

MultiGWAS report for Naive GWAS model

- [1 Input Parameters](#)
- [2 Best-ranked SNPs](#)
 - [2.1 Table of best-ranked SNPs](#)
 - [2.2 Venn diagram of best-ranked SNPs](#)
- [3 Significant SNPs](#)
 - [3.1 Table of significant SNPs](#)
 - [3.2 Venn diagram of significant SNPs](#)
- [4 Manhattan and QQ plots](#)
- [5 Chord diagrams for SNPs by chromosome](#)
- [6 Profiles for common significant SNPs](#)

MultiGWAS is a tool to do GWAS analysis in tetraploid organisms by executing in parallel and integrating the results from four existing GWAS software: two available for polyploids (GWASpoly and SHEsis) and two frequently used for diploids (PLINK and TASSEL).

MultiGWAS executes two types of GWAS models: Full and Naive. The Full model with control for population structure and individual relatedness and the Naive model without any control. In both models, users can apply different control quality filters for the genomic data.

This report, created by the MultiGWAS tool, is a summary of the input parameters used to run the tool and its main outputs, including (1) Score tables with detailed information on the associations for each tool. (2) Venn diagrams of shared SNPs among the four tools. (3) Heatmaps of significant SNP profiles among the four tools. (4) Manhattan and QQ plots for the association found by each tool. And (5) Chord diagrams for the chromosomes vs. SNP by each tool.

1 Input Parameters

MultiGWAS uses as input a configuration file where the user specifies the input genomic data (genotype and phenotype files) along with different values for parameters used by the MultiGWAS tool and by the other four GWAS software (GWASpoly, SHEsis, PLINK, and TASSEL).

The following table shows the current input parameters specified by the user in the configuration file:

| PARAMETER | VALUE |
|---|----------------------------|
| Ploidy (4 or 2) | 4 |
| Genotype filename | example-genotype-tetra.csv |
| Phenotype filename | example-phenotype.csv |
| Significance level (Genome-wide significance level) | 0.05 |
| Correction method (Bonferroni or FDR) | Bonferroni |

| PARAMETER | VALUE |
|---|------------------------------|
| GWAS model (Full or Naive) | Naive |
| nBest (Number of best-ranked SNPs to be reported) | 8 |
| Filtering (TRUE or FALSE) | FALSE |
| MIND Filter (Individual with missing genotype) | 0.1 |
| GENO Filter (SNPs with missing genotype) | 0.1 |
| MAF Filter (Minor allele frequency) | 0.01 |
| HWE Filter (Hardy-Weinberg test) | 1e-10 |
| GWAS Tools | GWASpoly SHEsis PLINK TASSEL |

2 Best-ranked SNPs

This section shows a table and a Venn diagram for the best ranked SNPs (For this report, N=8). The configuration file allows defining the parameter N.

2.1 Table of best-ranked SNPs

The following table shows the associations between SNPs and the trait as transformed p-values to scores as $-\log_{10}(\text{p-value})$ sorted from the highest to the lowest scores based on each tool analysis and the model selected (Full or Naive). Each SNP has its associated information, such as the chromosome, position in the genome, and name. Also, based on the threshold calculated by a multiple testing correction (i.e., Bonferroni or FDR), the table marks the significance of each SNP as a TRUE or FALSE.

| TOOL | MODEL | GC | SNP | CHROM | POSITION | PVALUE | SCORE | THRESHOLD | SIGNIFICANCE |
|----------|----------|-------|--------------|-------|----------|----------|--------|-----------|--------------|
| GWASpoly | additive | 0.973 | c1_80 19 | 10 | 48863165 | 0.000000 | 9.4700 | 4.8200 | TRUE |
| GWASpoly | additive | 0.973 | c2_25 471 | 10 | 48808404 | 0.000001 | 6.1900 | 4.8200 | TRUE |
| GWASpoly | additive | 0.973 | c1_80 21 | 10 | 48862950 | 0.000034 | 4.4700 | 4.8200 | FALSE |
| GWASpoly | additive | 0.973 | c2_45 606 | 10 | 48218826 | 0.000041 | 4.3900 | 4.8200 | FALSE |
| GWASpoly | additive | 0.973 | c2_45 | 10 | 48203431 | 0.000065 | 4.1900 | 4.8200 | FALSE |

