Genome-Enabled Discovery of Carbon Sequestration Genes in *Populus*
University of Florida
DOE Office of Science, Office of Biological and Environmental Research
Final Report, Award No. **DE–FG02–04ER63845**
October 1, 2007

A report submitted in January 2006 to the Office of Science was appended to our project renewal proposal, served as a report of the project's progress, and was used as a criterion for subsequent renewal of the award in January 2007. This report is appended as "Appendix A" in the accompanying file.

Since January 2006, the refereed works listed below have been either published or are in press from work conducted at the University of Florida based on this award:

Bocock PN, Dervinis C, Morse AM, Davis JM. 2007. Evolution and diversity of invertase genes in *Populus trichocarpa*. *Planta* (in press).

Ramirez-Carvajal G, Morse AM, Davis JM. 2007. Transcript profiles of the cytokinin response regulator gene family in *Populus* imply diverse roles in plant development. *New Phytologist* (in press).

Davis JM, Beewar MR. 2007. Tree Cloning. PIRA International, London, UK (in press).

Morse AM, Tschaplinksi TJ, Dervinis C, Pijut PM, Schmelz EA, Day W, Davis JM. 2007. Salicylate and catechol levels are maintained in *NahG* transgenic poplar. *Phytochemistry* 68: 2043-2052.

Filichkin SA, DiFazio SP, Brunner AM, Davis JM, Yang ZK, Kalluri UC, Arias RS, Etherington E, Tuskan GA, Strauss SH. 2007. Efficiency of gene silencing in *Arabidopsis*: direct inverted repeats vs. transitive RNAi vectors. *Plant Biotechnology Journal* 5: 615-626.

Huang L-F, Bocock PN, Davis JM, Koch KE. 2007. Regulation of invertase: a 'suite' of transcriptional and post-transcriptional mechanisms. *Functional Plant Biology* 34: 499-507.

Tuskan GA, DiFazio SP, Hellsten U, Jansson S, Rombauts S, Putnam N, Sterck L, Bohlmann J, Schein J, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen G, Cooper D, Coutinho PM, Couturier J, Covert SF, Cunningham R, **Davis J**, Degroeve S, dePamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grigoriev I, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Islam-Faridi N, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé JC, Déjardin A, Pilate G, Locascio P, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson CD, Nieminen KM, Nilsson O, Peter G, Philippe R, Poliakov A, Ralph S, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Salamov A, Schrader J, Segerman B, Sterky F, Souza C, Tsai C, Unneberg P, Wall K, Wessler S, Yang G, Yin T, Douglas C, Sandberg G, Van de Peer Y, Rokhsar D. 2006. The genome of western black cottonwood, *Populus trichocarpa* (Torr. & Gray ex Brayshaw). *Science* 313: 1596-1604.

Lawrence SD, Dervinis C, Novak N, Davis JM. 2006. Wound and insect herbivory responsive genes in poplar. *Biotechnology Letters* 28: 1493-1501.

COMPARATIVE GENOMIC STUDIES

Plants utilize carbon by partitioning the reduced carbon obtained through photosynthesis into different compartments and into different chemistries within a cell and subsequently allocating such carbon to sink tissues throughout the plant. Since the phytohormones auxin and cytokinin are known to influence sink strength in tissues such as roots (Skoog & Miller 1957, Nordstrom *et al.* 2004), we hypothesized that altering the expression of genes that regulate auxin-mediated (*e.g.*, *AUX/IAA* or *ARF* transcription factors) or cytokinin-mediated (*e.g.*, *RR* transcription factors) control of root growth and development would impact carbon allocation and partitioning belowground (**Fig. 1 – Renewal Proposal**). Specifically, the *ARF*, *AUX/IAA* and *RR* transcription factor gene families mediate the effects of the growth regulators auxin and cytokinin on cell expansion, cell division and differentiation into root primordia. Invertases (*IVR)*, whose transcript abundance is enhanced by both auxin and cytokinin, are critical components of carbon movement and therefore of carbon allocation. Thus, we initiated comparative genomic studies to identify the *AUX/IAA*, *ARF*, *RR* and *IVR* gene families in the *Populus* genome that could impact carbon allocation and partitioning. Bioinformatics searches using Arabidopsis gene sequences as queries identified regions with high degrees of sequence similarities in the *Populus* genome. These *Populus* sequences formed the basis of our transgenic experiments. Transgenic modification of gene expression involving members of these gene families was hypothesized to have profound effects on carbon allocation and partitioning.

AUX/IAA and *ARF* Gene Families

We used post-genome assembly and annotation gene models in a comprehensive comparative sequence analysis to yield information on the total number of genes in the *AUX/IAA* and *ARF* gene families, thus enabling a phylogenetic analysis of the closest ortholog to a known Arabidopsis gene. Because this effort was initiated prior to the assembly and annotation of the *Populus* genome sequence, we developed a bioinformatics pipeline for identification of sequence with homology to each gene, assembly of the into sequence contigs and prediction of coding sequences using GenScan (Burge & Karlin 1998), followed by manual editing using the Artemis program (Rutherford *et al.* 2000). Thereafter, nucleotide sequence homology tool, BLASTN, was used to identify a gene-specific region from each *AUX/IAA* and *ARF* gene.

Phylogenetic analysis, using 40 *Populus ARF* proteins and 23 Arabidopsis *ARF* proteins (Hagen & Guilfoyle 2002), indicated that *Populus* and Arabidopsis *ARF* genes have evolved distinctly since they diverged from their common ancestor (**Fig. 2**). The ratio of activator *ARF* (defined by the Q-rich middle domain) in Arabidopsis and *Populus* is 5:13, respectively, whereas the ratio of repressor and other *ARF* domains is 18:27, indicating an enrichment of activator *ARF* domains in *Populus*. Moreover, a pair of *Populus* genes, *PoptrARF7.3* and *PoptrARF7.4*, have no obvious Arabidopsis orthologs. Harper *et al.* (2000) has demonstrated that *AtARF7* plays a role in auxin dependent differential growth in aerial plant form; thus, we hypothesize that *PoptrARF7.3* and *PoptrARF7.4* may be involved in differential growth of woody stems. Since the number of *Populus ARF* gene family members is much larger than the number of *AtARF* genes, we conclude that the number of heterotypic *ARF* interactions will be several fold more complex in *Populus*. Thus, transgenic modification of *ARF* expression in *Populus* will most likely produce an intricate pattern of carbon allocation and partitioning.

The *Populus* genome has 35 *AUX/IAA* genes (henceforth referred to as *PoptrIAA*), while the Arabidopsis genome contains 29 *AUX/IAA* genes (henceforth referred to as *AtIAA*) (Hagen & Guilfoyle 2002). Phylogenetic reconstruction using all available and predicted *Populus* and Arabidopsis *AUX/IAA* protein sequences shows that four groups of *PoptrIAA* (*PoptrIAA3*, *29*, *16* and *27*) have expanded to contain 3 to 4 members each. RT-PCR analysis suggests that these groups are being differentially expressed in various tissues (**Fig. 3**). The predicted proteins of *PoptrIAA26.1* and *26.2* cluster closely with that of *AtIAA18* and *AtIAA26* genes (data not shown).

Appendix A: Progress Report:

Genome-Enabled Discovery of Carbon Sequestration Genes in *Populus*

DOE Office of Science
Office of Biological and Environmental Sciences

Principal Investigators:

Gerald Tuskan	Oak Ridge National Laboratory (ORNL)
John Davis	University of Florida (UF)
Steve Strauss	Oregon State University (OSU)

Co-Investigators and/or Subcontractors:

Bill Bergusson	University of Minnesota (NRRI)
Amy Brunner	Oregon State University
Stephen DiFazio	Oak Ridge National Laboratory
Lee Gunter	Oak Ridge National Laboratory
Frank Larimer	Oak Ridge National Laboratory
Timothy Martin	University of Florida
Richard Meilan	Oregon State University
Alison Morse	University of Florida
Tim Tschaplinski	Oak Ridge National Laboratory
Stan Wullschleger	Oak Ridge National Laboratory
Zamin Yang	Oak Ridge National Laboratory

January 2006

The *AtIAA18* and *26* genes express in shoot meristem, young leaves, senescing leaves, male

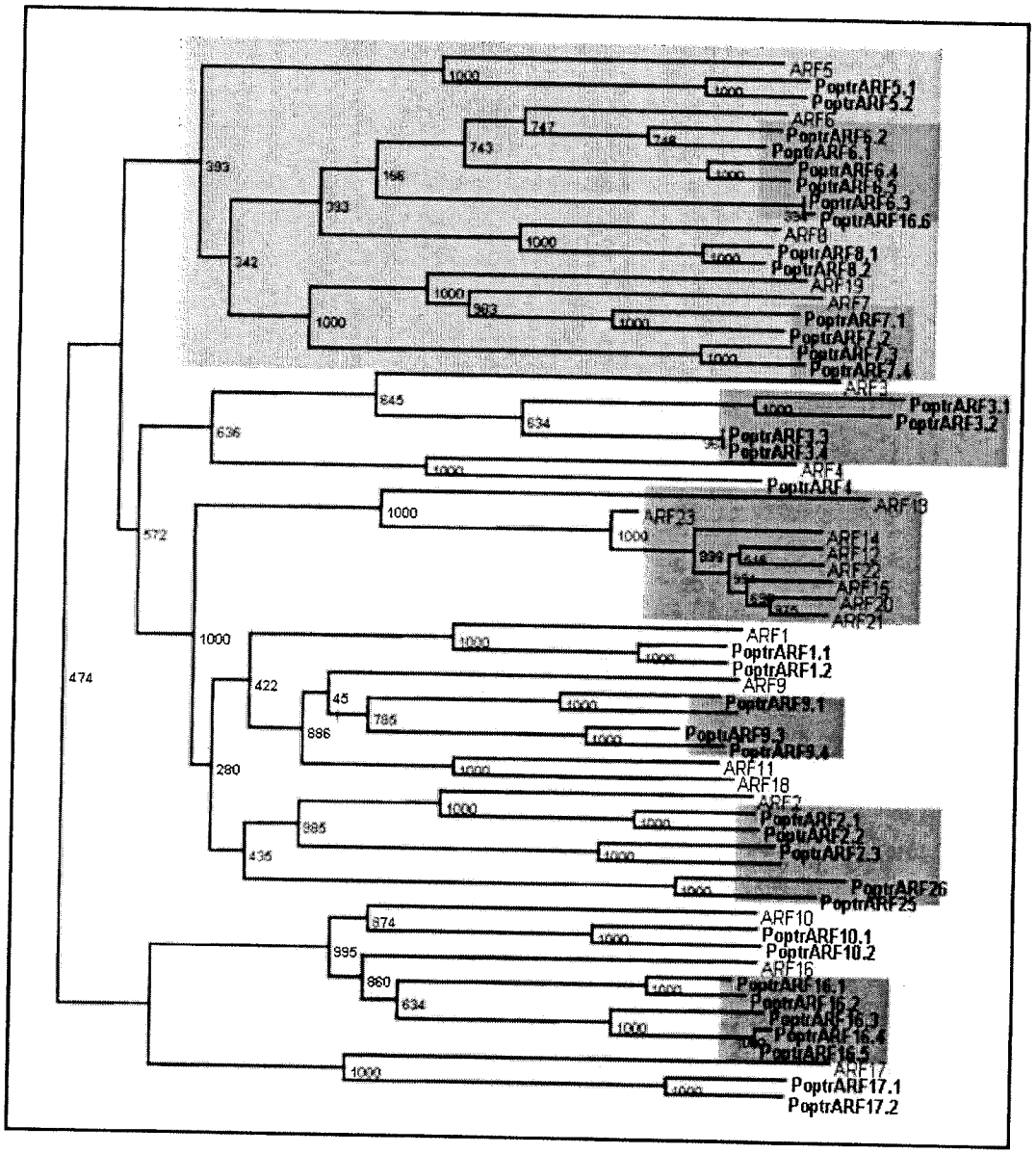

Figure 2. ClustalW phylogram (Bootstrap value 1000) of all predicted auxin response factor (*ARF*) amino acid sequences. Populus genes are designated as *PoptrARF* and Arabidopsis genes are designated as *ARF*. Divergences are highlighted as *Populus*-specific clades are represented in green boxes; Arabidopsis-specific clade is represented in the brown box. There are 40 *Populus ARF* genes and 23 Arabidopsis genes. Most of the difference between the two species can be attributed to an increase in the number of *Populus* gene containing the Activator domain (as indicated by the grey box).

catkin and flower buds. Since *AtIAA26* displays binding affinity towards *PHYA* (phytochrome A), it is likely that these *Populus* orthologs are involved in controlling the light-dependent activation of various tree development processes. Single and double mutants resulting from *AtIAA12* gain-of-function and *ARF5* loss-of-function mutations have similar phenotypic defects (Hamann *et al.* 2002). We proposed that *PoptrARF5* may act in a positive regulatory manner and *PoptrIAA12* in negative manner to control root meristem development in the presence of auxin. Our RT-PCR results show that *PoptrARF5.1* and *PoptrIAA12* genes have similar expression levels in leaf and stem tissue whereas they have contrasting expression patterns in roots (**Fig. 3B**). These results indicate that there is a profound need to determine high-resolution spatiotemporal expression patterns and to test create transgenic constructs/plants in carbon sequestration relevant species.

