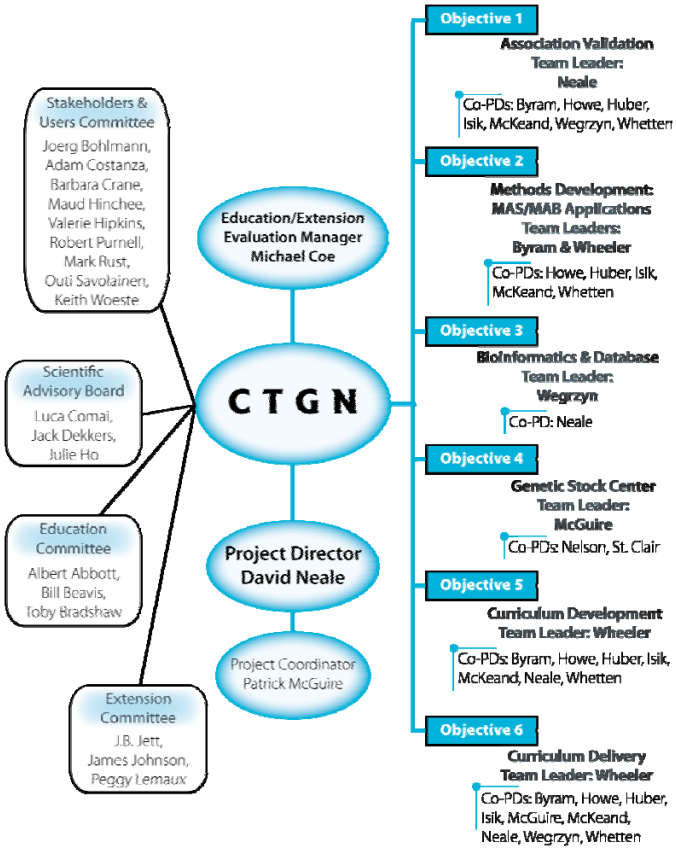
**Progress Report**

**Conifer Translational Genomics Network (CTGN) Coordinated Agricultural Project (CAP)**

**Award # 2009-85606-05680**

**Sept. 1, 2009-August 31, 2010**

This is the first year of the CTGN CAP as funded under USDA AFRI, and the third year of the four-year CTGN CAP originally funded under USDA NRI. **The overall goal of the CTGN CAP is to bring marker-based breeding to application for the major tree breeding cooperatives in the United States (**http://dendrome.ucdavis.edu/ctgn). Extensive efforts within the conifer genomics community have culminated in our ability to genetically dissect complex traits of economic and ecologic significance in forest trees, to understand the relationship between naturally occurring genetic and phenotypic variation in those traits, and to develop tools that accelerate the rate and improve the efficiency of tree breeding. We are using **association genetics** and single nucleotide polymorphisms (SNPs) in **candidate genes** to link genes to phenotypes in loblolly pine (*Pinus taeda* L.) and Douglas-fir (*Pseudotsuga menziesii* [Mirb] Franco). Using this approach, we identified dozens of significant associations between candidate genes and important phenotypes. To be used in breeding programs, associations discovered in experimental populations must be validated (see **Obj. 1.0**, below) in advanced generation breeding populations of loblolly pine, slash pine (*Pinus elliottii* Engelm.), and Douglas-fir. In addition, we are developing improved methods for detecting marker-trait associations and for finding complementary ways to combine phenotypic selection and MAS in conifers (see **Obj. 2.0** below). We continue to develop our extensive databases and web-based tools to serve the CTGN (Dendrome/TreeGenes: http://dendrome.ucdavis.edu)and the larger community (http://www.pinegenome.org; see **Obj. 3.0** below) and we are building an international genetic stock center (see **Obj. 4.0** below). An education program (see **Obj. 5** below) is targeted at undergraduate and graduate level curriculum development and instruction. In addition, this curriculum is being used to train practicing tree breeders in genomic-based breeding through an extension program (see **Obj. 6** below). By the end of this grant, genomic-based breeding strategies and protocols could be in place that influence over 1.3 billion pine and Douglas-fir seedlings planted annually in the US.




The current organizational chart is at right. One change with respect to Objectives 5 and 6 took place during this reporting year: Co-PD **David Harry**, who was team leader for Objective 5 left the Project. Building on the extensive curriculum he had developed,

co-PD **Nicholas Wheeler** took on the Obj. 5 team leader duties along with his role as team leader for Obj. 6. In addition, the chart reflects some transitions with respect to the advisory committees.

## Objective 1.0 *Validate SNP by quantitative trait associations discovered under prior USDA and NSF funding in operational tree improvement populations*

The goal of the CTGN is to validate marker-trait associations in four applied breeding programs that have previously been discovered in experimental populations under prior funding from NSF, USDA-IFAFS, and USDA-NRI.

### Objective 1.1 *Genotype ~2500 trees from each of four Cooperative breeding programs (~10,000 trees in total) for 7600 candidate gene SNPs (UC Davis)*

**Progress**

The Neale wet lab facility at UC Davis runs a high-throughput DNA isolation protocol with assistance from custom LIMS that directly interacts with the TreeGenes database. This system utilizes barcodes to track samples from field collection all the way through genotyping and final data collection. The CTGN genotyping targets are 6000 trees from loblolly populations via the NCSU Cooperative Tree Improvement Program (NCTIP), the University of Florida Cooperative Forest Genetics Research Program (CFGRP), and the Texas A&M University Western Gulf Forest Tree Improvement Program (WGFTIP); 2000 trees from slash pine via the University of Florida CFGRP; and 2,000 trees from Douglas-fir populations via the Pacific Northwest Tree Improvement Cooperative and the Northwest Tree Improvement Cooperative.

