

Rose GWAS browser

Survival guide

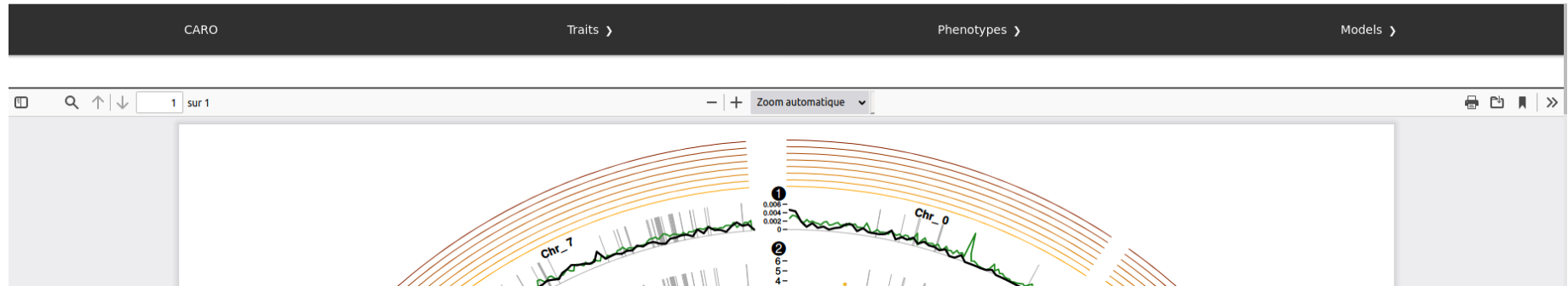


Leroy T., Albert E., Thouroude T., Baudino S., Caissard J-C., Chastellier A., Chameau J., Jeauffre J., Loubert T., Paramita S.N., Pernet A., Soufflet-Freslon A., Oghina-Pavie C., Foucher C., Hibrand-Saint Oyant L.* , Clotault J.*
Dark side of the honeymoon: reconstructing the Asian x European rose breeding history through the lens of genomics, bioRxiv

How to browse on this website?

1/ First **select a main trait** on the drop-down menu « Traits », e.g. « Black Spot »

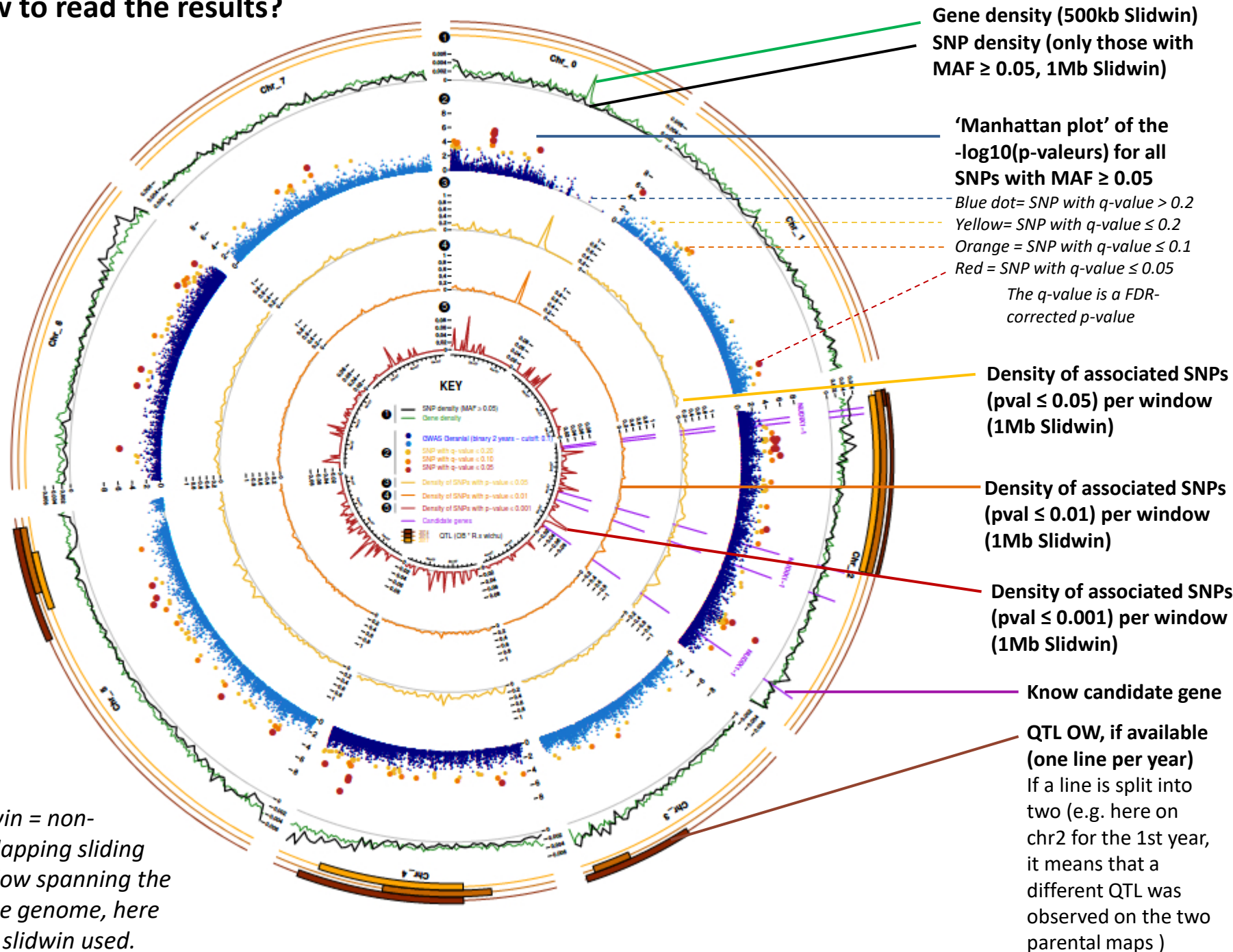
2/ Then you will **have access to a new drop-down menu « Phenotypes »**, « Parameters » or similar to more precisely select a phenotype of interest, *e.g.* « Blackspot - 3 years ». For some phenotypes, GWAS was also performed on (binary: 0/1 around the indicated threshold value)