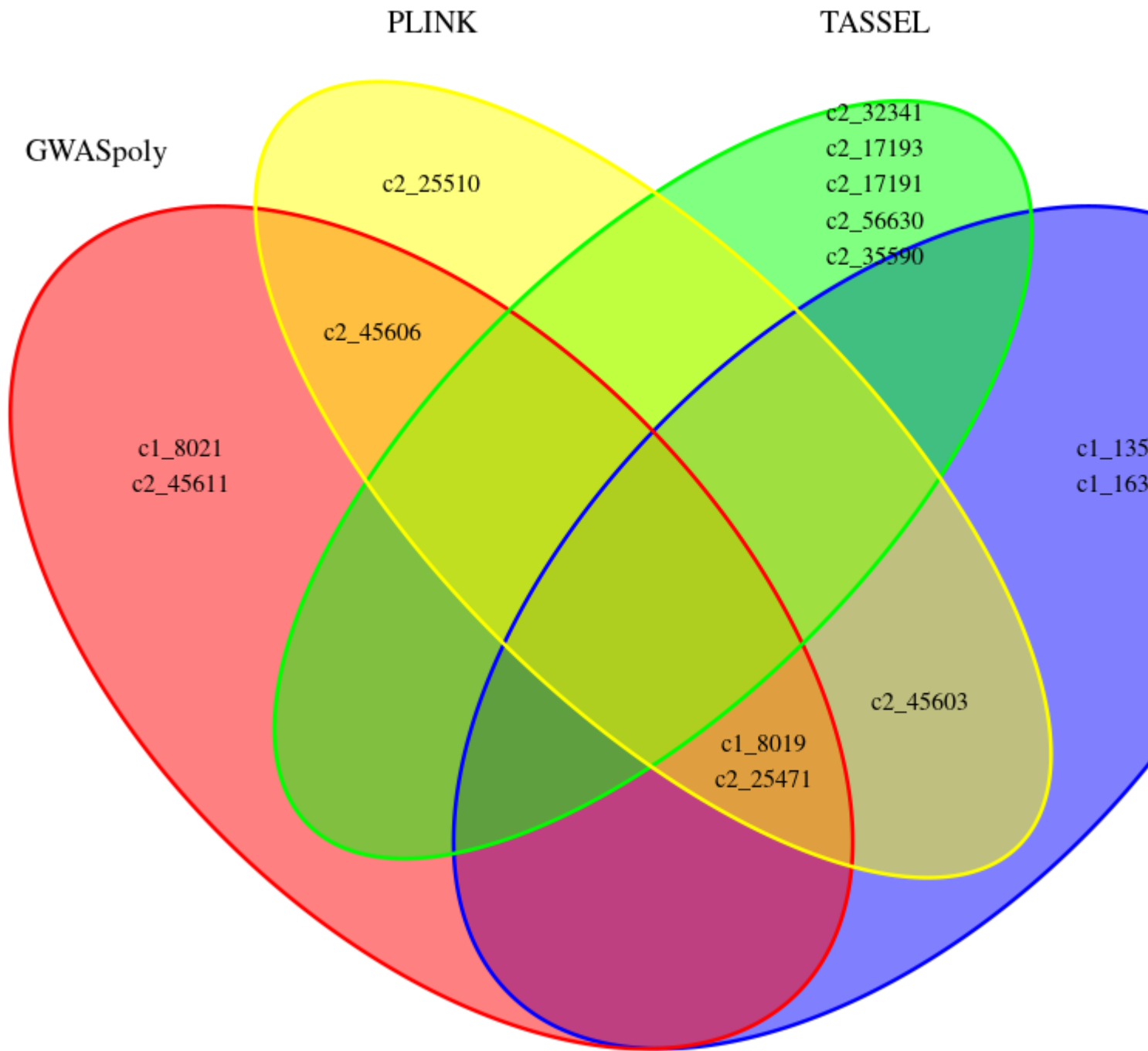
| TOOL | MODEL | GC | SNP | CHROM | POSITION | PVALUE | SCORE | THRESHOLD | SIGNIFICAN |
|------|-------|----|-----|-------|----------|--------|-------|-----------|------------|
|------|-------|----|-----|-------|----------|--------|-------|-----------|------------|

