0.86	0.22	0.79	-0.517	0.53	-0.389	0.402

① ; chr
② ; middle position
③ ; start position
④ ; end position
⑤ ; Low mean snp-index at 99% confidence interval
⑥ ; Higher mean snp-index at 99% confidence interval
⑦ ; Lower mean snp-index at 95% confidence interval
⑧ ; Higher mean snp-index at 95% confidence interval
⑨ ; Low delta mean snp-index at 99% confidence interval
⑩ ; High value of delta mean snp-index at 99% confidence interval
⑪ ; Low delta mean snp-index at 95% confidence interval
⑫ ; High value of delta mean snp-index at 95% confidence interval

txt

Abulk-Bbulk_95_peak.txt

This file contains the position and SNP-index values for each variety beyond the 95% confidence interval. Chromosome name, start position, end position, and peak position are described in tab-delimited format.

①	②	③	④	⑤	⑥	⑦
chr2	0	2050000	1050000	0.76	0.32	0.444
chr03	0	5000000	1300000	0.866	0.183	0.684
chr03	22200000	25000000	23500000	0.858	0.344	0.515
chr08	0	5100000	1800000	0.877	0.252	0.625

① ; chr
② ; start position
③ ; end position
④ ; peak position
⑤ ; Abulk snp-index
⑥ ; Bbulk snp-index
⑦ ; delta snp-index

txt

Abulk-Bbulk_99_peak.txt

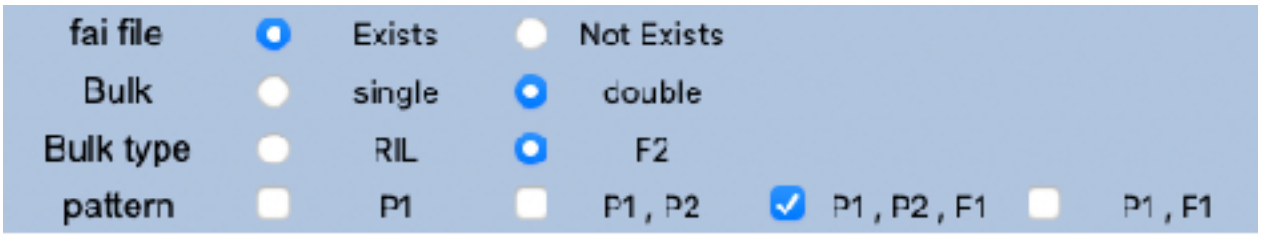
This file contains the position and SNP-index values for each variety beyond the 99% confidence interval. Chromosome name, start position, end position, and peak position are described in tab-delimited format.

①	②	③	④	⑤	⑥	⑦
chr2	0	2050000	1050000	0.76	0.32	0.444
chr03	0	5000000	1300000	0.866	0.183	0.684
chr03	22200000	25000000	23500000	0.858	0.344	0.515
chr08	0	5100000	1800000	0.877	0.252	0.625

① ; chr
② ; start position
③ ; end position
④ ; peak position
⑤ ; Abulk snp-index
⑥ ; Bbulk snp-index
⑦ ; delta snp-index