A total of 5,058 loblolly samples have been processed, DNA genotyped by Illumina, and the results stored in association with individual trees in the TreeGenes database (Table 1.1). A public interface available to the Cooperatives through the LIMS allows controlled access to the data. This information is also posted at the Dendrome Plone project website.

**Table 1.1.** Status of loblolly pine foliage samples.

| Institution | Samples collected | Samples plated for genotyping | Samples successfully genotyped |
|---|---|---|---|
| WGFTIP | 2,029 | 2,018 | 1,552 |
| NCSU-CTIP | 2,503 | 2,216 | 1,621 |
| CFGRP | 526 | 512 | 491 |
| Total | 5,058 | 4,746 | 3,664 |

Critically important loblolly pine individuals that failed to provide SNPs on the first attempt were resampled in the field and foliage again provided to UC Davis for DNA extraction. From WGFTIP, there were approximately 230 individuals combined from a clonal line trial and two progeny tests, from NCSU CTIP, there were an additional 177 individuals, and from the CFGRP, there were three individuals. Illumina genotyping runs on these have not yet been completed (See Table 1.2 for details on the WGFTIP samples).

The Douglas-fir and slash pine genotyping is well underway as can be seen from the following sections for Objectives 1.4 and 1.5. OPAs have been developed for slash pine.

**Plans**

Completion of genotyping of the slash pine and Douglas fir samples with all genotyping information available through the CTGN CAP Plone website.

**Table 1.2.** Status of loblolly pine foliage samples from WGFTIP.

| Population | Description | Total samples[1] | Samples genotyped | Samples resubmitted 2010 | Phenotype data for release |
|---|---|---|---|---|---|
| E TX 1st Gen | East Texas Breeding Population | 560 | 448 | | County of origin, BV for height, diameter, volume, specific gravity, rust, stem straightness, and forks |
| E TX 2nd Gen | East Texas Breeding Population – offspring form E TX 1st Gen | 527 | 388 | 1 | Parents, BV for height, diameter, volume, specific gravity, rust, stem straightness, and forks |
| MS | Out group of selections from MS parents | 192 | 141 | 51 | Differential breakage patterns when exposed to hurricane force winds. Expect growth to also be available |
| Clonal line trial | Entries in two locations of replicated line trials related to the E TX populations | 197 | 176 | 21 | BLUP values for 4 year height, diameter, volume |
| CP Test | Control-pollinated progeny test of E TX 1st Gen parents | 574 | 382 | 157 | Individual BLUP values for height, diameter, volume, specific gravity and stiffness (time of flight). |
| Total | | 1,858 | 1,535 | 230 | |

[1] Not available for all selections.

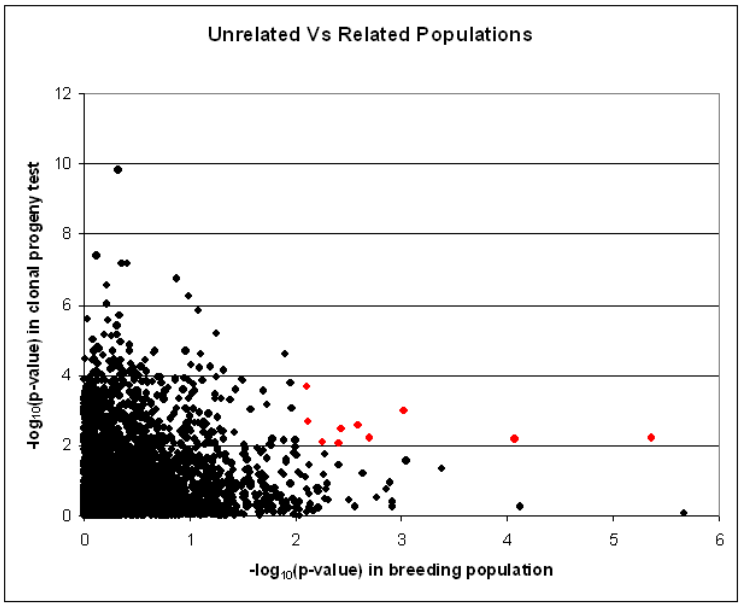
### Objective 1.2 *Validate SNP/Trait associations and estimate the effect of allelic substitutions in eastern loblolly pine breeding and testing populations (NCSU)*

**Progress**

The CTGN project has yielded genotypes for about 4,800 useful SNP loci in a set of about 1,500 trees from NC State, and we expect to receive some additional genotype data to bring the total to almost 1,700 trees. Multiple independent subsets of the genotyped trees are being used in independent association tests, to allow comparison of the results obtained from one analysis with those from an independent analysis.

To date these SNP genotypes have been used in three different types of populations: (1) Unrelated selections sampled from the breeding population, using breeding values estimated from progeny test data; (2) a set of full-sib and half-sib families in a structured mating design, using genetic values estimated from clonally-replicated field plantings; and (3) a large single full-sib family, using genetic values estimated from clonally-replicated field plantings. Each of these independent sets of trees contains 180 to 220 individuals, which is a relatively small number for traditional association genetics studies, but the availability of three separate populations will allow comparison of results among the separate analyses to further test hypotheses of marker-trait associations, and provide unbiased estimates of the proportion of genetic variation accounted for by each genetic marker. Mixed linear models are used to generate unbiased predictions of the genetic values, to minimize the effects of environmental variation on phenotypes.

