ANR-13-JSV6-0001

eQTLs are key players in the integration of genomic and transcriptomic data for phenotype prediction

Abdou Rahmane WADE[1], Harold Duruflé[1], Leopoldo Sanchez[1*] and Vincent Segura[2*]

[1] *INRAE, ONF, BioForA, UMR 0588, F-45075 Orleans, France*

[2] *UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, F-34398 Montpellier, France*

* *equal contribution*

Instituto Superior de Agronomia
20 – 23 June 2022, Lisbon, Portugal

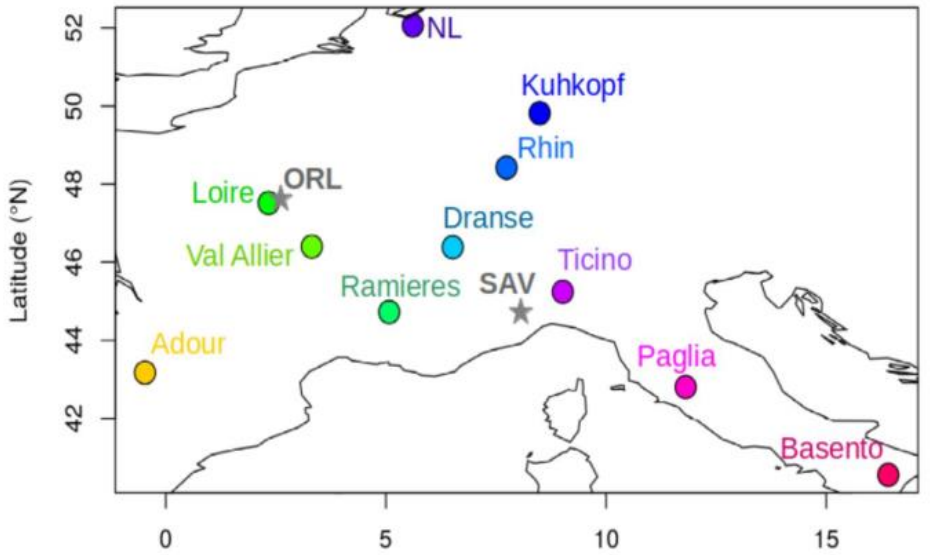
- Multi-omics represent the "***missing link***" between phenotypes and genome variation.
- The mechanisms by which integration is successful when predicting phenotypes are still not known, with **redundancy** between omics being one of the possible explanations. Redundancy reflects the interconnectivity from raw genomic sequence to the organismal phenotype.
- Both redundancy and interactivity are key to **understanding genetic architecture** beyond the simple list of effects that is typically provided by genomic approaches.

Phenotypes: 21 traits

- Growth
- Pathogene Tolerance
- Phenology
- Branch Architecture
- Biochemical

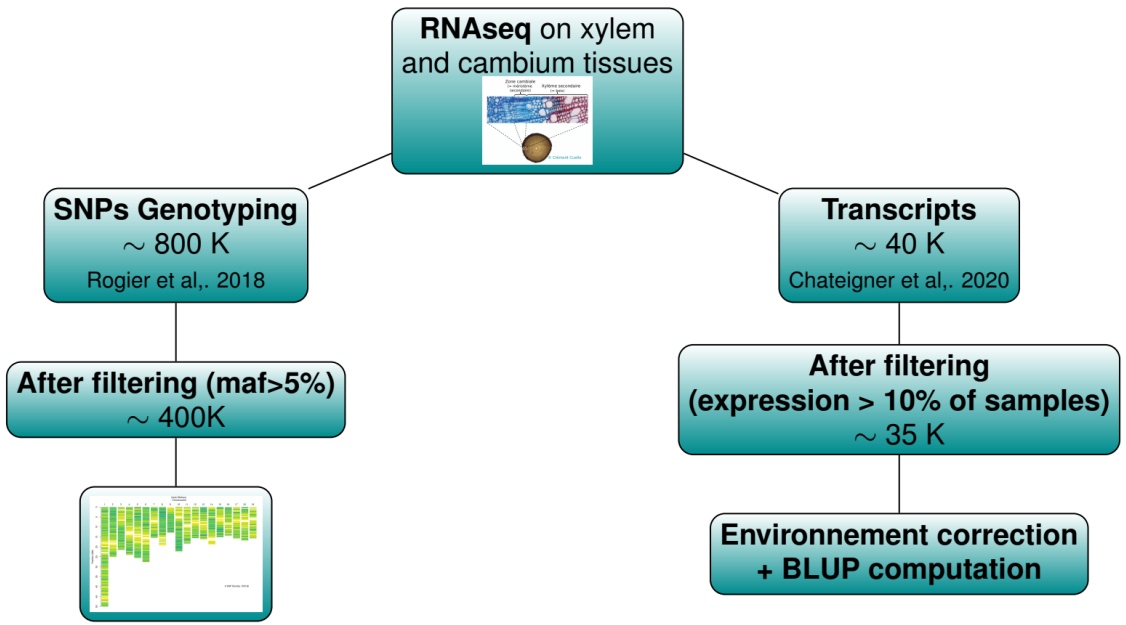
- In our study: we propose new insights on **data integration** for black poplar (Populus nigra L.), using one of the simplest integration alternatives (**concatenation**), combined with one of the most popular prediction approaches (**ridge regression**).
- We aimed to evaluate the factors affecting prediction accuracy when integrating **genomic** and **transcriptomic** data for phenotype prediction using a large number of diverse phenotypes collected in two common gardens.