3/ Then a pdf is expected to be open on the webpage. By default, the results shown are those from the general model of GWASpoly. For more information regarding the GWASpoly model, see slide 4 of this survival guide.

4/ To select any other model than the general model, a **new drop-down menu (« Models »)** is then available. Select the « additive » model for instance (see also slide 4).

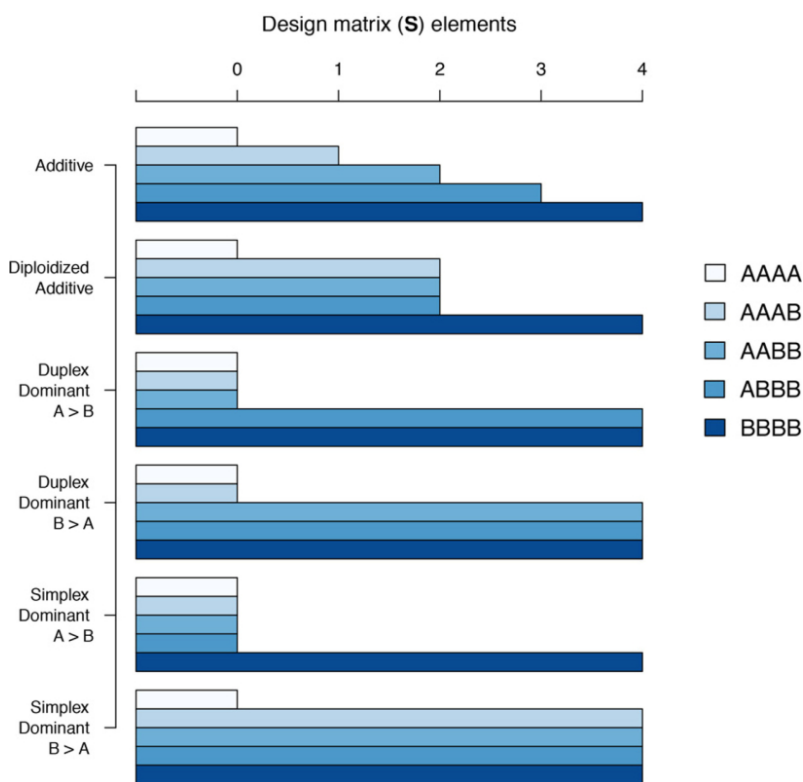
How to read the results?



Why so many different GWAS models ?

It is important to note that performing GWAS on a polyploid species generates an additional degree of complexity regarding the genetic dominance of a given trait.

Note that even if all our 204 rose genotypes are not tetraploids, all were genotyped as tetraploid (most are still truly tetraploids, the remaining genotypes are mostly diploid). At each SNP, each individual has 5 possible classes of alleles B {0, 1, 2, 3, 4} depending on the observed dosage levels of allele B, *i.e.* AAAA, AAAB, AABB, ABBB or BBBB.



The general model is a simple model where the SNP effect can be on any of the classes. This model allows to rapidly explore the results, but sometimes highlights too many associated SNPs (and can provide a quite noisy signal)

The other models are more specific regarding the dominance (see opposite and the GWASpoly paper for details)

How to rapidly identify the most interesting genomic regions for each trait to focus on?