611

| | | | | | | | | | |
|--------|----------|-------|--------------|----|----------|----------|---------|--------|-------|
| PLINK | additive | 4.162 | c1_80 19 | 10 | 48863165 | 0.000000 | 16.0598 | 4.8402 | TRUE |
| PLINK | additive | 4.162 | c2_25 510 | 10 | 48679881 | 0.000000 | 13.8935 | 4.8402 | TRUE |
| PLINK | additive | 4.162 | c2_25 471 | 10 | 48808404 | 0.000000 | 13.4407 | 4.8402 | TRUE |
| PLINK | additive | 4.162 | c2_45 606 | 10 | 48218826 | 0.000000 | 12.7625 | 4.8402 | TRUE |
| PLINK | additive | 4.162 | c2_45 603 | 10 | 48073593 | 0.000000 | 12.3502 | 4.8402 | TRUE |
| SHEsis | general | 5.878 | c1_80 19 | 10 | 48863165 | 0.000000 | 22.0297 | 4.8477 | TRUE |
| SHEsis | general | 5.878 | c2_25 471 | 10 | 48808404 | 0.000000 | 17.2581 | 4.8477 | TRUE |
| SHEsis | general | 5.878 | c2_45 603 | 10 | 48073593 | 0.000000 | 15.9547 | 4.8477 | TRUE |
| SHEsis | general | 5.878 | c1_13 526 | 10 | 48020996 | 0.000000 | 15.6576 | 4.8477 | TRUE |
| SHEsis | general | 5.878 | c1_16 351 | 10 | 48761642 | 0.000000 | 15.2351 | 4.8477 | TRUE |
| TASSEL | dominant | 1.289 | c2_32 341 | 11 | 10396503 | 0.000967 | 3.0147 | 4.5240 | FALSE |
| TASSEL | dominant | 1.289 | c2_17 193 | 1 | 70472380 | 0.001770 | 2.7520 | 4.5240 | FALSE |
| TASSEL | dominant | 1.289 | c2_17 191 | 1 | 70474651 | 0.001820 | 2.7399 | 4.5240 | FALSE |
| TASSEL | dominant | 1.289 | c2_56 630 | 11 | 10340256 | 0.003760 | 2.4248 | 4.5240 | FALSE |
| TASSEL | dominant | 1.289 | c2_35 590 | 6 | 28621753 | 0.003850 | 2.4145 | 4.5240 | FALSE |

2.2 Venn diagram of best-ranked SNPs

The Venn diagram shows the best-ranked SNPs with high scores that are either unique or shared by the four software. Shared SNPs appear in the intersection areas of the diagram.



3 Significant SNPs

This section shows a table and a Venn diagram for the signifcative SNPs (score is above the significance threshold for each tool).

