**Project Title: Discovery of Genes Controlling Wood Property Traits in Douglas-fir**

**Investigators:** David Neale, Institute of Forest Genetics, PSW; Brad St. Clair, Genetics Team, PNW; Glenn Howe, PNWTIRC, Oregon State University

**Issue:** This project represents Phase II of a collaborative effort to discover genes controlling complex traits of economic and ecological importance in Douglas-fir and to develop DNA marker-based approaches to tree improvement and gene resource conservation. This project is one component of our Douglas-fir Genome Project (http://dendrome.ucdavis.edu/dfgp). The proposed project responds directly to research pathway Biotechnology and Tree Improvement – Gene Discovery.

The primary objectives of Phase I of the Agenda 2020 funded project were to identify candidate genes for important adaptive traits through expressed sequence tag (EST) sequencing and to begin single nucleotide polymorphism (SNP) discovery in candidate genes. This work is completed and was reported at the Agenda 2020 meeting in Boise in April 2004. The Agenda 2020 funding was seed money that led to the funding of a large USDA/NRI Plant Genome competitive grant (Title: Association Mapping of Adaptive Traits in Douglas-fir; $490,000 FY05-07). However, because two important objectives are not covered under the USDA/NRI grant, we now request funding for these additional objectives:

(1) Discover SNPs in candidate genes controlling wood properties in Douglas-fir.
(2) Verify SNP x phenotype associations discovered in experimental populations in applied tree improvement populations for wood property traits (wood specific gravity and microfibril angle).

*Our long-term goal is to genetically dissect complex traits and understand the molecular basis of adaptation and wood quality in Douglas-fir*

We propose to use a population genomic approach called association mapping to dissect complex adaptive traits in Douglas-fir (*Pseudotsuga menziesii*) and to identify the genes (i.e., loci and alleles) that are responsible for phenotypic differences among trees. We have already used traditional quantitative trait locus (QTL) mapping to characterize the genetic control of adaptive traits in Douglas-fir (Jermstad *et al*. 2001a,b, 2003; Wheeler *et al*. 2004) and to study wood property traits in loblolly pine (Groover *et al*. 1994; Sewell *et al*. 2000, 2002; Brown *et al*. 2003). Through a series of large QTL studies in Douglas-fir, we developed a comprehensive understanding of important adaptive traits, including bud flush, bud set, lammas flush, and cold hardiness. The number, size of effect, and approximate location of QTL controlling these traits were estimated, as was the magnitude of QTL interactions with site and environmental treatments. These studies provide a solid foundation for selecting positional candidate genes and mapping selected functional and expression-derived candidate genes.

Under Phase I of Agenda 2020 funding, 11,000 ESTs from Douglas-fir seedlings (root and shoot) were sequenced in collaboration with researchers at Oregon State University (Brunner, Howe, Bond, Meilan, Strauss). These sequences are publicly available at GenBank (http://www.ncbi.nlm.nih.gov) and TreeGenes (http://dendrome.ucdavis.edu/treegenes). The EST database was then used to identify candidate genes for adaptive traits (Table 1). SNP discovery was completed for 18 candidate genes and nucleotide diversity and linkage disequilibrium were

estimated. These SNPs will be used for genotyping and association testing under the USDA/NRI Plant Genome grant.

**Table 1.** Nucleotide diversity in cold resistance, drought tolerance and wood quality positional and expressional candidate genes in Douglas-fir.

| Gene | Biological process and molecular function | Total sites, bp | Total SNP sites (S) | Θ ± SD (per site) | Average frequency of SNPs (bp per SNP) |
|---|---|---|---|---|---|
| 40S-RPS3a | protein biosynthesis | 500 | 12 | 0.0062 ± 0.0026 | 42 |
| 60S-RPL31a | protein biosynthesis | 609 | 21 | 0.0089 ± 0.0033 | 29 |
| ABA-WDS | dehydrin | 344 | 9 | 0.0067 ± 0.0030 | 38 |
| APX | stress response, detoxification | 867 | 26 | 0.0079 ± 0.0029 | 33 |
| EF1A | translational elongation | 1072 | 14 | 0.0034 ± 0.0014 | 77 |
| ERD15 | n/a | | | in progress | |
| Formin | actin cytoskeletal organization | | | in progress | |
| LEA-EMB11 | stabilizing membranes | 545 | 33 | 0.0159 ± 0.0056 | 17 |
| LEA-II | stress response | 504 | 18 | 0.0088 ± 0.0032 | 28 |
| LP3-like | dehydrin | 481 | 16 | 0.0085 ± 0.0032 | 30 |
| MT-like | metal ion binding, detoxification | 579 | 20 | 0.0091 ± 0.0034 | 29 |
| PolyUBQ | protein degradation | | | in progress | |
| TBE | DNA damage tolerance | | | in progress | |
| 4CL-1 | monolignol synthesis in xylem | 628 | 8 | 0.0032 ± 0.0014 | 79 |
| 4CL-2 | monolignol synthesis in xylem | 629 | 10 | 0.0038 ± 0.0016 | 63 |
| AT | cytoskeleton organization | | | in progress | |
| F3H1 | flavonoid pathway | 365 | 14 | 0.0099 ± 0.0040 | 26 |
| F3H2 | flavonoid pathway | 647 | 14 | 0.0056 ± 0.0023 | 46 |
| **Mean** | **31** | **598** | **17** | **0.0075 ± 0.0029** | 41 |