Main Trait (benzylalcohol)

Quantitative/qualitative phenotypic variation

GWAS model

compounds	traits	models	chr0	chr1	chr2	chr3	chr4	chr5	chr6	chr7
Benzylalcohol	quantitative	general	NA	NA	NA	NA+	NA+/-	NA	NA	NA
		diploidized general		+		++				++
		additive				+++	+			++
		diploidized additive		+		+++				++
		simplex dominant F				+++				+++
		simplex dominant A		++		++				++
		duplex dominant RNA		NA	NA	NA+	NA+/-	NA	NA	NA
		duplex dominant A NA		NA	NA	NA+	NA+/-	NA	NA	NA+/-
	qualitative 0.1	general	NA+/-	NA+/-	NA+/-	NA++	NA+/-	NA++	NA+	NA
		diploidized general				+++			+	
		additive			+	+			++	
	job en ligne	diploidized additive				+++			++	
		simplex dominant F				+++			++	
		simplex dominant A				+++			++	+
		duplex dominant R				++		+	++	
		duplex dominant A				++			++	
	Qualitative 5	general	+/-		+/-	+++				+
		diploidized general	+/-			+++				++
		additive				+++	+/-			++
		diploidized additive	+/-			+++				++
		simplex dominant F				+++				+/-
		simplex dominant A	+/-			+++				+/-
		duplex dominant R				++				
	job en ligne PD	duplex dominant A				++				
Betacitronellol	quantitative	general			++			+	++	
		diploidized general						+/-	+/-	

Key

+++	Clear signal, main peak
++	Likely true signal, 2 nd peak
+	weaker local signal
+/-	Even weaker local signal
NA	background noise too strong
NA+/-	background noise too strong, weak evidence
NA+	background noise too strong, but likely signal
NA++	background noise too strong, but clear signal
	none (or evidence considered to be too weak)

Whatever the analysis done (quali/quant) and the model used, the main signal is on chr3

Secondary signal on chr6 if the variation is qualitative for genotypes with very low (<0.1) vs. at least moderate (>0.1) content in Benzylalcohol

Secondary signal on chr7 if the variation is quantitative or qualitative for genotypes with low (<5) vs. high (>5) content in Benzylalcohol

This spreadsheet is available for download here:

https://github.com/roseGWASbrowser/rosegwasbrowser.github.io/blob/master/Summary_GWAS_CARO_100322.xlsx

How to have access to more information that the circlize view shown on the website?

All the results are available on a dedicated github repository:

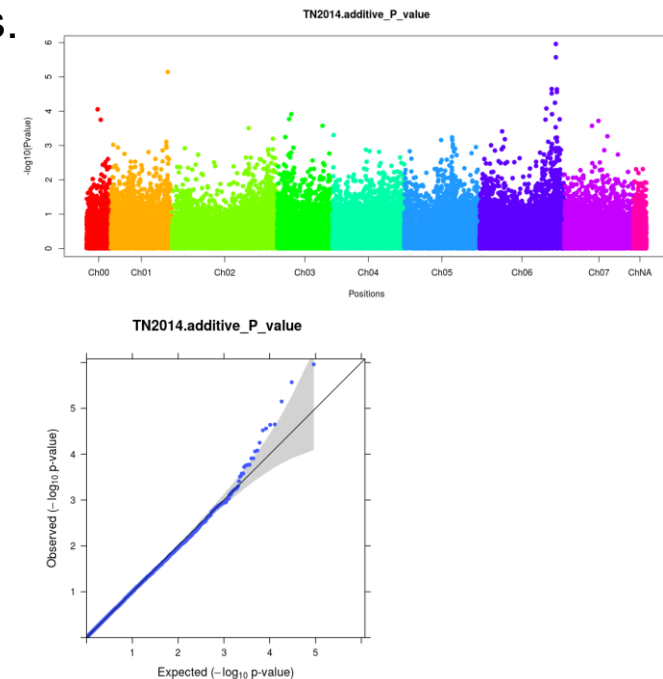
https://github.com/roseGWASbrowser/PopGen_GWAS_19thcentury_roses/tree/main/GWAS/ALL_GWAS_RESULTS/

The organization of result directories mirrors the structure of the website making it easy to locate the different zip files containing the results.

Concretely, each zip file includes:

- Linear Manhattan plots for all models
- QQplots for all models
- Raw values (p-values, q-values, ...)

For most of the traits, this last repository also contains local density in associated SNPs (./raw_results/res_density...)



Given the large amount of information available on the website and the GitHub repository, some details may be missing and occasional errors may be present. **For any questions regarding the results, please feel free to email me: [thibault.leroy\[AT\]inrae\[DOT\]com](mailto:thibault.leroy@inrae.fr)**