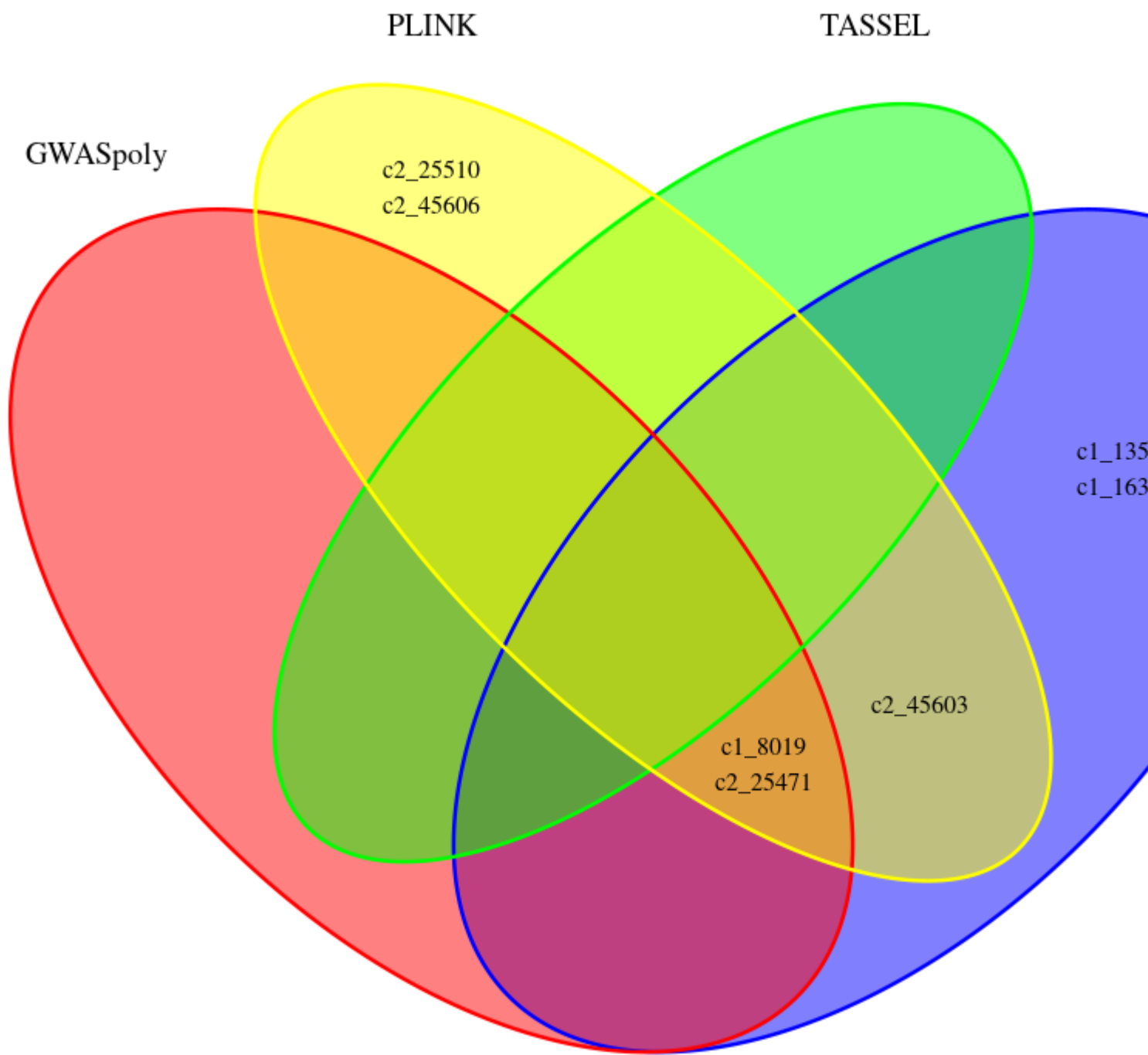
3.1 Table of significant SNPs

The following table shows the associations between SNPs and the trait as transformed p-values to scores as -log 10 (p-value) sorted from the highest to the lowest scores based on each tool analysis and the model selected (Full or Naive). Each SNP has its associated information, such as the chromosome, position in the genome, and name.

