Naive report

This is the report created by the MultiGWAS tool that includes the input parameters used to run the tool along with its main outputs. The outputs include the best and significative SNPs found for each tool which are showed as Venn diagrams and score tables, and also include a particular visualization of the SNPs found by more than one tool that we have called as the "SNP profile" as it is a unique representation for a SNP.

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 tools (GWASpoly, SHEsis, PLINK, and TASSEL).

This section shows the current input parameters specified by the user in the configuration file:

PARAMETER	VALUE
Genotype filename	example-genotype.tbl
Phenotype filename	example-phenotype.tbl
Significance level	0.05
Correction method	Bonferroni
Trait	
GWAS model	Full
Filtering	TRUE
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

Best Ranked SNPs

GWAS tools report the associations between SNPs and the traits as p-values. MultiGWAS transforms these p-values to scores as -log10(p-value), sorts them from high to low, calculates a significance threshold according to the method used for multiple testing correction (Bonferroni or FDR), and marks each SNP as significative if their score is greater than the threshold.

Here, MultiGWAS report the best ranked SNPs for each tool using Venn diagrams and score tables. Venn diagrams show in an intuitive manner the best ranked SNPs shared by the four tools. Shared SNPs (high scores in more than one tool) appear in the intersection areas of the diagram, while individual SNPs (high scores only in a specific tool) appear in the ellipse of each tool. Score tables show more detailed information of each SNP as: the GWAS tool which scores the SNP, the GWAS model used by the tool (Full or Naive), the chromosome and its genetic location, the SNP name, the p-value and its score as -log10 (p-value), the significative threshold, and a TRUE or FALSE flag wheter the SNP is significative or not.