result

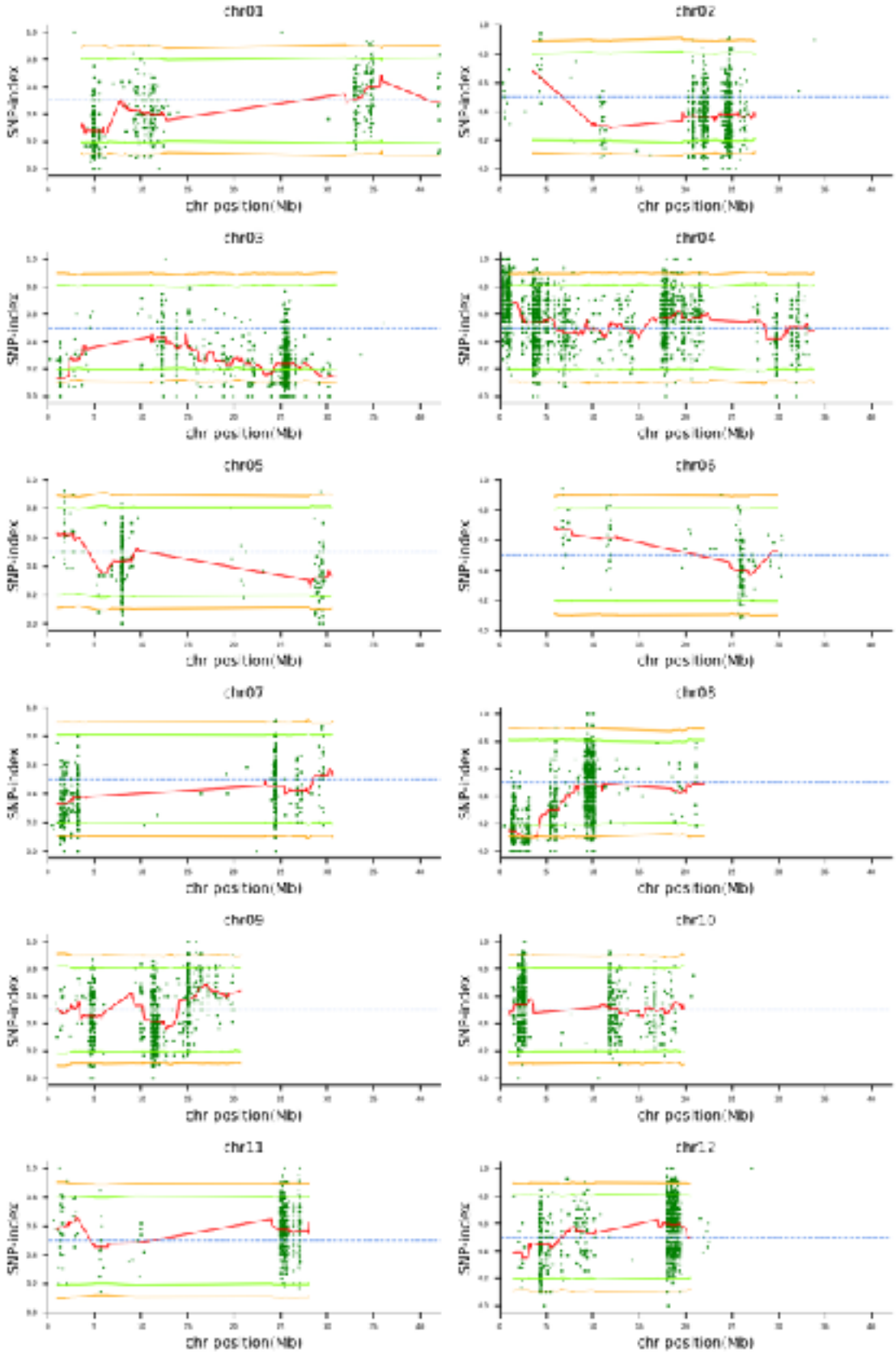
Example of output file when the following contents are selected




```
result
 |---Abulk.png
 |---Bbulk.png
 |---Abulk-Bbulk.png
```