Use of a single large full-sib family or progeny from a multi-family mating design to test for associations among phenotypes and genetic markers combines linkage analysis with association genetics, which provides greater sensitivity to detect effects that are segregating in the tested families and greater specificity to identify the relevant SNPs in the unstructured sample of the breeding population. The phenotypes for which breeding values are available in the cooperative database are height, volume, straightness, forking, and rust disease incidence. Comparison of p-values generated by TASSEL for association of SNPs with the same trait in two different populations identifies SNPs (red dots in the plot at right) with $p < 0.01$ in both experiments. The expected number of chance associations in both populations is less than one, so these ten SNPs are likely to be true associations.




***Validation:*** DNA sampling and SNP genotyping for the last 192 individuals (superior clones from a number of elite populations) were completed and allowed Jaime Zapata, a PhD student supported by the CTGN at NC State, to finish statistical analysis of marker-phenotype validation. Jaime has been analyzing SNP marker data using TASSEL and ASReml software. Early results are very encouraging. An index to predict sawtimber potential for loblolly pine parents was recently developed. The index value was generated using BLUP breeding values for over 2000 progeny-tested parents; this method removes the environmental effects from field trials and is based on many progeny per parent. Approximately 220 parents had 4800 SNP genotypes available for association testing, which was performed in TASSEL using the parental index value for predicted sawtimber value as the phenotype. Results show 69 SNPs were significantly associated with predicted sawtimber value at the $p<0.01$ level prior to multiple testing correction. R-square values from the association tests were as high as 0.12 for the most significant SNPs. In addition, six field trials from elite clonally replicated test populations have been planted. This population consists of progeny from parents that have been genotyped through the CTGN. This population will provide a recruitment population from which selections can be made using a combination of phenotypic data and marker-trait associations with an operational population.

***Molecular marker fingerprinting to establish parentage and clone identity.*** We used DNA from 118 foliage samples to obtain genotypes for 53 SNP markers on the MassArray platform, to investigate an uncertain identity of a clone in our breeding population. Analysis of genotype data using Cervus software unambiguously identified the clone, and confirmed the parentage of all progeny descended from that clone.

### Objective 1.3 *Validate SNP/Trait associations and estimate the effect of allelic substitutions in western loblolly pine breeding and testing populations, and test the predictive power of markers in a forward selection population (TAMU)*

**Progress**

Associations between parental breeding values and SNPs in the first and second generation populations are being currently analyzed. As the East Texas breeding population is subdivided into breeding groups based purely on administrative considerations, subsequent efforts were therefore shifted to characterizing the distribution of genetic variation within the breeding population and between individual trees and families. While East Texas represents a limited part of the loblolly range, it has substantial environmental variation in drought-inducing rainfall patterns and soil characteristics. Important factors we are attempting to detect with the SNP data from the first and second generation populations include: 1) geographic structure within the breeding population that might impact subsequent breeding and deployment plans, 2) evidence of previous natural selection, 3) partitioning of genetic variation among administratively determined breeding groups, 4) changes in SNP frequencies between generations and among progeny selected from the same set of original parents for growth or drought, 5) relationship of heterozygosity to performance, and 6) identify SNPs under selection using genetic differentiation outlier approach and population neutrality tests. SNP data is also being used for quality control in the applied program by validating pedigrees to determine error rates in record keeping within the breeding population.

**Plans**

With completion of the genotyping data for the clonal line trial and the progeny test, we will begin a more sophisticated attempt to relate SNPs to traits and to evaluate kinship matrices combined with mixed-model analysis as a tool for improving estimates of breeding values.

### Objective 1.4 *Verify SNP/Trait associations in slash pine breeding and testing populations, and demonstrate response to selection of markers in a forward selection population (UF)*

**Progress**

The CFGRP works primarily with slash pine (*Pinus elliottii*), secondarily with loblolly pine (*P. taeda*), and recently, has developed a hybridization program between the two species. We are using SNP markers developed in the CTGN project for an array of applications in our tree improvement program.

For instance, in our species hybridization and introgression program, markers will be used for foreground and background selection to significantly speed up the process. Additionally, markers are finding application here for identifying genes/alleles from parental species that are controlling QTLs. Using genotypes for 261 individuals in a BC1 (pseudo-backcross) population, preliminary analyses were performed on 980 SNP loci whose alleles were segregating approximately 1:1. First the 980 loci were mapped and a scaffold of 180 loci selected and used in QTL searches. Among 1st year traits, QTLs were found for total height and basal diameter. More refined analyses await genotypes for a larger number of individuals. In addition to these traits we have measured tip moth incidence, vegetative phenology (bud burst, bud set), and crown and foliar characteristics, all traits that are of economic or adaptational significance. A paper was submitted for publication from the analysis of 1st growing season traits, part of the MS thesis of Patricio Muñoz, a student supported by this project.

***BC1 population.*** The BC1, while measured repeatedly during the first growing season in the field, was measured only once this year at end of the second growth year. SNP marker data became available for the population through evaluation on the loblolly platform. The purpose of using the loblolly chip was to follow the effects of loblolly alleles in a slash background. Analysis of the genotypic classes for the 5,380 markers provided unexpected results. Sixty-eight percent of the markers were monomorphic, 19% had the expected two genotypic classes and 2% had three genotypic classes. In total, 802 markers were mapped onto 12 linkage groups. An evaluation of the QTL effects of loblolly alleles on second-year height and branches per unit height showed both positive and negative effects. These effects accounted for 20% of the phenotypic variation in height.

***Slash pine populations.*** The wood coring and foliage sampling in the full-sib block plots has been completed with approximately 1,000 trees sampled. The foliage samples have been freeze-dried and the cores are currently being phenotyped.

Two full-sib families with approximately 1,500 seed each were sown in the greenhouse in December 2009. The seedlings have since moved from greenhouse to shade and then to full sun and should be planted by the end of May.

Selection of contigs (1,000) for resequencing in slash pine nears completion. The SNPs discovered will provide the basis for analyses of the two slash populations.