Venn diagram of best ranked SNPs

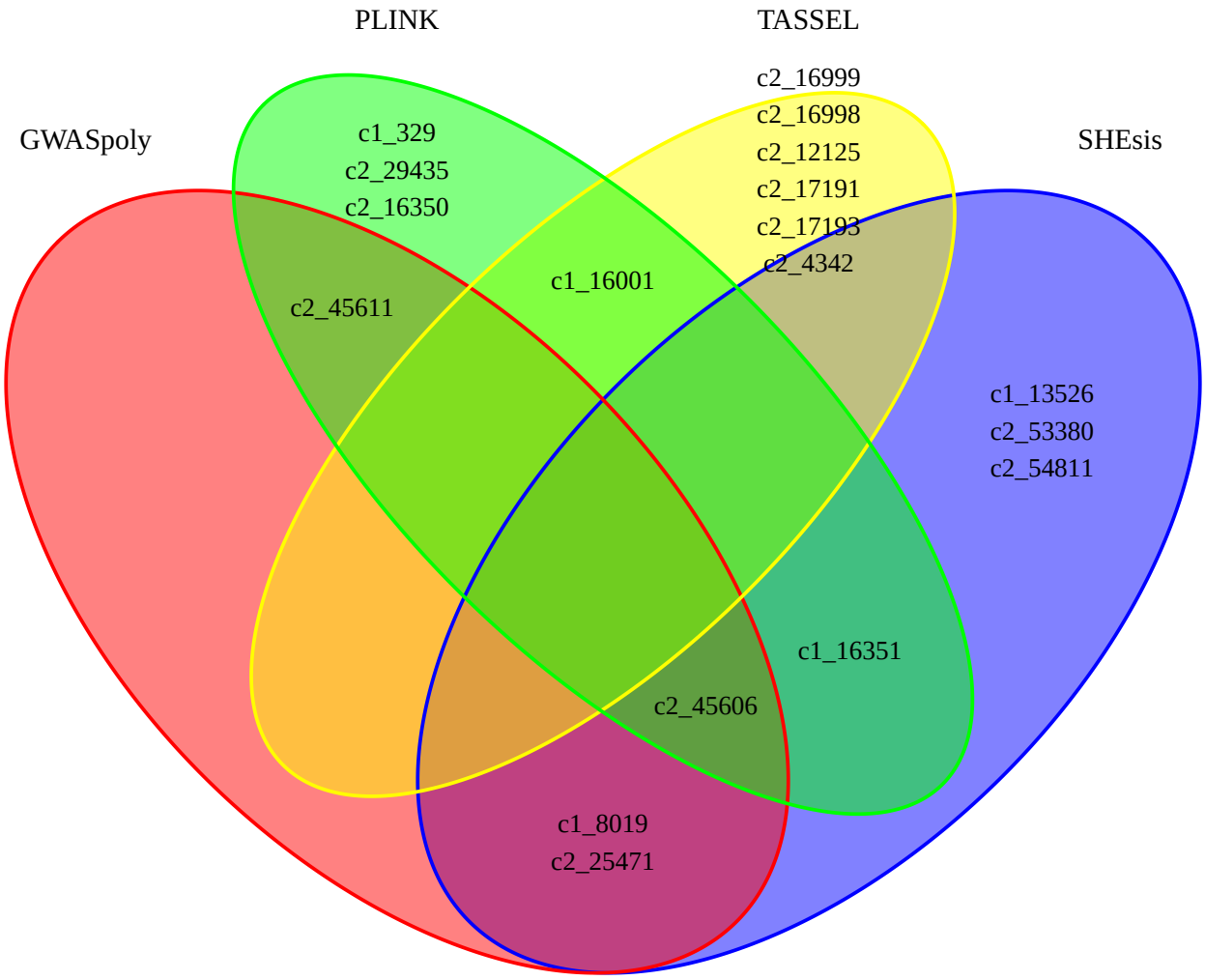


Table of best ranked SNPs

TOOL	MODEL	CHR	POS	SNP	P	SCORE	THRESHOLD	SIGNF
GWASpoly	Full	10	48863165	c1_8019	0.000017	4.780000	4.250000	TRUE
GWASpoly	Full	10	48808404	c2_25471	0.000027	4.570000	4.270000	TRUE
GWASpoly	Full	10	48203431	c2_45611	0.000044	4.360000	4.270000	TRUE
GWASpoly	Full	10	48218826	c2_45606	0.000021	4.680000	4.500000	TRUE
PLINK	Full	10	67293176	c1_16001	0.000187	1.769349	3.260071	FALSE
PLINK	Full	10	77351069	c1_329	0.000662	1.179470	3.301030	FALSE
PLINK	Full	11	51404231	c2_29435	0.000845	1.118849	3.255273	FALSE
PLINK	Full	10	69323144	c2_45611	0.001054	1.022917	3.255273	FALSE
PLINK	Full	2	41814861	c2_16350	0.001097	0.959793	3.301030	FALSE
PLINK	Full	10	69311500	c2_45606	0.001445	0.848906	3.292256	FALSE
PLINK	Full	10	69809843	c1_16351	0.002539	0.613066	3.283301	FALSE
SHEsis	Full	2	13697423	c1_8019	0.000000	9.471083	3.301030	TRUE
SHEsis	Full	1	30837971	c1_13526	0.000000	8.450065	3.292256	TRUE