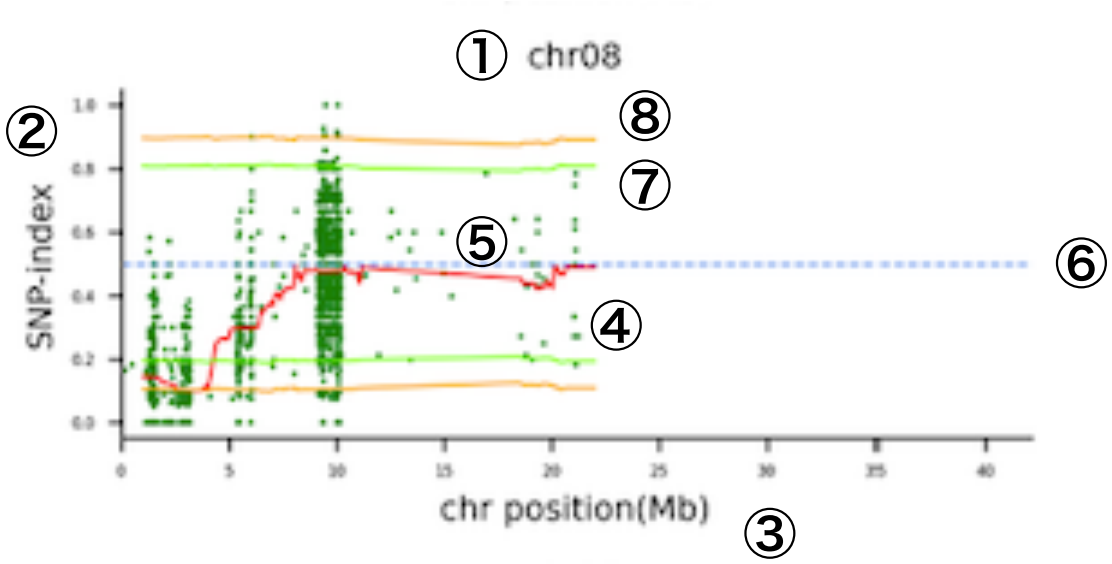
result

Abulk.png



result

Abulk.png



① : Chromosome name

② : SNP-index (0.0 < SNP-index < 1.0)

③ : position (Mb)