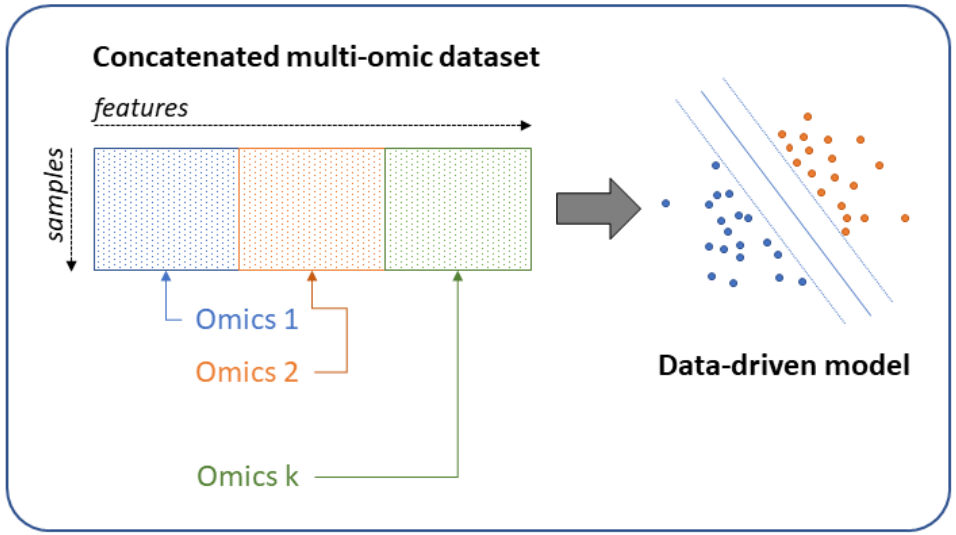
An original experimental design

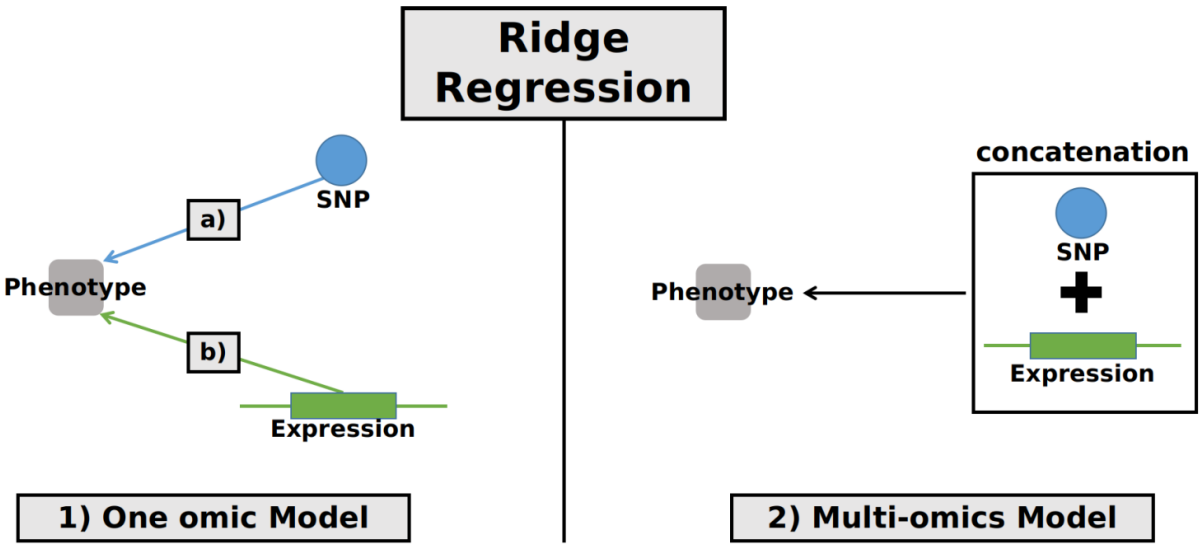


241 genotypes representing the genetic diversity of the West Europe

