Genomics Study of Western White Pine (*Pinus monticola*) Genetic Resistance Against White Pine Blister Rust

J.-J. Liu[1, *] and R.A. Sniezko[2]

[1] *Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, British Columbia, Canada*

[2] *USDA Forest Service, Dorena Genetic Resource Center, Cottage Grove, Oregon, USA*

[*] *Jun-jun.liu@nrcan.gc.ca*

Western white pine (WWP, *Pinus monticola*) is a long-lived conifer with an extensive geographic range in western North America. It is of high interest in ecological studies and forest breeding because of its high susceptibility to the invasive disease white pine blister rust (WPBR, caused by the fungus *Cronartium ribicola*). However, *P. monticola* lacks genomic resources and is evolutionarily far away from plants with available draft genome sequences. Use of high-throughput RNA-sequencing (RNA-seq) technology is a cost-effective strategy to generate substantial transcriptome data for global transcript profiling and DNA marker discovery. We report here the RNA-seq analysis results using Illumina® (San Diego, California, USA) deep sequencing of *P. monticola* infected with *C. ribicola*. De novo gene assembly was used to generate the first *P. monticola* consensus transcriptome of primary needles, which contained 39,439 unique transcripts with an average length of 1,303 bp and a total length of 51.4 Mb. About 23,000 *Pinus orthologous* genes (POGs) and 200 disease resistance gene analogs (RGAs) encoding NBS-LRR proteins were identified by BLAST search against the *Pinus* gene index database and plant R gene families, respectively. Comparison of transcriptomes from WPBR-susceptible and -resistant white pine genotypes revealed about 1,000 differentially expressed genes (DEGs) with statistical significance during early stages of the compatible and incompatible *P. monticola–C. ribicola* interactions.

In silico single nucleotide polymorphism (SNP) detection identified more than 100,000 high quality SNPs in the above three groups of candidate genes (POGs, RGAs, and DEGs) by a bulked segregation-based RNA-seq analysis. To estimate efficiency of in silico SNP discovery, 432 SNPs were selected to develop genotyping assays using Sequenom® iPlex (Sequenom, San Diego, California) technology. About 70 percent of in silico SNPs were successfully genotyped and 53 percent of the total SNPs were detected with predicted nucleotide variations in a collection of resistance germplasm. SNP clustering analyses consistently identified distinct populations, each composed of multiple full-sib seed families by parentage assignment. Linkage disequilibrium analysis identified 21 genes in significant association with major gene (Cr2) resistance. Genotyping verification by TaqMan™ (Thermo Fisher Scientific, Waltham, Massachusetts, USA) qPCR indicated that at least one SNP locus provided an excellent marker for Cr2 selection among populations across western North America (Oregon and Washington [USA] and British Columbia [Canada]).

The validated SNP markers can be used as novel genomic resources for genetic, evolutionary, and ecological studies in *P. monticola* and related five-needle pines. Our results demonstrate that integration of RNA-seq-based transcriptome analysis and high-throughput genotyping is an effective approach for discovery of a large number of nucleotide variations and for identification of functional gene variants associated with adaptive traits in a non-model species.

In: Schoettle, Anna W.; Sniezko, Richard A.; Kliejunas, John T., eds. 2018. Proceedings of the IUFRO joint conference: Genetics of five-needle pines, rusts of forest trees, and Strobusphere; 2014 June 15–20; Fort Collins, CO. Proc. RMRS-P-76. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 245 p.