SHEsis	Full	5	46046095	c2_53380	0.000000	8.240929	3.260071	TRUE
SHEsis	Full	3	39255236	c2_25471	0.000000	7.824082	3.292256	TRUE
SHEsis	Full	5	49804489	c2_54811	0.000000	6.963331	3.269513	TRUE
SHEsis	Full	1	69809843	c1_16351	0.000000	6.024734	3.283301	TRUE
SHEsis	Full	4	69311500	c2_45606	0.000000	5.955695	3.292256	TRUE
TASSEL	Full	8	54838024	c2_16999	0.000247	3.607621	3.894316	FALSE
TASSEL	Full	8	54838005	c2_16998	0.000329	3.482989	3.894316	FALSE
TASSEL	Full	1	71450400	c2_12125	0.003287	2.483226	3.894316	FALSE
TASSEL	Full	1	70474651	c2_17191	0.003548	2.449995	3.894316	FALSE
TASSEL	Full	1	70472380	c2_17193	0.005137	2.289293	3.894316	FALSE
TASSEL	Full	10	47539878	c1_16001	0.001230	2.910131	4.551206	FALSE
TASSEL	Full	7	14924207	c2_4342	0.001320	2.879298	4.551206	FALSE

Significative SNPs

MultiGWAS shows the significative SNPs for each tool corresponding to the SNPs above the significance threshold (see above). As shown earlier, these SNPs are shown by two ways: a Venn diagram showing the shared SNPs, and a table showing the SNPs details.