| TOOL | MODEL | GC | SNP | CHROM | POSITION | PVALUE | SCORE | THRESHOLD | SIGNIFICAN |
|----------|----------|-------|--------------|-------|----------|--------|---------|-----------|------------|
| GWASpoly | additive | 0.973 | c1_80 19 | 10 | 48863165 | 0e+00 | 9.4700 | 4.8200 | TRUE |
| GWASpoly | additive | 0.973 | c2_25 471 | 10 | 48808404 | 1e-06 | 6.1900 | 4.8200 | TRUE |
| PLINK | additive | 4.162 | c1_80 19 | 10 | 48863165 | 0e+00 | 16.0598 | 4.8402 | TRUE |
| PLINK | additive | 4.162 | c2_25 510 | 10 | 48679881 | 0e+00 | 13.8935 | 4.8402 | TRUE |
| PLINK | additive | 4.162 | c2_25 471 | 10 | 48808404 | 0e+00 | 13.4407 | 4.8402 | TRUE |
| PLINK | additive | 4.162 | c2_45 606 | 10 | 48218826 | 0e+00 | 12.7625 | 4.8402 | TRUE |
| PLINK | additive | 4.162 | c2_45 603 | 10 | 48073593 | 0e+00 | 12.3502 | 4.8402 | TRUE |
| SHEsis | general | 5.878 | c1_80 19 | 10 | 48863165 | 0e+00 | 22.0297 | 4.8477 | TRUE |
| SHEsis | general | 5.878 | c2_25 471 | 10 | 48808404 | 0e+00 | 17.2581 | 4.8477 | TRUE |
| SHEsis | general | 5.878 | c2_45 603 | 10 | 48073593 | 0e+00 | 15.9547 | 4.8477 | TRUE |
| SHEsis | general | 5.878 | c1_13 526 | 10 | 48020996 | 0e+00 | 15.6576 | 4.8477 | TRUE |
| SHEsis | general | 5.878 | c1_16 | 10 | 48761642 | 0e+00 | 15.2351 | 4.8477 | TRUE |

3.2 Venn diagram of significative SNPs

The Venn diagram shows the significative SNPs that are either unique or shared by the four software. Shared SNPs appear in the intersection areas of the diagram.



4 Manhattan and QQ plots

MultiGWAS uses classical Manhattan and Quantile–Quantile (QQ) plots to visually summarize GWAS results and identify both best-ranked and significant SNPs for each GWAS software. The Manhattan plot

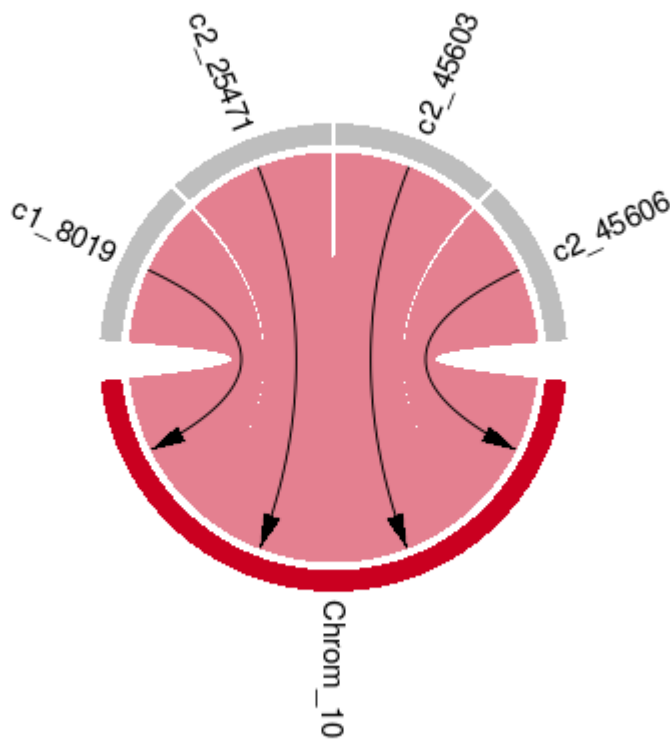
shows each SNPs distributed in the genetic location (x-axis) vs. the p-value (y-axis). The best-ranked SNPs are above the blue line and the significant SNPs above the blue line. In green are the SNPs that got high scores in more than one software.