### Objective 1.5 *Validate SNP/trait associations in Douglas-fir (OSU)*

**Progress**

***SNP markers –*** Douglas-fir genomic resources are not as well developed as are the resources for loblolly pine. Consequently, we are using transcriptome sequencing (454 and Illumina Genome Analyzer) to dramatically increase the number of SNP markers for Douglas-fir. We isolated RNAs from multiple tissues from a single genotype and then had them sequenced by the DOE Joint Genome Institute (JGI) using the Roche 454 GS-FLX Titanium approach. In the spring of 2010, we obtained 1.26 million sequence reads with an average length of 350 nucleotides (nts). These sequences have been assembled into a draft reference transcriptome that we will use as a foundation for SNP discovery. We will continue to (1) evaluate and refine the reference transcriptome and (2) identify high-quality SNPs using the additional 454 and Illumina sequences described below.

Other RNAs were isolated from multiple tissues collected from 82 seedlots of coastal Douglas-fir (var. *menziesii*) harvested on 11 dates throughout the year. We used this "genetic diversity" sample for additional 454 sequencing at the University of Illinois Carver Biotechnology Center, ultimately obtaining 1.7 million sequence reads with an average length of 395 nts. We will also use this "genetic diversity" sample for sequencing by the Illumina Genome Analyzer. The Illumina sequencing is scheduled for the first week of June 2010, so we should receive these data by the middle of June, 2010. We collaborated with Rich Cronn (USFS) to obtain Illumina transcriptome sequence from additional Douglas-fir genotypes that will also be used for SNP discovery. We collaborated with Samuel Cushman and Andrew Shirk of the USFS, and Barry Jaquish of the British Columbia Ministry of Forests to collect tissues and isolate RNA from a provenance test of Interior Douglas-fir (var. *glauca*) so that we can begin developing SNP markers for Douglas-fir breeders in the Interior West.

***New phenotypes –*** Douglas-fir breeders are very interested in using SNP markers to estimate breeding values for wood stiffness and adaptability (e.g., cold hardiness and bud phenology), but

these traits have not been included in the regular progeny test measurement protocol. Furthermore, past SNP development efforts have focused on candidate genes for these traits. Therefore, we are measuring wood stiffness and cold adaptation traits in breeding populations. This will allow us to test SNP-trait associations and evaluate the potential for MAS for these important traits.

This past year, we continued to collect phenotypic data from 830 clonally replicated parents in nine grafted seed orchards and some of their offspring in progeny test plantations. Measurements during the past year included wood stiffness, reproductive bud phenology, vegetative bud flush, and cold hardiness (e.g., 160 families in the fall of 2009). Measurements completed since last summer (not previously reported) include acoustic velocity (wood stiffness) data recorded on parents and progeny belonging to the Bureau of Land Management's (BLM) Lorane breeding unit. We obtained stiffness estimates for parents by measuring acoustic velocity on 528 standing trees representing 167 parental clones in two seed orchards using the TreeSonic tool. As these trees were being harvested, we measured acoustic velocity on logs derived from 407 of these trees using the HM200 tool. In addition to the parents, we also measured acoustic velocity on their progeny as trees were thinned from two first-generation progeny sites. We measured 437 progeny from 114 families at the Carpenter Bypass site and 427 progeny from 119 families at Hawley Creek. Standing-tree acoustic measurements will be made on the remaining trees at both progeny test sites in 2010. Finally, we collected foliage from the parents in the seed orchards so that we can include these genetic materials in our analyses of SNP-trait associations for wood stiffness.

Now that most phenotypic measurements are completed (i.e., with the completion of the bud flush measurements in the spring of 2010), we can compete our quantitative genetic analyses of these data, our evaluation of alternative sampling strategies using previously developed simulation software, and then choose the 2,500 Douglas-fir trees to genotype. The trees will be chosen based on the number of phenotypic traits available (new and existing), quality of the phenotypic data (e.g., heritabilities), and amount of information from relatives (i.e., extent of the measured pedigree).

***Data analysis*** – We tested our previously developed data simulation program using Tassel and ASReml, and have began developing programs for combined linkage/linkage disequilibrium analysis.

**Plans**

***SNP markers*** – We will continue to evaluate and refine the reference transcriptome that we developed using the 454 sequences from the single genotype of coastal Douglas-fir. We will then identify high quality SNPs by comparing the "genetic diversity" sequence data (454 and Illumina) to the reference transcriptome. We will also isolate RNA from the Interior Douglas-fir samples and conduct additional 454 sequencing to identify high-quality SNPs in this variety. We will add these new SNPs to our existing set of Douglas-fir SNPs to construct the next-generation SNP genotyping chip (1536 SNPs). This chip will be designed to identify SNPs in coastal Douglas-fir, with a preference for SNPs that target nucleotides that are also polymorphic in Interior Douglas-fir.

***New phenotypes –*** We will complete our phenotypic measurements this summer by recording standing-tree wood stiffness (acoustic) measurements on the remaining trees at two BLM progeny test sites. If another BLM seed orchard is harvested as expected, we will attempt to collect

more wood stiffness data from the harvested logs to augment previous standing-tree wood stiffness data from these orchards.

***Data analysis.*** We will use our simulation software to test alternative sampling strategies for association tests. In particular, we will evaluate alternative numbers of parents and progeny to genotype. These sampling strategies will be evaluated using analyses that focus on combined analyses of within-family linkage and population-level linkage disequilibrium. Once optimal distributions of parents and progeny are determined, we will choose the trees to genotype based on the criteria listed above.

***SNP genotyping.*** We will collect foliage from selected trees and isolate DNA in the first half of the year, and then begin SNP genotyping as soon as our 1536 SNP-chip is completed.