Venn diagram of signficative SNPs

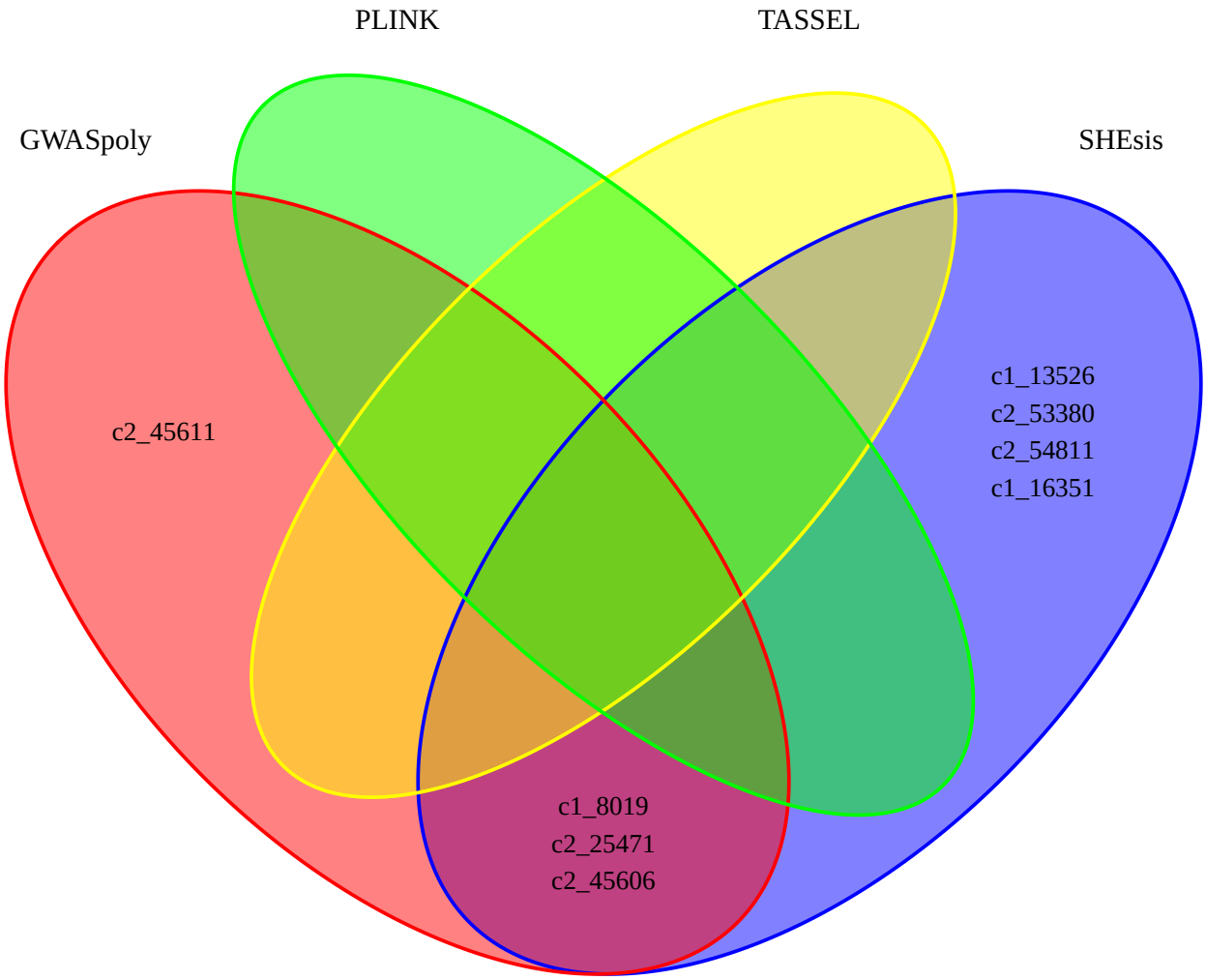


Table of significative SNPs

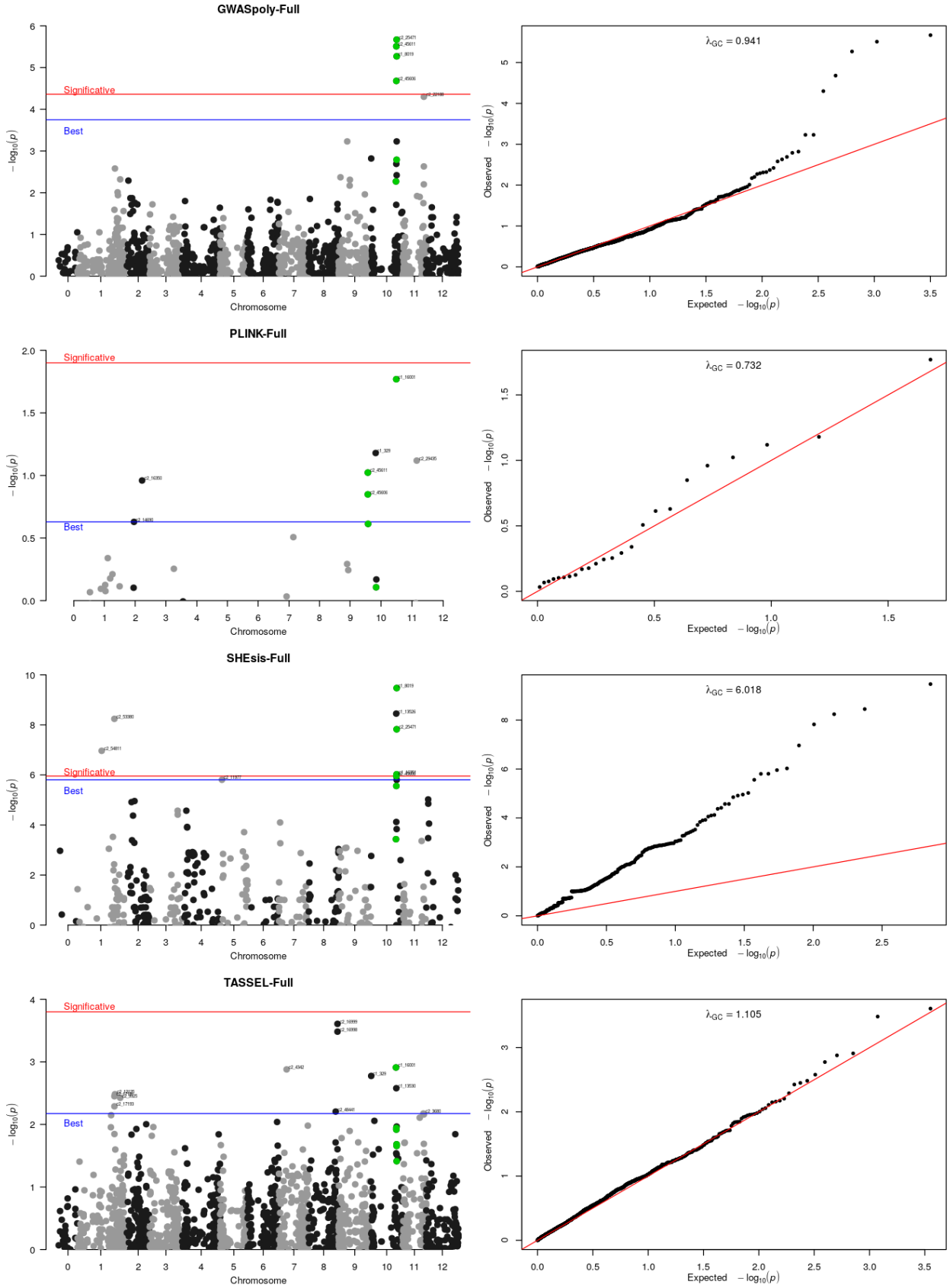
TOOL	MODEL	CHR	POS	SNP	P	SCORE	THRESHOLD	SIGNF
GWASpoly	Full	10	48863165	c1_8019	1.7e-05	4.780000	4.250000	TRUE
GWASpoly	Full	10	48808404	c2_25471	2.7e-05	4.570000	4.270000	TRUE
GWASpoly	Full	10	48203431	c2_45611	4.4e-05	4.360000	4.270000	TRUE
GWASpoly	Full	10	48218826	c2_45606	2.1e-05	4.680000	4.500000	TRUE
SHEsis	Full	2	13697423	c1_8019	0.0e+00	9.471083	3.301030	TRUE
SHEsis	Full	1	30837971	c1_13526	0.0e+00	8.450065	3.292256	TRUE
SHEsis	Full	5	46046095	c2_53380	0.0e+00	8.240929	3.260071	TRUE
SHEsis	Full	3	39255236	c2_25471	0.0e+00	7.824082	3.292256	TRUE
SHEsis	Full	5	49804489	c2_54811	0.0e+00	6.963331	3.269513	TRUE
SHEsis	Full	1	69809843	c1_16351	0.0e+00	6.024734	3.283301	TRUE
SHEsis	Full	4	69311500	c2_45606	0.0e+00	5.955695	3.292256	TRUE