Cytokinin Response Regulators (*RR*) Gene Family

As a first step in identifying specific *Populus RR* with important roles in belowground processes related to carbon allocation, we identified all members of the *Populus RR* gene family. This was achieved through protein and nucleic acid sequence similarity searches carried out against the *Populus* genome sequence database using known Arabidopsis *RR* genes. These searches yielded 24 *Populus RR*; 11 type As and 13 type Bs (**Fig. 4**). Publicly available databases

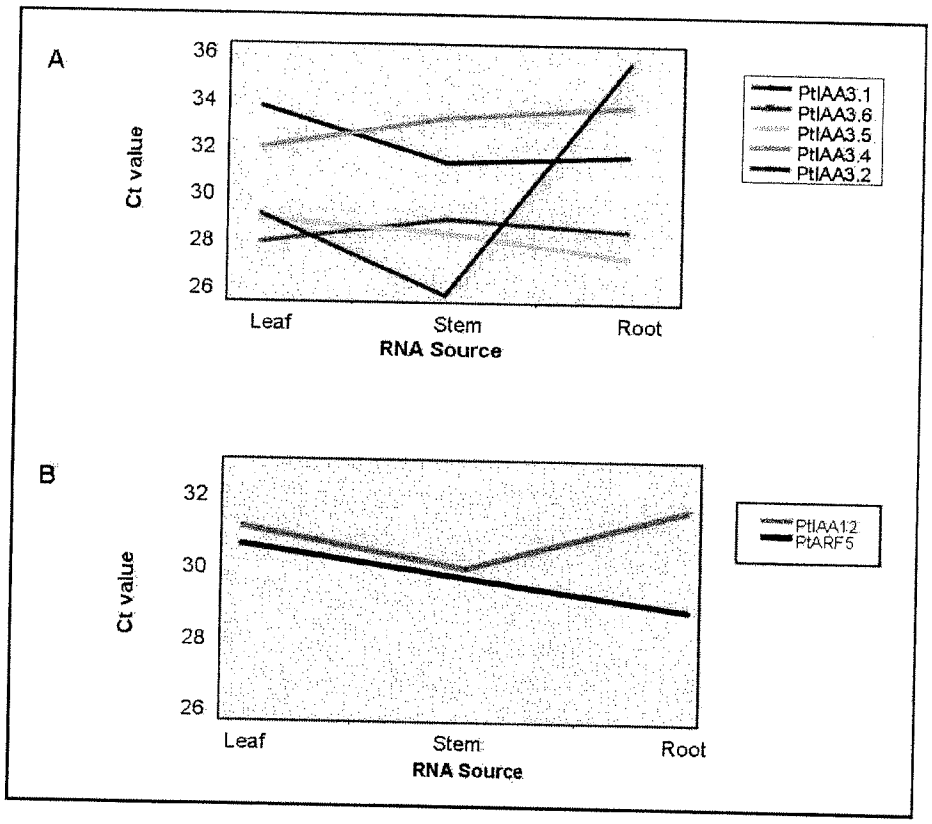

Figure 3. A) Normalized real-time PCR cycle values for genes in *PoptrIAA3* subgroup obtained using leaf, stem and root cDNAs. B) Normalized real-time PCR values obtained for *PoptrIAA12* and *PoptrARF5* genes obtained using leaf, stem and root cDNA. A lower threshold cycle (Ct) value indicates a higher expression of the gene in the corresponding cDNA sample.

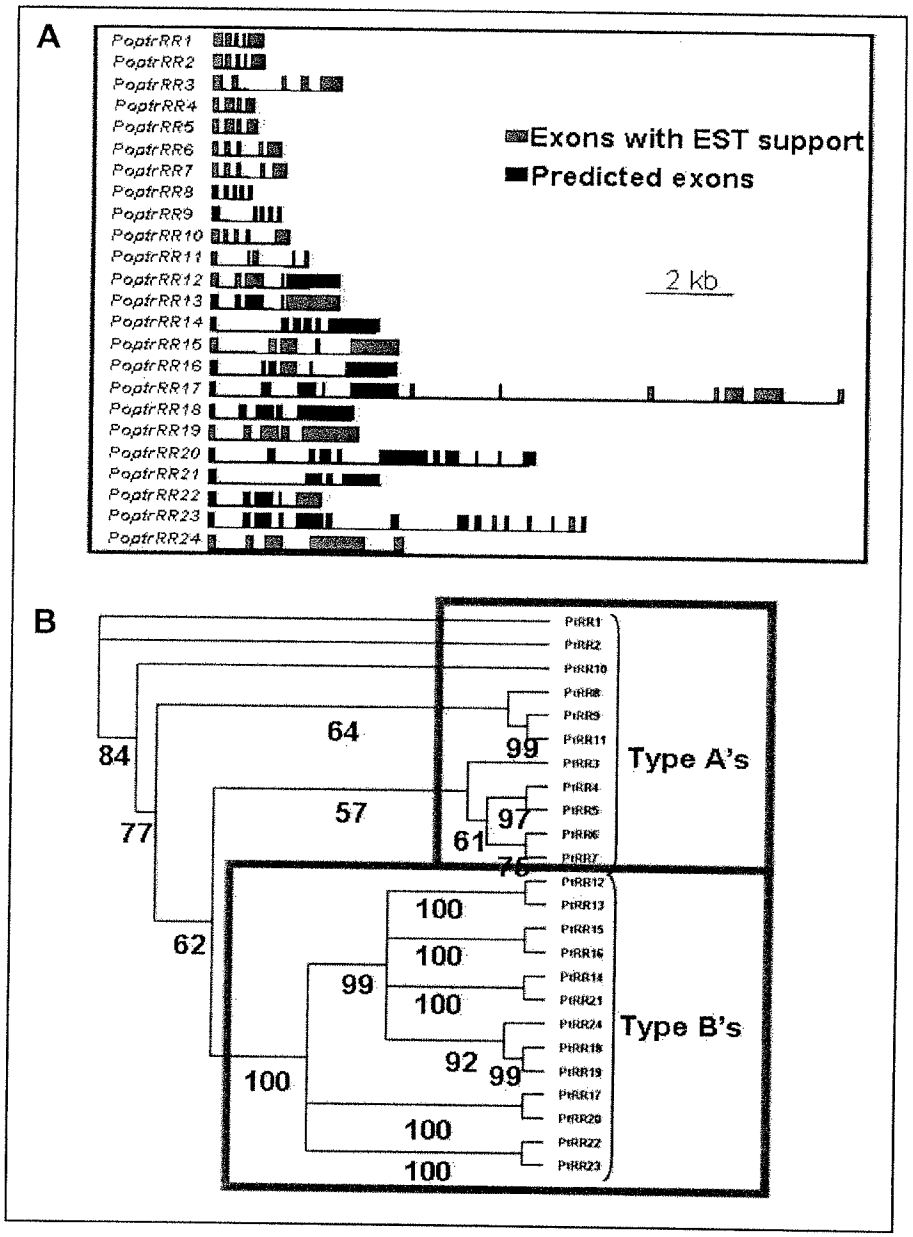

Figure 4. The *Populus* cytokinin *RR* gene family. A) Gene structure and EST coverage. Type A *RR* are enclosed within the inset box. B) Protein similarity tree of poplar response regulators using only the receiver domain for the alignments. Bootstrap values were obtained using PAUP 4.0 (10,000 reps).

were screened for EST that corresponded to the annotated *Populus RR* to: 1) confirm the annotation and 2) provide evidence for tissue expression. The annotation data obtained on the *Populus RR* gene family was used to create constructs for the creation of knock-down (via RNAi) and gain-of-function (via full-length over-expressors and constitutively active △DDK)

phenotypes. It was envisioned that these transgenics would allow us to obtain phenotypes important for elucidating the roles of individual *Populus RR* family members in processes related to carbon allocation and partitioning. For each of the 11 type A *RR*, we generated both a full-length and an RNAi construct. No constitutively active △DDK constructs were created for the type As since they lack the DNA binding output domain with which the DDK (receiver) domain interacts. All 22 constructs were created, half of which were sent to OSU for transgenesis. For each of the type B *RR* for which we were able to obtain transcript data (10 out of 13), we generated an RNAi, a full-length and a △DDK construct (total of 30 constructs). To date, all RNAi constructs for type B *RR* have been completed, along with half of the full-length constructs and half of the △DDK constructs were completed. Of the completed constructs for the type B *RR*, 11 were sent to OSU for transformation. As of June of 2005, we received multiple transformation lines for each of 2 constructs, both for *PoptrRR13* (one set of transgenic lines with the full-length construct and one set with the △DDK construct).

Invertase Gene Family

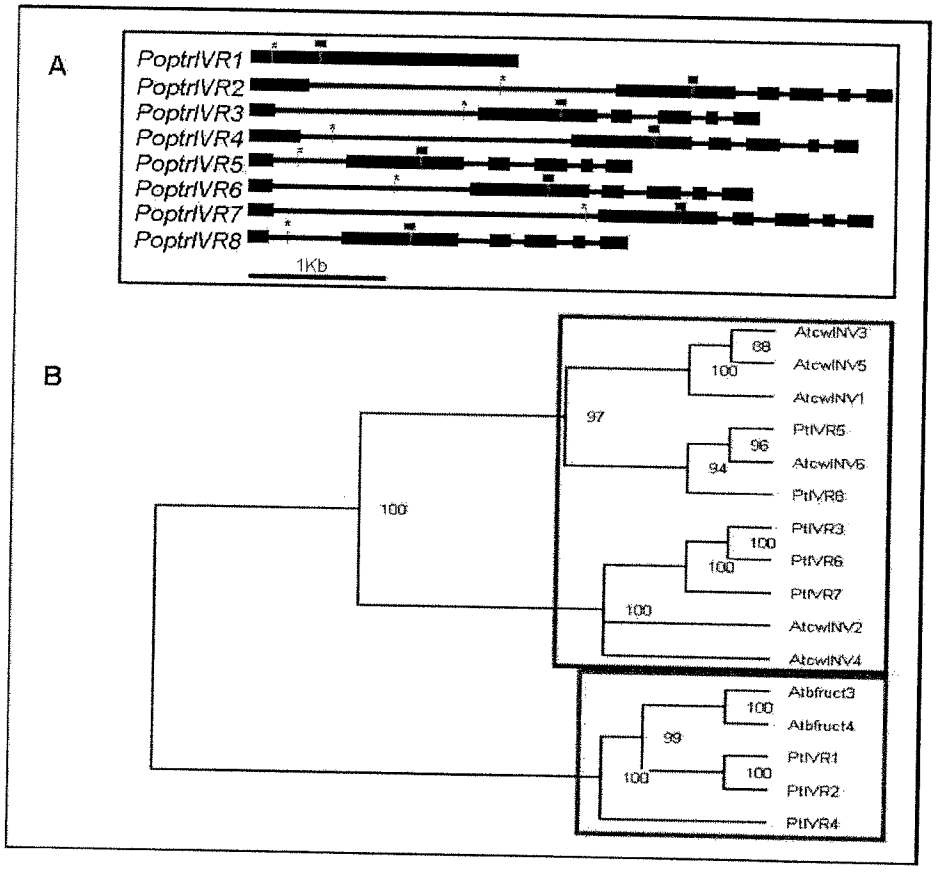

Figure 5. The *Populus* acid invertase family. A) Asterisks depict the β-fructosidase domain, squares depict the catalytic domain. All genes contain the conserved catalytic and β-fructosidase domains. B) ClustalW phylogram (Bootstrap value 1000) of all predicted *PoptrIVR* and *AtIVR* amino acid sequences. Red box = vacuolar invertase, blue box = apoplastic invertase.