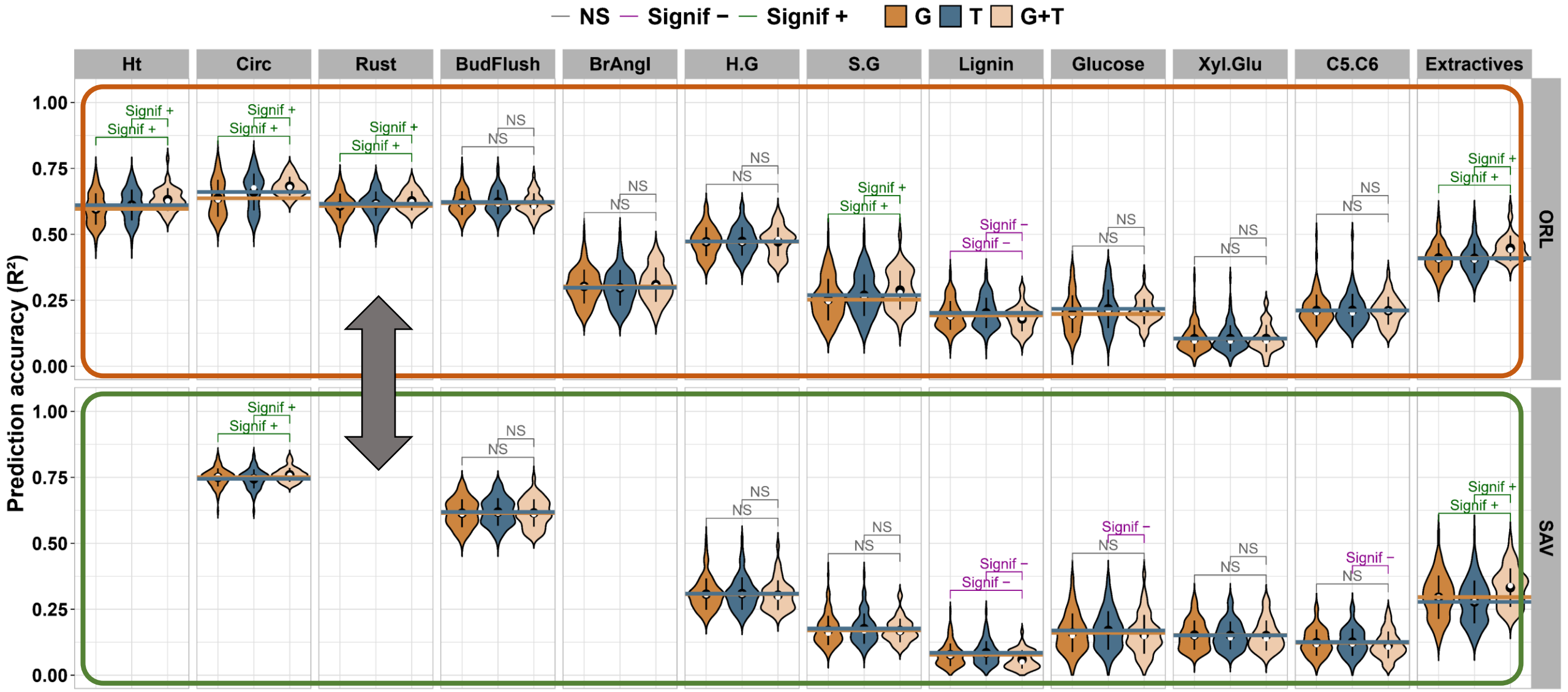
Strategies & Objectives



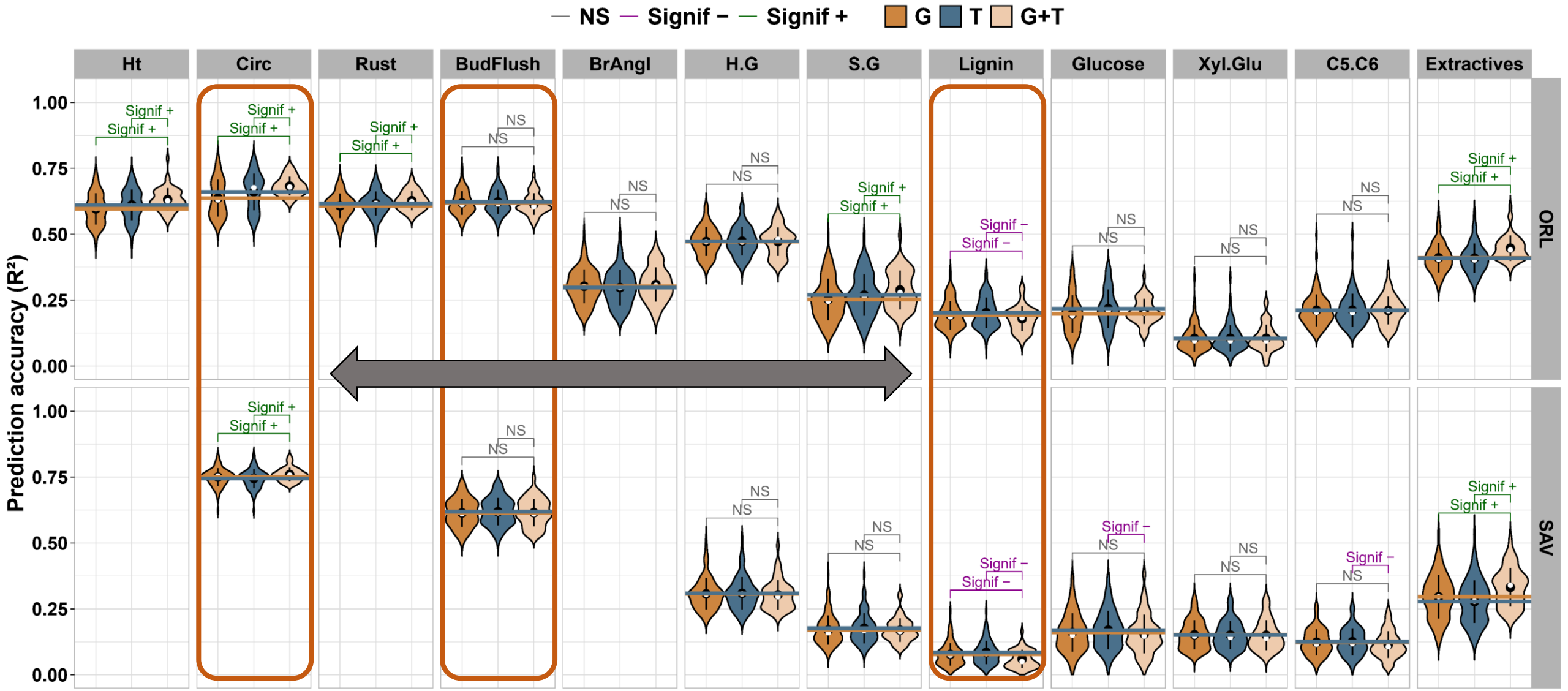




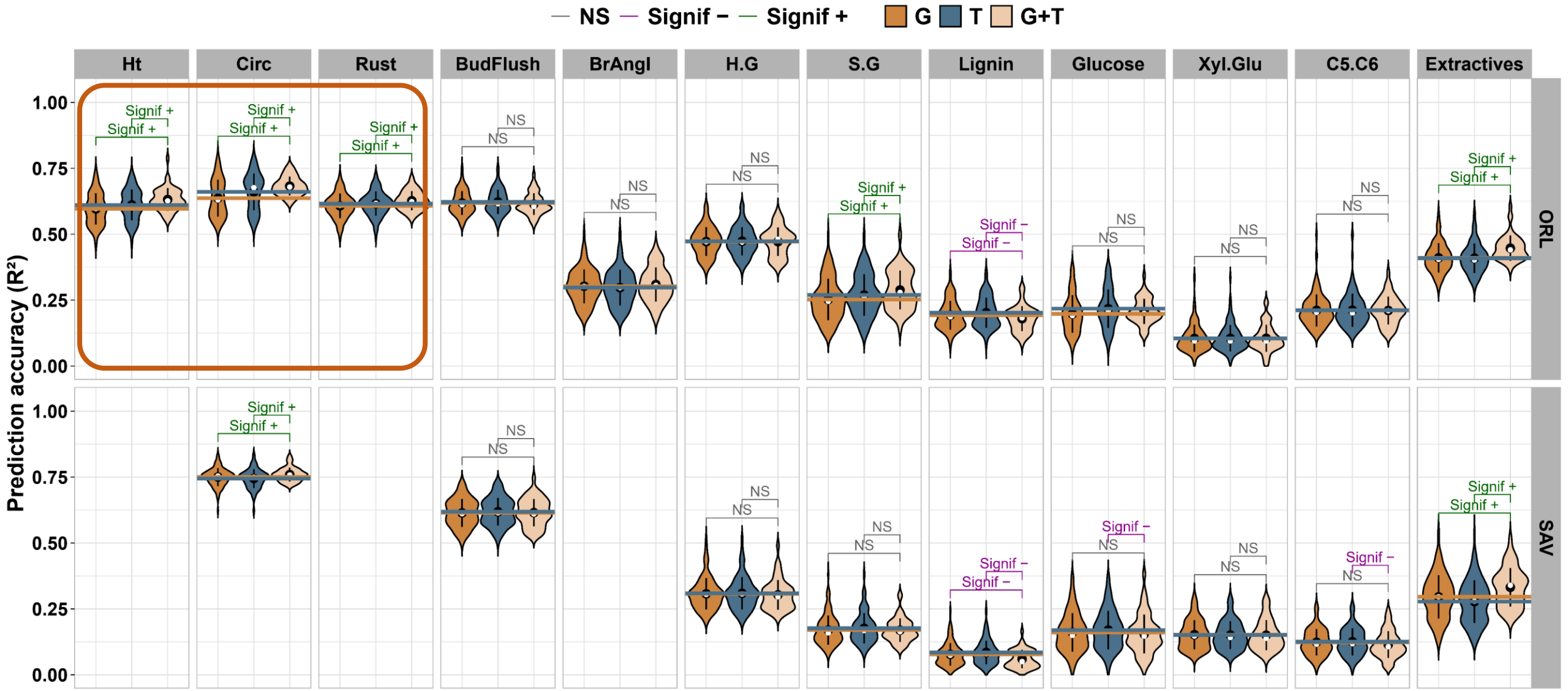
Results


NS
Signif −
Signif +
G
T
G+T
Ht
Circ
Rust
BudFlush