Manhattan and QQ plots

MultiGWAS utilizes classical Manhattan and Quantile–Quantile (QQ) plots to visually summarize GWAS results and identify both best ranked and significative SNPs. These plots are presented for each GWAS tool, the QQ plot plots the observed vs. expected -log10(p-value), and the Manhattan plot plots each of the SNPs as a dot with its genetic location on the x-axis and its p-value on the y-axis.

Furthermore, MultiGWAS shows the following details in the Manhattan plots:

- The best ranked SNPs, above the blue line.
- The significative SNPs, above the red line.
- The shared SNPs, marked in green (with high scores in more than one tool)

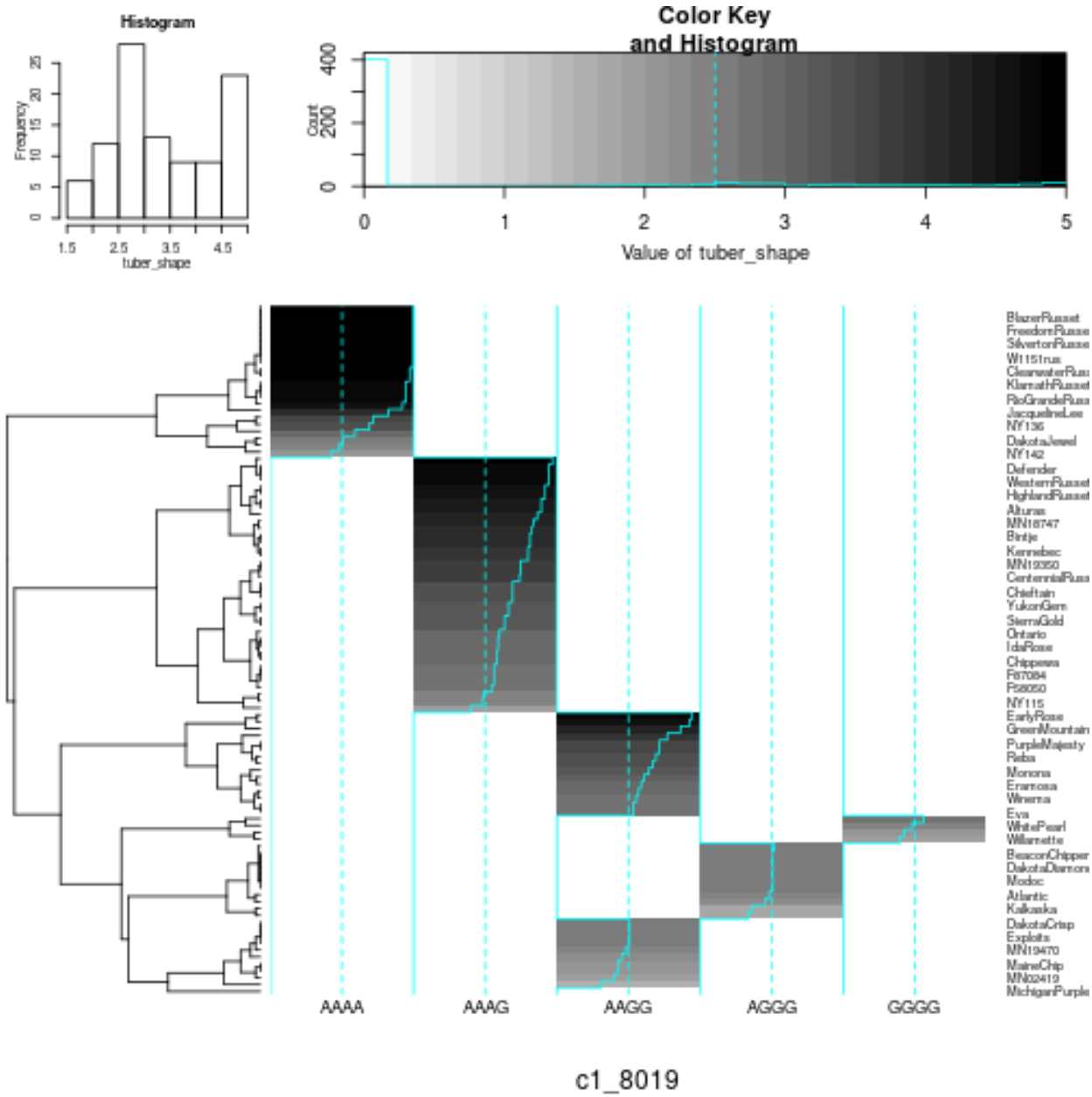

GWASpoly-Full
Significative
Best
Chromosome
$-\log_{10}(p)$
$\lambda_{GC} = 0.941$
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$
PLINK-Full
Significative
Best
Chromosome
$-\log_{10}(p)$
$\lambda_{GC} = 0.732$
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$
SHEsis-Full
Significative
Best
Chromosome
$-\log_{10}(p)$
$\lambda_{GC} = 6.018$
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$
TASSEL-Full
Significative
Best
Chromosome
$-\log_{10}(p)$
$\lambda_{GC} = 1.105$
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