Invertases (EC 3.2.1.26) play key roles in carbon allocation by regulating sink strength in tissues such as roots and regulating sucrose export from source tissues. These enzymes catalyze the irreversible cleavage of sucrose. Thus, invertases are ideal targets for manipulation to understand carbon allocation and partitioning (Sturm & Tang 1999). Since there is little, if any, information on the *Populus* invertase gene family, we used the *Populus* genome sequence to identify 8 acid invertase family members (*PoptrIVR*) (**Fig. 5**). Three of these family members encode invertases targeted to the vacuole while the other five are targeted to the apoplast. This is slightly different than what is found in Arabidopsis, where there are two vacuolar invertases and six cell wall invertases (Tymowska-Lalanne & Kreis 1998). *Populus* invertases share the intron/ exon structure generally conserved in plants (7 exons separated by 6 introns), with one exception, *PoptrIVR1*. *PoptrIVR1* encodes the first expressed intronless invertase found in any organism to date and appears to be evolutionarily derived from *PoptrIVR2*. The annotation data obtained on the *PoptrIVR* gene family has been then used to create constructs for RNAi. The goal was to obtain a knock-down phenotype to help elucidate the roles of individual *PoptrIVR* family members. Constructs directed at a unique ~200 base pair region at both the 5' and 3' ends of the all 8 genes have been made (16 constructs total). Half of the constructs were sent to OSU in Year 1 of the project and the other half have been constructed and are awaiting results from the first set. We have not received any *PoptrIVR* RNAi transgenics for phenotyping to date.

CANDIDATE GENE TRANSFORMATION

Gene silencing is one of the primary means of testing gene function. Previous work has shown that inverted repeat transgenes resulting in 'hairpin' double-stranded mRNA gave the most efficient silencing of all targeted genes (Waterhouse *et al.* 1998). This type of 'RNA interference' (RNAi) is one of the post-transcriptional gene silencing mechanisms in plants (Hamilton & Baulcombe 1999). It has subsequently been determined that this effect could be 'transitive' such that an inverted repeat in the 3' portion of a polycistronic mRNA molecule containing an unrelated gene 5' to the inverted repeat can silence the endogenous gene (Alder *et al.* 2003).

Researchers have taken advantage of this highly efficient and specific gene silencing mechanism by creating 'binary' vectors into which gene fragments could be inserted, followed by transformation of *Agrobacterium* spp. with the vector. Early attempts at RNAi utilized vectors containing inverted repeats of the target gene fragment. However, since this approach is quite laborious a group at CSIRO led by P. Waterhouse created vectors that utilize the 'Gateway™' transformation system to simultaneously introduce gene fragments in inverse orientation using topoisomerase-mediated cloning (Helliwell *et al.* 2002). This method also proved somewhat inefficient because the inverted repeat increases the complexity of the construct. Brummell *et al.* (2003) developed a vector that takes advantage of the transitive silencing effect. This vector consists of an inverted repeat of the Octopine Synthase (Koncz *et al.* 1983) terminator 3' sequence and a multicloning site where the endogene target could be inserted. However, this vector does not contain Gateway™ technology. Therefore, we created unique RNAi, Gateway-like transformation vectors for this project.

Creation of RNAi Vectors for *Populus*

We created a series of vectors containing the binary vector backbone pART27 (Gleave 1992), containing 1) the chloramphenicol resistance gene *CMR* to allow selection of *Agrobacterium* transformants and 2) a kanamycin resistance gene *NPTII* to allow for selection of transformed plant cells. We used the Gateway™ conversion kit to incorporate the proper recombination sites and genes for negative (*ccdB*, which prevents growth in most *E. coli* strains by interfering with DNA gyrase) and positive (Spectinomycin Resistance, *SpecR*) selection. In addition, the created cassette contains an insertion site flanked by specific sequences (attR sites)

for incorporation of a gene fragment flanked by alternative sites (attL sites) that are recognized by the LR Clonase™ enzyme.

Two such vectors were created. Vector pCAPD is a 'direct' RNAi silencing vector based on the Waterhouse *et al.* (1998) design, containing two Gateway-like cassettes in inverse orientation which flank the *piv2* intron from potato (Vancanneyt *et al.* 1990). The 'transitive' vector, pCAPT, has the insertion site for the target gene fragment upstream of an inverted repeat of the OCS terminator. Furthermore, 5' of the insertion site there is a small fragment of green fluorescent protein (GFP) which is predicted to silence GFP in lines with adequate expression of the transgene. In both pCAPD and pCAPT, the expression of hairpin RNA cassettes is under the control of a constitutive CaMV 35S promoter (Franck *et al.* 1980) comprising 857 bp of the promoter sequence. Transcription is terminated by the *Agrobacterium* OCS terminator.

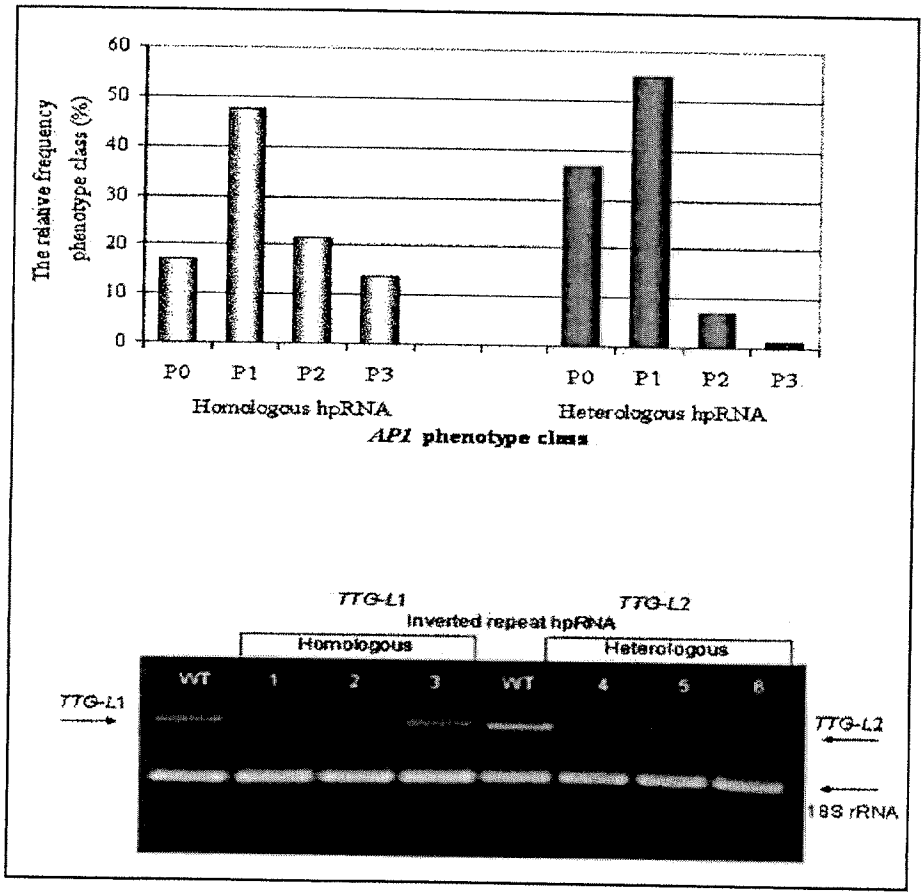

Figure 6. A) Efficacy of pCAPD (homologous hpRNA) and pCAPT (heterologous hpRNA) RNAi constructs in achieving severe mutant phenotypes (where the severity of the transgenic phenotype observed is classified into P0-P3, P0 representing least altered and P3 representing the most severe). B) Semi-quantitative RT-PCR showing reduction in expression of target gene (*TTG*) in transgenic RNAi plants as indicated by the brightness of the DNA band.

Evaluation of RNAi Vectors

To evaluate gene silencing efficacy of the transitive RNAi, we selected a set of Arabidopsis genes with well characterized phenotypes, including *APETALA1* (*AP1*), (At1g69120) (Mandel *et al.* 1992); *ETTIN* (*ETT*), (At2g33860) (Sessions *et al.* 1997) and *TRANSPARENT TESTA GLABRA 1* (*TTG1*) (At1g63650) (Walker *et al.* 1999). *pCAPD::AP1* (homologous hpRNA) and *pCAPT::AP1* (heterologous hpRNA) transgenics induced wide arrays of stable phenotypes (**Fig. 6**). The proportion of the *AP1* RNAi knock-down events was lower among lines transformed with the *pCAPD::AP1* as compared to the transitive *pCAPT::AP1* (**Fig. 6A**). Furthermore, the degree *ETT* suppression by RNAi showed that both transitive and direct RNAi constructs caused phenotypic changes in flower organs in approximately 50% of all events in T1 transformants. In *Populus*, both direct and transitive constructs showed an ability to suppress

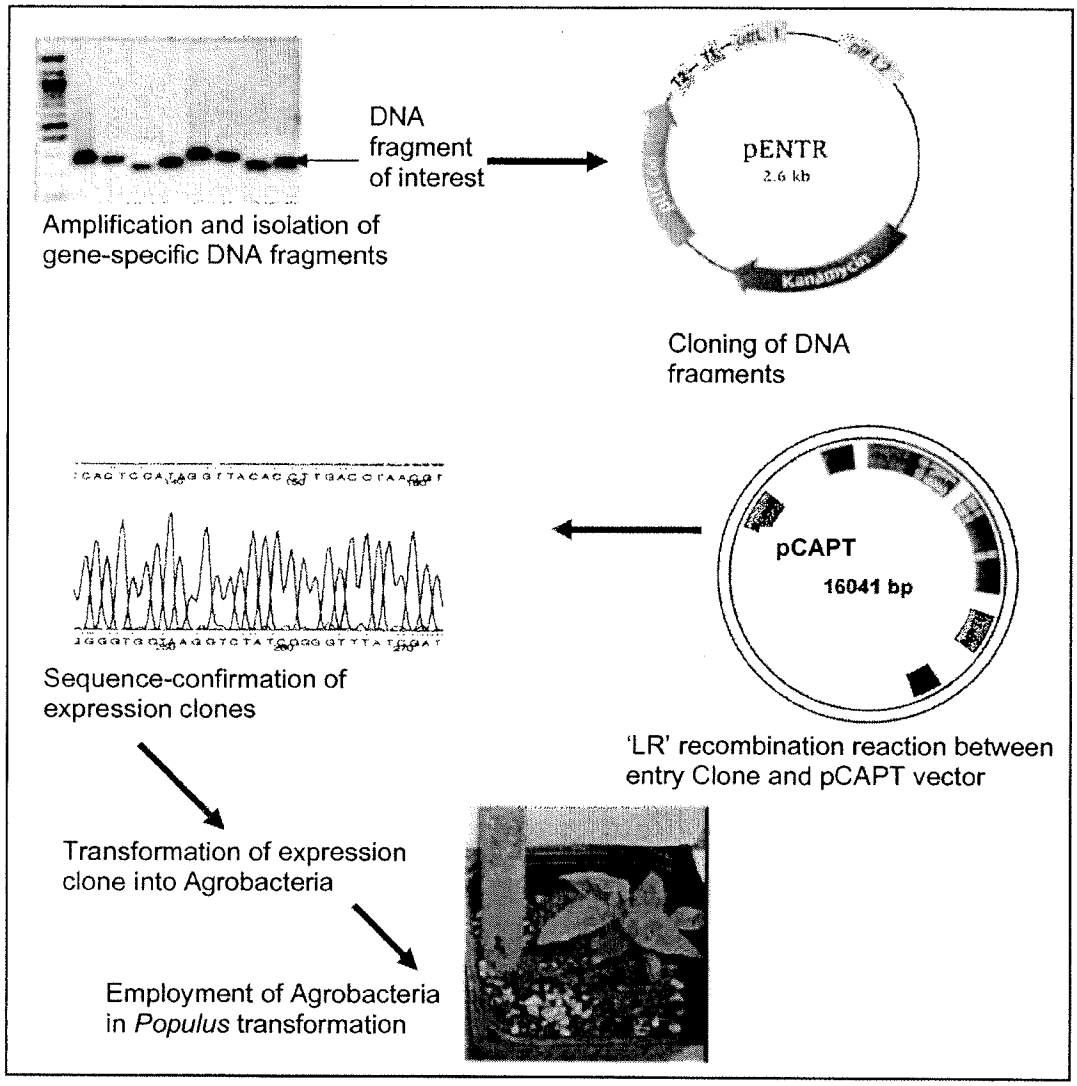

Figure 7. Flowchart of experimental steps involved in generation of an RNAi construct.

