

CHROMATIN ACCESSIBILITY IN THE LOBLOLLY PINE GENOME

Lilian P. Matallana¹, Sumaira Zaman², Jill Wegrzyn³, Ross W. Whetten¹

¹Department of Forestry and Environmental Resources, North Carolina State University, Raleigh NC;

²Department of Computer Science, University of Connecticut, Storrs CT; ³Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs CT

The nucleus of a diploid loblolly pine cell is less than 15 microns in diameter, but contains about 14 meters of double-stranded DNA. Packing that DNA into the cell nucleus requires efficient organization of DNA into chromatin, a protein-DNA complex with higher-order structure. Some regions of chromatin are highly condensed and relatively inaccessible to nuclear factors, while other regions are less condensed and more accessible. Results from several species indicate that both regulatory and transcribed DNA sequences are enriched in regions of accessible chromatin, and that DNA sequence variation in regions of accessible chromatin can be associated with observed phenotypic variation. We conducted two experiments to explore chromatin structure in the loblolly pine genome, using nuclei isolated from the same tissue type. One experiment used Assay for Transposase-Accessible Chromatin by Sequencing on six replicate samples, while a second experiment used digestion with micrococcal nuclease followed by sequencing on a single sample of isolated nuclei. Comparing and contrasting the results of these experiments shows that the transposase-based method seems to be more tightly focused on a smaller proportion of the entire genome, while the micrococcal-nuclease based method yields a broader distribution of sequence reads across a larger proportion of the genome. Regions with coverage above the 99th percentile within each experiment were combined and used to search for transcribed regions of the pine genome that do not contain predicted gene models in the current version of the genome annotation. The recovered candidate regions were examined for the presence of conserved protein coding domains and putative single-copy orthologs of genes described in other Embryophyta. The results of these analyses suggest that chromatin accessibility can add value to genome annotation efforts, improving the ability to discriminate between functional genes and processed pseudogenes in conifer genomes.