### Objective 2.0 *Develop and economically evaluate new methods incorporating marker-assisted selection into conifer tree breeding programs*

**Progress**

***Modeling marker applications:*** The NCSU team has been working on several analytical tools to incorporate SNP markers in estimations of breeding values using routine BLUP analyses. Protocols have been developed for TASSEL, ASReml, and SAS software for processing large numbers of SNP markers using linear mixed models. The mixed models validate the association of SNP markers with the phenotypes (growth) while taking into account the population structure (relatedness in the data) and correcting for common and specific environmental variation. F-test probability values of markers are corrected for false discovery rates (q-values) and the significant markers are treated as random effects in mixed models to evaluate their efficiency in improving average prediction error of estimated breeding values. We have also made progress on a simulation project. The objective is to understand the effect of population parameters (heritability, effective population size, number of markers in LD with the trait loci) on the accuracy of breeding value estimates, in simulated forest tree breeding populations. Co-PD Ross Whetten wrote simulator software in R to create mating populations with different levels of heritability and differing numbers of QTLs in LD with SNP markers. The simulator software is being used to test alternative strategies for application of marker-trait associations in forward selection.

Preparation of a manuscript (submitted 2/2010, see the first item under Publications reported for Objective 6, below) on the implementation of marker-informed breeding in complex recurrent selection programs was a result of the collaboration and discussion involved in the work on this objective. The manuscript focuses on the logistics of implementation and is intended to be the first in a series. Progress has been made on two fronts: conceptualization and graphical description of the logistical nature of MAS implementation in the tree breeding cycle, and the creation of an analytical tool for conducting stochastic simulation of tree breeding alternatives using MAS, the objective of which is to provide economic insights on probability of success (financial return) for an array of tree breeding scenarios.

**Plans**

We plan to validate significant markers detected in one population (sawtimber elite) in a validation population (cloned population) for predictions. The question is can we use marker effect detected in a training population as phenotype (genome wide selection approach) or factors (for small number of markers) in a validation population for prediction of breeding values? An accuracy value based on the correlation of true and predicted BVs of both training and validation

populations will tell us about the reliability of markers for selection. We will explore BLUP, Bayes A (homogeneous variance) and Bayes B (different variances for markers to let uninformative marker effects shrink to zero) in linear mixed models for predictions.

### Objective 3.0 *Develop databases (TreeGenes) and web-based tools (Dendrome) to facilitate all aspects of the CTGN*

**Progress**

A Plone content management system (Dendrome Plone) has been implemented for CTGN project content. Included at this location are updates of data, presentations, and education and extension activities. The Plone allows real-time editing of content to enable users to modify and add content as needed without programming language knowledge.

The custom LIMS developed for the DNA isolation lab at UC Davis to guide activities of the wet lab including sample preparation and DNA isolation continues to be improved. The LIMS allows full tracking based on barcode identifiers from receipt of sample through genotyping and data delivery. The wet lab interface assists individuals with plate preparation and includes features to monitor concentrations and create inter- and intra-plate controls. The LIMS was designed for full integration with the laboratory equipment. There is full integration of the LIMS data into the overall TreeGenes Database. The database includes individual tree and genotyping data for all 5,058 loblolly pine samples. Extensive user-side interfaces exist to allow tree improvement cooperatives to directly participate in the data curation and procurement efforts through regular updates of the metadata and phenotypic data. CTGN project participants access the system via controlled logins to upload data tagged with specific tree identifiers.

Genotype data delivered from Illumina is accumulated in the TreeGenes relational database, integrated with the available phenotypic data for these same individuals, and made available to project directors and the cooperatives for downstream analyses.

Extensive search interfaces also exist to allow bulk download as well as itemized searches. Data types available for search by project participants now include EST, EST contig, amplicon, sequence assembly, phenotype, sample identifiers, and SNPs. A search can begin at any of these levels and can yield detailed information on tracefiles, FASTA sequences, SNP scores, annotations, and genotypes.

**Plans**

We will continue to collect phenotype data and metadata on the genotyped individuals. Collaboration with tree improvement cooperatives towards the standardization of phenotype descriptions and interpretations will allow further development of interfaces to support more robust and comprehensive searches. In addition, significant associations generated from the downstream analysis of genotypes and phenotypes will be imported into the TreeGenes database. Interfaces will be modified and expanded to allow for retrieval of relevant associations through genotype and phenotype queries.

### Objective 4.0 *Develop an international genetic stock center for conifers*

**Progress**

The United States Forest Service has committed to developing an international Genetic Stock Center as part of its role in the Conifer Translational Genomics Network. The primary goal of the stock center is to archive, curate, and make available to researchers valuable genetic material.

The Forest Service is an important partner in this effort due to their commitment to long-term research.

There are three components to the stock center at this time: (1) the **Loblolly Pine Genetic Stock Center,** a clonal archive at the Southern Institute of Forest Genetics, Southern Research Station near Saucier, Mississippi, (2) the **Douglas-Fir Genetic Stock Center,** a clonal archive at the Pacific Northwest Research Station near Corvallis, Oregon, and (3) the **UC Davis Molecular Genetic Stocks Center, comprised of** DNA stored in -80°C freezers.

Funding from the USDA for this project is making it possible to establish the field archives, set up the databases, and build a system for public distribution for these valuable genetic resources. The archives of loblolly pine and Douglas-fir are genotypes from genetic mapping and association mapping populations from earlier studies. Approximately 2,000 clones are being propagated as grafts and rooted cuttings to be maintained by the Forest Service at those two field sites. The molecular biology resources include cDNA libraries and clones, BAC libraries and clones, mapping and association population DNAs, and PCR primer sets from earlier research.