TTG-1-like mRNA (as judged by a semi-quantitative RT-PCR, **Fig. 6B**). These results indicate that both *pCAPT* and *pCAPD* vectors were functional in achieving down-regulation of the target gene in the host plants. Because of the relative ease of creating RNAi vectors using the *pCAPT* vector relative to the *pCAPD* vector we choose the transitive approach for all transformation experiments.

Generation of RNAi Constructs

Thus, *Populus*-centric RNAi constructs were created through a series of cloning steps (**Fig.** 7). Briefly, gene fragments were amplified using *Populus* genomic DNA and gene-specific primers cloned into D-TOPO pENTR vector (Invitrogen) to generate entry clones. PCR-verified entry clones were recombined with a destination vector, *pCAPT*, to obtain expression vector for transformation. Expression vectors were sequenced to confirm identity as well as orientation of the cloned gene fragment. Detailed backbone description, vector maps and sequences are available via the ORNL WEB site at: http://www.esd.ornl.gov/PGG/cap/belowground_Populus_project_vectors.htm.

Development of the *Populus* transformation toolkit

As a part of a transformation improvement effort OSU developed transformation protocols for *P. trichocarpa* clone 'Nis-1' and early flowering *P. alba* clone, '6K10'. A total of 21 'Nis-1' transgenic plants have been generated and a complete description of transformation system for 'Nis-1' was recently published (Ma *et al.* 2004). A total of 286 transgenic plants were generated using '6K10'. Four major classes of constructs were used for 6K10 transformations: a) RNAi constructs (a total of 194 transgenic plants); b) overexpression±constructs (53 transgenics), c) ablation constructs (13 transgenics) and d) promoter-reporter gene constructs (26 transgenics).

In an effort to identify root predominant promoters a set of three promoters were tested for their tissue-specific expression patterns in *Populus*, including: 1) a root-predominant promoter from a *Casuarina glauca* (*pCgMT1*), 2) a promoter from Arabidopsis root specific *pin2* gene (*pAtPin2*) and 3) a putative root-specific promoter from a tagged enhancer trap *Populus* line (*pET304*). *pCgMT1* promoter showed high levels of activity in roots of multiple *Populus* transgenic lines. Both *pAtPin2* and *pET304* promoters directed expression of the reporter beta-glucuronidase gene (*GUS*) in a tissue-specific manner in roots of transgenic *Populus*. *GUS* expression directed by the newly discovered *pET304* promoter was highly specific for the lateral roots and their primordia in *Populus*. This raised a possibility that *pET304* gene may play an important role in formation of adventitious roots. The lateral root-specific expression pattern of *pET304* was confirmed by the analysis of T1 and T2 generations of transgenic Arabidopsis. Both *pAtPin2* and *pET304* promoters provide excellent candidates for utilization in carbon sequestration research. Further characterization of *pET304* response to auxin regulation and its role in lateral root formation and root biomass accumulation is on-going.

To evaluate the efficiency of two chemically regulated promoter systems -- alcohol- and estrogen-inducible -- OSU tested these systems in multiple *Populus* transgenic lines. The alcohol-inducible "gene switch" system, carrying the *GUS* reporter gene driven by the alcohol-inducible *alcA* promoter, was functional in *Populus*. Different induction methods and concentrations of inducers were tested. Ethanol under tested conditions was the most effective inducer when compared with acetaldehyde or 2-butanone. Up to 8% ethanol had no apparent toxic effects on the plants both in tissue culture and greenhouse conditions. The *alcA* promoter system was sensitive to induction by ethanol at concentrations as low as 2%. Long-term continuous induction showed significant increases of *GUS* levels up to 30 days. Vapor treatment was an effective induction procedure resulting in high levels of *GUS* activity in leaves. The induction by ethanol was more efficient in plants grown in tissue culture than those grown under greenhouse conditions. No significant *GUS* activity was detected in roots of greenhouse-induced transgenic

trees. The alcohol-inducible "gene switch" system was functional in transgenic *Populus* and potentially allows temporal control of gene expression.

TRANSGENIC GROWTH AND PHENOTYPING

Acclimation of Transgenics to Soil and Greenhouse Conditions

Following transformation, it is essential to devise growth conditions to acclimate transgenic plants to greenhouse conditions and remove artifactual effects of tissue culture. The first batch of 3 *AUX/IAA* transgenic lines was shipped to ORNL from OSU in June 2005 and consisted of control plants and transgenic plants transformed with *PoptrIAA3.2*, *PoptrIAA3.4* or *PoptrIAA3.6* RNAi constructs. Upon receipt, the tissue-culture propagated transgenic RNAi and control plantlets were transplanted in soil (autoclaved soil [85% peat] and Perlite mixed in 1:2 ratio) and acclimated in a growth chamber for 30 days under controlled conditions (high humidity, long-day conditions (18 hr photoperiod) at 22°C temperature). Ramets of plants from a single transformation event were potted in soil and placed in flats sealed with parafilm to retain moisture. Plants were treated with Gnatrol once a week and fertilized (20-10-20) every two weeks. Differences in overall plant health between transgenic and control plants were noted. To acclimate transgenic plants to greenhouse conditions (greater light intensity and decreased

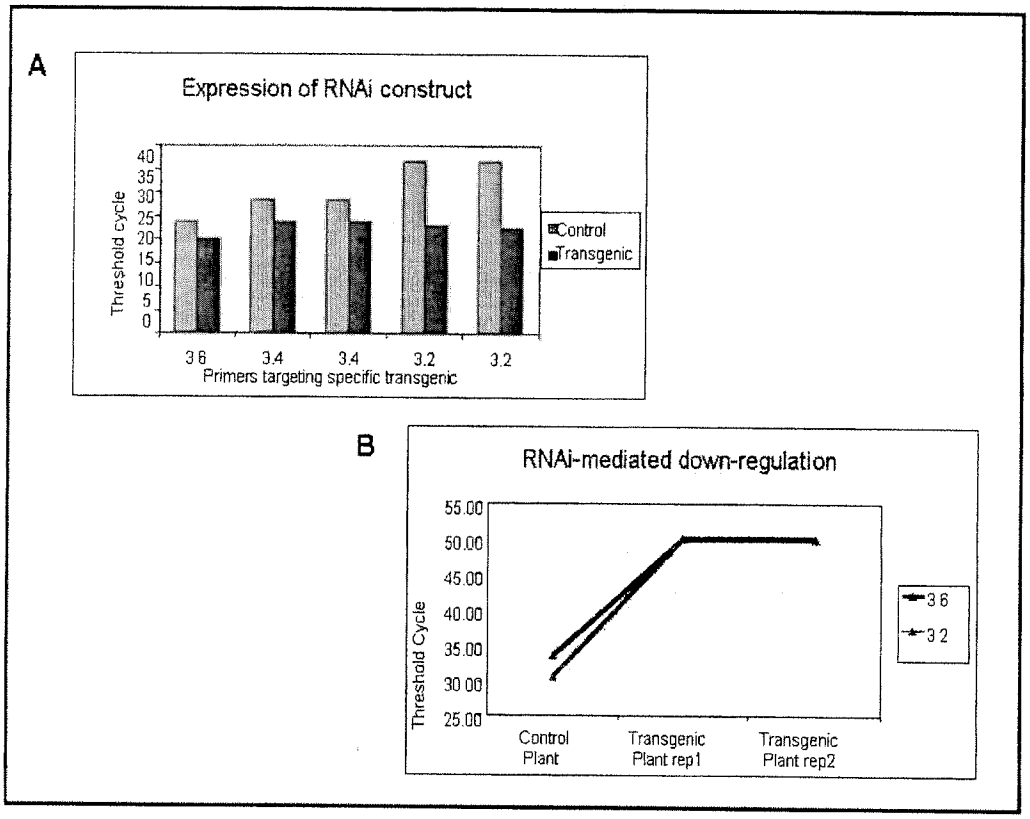

Figure 8. Normalized RT-PCR cycle values indicating A) High expression of RNAi construct in transgenic plant. B) Low expression of target gene in transgenic RNAi *Populus* plants. A lower threshold cycle value indicates a higher expression of the gene in the corresponding sample.

humidity), plants were placed under a shade cloth and transferred to an unshaded area two weeks later. This first set of transgenic plants were successfully moved to the greenhouse in September of 2005

Verification of Down-Regulation of Target Gene

Following transformation and selection on media containing antibiotics, it is essential to confirm the presence of the transgene to identify and eliminate 'escapes' or plants that survive selection that lack the selectable marker or transgene. This is typically a bottleneck in the production of transgenic plants, so we developed a highly efficient PCR screening method. For PCR confirmation of transgene integration, a simple, fast, high-throughput DNA extraction method was developed based on two earlier methods (Dilworth & Frey 2000, Burr *et al.* 2001). Briefly, young leaves were cut into small pieces (<1 mm^2) and coarsely ground in a DNA extraction buffer (0.5% tritonX 100, 2mg mL^{-1} Proteinase K and TE buffer pH 8.0). Lysis of leaf cells was performed using a thermocycler program set at: 65°C for 5min, 96°C for 2min, 65°C for 4min, 96°C for 1min, 65°C for 30sec, 96°C for 20sec. Target DNA was amplified using 35S forward and gene-specific reverse (*PoptrIAA3.6*, *PoptrIAA3.4*, *PoptrIAA3.2*) primers to verify integration of the RNAi cassette in the host genome. RT-PCR was conducted using RNA extracted from leaves of transgenic and control plants in order to test expression of the RNAi cassette and reveal the extent of down-regulation. Our results show that the RNAi cassette of three constructs (*PoptrIAA3.6*, *PoptrIAA3.4* and *PoptrIAA3.2*) were expressed very strongly in

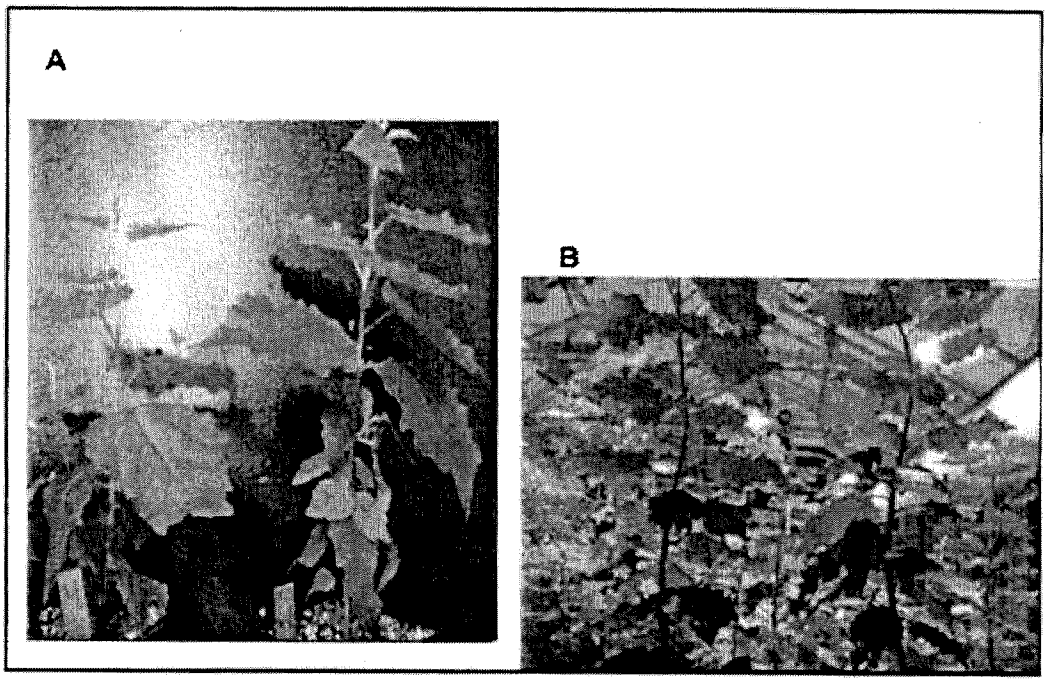

Figure 9. A) A taller *PoptrIAA3.2* transgenic plant with lateral bud growth (right) placed next to control plant (left). B) Characteristic curvy stem structure observed in three transgenic plants (two independent events and a copy) obtained using *PoptrIAA16.3* RNAi constructs.

transgenic plants (**Fig. 8A**) and that high levels of target gene down-regulation for *PoptrIAA3.2* and *PoptrIAA3.6* were achieved (**Fig. 8B**). We also found that, in all investigated plants, there were negligible impacts on the expression of non-target *AUX/IAA* genes.