*Association genetics of wood property traits in loblolly pine*

Association genetics of wood property traits in loblolly pine has been completed under funding from the NSF Pine Genomics Project. Eighteen wood property candidate genes were identified: (1) nine genes coding for enzymes in the phenylpropanoid pathway leading to the synthesis of lignin monomers (*pal, c4h-1, c4h-2, 4cl, c3h-2, ccoaomt, ccr, comt-2, cad*), (2) three genes coding for enzymes that are precursors to the phenylpropanoid pathway (*sam-1, sam-2, glyhmt*), (3) two genes coding for transcription factors (*ptlim-1, ptlim-2*), (4) a gene coding for a cellulose synthase (*cesA3*) and (5) three genes coding for arabinogalactan proteins (*agp-like, agp-X, agp-Y*).

Table 2. Statistical significance of associations between specific candidate gene SNPs and wood property traits in loblolly pine.

| Trait | Gene | SNP | F | Threshold | |
|---|---|---|---|---|---|
| | | | | 0.05 | 0.01 |
| Cellulose | 4cl | 05 | 3.06 | 2.82 | 3.79 |
| | c3h-1 | 17 | 4.03 | 3.17 | 4.31 |

SNPs were discovered by directly sequencing PCR amplicons from each candidate gene using a panel of megagametophyte DNA samples, one from each of 32 trees. A total of 286 SNPs were identified among the 18 wood property candidate genes. A subset of the 50 most informative and potentially functional SNPs was typed in 425 clones of a clonal association population. Wood property phenotypes were also evaluated for all 425 clones (~850 trees in total). Estimates of nucleotide diversity were moderate to high ($\theta_T = 0.00467$; range 0.00175-0.01633) and LD decayed rapidly ($r^2 = 0.20$ within 2,500 bp).

Statistical tests (ANOVA) have identified an array of candidate genes that have significant associations with wood property phenotypes (Table 2). Associations such as those between cesA3 alleles and cellulose content—and between tubulin alleles and microfibril angle make biological sense and begin to show that complex traits can now be dissected to some of their individual gene components. A number of alleles for lignin biosynthetic pathway genes appear to associate with wood density. In short, we have clearly demonstrated the ability to conduct all aspects of association testing in loblolly pine and have developed the infrastructure for high-throughput candidate gene characterization and testing. In this proposal, we seek to verify these associations in a second species—namely Douglas-fir. Verification of these associations in a second conifer species would provide convincing evidence that the individual genes controlling wood property traits in conifers have been identified.

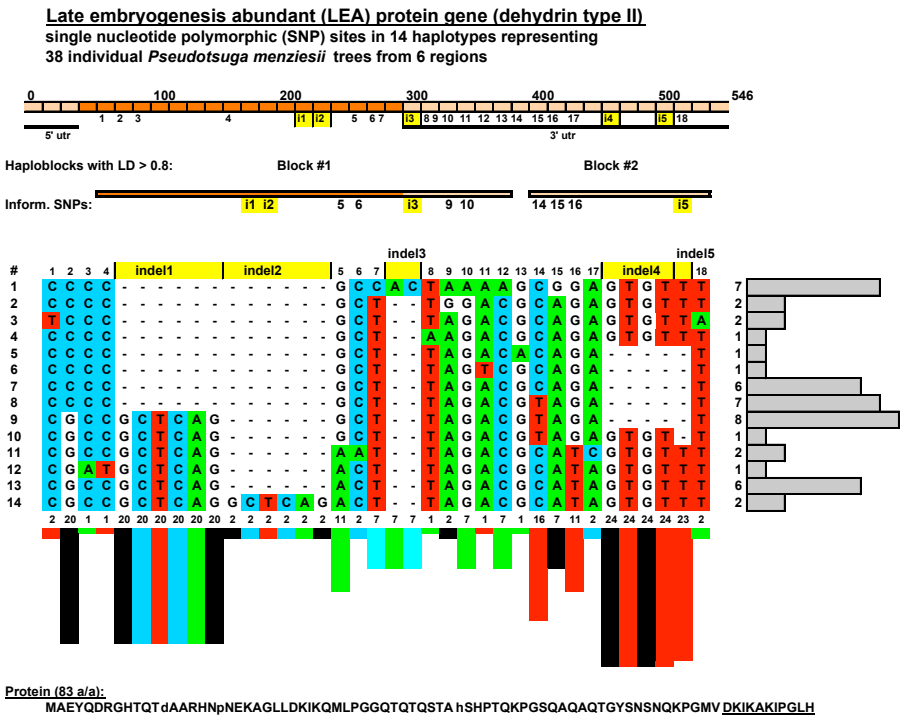
**Research Objectives:** This study has two primary objectives that will be completed through a series of tasks:

(1) Discover SNPs in candidate genes controlling wood properties in Douglas-fir.
(2) Verify SNP x phenotype associations discovered in experimental populations in applied tree improvement populations for wood property traits (wood specific gravity and microfibril angle).