BrAngl
H.G
S.G
Lignin
Glucose
Xyl.Glu
C5.C6
Extractives
ORL
SAV
Prediction accuracy (R²)
1.00
0.75
0.50
0.25
0.00

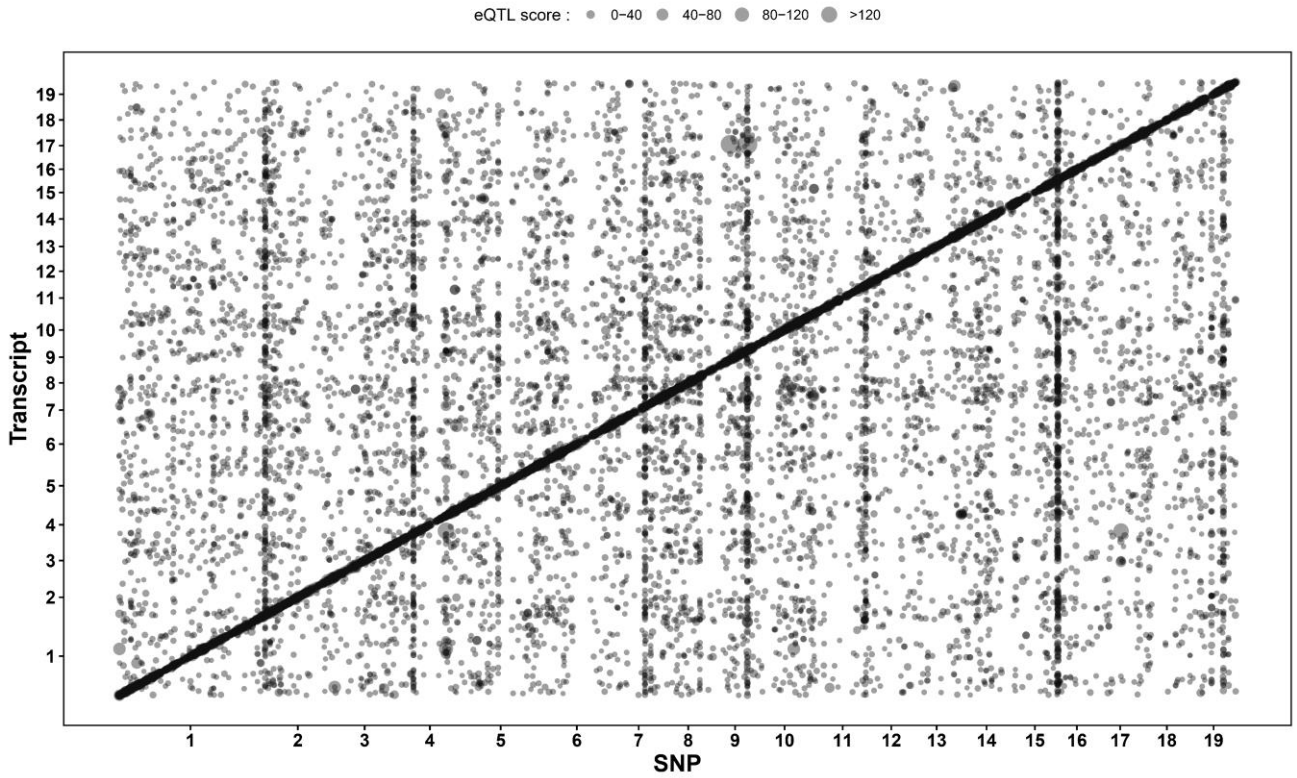
Results

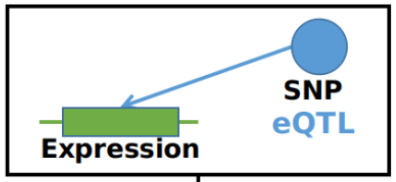

NS
Signif −
Signif +
G
T
G+T
Ht
Circ
Rust
BudFlush
BrAngl
H.G
S.G
Lignin
Glucose
Xyl.Glu
C5.C6
Extractives
ORL
SAV
Prediction accuracy (R²)
1.00
0.75
0.50
0.25
0.00
Signif +
Signif −
NS