Preliminary Morphological Phenotyping

As part of an initial transgenic characterization effort, we quantified the aboveground morphology of rooted stock plants in order to identify any observable phenotypic changes associated with RNAi-mediated down-regulation of the candidate gene. *PoptrIAA3.2* plants grew faster, were taller and had thicker stems than control plants (**Fig. 9A**). We also observed that the faster growing *PoptrIAA3.2* plants initiated bud elongation when subjected to supplemental fertilizer treatment compared to the control. Our results indicate that *PoptrIAA3.2* acts as a repressor of genes that play a role in facilitating bud growth and that RNAi-mediated down-regulation of this gene relieves this repression. Multiple independent transgenic plants lines carrying the *PoptrIAA3.4* RNAi construct displayed slower growth resulting in shorter heights than the wild-type plants. Transgenic plants carrying the RNAi construct targeting *PoptrIAA16.3* produced a visible 'S-shaped' stem structure (**Fig. 9B**). These stems were also thicker than any of the other transgenic or control plant of the same height. In addition, its leaves were greener than those of the control plants.

Table 1. Data on number and length of primary roots and total root dry weight obtained from 4-week growth of transgenic as well as control *Populus* plants.

Construct	Treatment	Avg. # Primary	Avg. Length (cm)	Avg. Dry Wt (mg)
IAA3.6	Hydroponics + NAA	8	7.3	34.6
IAA3.4	Hydroponics + NAA	6	3.7	20.2
IAA3.2	Hydroponics + NAA	6	1.6	12.3
Control	Hydroponics + NAA	7	7.1	38.5
IAA3.6	Soil + IBA	0	0.0	0.0
IAA3.4	Soil + IBA	1	5.7	3.4
IAA3.2	Soil + IBA	0	0.0	0.0
Control	Soil + IBA	0	0.0	0.0
IAA3.6	Root Cube + IBA	3	1.7	3.0
IAA3.4	Root Cube + IBA	2	2.5	3.6
IAA3.2	Root Cube + IBA	1	2.3	2.8
Control	Root Cube + IBA	3	4.8	7.0
IAA3.6	Root Cube - IBA	5	6.1	21.3
IAA3.4	Root Cube - IBA	3	3.1	4.0
IAA3.2	Root Cube - IBA	4	6.0	17.0
Control	Root Cube - IBA	2	0.8	1.5

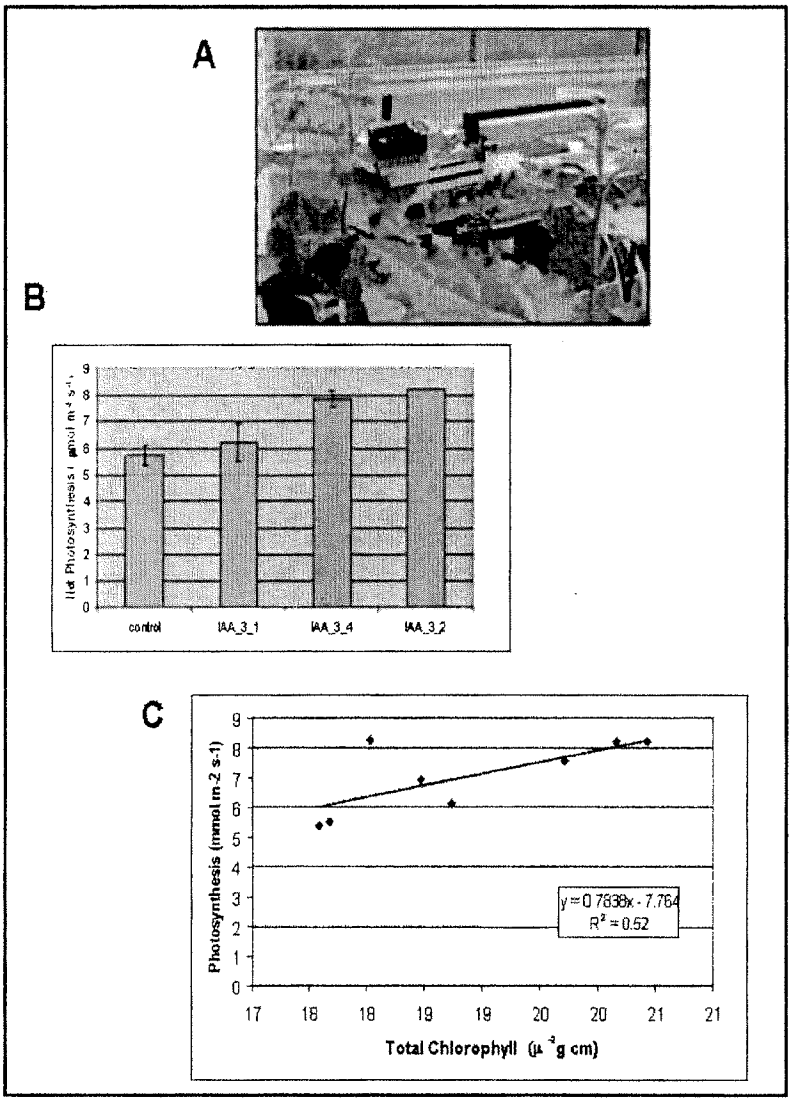

Figure 10. A) Set up showing measurement using LI-6400 instrument and *Populus* leaves. B) Mean photosynthetic rates for the third fully expanded leaf of each plant. Error bars represent standard error of the mean (n=2). C) Mean photosynthetic rates as a function of leaf chlorophyll content.

Preliminary assessment of root growth in the transgenics, based on stem cuttings placed in: 1) a 75 ppm nitrogen (20-10-20) hydroponic solution containing 50 µM NAA with constant aeration, 2) rooting powder (0.1% IBA) and moved to soil (Fafard 4M mix) or 3) pre-soaked Oasis® Hortitcubes® rooting cubes after treatment with IBA, indicated that Treatment 1 consistently yielded greater amounts of roots than the control (**Table 1**). The cuttings grown in the Hortitcubes® system without hormone treatment had better root growth than those treated with rooting powder, suggesting that the hormone treatment may have been inhibitory rather than stimulatory to root initiation in *AUX/IAA* transgenic cuttings whereas the effect of auxin (IBA/NAA) on root growth was positive for control cuttings. Of the rooting procedures tested,

hydroponic system seems to be the most suitable for rapid root initiation, as well as ease of accessibility for phenotyping transgenic root architecture. This preliminary experiment suggests that *AUX/IAA* genes (*PoptrIAA3.2, 3.4* and *3.6*) have a functional role in facilitating auxin-dependent root growth.

Preliminary Physiological Characterization

As part of an initial transgenic characterization effort, we quantified the aboveground physiology of the rooted stock plants in an attempt to illuminate the underlying physiological changes associated with transgenesis and to identify abnormal phenotypes that would subsequently be removed from further evaluation (**Fig. 10**). Leaf gas exchange measurements were taken using a portable steady-state photosynthesis system, model LI-6400. Differences among transgenic lines were evaluated using ANOVA followed by an lsd mean separation test. We found that net photosynthetic rates were significantly different among the three transgenic lines ($P=0.04$) and the control. None of the other measured variables produced significant different at $P\geq0.2$. All of the transgenic plants had slightly higher photosynthetic rates than the controls; rates in *PoptrIAA3.4* plants were 43% higher than in the controls (**Fig. 10B**). Although chlorophyll content did not differ among lines, photosynthesis in individual plants was positively correlated with chlorophyll content ($R^2 = 0.52$). Enhanced photosynthetic rates may be a transgenic phenotype that merits further study.

In total, these efforts have led to the establishment of a multi-level transgenic plant evaluation pipeline. Transgenic plants representing the following transgenic constructs were shipped to ORNL by December 2005 -- *PoptrIAA3.1, 3.2, 3.4, 3.6, 11, 12.2, 12.1, 16.2, 16.3, 7.2, 15, 20.2, 28.1, 26.2* and *26.1*. These plants have been established as rooted stock plants and will be maintained as hedges to allow uniform propagation of clonally replicated rooted cuttings. Reliable morphologic and physiologic data will emerge from above- and belowground observations based on replicates of propagated plants, enhancing our understanding of the functional roles of *AUX/IAA* and *ARF* genes in plant development in general and in root development specifically. The production and acclimation of these transgenic plants is a major accomplishment of this project. We have created a unique biological resource for deconvoluting the functions of these complex, inter-related gene families and we have developed the protocols and pipelines to allow efficient genotyping and rapid progress in functional characterization of the candidate genes in the next phase of the project.

WHOLE-GENOME MICROARRAY DESIGN AND CONSTRUCTION

In order to be able to characterize comprehensive changes in gene expression associated with transgenesis, we developed a whole-genome microarray for *Populus*. Gene expression analyses using the whole-genome microarray will lead us to a more mechanistic understanding of how the *AUX/IAA*, *ARF*, *RR* and *INV* genes influence carbon allocation and partitioning processes in *Populus*. The *Populus* microarray was designed to target all predicted genes in the *Populus* nuclear and organellar genomes, as well as divergent aspen transcripts and predicted miRNA.

Briefly, we created the *Populus* whole-genome microarray by examining *P. trichocarpa* nuclear gene models predicted from the whole-genome sequence (Tuskan *et al.* 2006) using four methods, Genewise, FgenesH, GrailEXP6 and EuGène (http://genome.jgi-psf.org/cgi-bin/searchGM?db=Poptr1). We selected gene models based on EST support and completeness and homology to a curated set of proteins. Chloroplast and mitochondrial genes were predicted using the ORNL prokaryotic pipeline, which includes gene predictions based on Glimmer (Delcher *et al.* 1999) and Critica (Badger & Olsen 1999) algorithms. Aspen unigenes were selected using TGICL (Pertea *et al.* 2003). These methodologies identified 15,245 unigene clusters and 6318 singletons. This unigene set was compared to the predicted *Populus* transcripts using BLASTN. Aspen unigenes and singletons that failed one or more of the following criteria

Table 2. Number of sequences used from each source at different stages of the microarray design. Sources: Aspen: unigenes and singletons from the EST clustering that did not match predicted *Populus trichocarpa* genes; Chloroplast: genes predicted from the *P. trichocarpa* chloroplast sequence; EuGène, FgenesH, Grail, and Genewise: genes predicted from the *P. trichocarpa* genome sequence using one or more of these four gene prediction algorithms; miRNA: microRNAs predicted from the *P. trichocarpa* genome sequence; Mitochondria: genes predicted from the *P. trichocarpa* mitochondrial genome sequence.

Source	28-Oct	17-Nov	Final	Excluded
Aspen	10390	9995	9995	0
Chloroplast	71	71	69	2
EuGène	23580	17933	15271	2662
FgenesH	22002	19087	19042	45
Grail	7034	5865	5308	557
Genewise	3068	16331	16173	158
miRNA	49	49	49	0
Mitochondria	38	59	58	1
Total	66232	69390	65965	3425
Total Nuclear	55684	59216	55794	6850

were identified based on: 1) expectation score > $1e^{-10}$ (1325 clusters, 472 singletons), 2) hit length < 200 nt (690 clusters, 1625 singletons), 3) hit covers less than 50% of EST length (716 clusters, 121 singletons) and 4) identity (average weighted by HSP length) <95% (4140 clusters, 1548 singletons). An additional 9,995 unique aspen sequences were included in the design of the microarray (these sequence names begin with "aspen").