At the Loblolly Pine Center, additional scions for individuals that had been missing in the collection for the Base and QTL Three-generation populations were obtained and grafted this year for the collection, completing the grafting effort for these populations. For the NCSU Association population, cuttings were planted in the archive on the Harrison Experimental Forest near Saucier (from 1 to 3 ramets were planted per clone for a total of 413 clones). Some mortality was observed due to the relatively few poorly rooted plants. Lists of needed ramets have been provided to NCSU for replacement rooted cuttings. The field inventory will be updated this year and the replacement list will be updated. If rooted cutting material is available this year, we will plan to interplant. Otherwise we will request scions and graft and plant in subsequent years.

**Plans**

For the loblolly pine Base and QTL Three-generation populations, the final grafted ramets (from the Jan.–Feb. 2010 grafting) will be planted in spring 2011. The inventory will be updated at that time. For the NSCU Association population, we will work with NCSU to achieve 3 ramets per clone as well as establish a backup clonal archive on the Erambert Seed Orchard site, about 40 miles north of the Harrison. Some of the ramets at this backup archive site will likely be developed by grafting as the rooting success of loblolly pine falls off considerably after age 4 or 5 years. The two archives (Harrison and Erambert) will be maintained and tissue samples will be provided/coordinated upon request.

Much material from which DNA samples for the Molecular Genetic Stocks Center are needed currently exists at numerous other locations, as well as at UC Davis. Thus efforts at location and transportation of relevant tissues to the UC Davis will be ongoing. Tissue from these stocks will be transferred into 15 or 50mL vials and labeled with standardized barcode IDs assigned by the TreeGenes database and LIMS. DNA isolation of tissue samples will occur as needed for specific projects. Currently, samples from the loblolly pine Base and QTL Three-generation populations have been acquired, sampled for DNA, and stored. The NCSU loblolly pine population is currently queued for DNA isolation.

Dendrome and TreeGenes will be further developed to serve as the web resource and database components of the genetic stock center, to be maintained by CTGN staff at UC Davis. Although the initial emphasis is on Douglas-fir and loblolly pine genotypes, our vision includes other conifers and perhaps also broad-leaved species. Our long-term goal is the complete integration and curation of the biological and information resources for forest tree genomics.

## Objective 5.0 *Develop an education plan for undergraduate and graduate curriculum development in genomics-based breeding in forest trees*

**Progress**

***Special topics courses***

"*Quantitative Forest Genetics*", a graduate course at NCSU, was taught by **Fikret Isik** in Fall 2009. Genomics-based selection using mixed models was incorporated and graduate students working with the NCSU Tree Improved Cooperative attended the course and learned the theory and application of mixed models for genomics-based selection and creating kinship matrices from markers.

**David Neale** (UC Davis) taught a molecular breeding course (Genetics Graduate Group GGG 298) at UC Davis spring quarter 2010.

***Planned 2010 shortcourse cancelled***

"*Genomics in Tree Breeding and Forest Ecosystems*" had been planned and advertised for June 21–25, 2010. It was designed for graduate students and practicing professional tree breeders and genetic resource managers. Though considerable effort had been invested in revising the 2009 course and planning the new one, we were concerned that the low projected enrollment (5 US, 8 international applicants) would not justify the time and expense of holding it at this time. We anticipate offering the shortcourse again in 2011. For alternatives, we referred the applicants to the July workshop (see Plans) and an August workshop in Italy (*Landscape Genomics of Forest Trees and Applications to Climate Change*, August 23–27, 2010, Trento Italy). In addition, the content of the modules developed for the 2009 course are available at the Project website (http://dendrome.ucdavis.edu/ctgn/educationextension/ shortcourse2009.php). Applicants were also encouraged to take advantage of CTGN internships (http://dendrome.ucdavis.edu/ctgn/educationextension/internships.php).

Among the organizers of the Italian workshop is CTGN Project Director Neale. He and other CTGN personnel are among the presenters (http://conferences.fem-environment.eu/ conferenceDisplay.py?confId=4). It will feature computer exercises of many of the genomic applications relevant to breeders or gene resource managers (roughly 50% of the workshop activities).

**Plans**

"*Genomics for Applied Tree Breeding and Ecosystem Management – With Application to Gene Conservation and Climate Change*"—July 13–14, 2010

This one- to two-day workshop will focus primarily on the application of genomics in the characterization and management of natural populations, though there will be an applied breeding element. The workshop will be hosted by the USDA Forest Service's Dorena Genetic Resource Center in Cottage Grove, Oregon. It will deliver information for geneticists and natural resource managers from regional Federal Agencies and others who may have wished to take the five-day shortcourse previously scheduled for June 2010 (described above). The workshop announcement may be viewed at http://dendrome.ucdavis.edu/ctgn/files/2010Workshop_OR.pdf.

"*Genomics-Based Breeding in Forest Trees*"—June 22–24, 2011

This international symposium will focus on the application of markers in tree breeding. The Symposium will feature plenary speakers from leading agricultural, horticultural, and animal improvement programs utilizing molecular marker technologies and a series of invited speakers working with markers in forestry. Facilitated discussion sessions engaging forest industry representatives will continue our theme of moving this technology from academia to the private sec-

tor. An international advisory committee has been enlisted to assist with Symposium planning. The committee membership and updates on Symposium progress are posted at our Project website: http://dendrome.ucdavis.edu/ctgn/educationextension/symposium.php.

