

Validation of a Metabolite-GWAS Network for *Populus trichocarpa* Family 1 UDP-Glycosyltransferases

Background

- Metabolite genome wide association studies (mGWAS) are increasingly used in plant genomics to predict metabolite-gene associations, but high-throughput experimental validation is lacking.

Approach

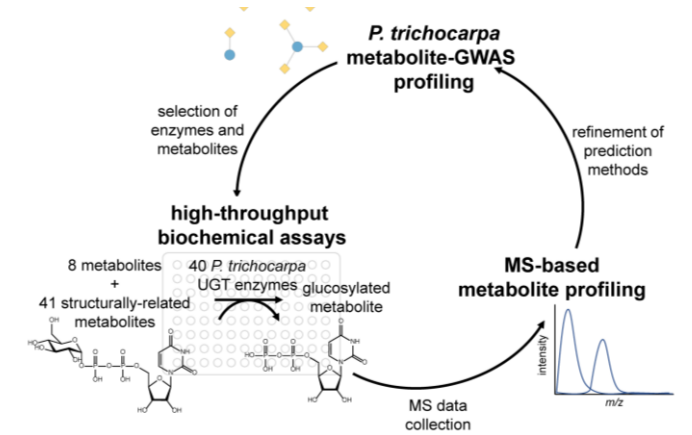
- A functional genomics workflow for validating mGWAS-predicted enzyme-substrate relationships was developed for uridine diphosphate glycosyltransferases (UGTs), which are critical for metabolite functionalization. The mGWAS-predicted UGT-metabolite associations were interrogated using high-throughput expression, biochemical assays, and mass spectrometry-based reaction analysis.

Results

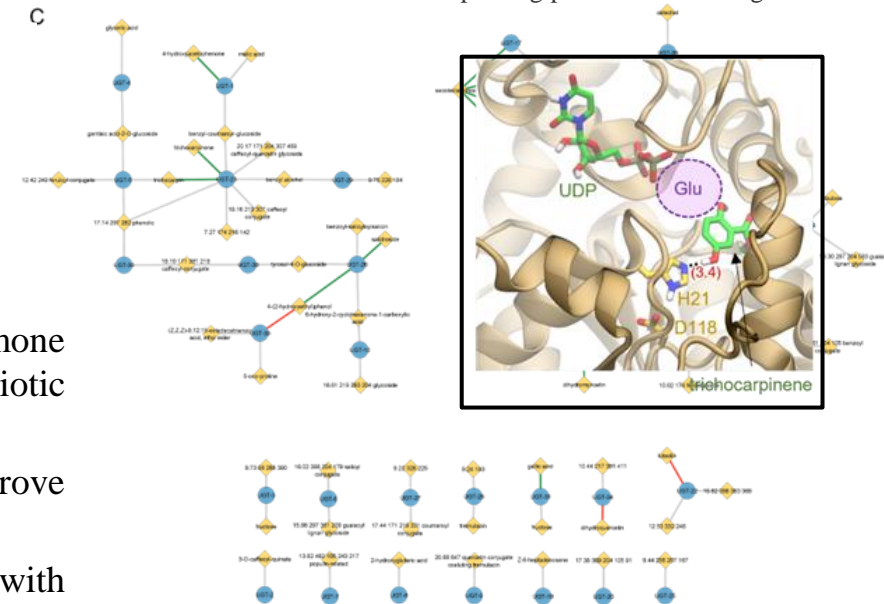
- UGTs displayed relaxed substrate specificity.
- Ten of 13 metabolite-gene associations represented in the biochemical assays were confirmed.
- AlphaFold docking confirmed that trichocarpinene, a metabolite in the mGWAS network, would properly orient in the binding pocket of UGT-23 to produce trichocarpin, a major higher-order salicylate in *Populus trichocarpa*.

Significance

- The identified UGTs are putatively involved in lignan, flavonoid, salicylate, and phytohormone metabolism, with potential implications for cell wall biosynthesis, nitrogen uptake, and biotic and abiotic stress response.
- Validating predicted relationships in metabolite-GWAS networks in biochemical assays will improve prediction models and identify gene targets for manipulating metabolite production.
- These studies have implications for identifying enzymes that can transform secondary metabolites with utility in biomedical and bioenergy applications.



Functional genomics workflow for experimentally validating mGWAS associations and improving predictive modeling



AlphaFold docking of the UGT-23 and trichocarpinene, whose mGWAS-predicted relationship is boxed in the network