In total, 66,232 targets were initially used in the design of the microarray (**Table 2**). The three best probes for each gene were identified using a NimbleGen design algorithm, based on the 1) uniqueness of all 24-mers comprising the oligonucleotide, 2) the composite scores based on position of the probe relative to other selected probes, 3) self-complementarity and 4) base composition. Probes were rejected only if they fail the self-complementarity and base composition criteria. We screened the full set of 198,696 probes (= 66,232 targets x 3 probes per target) for uniqueness by directly comparing probe composition to other selected probes and for similarity to non-target sequences by comparing all probes to the predicted transcript set using unfiltered WU-BLASTN. 40,207 probes had 85% or greater identity over 95% of the probe. The predicted mRNA were appended to include up to150 bp of 3' genomic sequence ('pseudoUTR') in an effort to create at least one unique probe sequence per gene model. This was necessary for 15,118 gene models.

NimbleGen then synthesized probes for the 69,390 targets. This resulted in a set of 208,158 probes. We removed 3,722 probes that exactly matched other probes on the array and

9,460 probes corresponding to targets with matches to transposable elements. We annotated all array targets using best BLAST hits using a WU-BLASTX comparison of the targets against the Arabidopsis peptide database (ATH1_pep_20020731 from TAIR), followed by a WU-BLASTX against the non-redundant protein database at GenBank (downloaded from NCBI on 1/30/2005). The cutoff for 'No Hit' was an E-score$\leq 1e^{-10}$. Also, to facilitate comparison of the aspen transcript results to those of the closest matching poplar gene, the best WU-BLASTN hit for the aspen probes was determined and compared to the predicted poplar transcripts.

RT-PCR AND MICROARRAY-BASED EXPRESSION STUDIES

Bioinformatics is one path towards predicting the functional roles *AUX/IAA*, *ARF* and *RR* transcription factors play during growth, development and environmental adaptation of *Populus*. Support from experimental explorations is, however, required to unravel the functional contexts of these genes. Our experimental approach included gene-specific and global spatiotemporal expression profiling.

AUX/IAA and *ARF* Gene Expression

We undertook RT-PCR studies using leaf, stem and root tissues because expression patterns of *AUX/IAA* and *ARF* genes were largely unknown in wild-type *Populus* tissues. *Populus* genomic DNA was used to determine appropriate PCR annealing temperatures for the gene-specific primers. RNA was extracted using RNeasy Plant Mini Kit (Qiagen) and the cDNA was synthesized using Invitrogen's SuperScript First-Strand Synthesis system. RT-PCR reactions were set up in triplicates using cDNA samples, the gene-specific primers and iQ supermix (BioRad). Primers specific to *Populus* Actin gene were included as control on every PCR plate. 40 cycles of PCR were run using an icycler real-time PCR machine (BioRad). A difference of 1-cycle implies a 2-fold difference in expression level.

Our RT-PCR and EST distribution studies show that *PoptrIAA3* co-orthologs display differential expression between leaf, stem and root tissues. For example, *PoptrIAA3.1* and *PoptrIAA3.2* appear to express in cambial zone and during wood formation. *PoptrIAA3.2* had a 32-fold higher expression level in stem than in roots. *PoptrIAA3.5* and *PoptrIAA3.6* produced the highest expression in roots, suggesting that these genes may have functional roles similar to *AtIAA3* (*AtIAA3* has a gain-of-function mutant, *shy2*, whose phenotypic characteristics include fewer lateral roots (Tian *et al.* 2002)). *PoptrIAA3.4* has a distinct expression pattern compared to other *PoptrIAA3* genes, yielding the weakest expression in leaf, then stem and then root tissues. RT-PCR studies have provided expression support for several predicted genes and provided clues to their functional contexts which will be duly considered in our renewal proposal.

Expressed Sequence Tag Analysis

Prior to the development of whole-genome microarrays, our greatest opportunity for identifying the suits of genes turned on or off in response to auxin treatment was through EST sequence analysis. Analysis based on the Munich Information Center for Protein Sequences (MIPS) functional categorization revealed that the highest majority of up-regulated (28%) and down-regulated (28%) EST represented unclassified proteins (**Fig. 11**). Eight of the resulting 26 functional categories differed significantly between the up- and down-regulated EST. For example, significantly higher number of up-regulated EST were represented in the functional categories, protein synthesis, interaction with the cell environment, interaction with the environment, cell fate and cell type localization. There was significant correlation between differences in expression levels as indicated by frequency of EST representation and cycle (fold) differences observed in follow-on RT-PCR. Because this experiment was based on RNA obtained after prolonged (9-day) auxin exposure, we predict that much of the expression data represents

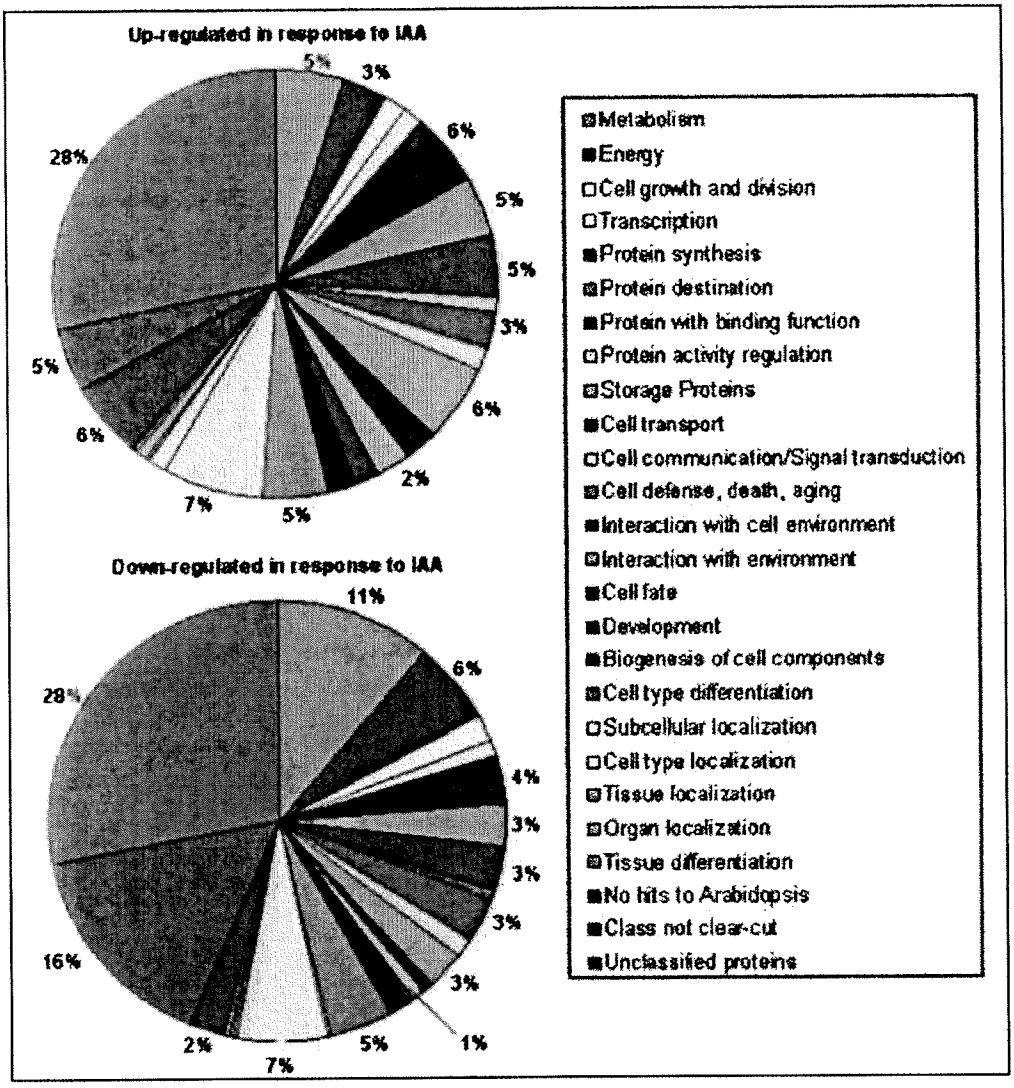

Figure 11. Pie chart depicting Munich Information Center for Protein Sequences (MIPS) classification of EST differentially expressed between auxin-treated and control libraries. The categorization of EST into various biological functional categories (*e.g.*, metabolism, energy, etc) is based on information obtained from querying the MIPS database using EST from auxin study. The various functional categories are color coded.

downstream mediators of auxin response or stress response and may preview RNAi transgenic phenotypes for the *AUX/IAA* and *ARF* genes.

Auxin Treatment Studies: Nis-1 Microarray Experiments

The purpose of this sub-task was to identify auxin regulated genes using whole-genome microarrays. *Populus* plants were grown, treated and sampled at UF by both UF and ORNL personnel. Plants were harvested at 0 hr, 1 hr and 24 hr time points. Total RNA was extracted

from young leaves, mature leaves, nodes, internodes and roots using a CTAB extraction protocol. RNA samples (representing two biological replicates) were shipped to NimbleGen for microarray hybridization.

In the absence of *AUX/IAA/ARF* transgenic plants, we conducted a pharmacological experiment to understand dynamic qualitative and quantitative changes in the transcriptome of auxin-treated plants and to phenocopy the predicted phenotypes of the *AUX/IAA/ARF* transgenics. Prior to analysis, expression data for non-plant-related elements used for within-array quality control were discarded. The log2 transformed data (Y_{ijklm}) were normalized by microarray to a mean of 0 and a standard deviation of 1. Gene significance was estimated using the following model:

$$Y_{ijklm} = \mu + C_i + T_j + I_k + P_l + CT_{il} + CI_{jl} + TI_{kl} + CTI_{ijk} + N(CT)_{mij} + IN(CT)_{ijkm} + \epsilon_{ijklm}$$

where, μ is the mean, C, T and I are the fixed concentration (i=1-4), time (j=1-2) and tissue effects (k=1-5), respectively, and CT, CI, TI and CTI are the fixed interactions among tissue, time and concentration. P, N(CT) and IN(CT) are the random probe (l=1-2), plant (m=1-3) and chip effects, respectively. ϵ denotes the random variable error within the experiment. The model was run using the PROC MIX procedure in SAS (SAS Institute Inc. SAS/STAT Software version 9.1). Significant gene differences were generated using a false discovery rate of 1% (Storey & Tibshirani 2003).

Of the 75 *AUX/IAA/ARF* genes predicted from the *Populus* genome sequence, 15 are regulated in one or more of our treatment combinations; 7 show significant regulation 1 hr after treatment, 12 after 24 hr of treatment and 4 after both 1 and 24 hr of auxin treatment (**Table 3**, parsed from list of significant genes with false detection rate of 1%). To further parse this gene list, we filtered the SAS output by eliminating genes whose expression was inconsistent with the expected concentration-dependence of the response to auxin, *i.e.*, significantly regulated in both 0 vs.1 μM NAA and 0 vs.100 μM NAA contrasts. At the 1 hr time point, young leaves were most

Table 3. Regulation of *AUX/IAA* and *ARF* gene in *Populus* by auxin treatment.

Gene	Regulated 1hr	Regulated 24 hr
PoptrARF2.1	x	
PoptrARF2.2		x
PoptrARF3.1	x	
PoptrARF4	x	x
PoptrARF6.5		x
PoptrARF8.1	x	x
PoptrARF9.1		x
PoptrARF9.4		x
PoptrIAA12.2	x	x
PoptrIAA19.1		x
PoptrIAA27.1	x	
PoptrIAA3.5		x
PoptrIAA3.6		x
PoptrIAA33.1		x
PoptrIAA9	x	x

responsive, which is consistent with the expectation that auxin uptake would be greatest in tissues lacking a well-developed cuticle. At the 24 hr time point, mature leaves had the highest number of significantly regulated transcripts over 1, 10 and 100 μM NAA concentrations (604, 883 and 688, respectively). Interestingly, roots were the next highest in terms of the number of significantly regulated genes, indicating uptake of the foliar applied auxin and gene expression responses distal to the site of application. Thus, it appears that exogenous auxin can at least partially phenocopy the phenotypes expected from transgenesis, suggesting that transgenesis offers the opportunity to deconvolute the transitional imperceptible steps between genes and root phenotypes.