The QQ plot plots the observed (black dot lines) vs. expected $-\log_{10}$ (p-value) (red line).

5 Chord diagrams for SNPs by chromosome

Chord diagrams are used in MultiGWAS to highlight nearby SNPs found on the same chromosome. These SNPs may have been identified simultaneously by multiple GWAS packets, which can be interpreted as a possible strong association, as in the Manhattan charts, where these associations form clean peaks with multiple SNPs displaying the same signal.

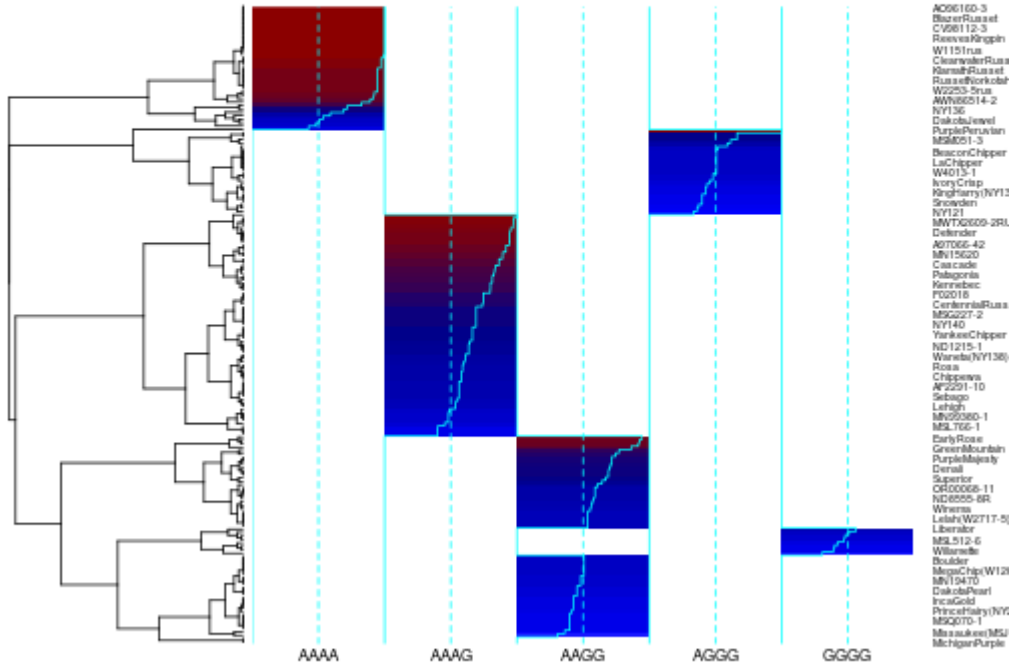
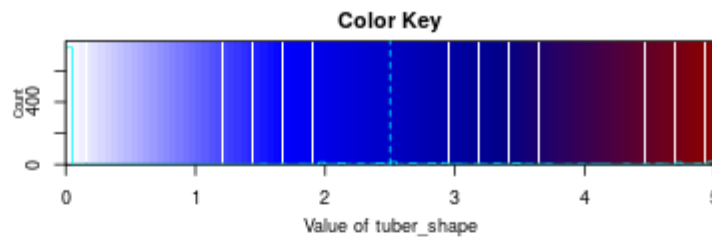
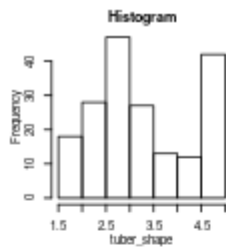
Next, we show the chord diagram for the shared SNPs shown in the Venn diagram of the best-ranked SNPs (Section 2.2).