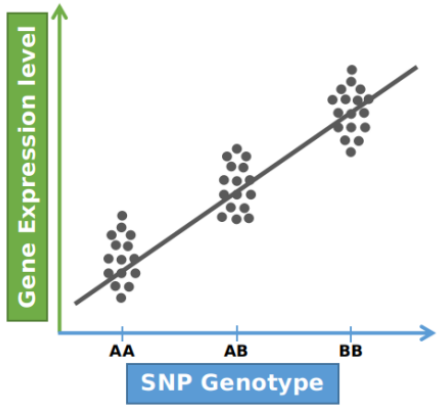
Results

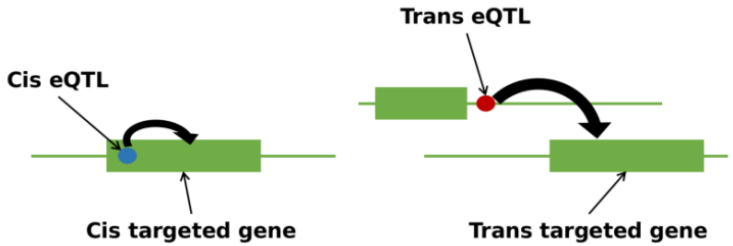

NS
Signif −
Signif +
G
T
G+T
Ht
Circ
Rust
BudFlush
BrAngl
H.G
S.G
Lignin
Glucose
Xyl.Glu
C5.C6
Extractives
ORL
SAV
Prediction accuracy (R²)
1.00
0.75
0.50
0.25
0.00
Signif +
Signif −
NS