④ : SNP

Represents SNPs for position

⑤ : Average SNP-index

Average SNP-index value for Window size

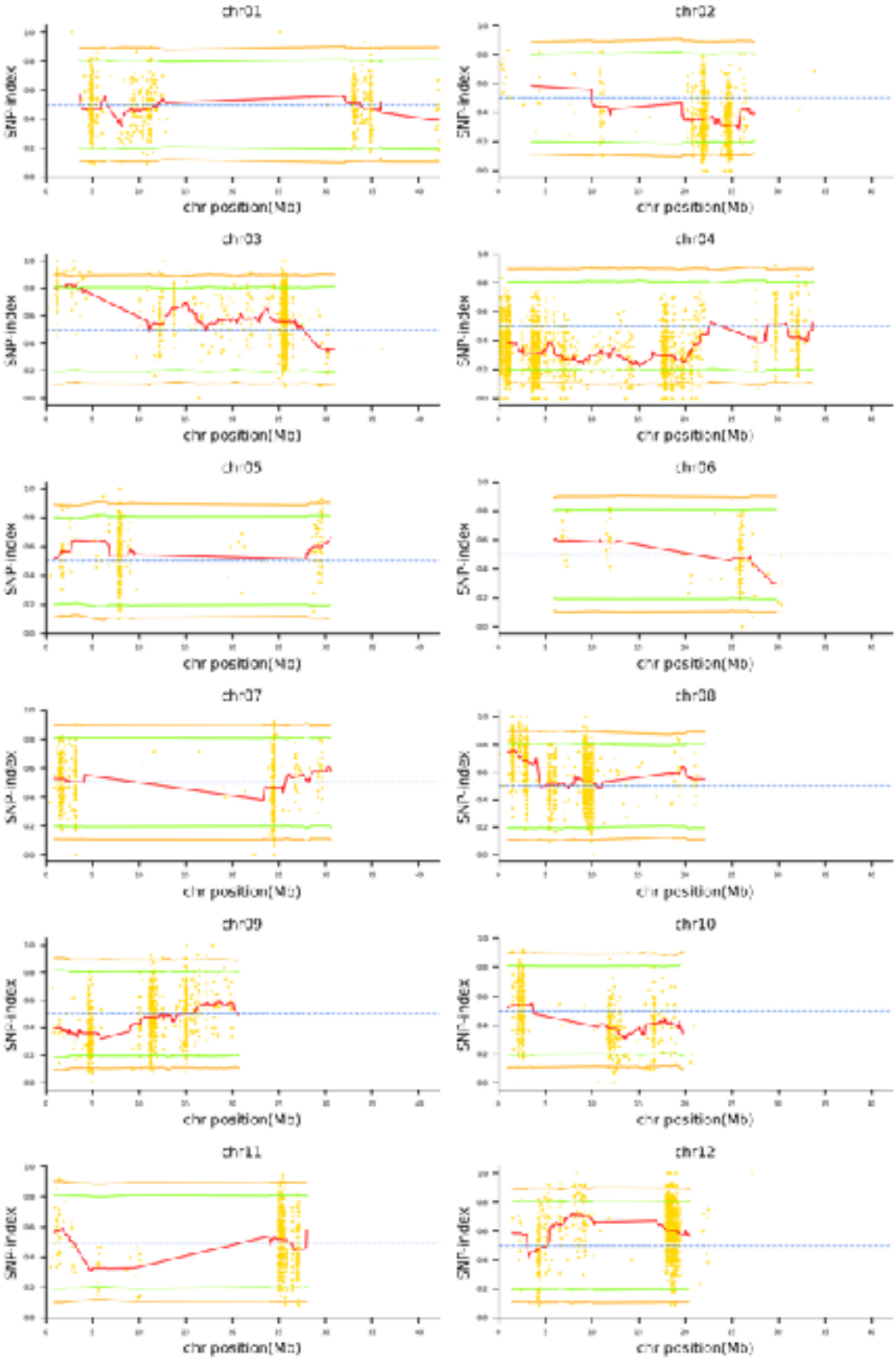
⑥ : SNP-index = 0.5

⑦ : 95% confidence interval for depth

⑧ : 99% confidence interval for depth

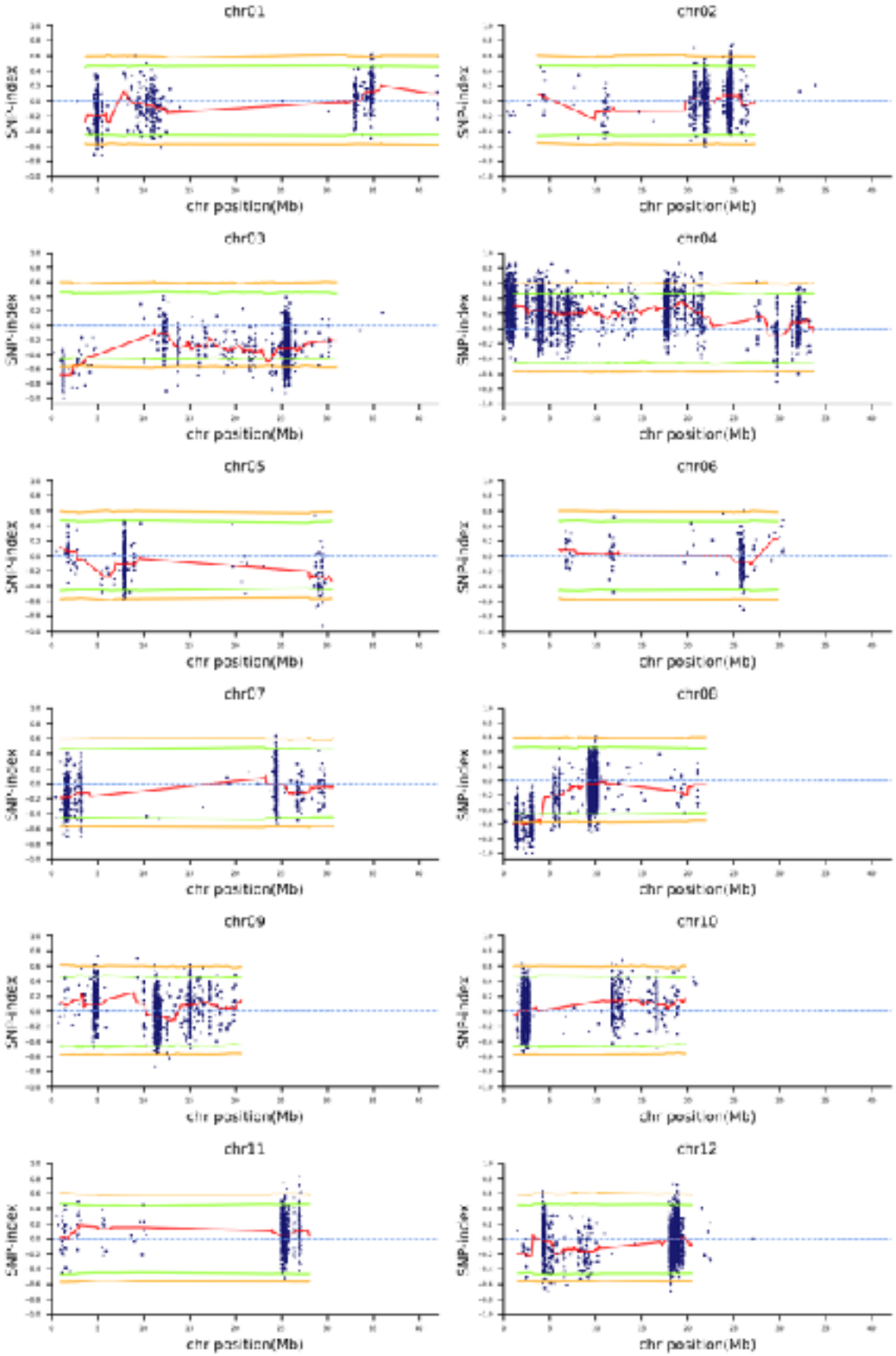
result

Bbulk.png



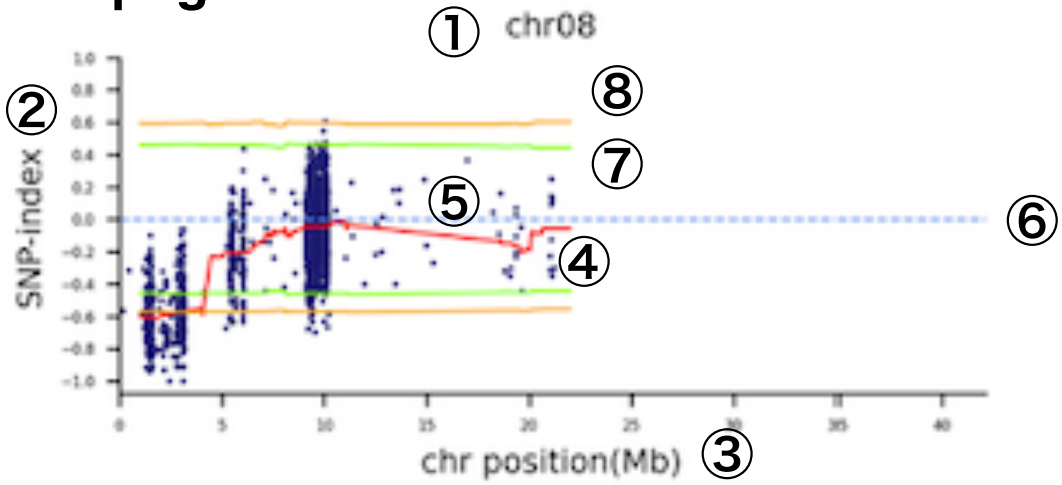
result

Abulk-Bbulk.png



result

Abulk-Bbulk.png



① : Chromosome name

② : SNP-index (-1.0 < SNP-index < 1.0)

③ : position (Mb)

④ : SNP

Represents SNPs for position

⑤ : Average SNP-index

Average SNP-index value for Window size

⑥ : SNP-index = 0.0s

⑦ : 95% confidence interval for depth

⑧ : 99% confidence interval for depth