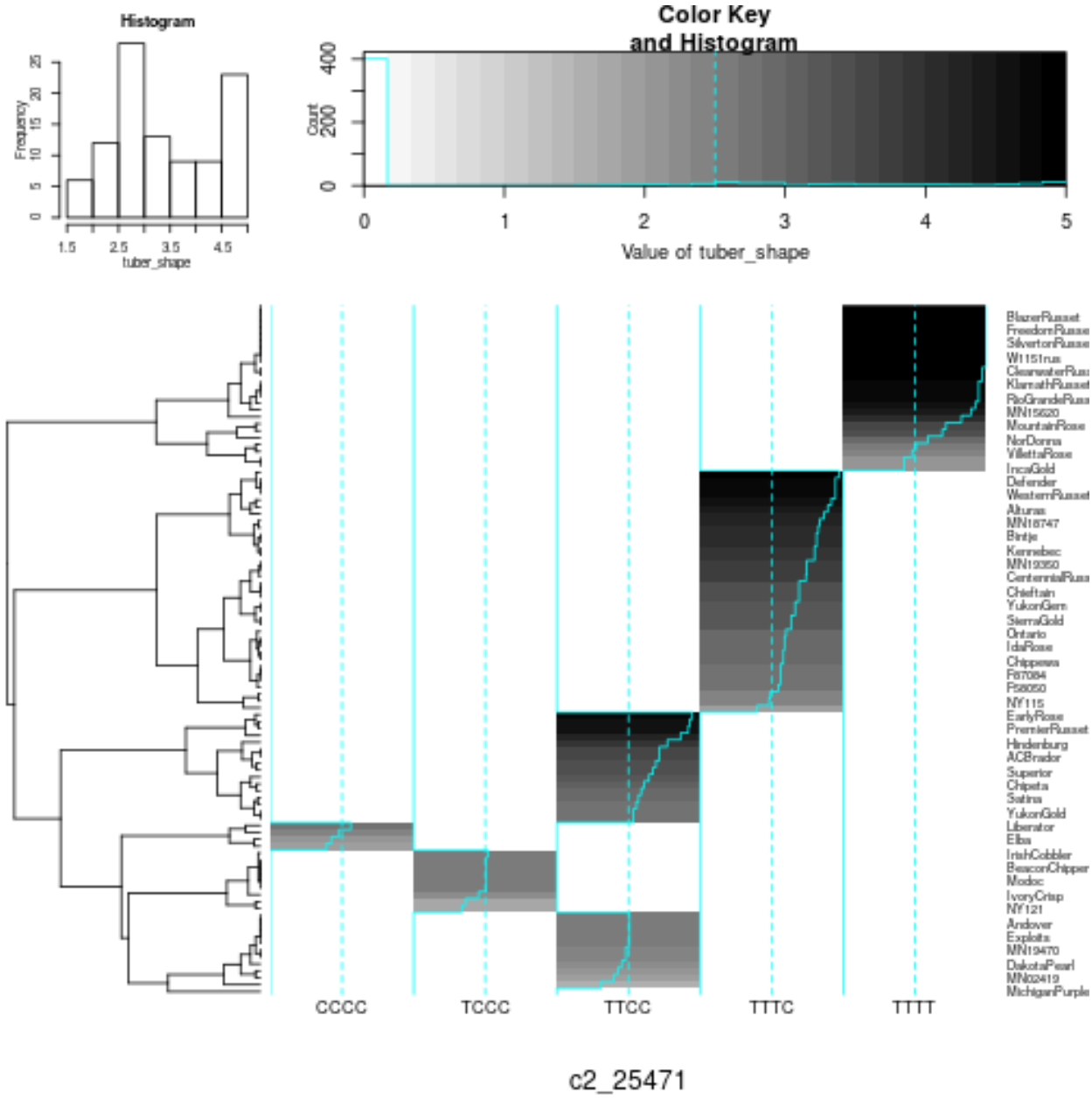
Profiles for common significative SNPs

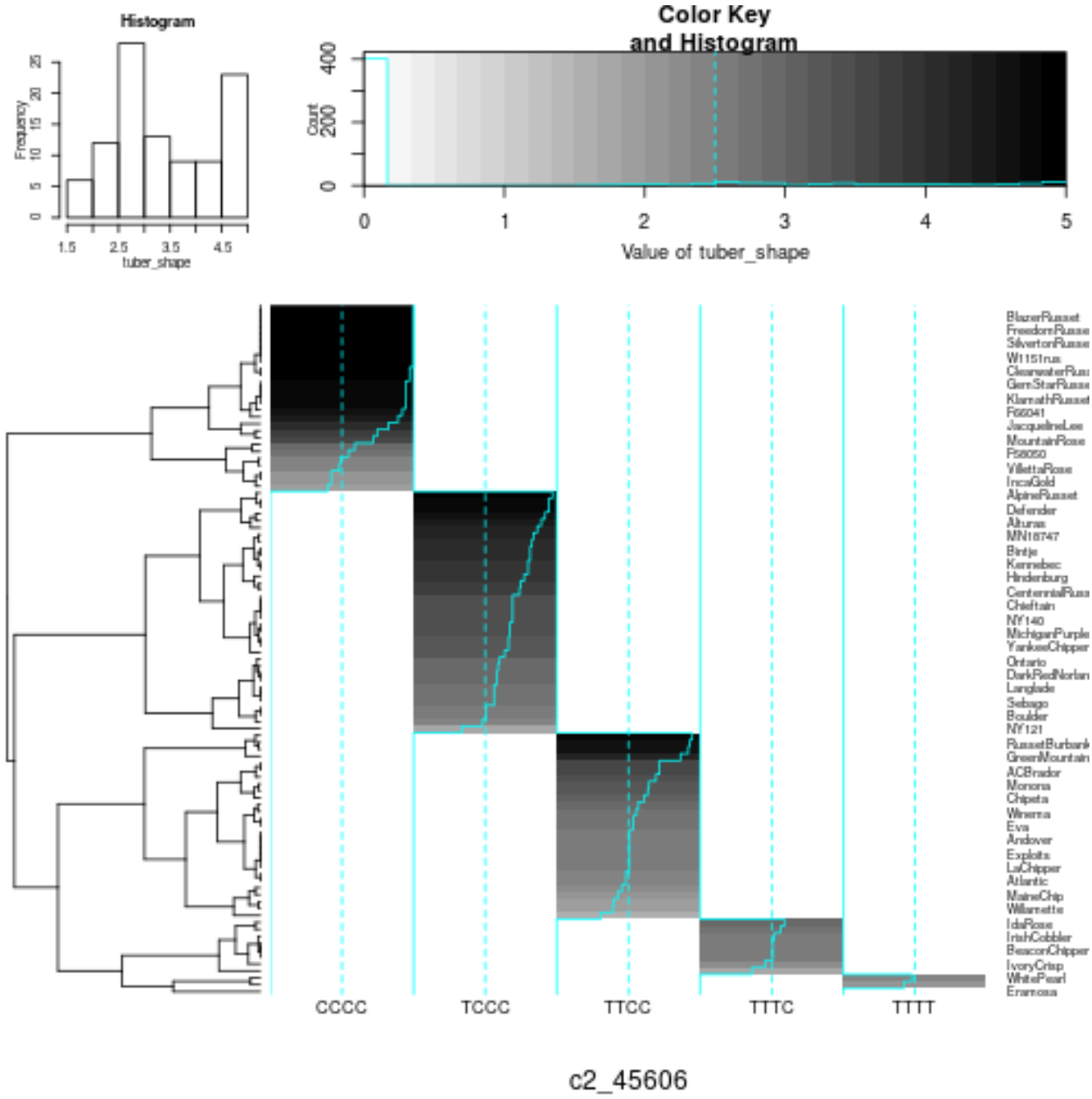
For the common significative SNPs, MultiGWAS provides a figure to visualize each trait by individuals (rows) and genotypes (columns). At the left, the individuals are grouped in a dendrogram by their genotype. At the right, there is the name or ID of each individual. At the bottom, the genotypes are ordered from left to right, starting from the major to the minor allele (i.e., AAAA, AAAB, AABB, ABBB, 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.




Histogram
Frequency
tuber_shape
Color Key
and Histogram
Count
Value of tuber_shape
BlazerRusset
FreedomRusse
SilvertonRusse
W1151rus
ClearwaterRus
KlamathRusset
RioGrandeRus
MN15620
MountainRose
NorDonna
VillettaRose
IncaGold
Defender
WesternRusset
Alturas
MN18747
Bintje
Kennebec
MN19350
CenterntialRus
Chieftain
YukonGem
SierraGold
Ontario
IdaRose
Chippewa
F87084
F58050
NY115
EarlyRose
PremierRusset
Hindenburg
ACBrador
Superior
Chipeta
Satina
YukonGold
Liberator
Elba
IrishCobbler
BeaconChipper
Modoc
IvoryCrisp
NY121
Andover
Exploita
MN19470
DakotaPearl
MN02419
MichiganPurple
CCCC
TCCC
TTCC
TTTC
TTTT
c2_25471


Histogram
Frequency
tuber_shape
Color Key
and Histogram
Count
Value of tuber_shape
CCCC
TCCC
TTCC
TTTC
TTTT

c2_45606