Results



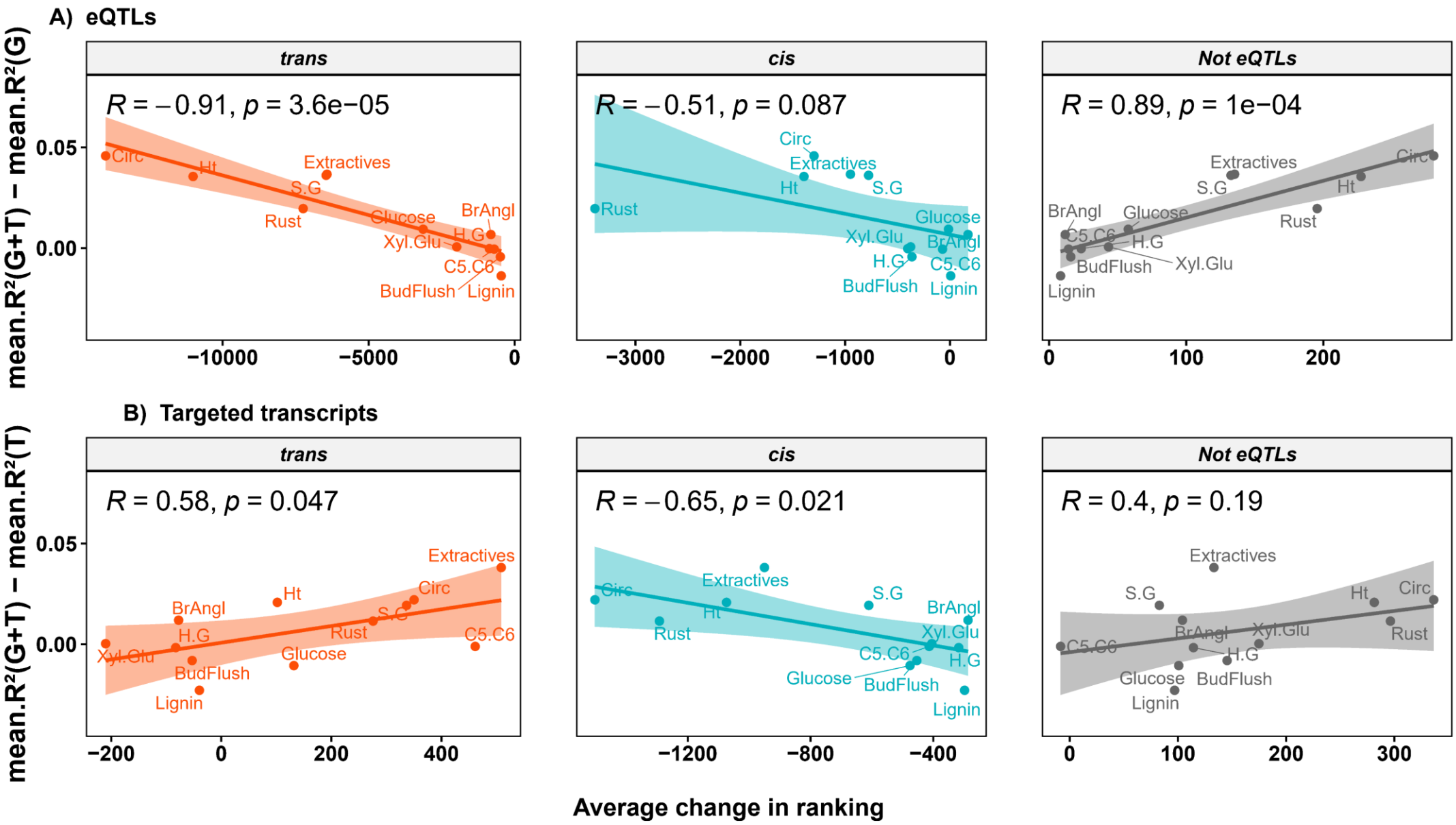






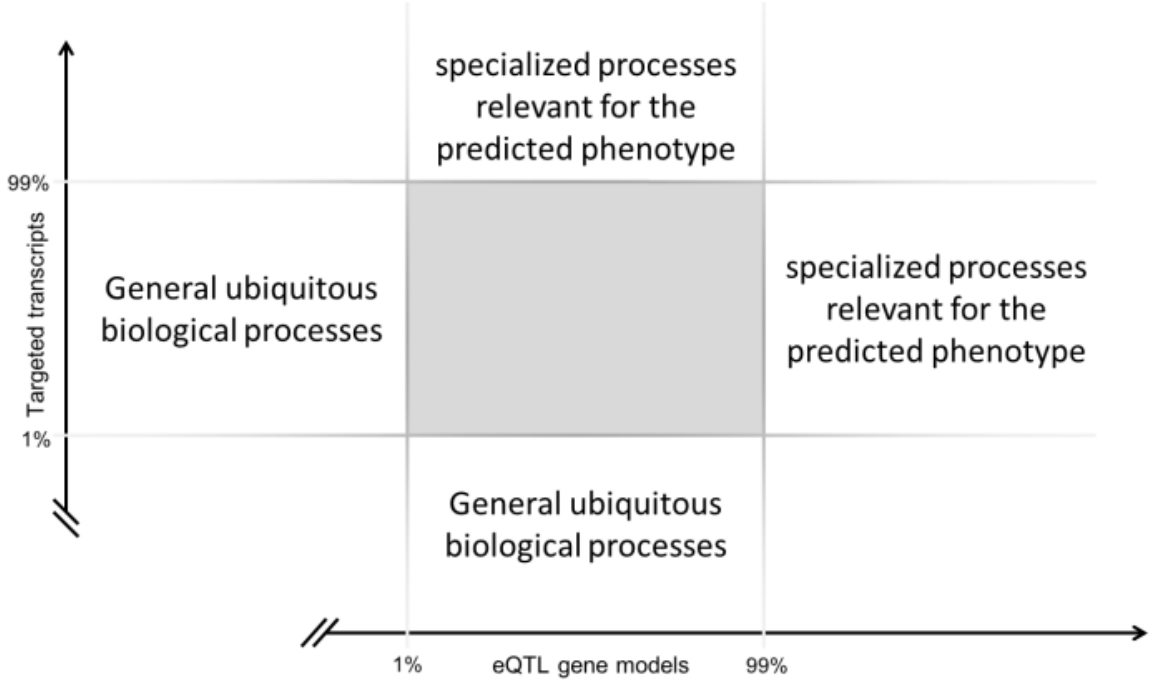
- Classification of predictors : TRANS, CIS , and Not-eQTLs