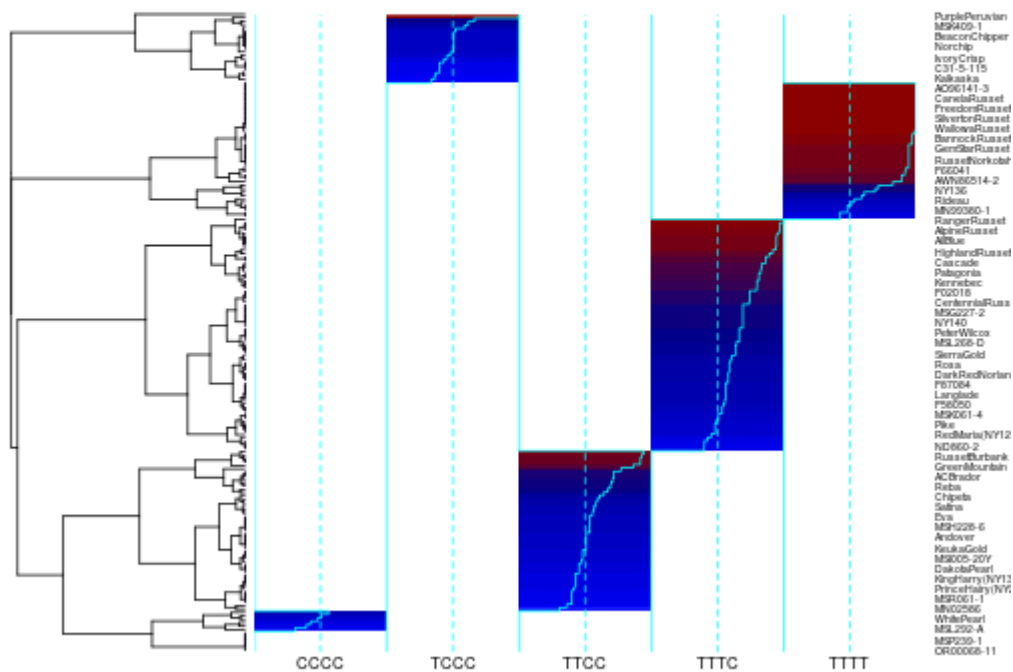
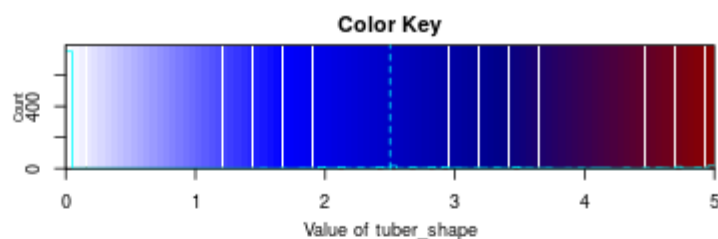
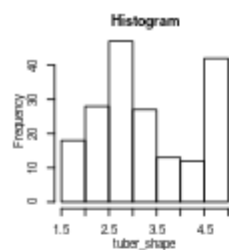
6 Profiles for common significant SNPs

For the common significant SNPs, MultiGWAS provides a figure to visualize each trait by individuals (rows) and genotypes (columns). At the left, the individuals are grouped in a dendrogram by their genotype. At the right, there is the name or ID of each individual. At the bottom, the genotypes are ordered from left to right, starting from the major to the minor allele (i.e., AAAA, AAAB, AABB, AB BB, BBBB). At the top, there is a description of the trait based on a histogram of frequency (top left) and by an assigned color for each numerical phenotype value using a grayscale (top right). Thus, each individual appears as a colored line by its phenotype value on its genotype column. For each column, there is a solid cyan line with the mean of each column and a broken cyan line that indicates how far the cell deviates from the mean.

Because each multiGWAS report shows one specific trait at a time, the histogram and color key will remain the same for all the best-ranked SNPs.



c1_8019



c2_25471