### Objective 6.0 *Develop and deploy an extension curriculum for continuing education in genomics-based breeding for practicing tree breeders and forest tree gene resource managers*

**Progress**

***Presentations***

Co-PDs continued to deliver presentations, including invited talks, class lectures, posters, talks to professional societies, and conference talks. We anticipate adding presentations for public access at the Project website. Presentations include:

- **D. Harry / N. Wheeler**. June 25, 2009. "New developments in genetic markers and their applications" Invited talk: Northwest Seed Orchard Managers Meeting, Aurora, Oregon.
- **N. Wheeler / T. Byram**. Aug. 10-13, 2009. "Marker informed breeding in tree improvement: How does that work?" Volunteer Talk: Western Forest Genetics Association Conference, Asilomar, CA.
- **David Neale**. Oct. 7-9, 2009. Two invited talks: "Development and application of genomic-based tools to manage forests in response to climate change" and "Association studies for complex traits in conifers" Workshop on Opportunities, Challenges, and Limitiations of Genomics-Based Technologies in Forest Tree Breeding and Forest Genetics. Forest Resesarch Institute, Freiburg, Germany.
- **Glenn Howe**. Oct. 21, 2009. Conifer Translation Genomics Network: Results for Douglas-fir, loblolly, and slash pine, and their implications for applied tree improvement. Presentation at the Northwest Tree Improvement Cooperative Annual Meeting, Aurora, OR.
- **Ross Whetten**. Dec. 3–5, 2009. Conifer Translational Genomics Network (CTGN). Using markers in tree breeding. A presentation given to the NC State University Cooperative Tree Improvement Program Contact Meeting. Tuscaloosa AL.
- Andrew J. Eckert, **Jill L. Wegrzyn**, Jennifer M. Lee, John D. Liechty, Kristian Stevens, Kathleen D. Jermstad, Betty Woolf, Wei Tao, **C. Dana Nelson**, Santiago C. Gonzalez-Martinez, Charles H. Langley, and **David B. Neale**. January 9-13, 2010. Patterns of nucleotide diversity and associations to environmental heterogeneity across the functional gene space of loblolly pine (*Pinus taeda*). Plant and Animal Genome XVIII, Forestry Workshop.
- **Jill L. Wegrzyn**, Ben Figueroa, Minyoung Choi, John D. Liechty, Andrew J. Eckert, and **David B. Neale**. January 9-13, 2010. Bioinformatic Solutions for data storage, analysis, and interpretation in forest genomics: Overview of the tools and resources from the Dendrome Project. Plant and Animal Genome XVIII, Forestry Workshop.
- **Fikret Isik** and **Ross Whetten**. January 9-13, 2010. Effects of marker numbers and population parameters on the accuracy of predicted breeding values in forest trees. Plant and Animal Genome XVIII, Forestry Workshop.
- Anna Stambolia-Kovach, Jill L. Wegrzyn, Genis Parra, Carson Holt, James Hartigan, Charles M. Nicolet, George E. Bruening, Michela Troggio, Carol Loopstra, Mark Yandell, Ian Korf, Charles H. Langley, and **David B. Neale**. January 9-13, 2010. Whole genome shotgun and BAC sequences in loblolly pine (*Pinus taeda* L.): The majority of the

22-Gb genome appears to be highly diverged and nested repetitive elements. Plant and Animal Genome XVIII, Forestry Workshop.

- **Ross Whetten**. 2010. Bringing marker-informed breeding to application in the NCSU Cooperative Tree Improvement Program.
  Jan. 15—Beijing Forestry University, Beijing, China
  Jan. 16—Chinese Academy of Forestry, Beijing, China
  Jan. 19—Research Institute of Tropical Forestry, Guangzhou, China
- **David Neale**. March 4, 2010. Forest tree genomics research in North America. FoResT-Trac Kick-off Meeting. Institut National de la Recherche Agronomique (INRA). Paris, France.
- **David Neale**. March 7–12, 2010. Keynote: Development and application of genomic-based tools to manage forests in response to climate change. —Sustainable Utilisation and Conservation of Forests in the Genomics Era. IUFRO Kuala Lumpur, Malaysia.
- **David Neale**. March 7–12, 2010. Workshop: Association genetics of complex traits in forest trees. Sustainable Utilisation and Conservation of Forests in the Genomics Era. IUFRO Kuala Lumpur, Malaysia.
- **David Neale**. 2010. Development and application of genomic-based tools to manage forests in response to climate change.
  April 29–30—Harvard's 6th Annual Plant Biology Initiative Symposium. Trees and the Global Environment. Cambridge MA USA.
  May 26–29—International Symposium on Biology of Rare and Endemic Plant Species. Muğla, Turkey.
- **Nicholas Wheeler**. May 12, 2010. Genomics in tree breeding: Past, present, and future. Presentation to the University of Florida Cooperative Forest Genetics Research Program.
- May 18–19, 2010. WGFTIP Contact Meeting. DeGray Lake Resort State Park, Bismarck, AR
  **Nicholas Wheeler** / **Tom Byram**. Genomics in tree breeding: Past, present, and future.
  **Tom Byram**. State of the conifer genome.
  **Vikram Chhatre**. Genetic structure in loblolly pine breeding populations: What can we learn from molecular markers?
- Elena Mosca, Erica Di Pierro, Nicola La Porta, Giovanni G. Vendramin, Piero Belletti, and **David Neale**. May 19-20, 2010. Adaptation in alpine conifers. Alpine Ecosystems in a Changing Environment: Biodiversity Sensitivity and Adaptive Potential (ACE-SAP) 3rd Annual General Meeting. Trento, Italy.
- **Nicholas Wheeler**. July 7, 2010. Genomics in tree breeding: Past, present, and future. Presentation at the Annual Meeting of the Pacific Northwest Tree Improvement Research Cooperative, Vancouver, WA.

***Website redesigned***

The CTGN website (http://dendrome.ucdavis.edu/ctgn) has undergone a notable format change and acquired considerable content. The CTGN CAP is thoroughly described in the material found at the *Home*, *Description*, and *Organization* links while details on the CTGN team, including quick email access, may be found at the *People* link.