Results

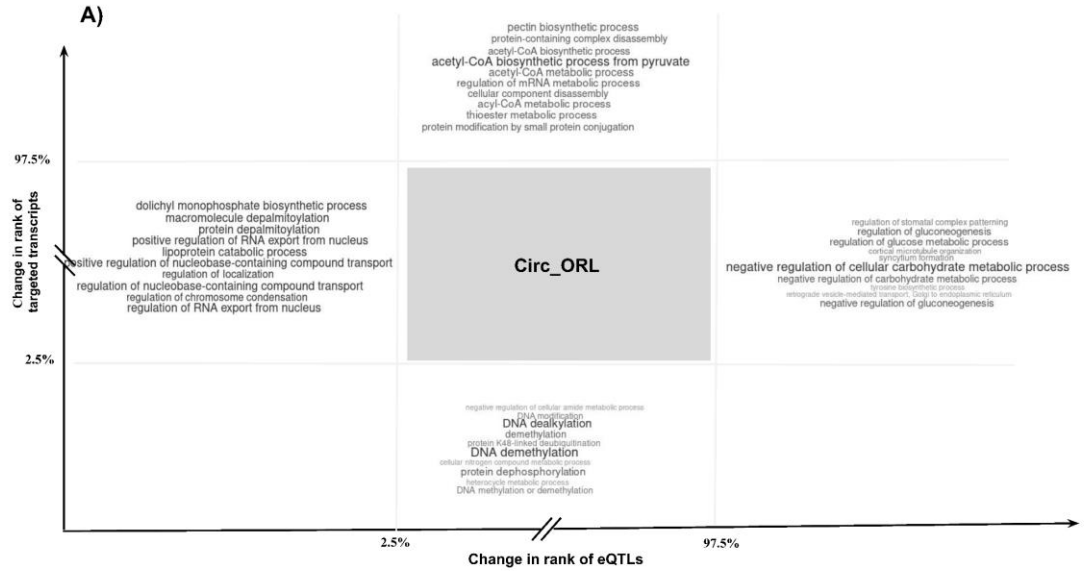

A) eQTLs
mean.R²(G+T) − mean.R²(G)
trans
R = −0.91, p = 3.6e−05
Circ
Ht
Extractives
S.G
Rust
Glucose
BrAngl
Xyl.Glu
H.G
C5.C6
BudFlush
Lignin
0.05
0.00
−10000
−5000
0
cis
R = −0.51, p = 0.087
Circ
Extractives
Ht
S.G
Rust
Glucose
Xyl.Glu
BrAngl
H.G
C5.C6
BudFlush
Lignin
−3000
−2000
−1000
0
Not eQTLs
R = 0.89, p = 1e−04
Extractives
S.G
Circ
Ht
BrAngl
Glucose
Rust
C5.C6
H.G
BudFlush
Xyl.Glu
Lignin
0
100
200
B) Targeted transcripts
mean.R²(G+T) − mean.R²(T)
trans
R = 0.58, p = 0.047
Extractives
Ht
Circ
BrAngl
S.G
H.G
Rust
C5.C6
Xyl.Glu
Glucose
BudFlush
Lignin
0.05
0.00
−200
0
200
400
cis
R = −0.65, p = 0.021
Circ
Extractives
S.G
BrAngl
Ht
Xyl.Glu
Rust
C5.C6
H.G
Glucose
BudFlush
Lignin
−1200
−800
−400
Not eQTLs
R = 0.4, p = 0.19
Extractives
S.G
Ht
Circ
C5.C6
BrAngl
Xyl.Glu
Rust
H.G
Glucose
BudFlush
Lignin
0
100
200
300
Average change in ranking

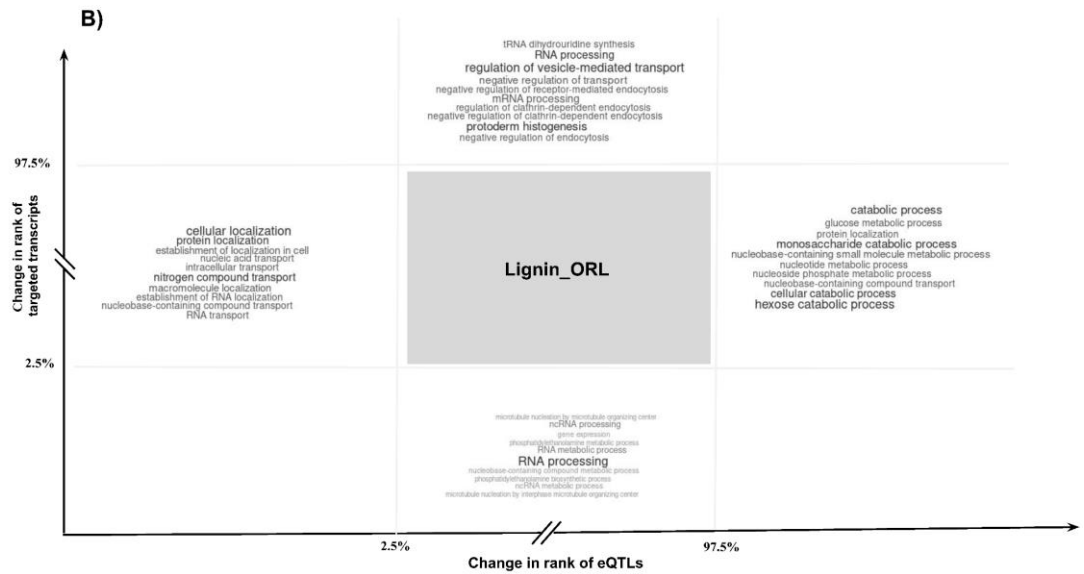
Results

5. Gene ontology analysis suggests that top targeted transcripts or eQTLs are trait specific



Schematic representation of the enriched GO terms among the top targeted transcripts or eQTL gene models list for the circumference of the tree trunk.





Conclusion

- Consequently, beneficial integration happens when the redundancy of predictors is decreased, leaving the stage to other less prominent but complementary predictors.

- An additional gene ontology (GO) enrichment analysis appeared to corroborate such statistical output.

- These two complementary approaches showed empirically over a series of traits how the best predicting scenarios are built, excluding certain features while promoting others according to their redundancies within the data.

- To our knowledge, this is a novel finding delineating a promising method to explore data integration.