**A Transcription Factor Network for Populus Trichocarpa**

Carissa Bleker, Deborah Weighill, Jay Chen, Wellington Muchero, Gerald Tuskan, and
Daniel Jacobson

Oak Ridge National Laboratory

blekercr@ornl.gov

**Abstract**

The phenotype of an organism is the result of the interaction between its genome and the environment. The expression of a gene is partly dependent on its transcription factors: proteins that bind to specific parts of the genome and thereby play a role in transcriptional regulation. An expression Quantitative Trait Nucleotide (eQTN) is a Single Nucleotide Polymorphism (SNP) that is responsible for the association between the genotype and a gene expression event. Finding these SNPs can help identify transcription factors controlling certain genes, thus determining the transcription factor network. Gene expression data was collected from the xylem of over 530 Populus trichocarpa genotypes and for almost 40,000 genes. To find SNPs associated with the expression of a particular gene, we performed a Genome Wide Association Study (GWAS) using EMMAX. EMMAX finds associations between SNPs and a measured phenotype using a linear mixed model, and using a kinship matrix to correct for population structure amongst the genotypes. After controlling for multiple hypothesis testing, the resulting significant associations were used to create a SNP-to-gene (expression) network. By mapping these SNPs to known genes by chromosomal position, we created a gene-to-gene directed network in which transcription factors can be identified using existing annotation. Through clustering of the transcription factors, we can find strongly associated modules for hypothesis generation. The creation of a transcription factor network (and the latter inclusion of the downstream target genes) will allow us to better understand the regulatory architecture of Populus trichocarpa and to determine which of the transcription factors have a high probability of being master regulators.