

Integrating mRNA and protein sequence data to enable the detection and quantitative profiling of natural protein sequence variants in *Populus*

Background and Approach

- *Populus* genome wide association studies have transformed the ability to link genotypes to phenotypes and facilitate the dissection of the genetic contribution to complex traits.
- There is a significant need to characterize the impact genetic variants at the functional level (i.e., protein level).
- In this study, we used next-generation RNA-sequencing data to generate genotype-specific protein databases, which allowed the identification and quantification of sequence variations at the protein level.

Outcomes

- A variety of natural sequence protein variants were quantified in two *Populus trichocarpa* genotypes.
- In addition to expected, neutral single amino acid polymorphisms, investigators characterized polymorphisms predicted to be non-neutral and located in regions of the genome predicted to have undergone recent positive and/or divergent selection and therefore represent a candidate list of protein variants relevant to plant adaptability.

Significance

- As demonstrated, profiling genotype-specific proteomes derived from RNA sequencing data better defines the link between genotypes and phenotypes.
- This will enable future studies to detect and quantitatively profile non-neutral variants underpinning bioenergy-relevant traits.