Genes down-regulated by auxin in roots include genes encoding the putative cell wall biosynthetic enzymes O-methyltransferase, caffeoyl-CoA-O-methyltranserase, and cinnamoyl CoA reductase, suggesting a trade-off between root growth and secondary cell wall synthesis. Also regulated are genes encoding enzymes that modify cell walls, such as pectinesterase (repressed) and xyloglucan endotransglycosylase (induced). The apparent modulation of secondary cell wall biosynthesis in response to exogenous auxin in *Populus* is intriguing with respect to potential carbon partitioning.

Response Regulator (*RR*) Gene Expression

With 24 members, it is likely that *Populus RR* gene family members would reveal varying degrees of tissue specificities. We used both whole-genome microarray and EST profiling data to gain a better understanding of which members of the *RR* gene family were being expressed in root tissues. The whole-genome microarray data revealed that 7 of the 11 type A *RR* were expressed above background in one or more of the tissues assayed. Tissue specificity was evident for specific *Populus* type A *RR*; *e.g.*, *PoptrRR7* does not appear expressed in young leaves. Transcripts were not detected in the microarrays for 4 type A *RR*; 2 that had EST support and 2 did not. For the *Populus* type B *RR*, microarray expression data identified 9 of 13 *Populus* genes with expression levels above background and with variable levels of expression across tissues. For the 4 *Populus* type B *RR* in which microarray expression data was not obtained; 3 also lacked EST support (**Fig. 12**). Using Northern blot analysis, we were able to confirm the lack

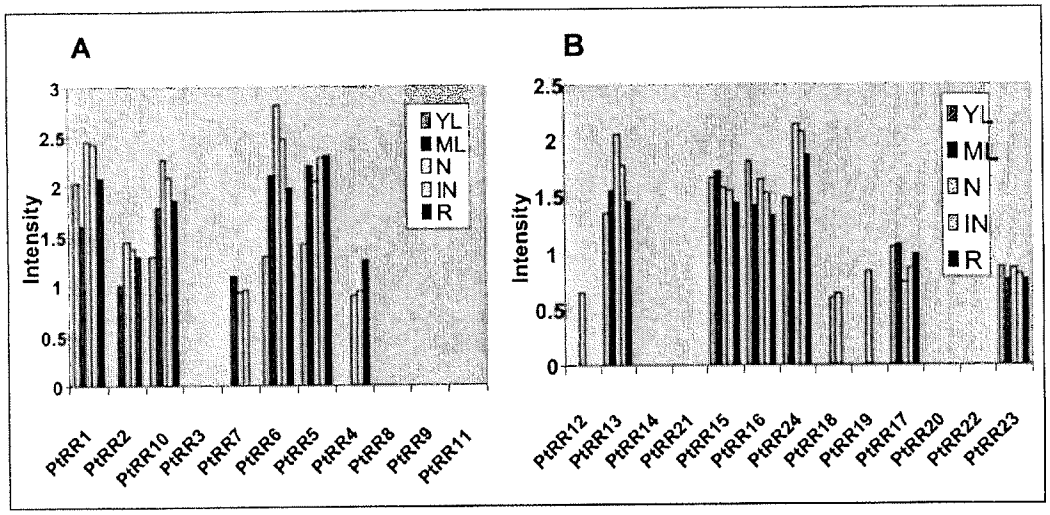

Figure 12. Transcript abundance for *Populus RR* genes. Expression data for (A) type A *RR* and (B) type B *RR* from 5 different tissues nodes (N), internodes (IN), roots (R), mature leaves (ML) and young leaves (YL).

of transcript abundance for 1 of the 4 type B *RR* that were not detected in the microarrays.

Invertase Gene Expression

The *Populus* acid invertase gene family consists of 8 members, suggesting that individual family members may have specific roles in different tissues or under different conditions. This hypothesized specificity could, in part, be explained by differential gene expression. To look for this specificity, we exploited the control plants in the microarray experiment to determine basal expression levels for each of the *PoptrIVR* genes in 5 tissue complements. Three of the *PoptrIVR* (*PoptrIVR2*, *-4* and *-7*) are expressed in all tissues assayed. *PoptrIVR5* is significantly expressed in all tissues except young leaves while *PoptrIVR8* is significantly expressed only in mature leaves ($P \leq 0.05$). The genes not detected by the arrays (*PoptrIVR1*, *-3*, and *-6*) are either restricted to tissues not assayed, such as floral, or expressed below the level of sensitivity of the arrays.

A major goal of this sub-task was to alter carbon allocation by over-expressing a yeast invertase (*suc2*) in the apoplast, cytosol and vacuole. We selected *suc2* because it is not susceptible to plant-specific mechanisms of invertase regulation and we predicted that ectopic

Table 4. Data are ratios of μg g^{-1} fresh weight glucose equivalents of the specified transgenic line relative to the non-transgenic control. Mean metabolite concentrations were calculated from leaves at LPI 11-13 collected from each transgenic line (3 biological replications) and non-transgenic controls (6 biological replications). Means were contrasted by t-tests and presented on the table as mean ratios (= fold change). **Black** font denotes no significant difference, **blue** is significant at $P \leq 0.1$ and **red** is significant at $P \leq 0.05$. The lines with the greatest number of significant metabolite shifts also have the highest *suc2* transcript abundance.

	Vac20/WT	Vac30/WT	CW2/WT	CW14/WT
5C-sugar alcohol	1.08	1.18	1.17	1.18
Glycerol-3-phosphate (+Gly-1-P)	0.79	1.13	1.08	1.19
Iditol?	1.11	1.29	1.21	1.16
Glucoside (m/z 434)	1.08	1.55	1.21	1.37
16.25-273	0.74	0.56	0.52	0.45
alpha-keto-Glutaric acid	1.11	1.27	1.63	2.43
Maleic acid	1.13	1.25	1.42	1.33
Citric acid	1.21	1.25	1.39	0.79
10.79-184	1.08	0.58	0.35	0.28
Fumaric acid	1.43	2.48	1.58	1.32
Serine	0.88	0.87	0.89	0.69
Carbamoyl-phosphate	0.97	0.62	0.35	0.35
15.7-221 327	0.5	0.62	0.51	0.42
Galactosyl-glycerol	0.93	2.26	2.07	1.5

expression would produce a dominant effect on carbon allocation (*e.g.*, altered root:shoot ratio). To target *suc2* to relevant cellular compartments, the following peptides were fused to *suc2*: 1) a secretory transit peptide and a vacuolar targeting domain for vacuolar localization, 2) a secretory transit peptide alone for cell wall localization, and 3) neither for cytosolic localization. A double *35S* promoter was used to drive ectopic expression. Transformation of clone '717' resulted in 25 independent lines for the vacuolar targeted *suc2*, 10 lines for the cell wall localized *suc2* and ~30 for the cytosolic *suc2*. The two RT-PRC based highest vacuolar and cell wall invertase transgenics have been identified; these transgenic plants have no obvious growth phenotypes. We did observe metabolic phenotypes when GC-MS was performed on leaf tissue from these lines (**Table 4**). Overall, the samples are carbohydrate-rich and relatively secondary metabolite-poor. Transgene activity also appears correlated to an accumulation of low molecular weight organic acids from the Krebs cycle. This suggests the products of invertase activity are being shunted into primary metabolism and that the pool of sucrose in the secretory system of *Populus* is relatively low. This is consistent with our emerging hypothesis that *Populus* may primarily utilize symplastic routes for sucrose movement, which in turn impacts carbon allocation in *Populus*.

METABOLOMICS AND QTL ANALYSIS

Metabolomics or metabolic profiling offers tremendous potential to discover novel genes, and assign function to those genes. For gene discovery, we have combined the characterization of the broad spectrum metabolite profiles with quantitative trait loci analysis to identify mQTL.

Construction of Consensus Genetic Map

One of our goals was to build a consensus genetic map that would fuse the genus *Populus* into a macro-pedigree, thus supplying a platform for uniting phenotyping studies with the *Populus* genome. Towards this goal, we built the consensus map of *Populus*, involving four pedigrees and three species, *P. trichocarpa, P. nigra and P. deltoides*. The consensus map consists of 692 SSR markers, with a 3.5 cM average space between markers along the complete genome (**Fig. 13**). This map will be used to facilitate the rapid decomposition of detected QTL into candidate carbon allocation and partitioning genes.

Metabolite Quantitative Trait Loci (mQTL) Analysis

In order to identify genes that control secondary metabolites associated with reduced microbial degradation of *Populus* root biomass we conducted an mQTL analysis. The fine roots of ~190 progeny from Family 13 were sampled during the spring of 2004 at Thief River Falls, MN. The metabolites of the fine roots were subsequently characterized by gas chromatography-mass spectrometry (GC-MS) using electron impact ionization of their trimethylsilyl derivatives. The concentrations of 20 secondary carbon metabolites that included both known and unknown compounds were then subjected to QTL analysis, resulting in the identification of 25 mQTL that are associated with the production of 13 phenolic compounds and phenolic glycoside conjugates. These mQTL explained 6.4-17.7% of the phenotypic variation in the concentration of these metabolites. Furthermore, mQTL were identified for several unique metabolites, including trichocarpinine and its glucoside, trichocarpin. The three mQTL identified for trichocarpinine co-located with mQTL identified for trichocarpin. Similarly, two mQTL associated with the production of three related unknowns shared a 297 m/z fragment that are typical of methoxy-phenolic metabolites (*e.g.*, methoxysalicylates). A methylase reaction (methyl addition to a hydroxyl group via a methyl donor substitution) is a probable candidate for the production of phenolic metabolites with methoxy functional groups. Another possibility is a gene that codes for the glucosyl conjugation of these methoxy-phenolics.

We also identified mQTL for other major secondary carbon metabolites, including salicin and its putative precursor, helicin, (+)-catechin and salireposide, the latter being a terminal

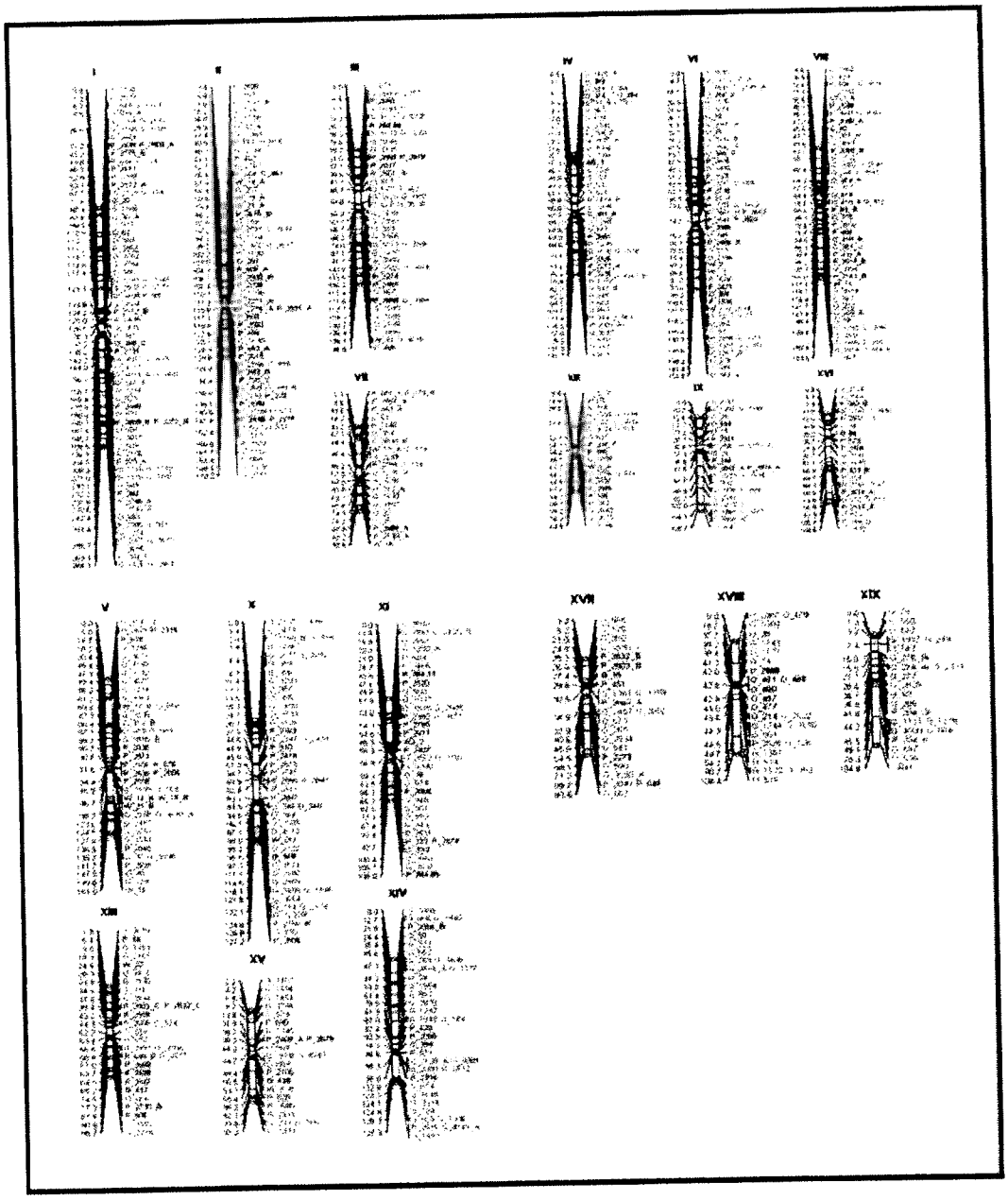

Figure 13. Microsatellites consensus map of *Populus*. All 19 linkage group have been identified and designated with roman numerals. The framework map consists of microsatellite (SSR) and AFLP markers at a 20 cM resolution.

