**Centromere Wavelet Signatures and Co-evolution with CENH3 in Populus trichocarpa**

Deborah Weighill, David Macaya-Sanz, Stephen DiFazio, Gerald Tuskan and
Daniel Jacobson

Oak Ridge National Laboratory

weighillda@ornl.gov

**Abstract**

A vast collection of different data types is available for Populus trichocarpa. Approximately 1,000 genomes from different P. trichocarpa genotypes have been sequenced and 12 clonal replicates of each genotype have been propagated across common gardens in four different locations. This has provided a large set of ~28,000,000 Single Nucleotide Polymorphisms (SNPs) which have recently been publicly released. Many molecular and growth phenotypes have been measured across this population and provide an unparalleled resource for Genome Wide Association Studies. DNA methylation data in the form of MeDIP (Methyl-DNA immunoprecipitation) sequencing has been performed on 10 different P. trichocarpa tissues. Transposable elements (TE) and low complexity regions (leading to gaps in the chromosome scaffolds) have also been identified and included as independent layers of genomic information. We have used these different data types as signals, which vary across a chromosome, for example, the gene density, SNP density, SNP correlation (LD) density, TE density, low complexity sequence density, or methylation density across a chromosomes in P. trichocarpa genome. Applying the Continuous Wavelet Transform signal processing technique allowed us to characterize the variation in these signals at multiple scales, and consequently, allowed us to identify the centromere positions of each chromosome based on these various data signals.