*SNP discovery by direct DNA sequencing* (Neale) - We propose to find and characterize SNPs using a 2-step approach: 1) primer testing and sequencing using 3 megagametophytes from each of 2 parents in a segregating mapping population and 2) sequencing of the 18 candidates using a discovery panel of 24 megagametophytes (1 megagametophyte from each of 4 trees in each of 6 Douglas-fir regions). Amplification products will be sequenced in both directions using ABI PRISM® BigDye™ Primer Cycle Sequencing Kit v.3.1. and ABI 3730 Genetic Analyzer at the UC Davis Plant Genomics Facility.

*Sequence analysis and SNP identification* (Neale) - Raw sequences (ABI-sequencer chromatograms) will be analyzed using the base-caller computer program PHRED (Ewing *et al*. 1998; Ewing and Green 1998), the assembler PHRAP and the Unix-based graphical editor and automated finishing program for PHRAP sequence assemblies CONSED (Gordon *et al*. 1998, 2001). The SEQUENCHER software (Gene Codes Corporation, Ann Arbor, MI, USA) can be also used for analysis of alignments. SNPs will be visually detected and confirmed (e.g., see haplotypes, SNPs, and indels for the LEA gene in Figure 1).

**Figure 1. SNPs, haplotypes and haploblocks in the cold inducible late embryogenesis abundant (*LEA*) protein gene found in the SNP discovery panel representing 38 Douglas-fir trees from 6 regions.**




Nucleotide diversity and LD will be estimated using DNASP (Rozas and Rozas 1999). Haploblocks and minimum numbers of SNPs needed to genotype all haplotypes will be inferred using *HAPLOBLOCKFINDER* (http://cgi.uc.edu/~kzhang/, Zhang and Jin 2003).

*Experimental population* (DeBell) – A Douglas-fir clonal seed orchard belonging to the Washington State Department of Natural Resources has been identified as an excellent population for association genetics (see letter from Jeff DeBell). A subset of 63 clones (1-6 ramets per clone, 14-23 years old from grafting) will be harvested and increment cores will be taken for wood property analysis. In addition, open-pollinated progeny of these 63 parental clones are also found in genetic field tests where increment cores can also be taken, setting up the possibility for comparative analysis of results from clonal tests and progeny tests.

*Phenotype assessment* (Howe and St. Clair) – Wood specific gravity will be determined using the standard volumetric method at Oregon State University by students and research technicians under the supervision of Howe and St. Clair. Microfibril angle analysis will be done under a subcontract to Weyerhaeuser Company (Greg Leaf).

*SNP genotyping in association population* – (Neale) SNP genotyping will be done by single base extension and fluorescence polarization (FP-TDI; Chen *et al*. 1999; Kwok 2002) using the Acycloprime-FP SNP detection kit (Perkin Elmer). FP will be detected using a Perkin Elmer Wallac Victor2 microplate reader. We anticipate generating about 2,500 SNP genotypes.

*Association tests* – (Neale, Howe, St. Clair) Statistical tests for association between genotypes (SNPs) and phenotypes (quantitative traits) are similar to those used in QTL detection. Members of the association population are classified based on their genotype and standard statistical analyses (ANOVA/regression) can be used to test for differences in mean phenotype among genotypic classes. Genotypic classes can be based on individual- or multi-SNP genotypes (Long and Langley 1999). We will evaluate all types of tests, using methods that account for the problems of multiple comparison testing.

**Expected Products:** The expected products of the planned research are the discovery of genetic loci and alleles that are associated with phenotypic differences in adaptive traits (growth rhythm and cold-hardiness–Phase I) and wood property traits (wood specific gravity–Phase II).
Year 1. Allele discovery in Douglas-fir for 18 wood property candidate genes.

Years 2-3. Verification in Douglas-fir of candidate gene by wood property phenotype associations initially discovered in loblolly pine.

**Outcomes and Benefits:** The proposed research will deliver a set of DNA diagnostic tools that could be used by breeders and gene resource managers of Douglas-fir.

**Collaborators:** Financial support ($12,500/yr) will be provided by the Pacific Northwest Tree Improvement Research Cooperative (PNWTIRC; Glenn Howe, Director). The PNWTIRC has identified wood quality and genomics as being two of their high-priority research topics.

**Funding:** PSW - $50,000/yr for a graduate student stipend and supplies.
PNW - $10,000/yr for partial salary of research assistant and temporary assistants to conduct field sampling of increment cores and wood specific gravity analysis.
PNWTIRC - $12,500/yr for partial salary of a research assistant to coordinate field sampling of increment cores and wood specific gravity analysis; salary support for research technicians; subcontract to Weyerhaeuser Company for microfibril analysis.