The *Education and Extension* link provides easy access to a broad spectrum of information, including announcements for Project-supported internships and the international Sympo-

sium on Genomics in Tree Breeding that we will host in June of 2011. Access to teaching materials used in the 2009 Shortcourse and a graduate-level course offered at one of our participating institutions is also provided. Finally, we have enumerated all outreach presentations made by Project staff and provided detailed, professional evaluation reports on all our major education and extension activities. All past Project progress reports and newsletters are located at the *Reports* link and fact sheets and literature reviews are available at the *Resources* link.

***Newsletters***

A CTGN newsletter series was initiated with a first issue in October 2009. The second followed in February 2010 and the next will be out in mid June 2010. Each has been six pages long. They feature CTGN research progress notes, articles featuring the collaborating tree improvement cooperatives and their activities, reviews of current literature, and an events calendar. Distribution is by pdf file emailed to a ~300-count email database of researchers, forestry practitioners, students, and administrators and each issue is also posted for downloading at the CTGN website. The newsletters are compiled and edited by co-PDs **Nicholas Wheeler** and **Patrick McGuire**.

***Publications***

- **Nicholas Wheeler**, **Thomas Byram**, **David Harry**, **Glenn Howe**, **Dudley Huber**, **Fikret Isik**, **Steve McKeand**, **David Neale**, **Dana Nelson**, **Brad St. Clair**, and **Ross Whetten**. Marker informed breeding (MIB) in forest trees: I. Breeding and selection applications for recurrent selection programs. Submitted 2/2010. *Tree Genetics and Genomes*

  This manuscript is intended to function largely as an extension document for practicing tree breeders, combining theoretical discussion of the merits of association genetics and genomic selection with a discussion of the likely logistics of implementing marker-informed breeding in complex recurrent selection programs.
- Keith J.S. Jayawickrama, Terrence Z. Ye, and **Glenn T. Howe**. Heritabilities, intertrait genetic correlations, GxE interaction and predicted genetic gains for acoustic velocity in mid-rotation coastal Douglas-fir. Submitted. *Silvae Genetica*.

  This manuscript presents results of genetic analysis of the stiffness trait in preparation for studying the trait's association with CTGN SNPs.

***Evaluation***

Education and Extension activities conducted by CTGN are being evaluated on an ongoing basis by our independent evaluator, Dr. Michael Coe. As part of this study, participants in the 2009 CTGN week-long summer workshop in 'Genomics in Tree Breeding and Forest Ecosystems' answered a questionnaire two months before and again immediately after the course. The survey included questions about participants' academic background and focus; knowledge, skills, and interest in areas related to the use of genetic markers; ability to teach or communicate to others about these topics; and interest and ideas about further work in marker-assisted breeding and related subjects. After the workshop, participants were significantly more confident in their ability to perform key tasks related to the use of genetic markers; their scores on a 12-item efficacy scale were on average more than a standard deviation higher after the course. A similar increase was observed after the course in participants' confidence in their ability to teach or communicate about markers to various kinds of audiences.

Participants also answered a set of feedback questions about the course. Each of the ten course modules was rated for length as well as clarity and effectiveness, with average responses indicating general satisfaction but a need for some modules to be lengthened (or for students to do more preparatory work beforehand). Course participants rated the course materials and cur-

riculum highly, and most plan to apply what they learned in further academic or professional work. In addition to structured ratings, participants also provided open-ended feedback to help guide future CTGN educational activities.

**Plans**

***Project meetings***

The annual CTGN CAP meeting is scheduled for June 17–18, 2010 at Oregon State University, Corvallis OR USA. It will bring together CTGN personnel with advisors to review progress and plan research and analyses.

***Evaluation***

Data collected independently by the external evaluator will be combined with other Project documentation to form a record of the extent to which CTGN attained the outcomes and goals that were envisioned in the proposal, lessons learned that can be useful for similar projects in the future, and changes in the field that can inform further efforts.

## Additional activities in support of CTGN goals

The WGFTIP Established one of three field locations with clonally replicated material from the ADEPTII populations. The TFS planting contained approximately 300 clones established in 6 replications designed with incomplete blocks. Due to the irreplaceable nature of this material to provide data for future association studies, trickle irrigation was installed for the six-acre test.

## Training

At NCSU, research assistant **Joshua Steiger**, funded part-time by the CTGN Project, provides support for the field and lab research carried out by the NC Cooperative Tree Improvement Program and PhD student **Jaime Zapata**, also funded by the CTGN, carries out research on loblolly pine relevant to the CTIP mission and the CTGN Project.

At the University of Florida, the MS thesis of **Patricio Muñoz** involved analysis of the slash pine/loblolly pine hybrid population phenotypes for CTGN.

At TAMU, graduate students **Tomasz Koralewski** and **Vikram Chhatre** are supported by CTGN and are using portions of the SNP association data for their dissertations and are carrying out the WGFTIP population characterizations. This work is being done under the direction of Dr. Konstantin Krutovsky.

**Harold Mackin** (the 2006 Washington state and national Agriscience Teacher of the Year, as awarded by the National FFA Foundation) has participated since 2007 in the North Carolina State University Kenan Fellows Program (http://www.ncsu.edu/kenanfellows/) under the sponsorship of the NSF-funded ADEPT 2 project and this CTGN Project with mentoring by CTGN Co-PD Wheeler. Mackin developed study plans that integrate genetics and molecular genetic techniques. The study plans are publicly available at the Kenan Fellows Program website (http://www.ncsu.edu/kenanfellows/kfp-cp-sites/cp13/cp13/index.html).