pathway compound, whereas the former metabolites are potentially recalcitrant, higher-order conjugates. The presence of both salicin and helicin in fine roots, coupled with the identification of mQTL for each, without identifying an mQTL is for salicyl alcohol is remarkable. Despite the

common occurrence of salicin and its many higher-order conjugates, we know very little about its biosynthetic pathway. Given that helicin (salicyl aldehyde glucoside) is not abundant in any plant organ, but does accumulate to some degree in roots, the primary site of its production and conversion to salicin may indeed be in fine roots. The implication is that salicin production may occur in its conjugated form rather than in its monomeric form, and the gene responsible, likely code for a protein that reduces the aldehyde residue of helicin to an alcohol to produce salicin. If this proves to be the primary site and pathway of salicin production in *Populus*, then the manipulation of the whole salicylate pathway can be targeted for the overproduction of the complex array of secondary carbon conjugates characteristic of this species and its genus. Our early metabolic profiling analyses targeting the salicylate pathway confirmed the complexity and degree of integration of many of the phenolic pathways.

A preliminary bioinformatic search near the salicin mQTL interval (230 to 625 kb) indicates an aldehyde dehydrogenase homolog, putative/similar to an aldehyde dehydrogenase homolog GI:913941 from [*Brassica napus*]. If salicin is simply produced in fine roots by the conjugation of salicyl alcohol with glucose, we would expect to identify glycosyltransferases (hexosyltransferases). Again, a putative UDP-glucose:glycoprotein glucosyltransferase that is similar to UDP-glucose:glycoprotein glucosyltransferase precursor GB:Q09332 [*Drosophila melanogaster*] is identified near the salicin mQTL interval. It is actually tandemly repeated 4 times just outside the salicin interval (750 to 779 kb on LGIII). Additionally, there is a sterol glucosyltransferase-like protein/sterol glucosyltransferase (ugt52), *Dictyostelium discoideum*, EMBL:AF098916, which is again present in a tandem repeat. Near the helicin interval (8389 to 8832 kb), there are genes that code for glucosyltransferase-like proteins (at 7189 kb and at 9900 kb), a hypothetical protein that is similar to a glucose-repressible alcohol dehydrogenase transcriptional effector GI:3859723 from (*Candida albicans*) (10250 kb), and lactate dehydrogenase (LDH1); supported by full-length cDNA: Ceres:33509 (10687 kb). These are all candidate genes that merit further attention and scrutiny to determine if they have putative roles in the production of higher-order salicylate conjugates.

In summary, this is the first study to combine broad-based metabolic profiling with QTL analysis to successfully identify mQTL that can now be used to elucidate the function of genes that have carbon sequestration potential. The extensive biochemical characterization of unique metabolic phenotypes, as well as mutant *Populus* clones and segregating F_2 progeny from other pedigrees are together providing insights into how metabolic pathways of potentially recalcitrant secondary carbon metabolites are interconnected. Elucidating these biochemical networks will provide opportunities for tailoring metabolic engineering for the overproduction of secondary carbon metabolites in both leaves and fine roots, which can be exploited in carbon sequestration scenarios. Our future research will expand the mQTL identification, use bioinformatic tools to probe the identified marker intervals for putative candidate genes coding for the production of the metabolites and verify gene function through transgenesis.

SUPPLEMENTAL PHENOTYPING OF TRANSGENICS

In order to understand the general impacts of transgenic approaches to modifications of secondary metabolites on tissue biochemistry and overall plant morphology a series of supplemental studies were conducted looking at the salicylate pathway – a potential source of recalcitrant metabolites, and the gibberellin pathway – a potential source of altered carbon allocation genes.

Higher-order Salicylates Perturbed in *nahG* Transgenics *Populus*

Prior to receiving our *AUX/IAA* transgenics, we conducted broad-spectrum metabolic profiling on *Populus* mutants that were likely to have perturbed salicylate metabolism and/or perturbed patterns of carbon partitioning. Leaf material from four transgenic '717' clone

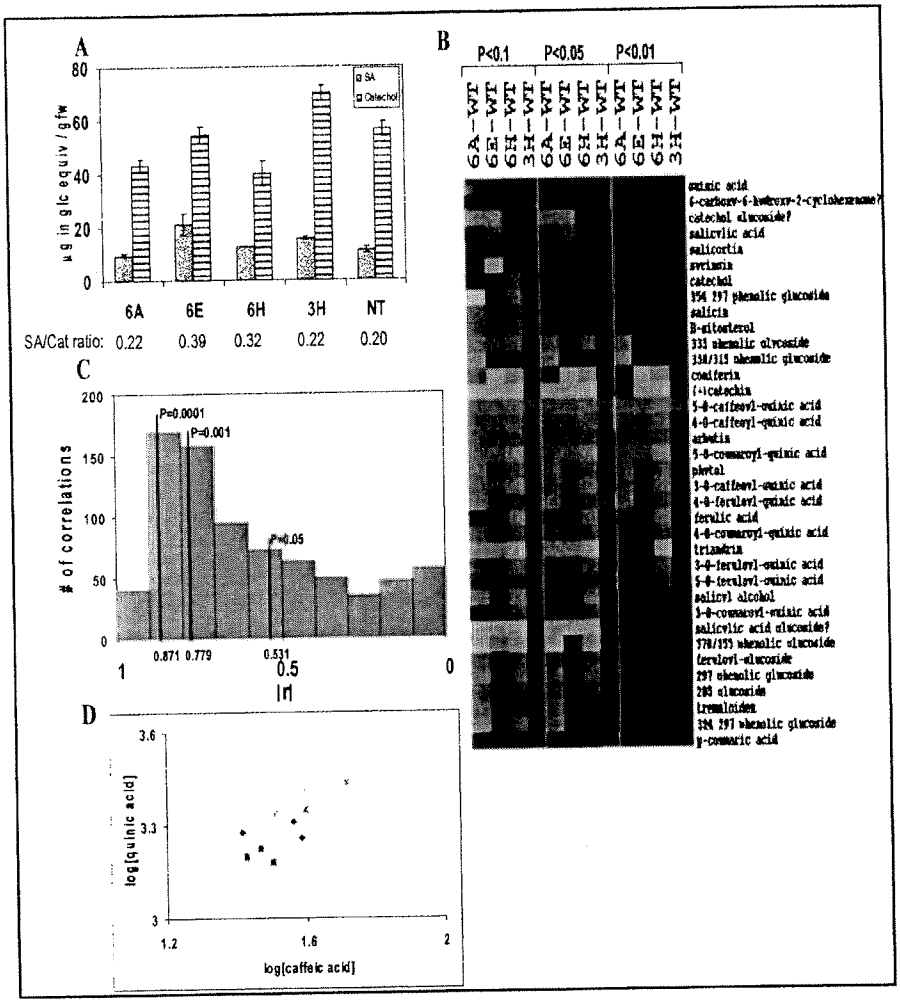

Figure 14. Metabolite levels, perturbation by *nahG*, and underlying correlation structure among metabolites. A) SA and catechol levels are not reduced and increased, respectively, in *nahG* transgenic *Populus*. B) Hierarchical clustergram of significant contrasts between metabolite concentrations in transgenic lines versus non-transgenic controls. Analysis of variance identified metabolites whose concentrations differed significantly, and significant contrasts are shown (red for higher concentration in the transgenic line, green for lower concentration in the transgenic line), at three increasingly stringent levels of significance from left to right ($P \leq 0.10$, $P \leq 0.05$ and $P \leq 0.01$). Thirty-six metabolites were significantly altered by *nahG* at $P \leq 0.10$, 31 at $P \leq 0.05$ and 17 at $P \leq 0.01$. C) Non-random distribution of Pearson's correlation coefficients among significantly regulated metabolites. D) Examples of correlated metabolites in transgenic lines and non-transgenic controls.

containing the microbial 35S:*nahG* gene construct, which were contrasted with the non-transformed control, were sampled. The *nahG* enzyme catalyzes the oxidative decarboxylation of salicylic acid (SA) to catechol (Gaffney *et al.* 1993). Surprisingly, a comparison of the transgenic clones with the non-transformed clone indicated that there were no statistically significant differences in either catechol or salicylic acid concentrations, but large changes in salicylate-derived metabolites, as well as many unexpected phenolic metabolites. Whereas salicylic acid was not impacted, its glucoside was, reducing catechol levels in transgenic lines 6A and 6E to 4.3 & 8.5% of the controls (**Fig. 14**). With stable salicylic acid concentrations, maintained at the expense of higher-order conjugates, we must conclude that the salicylate synthesis machinery in *Populus* is robust. The metabolic responses of the 3H transgenic line were not as pronounced as the 6A, 6E and 6H lines. Whereas there was no change in catechol (1, 2-dihydroxybenzene) or hydroquinone (1, 4-dihydroxybenzene) in 3H, there was a 5-fold accumulation in catechol glucoside, which typically does not accumulate in *Populus*. The less robust 3H transgenic clone exhibited an accumulation of salicylate derivatives, including benzylsalicylate (1.67X), salicortin (1.26X) and salicin (1.39X), but salicyl alcohol was not changed. In contrast, salicyl alcohol was reduced in 6A and 6H.

Additionally, all transgenics showed some declines in phenolic acids of the lignin biosynthetic pathway (**Fig. 14B**) and their organic acid and glucoside conjugates, minor in the case of 3H and much more pronounced in the case of the 6-series clones. The 3H transgenic clone only had two minor phenolic acid-quinic acid conjugates decline while quinic acid increased. In contrast, quinic acid and all of the phenolic acids (with the exception of caffeic acid) and their conjugates declined in 6A. Surprisingly, these declines extended to the glucoside conjugates of monolignols. Additionally, some high concentration phenolics, such as ±catechin and largely unrelated metabolites like phytol and B-sitosterol were reduced (markedly in the case of ±catechin). Perturbation of salicylate production clearly impacted phenolic metabolism, including the upstream precursors of the lignin biosynthetic pathway, and thus is a potentially fruitful avenue to manipulate carbon sequestration.

gai and *rgl1* dwarfs in Transgenic *Populus*

RNAi knockouts of *gai* and *rgl1*, which lack a functional DELLA domain that alters the response to gibberellin (GA), in clone '717', were generated. The *gai* and *rgl1* DELLA-less *Populus* mutants produced dramatic shoot dwarfing accompanied by a pronounced (2- to 3-fold) increase in carbon allocation to root growth. To gain insight into the biochemical changes that might be associated with the dramatic architectural changes of the dwarfed transgenic plants with extensive root systems, we performed metabolic profiling in different organs from plants grown under the same conditions in the greenhouse. Roots of both types of mutants had similar metabolic profiles, including monosaccharide (glucose, fructose and galactose) concentrations that were ≤50% that of the wild-type plants, coupled with two- to three-fold accumulations of citric acid and several amino acids, including Glu, Arg, GABA and Asp. In addition, roots of *rgl1* transgenics accumulated Asn (4.39-fold that of wild-type) and carbamoyl aspartic acid (3.42-fold). In contrast to *rgl1*, *gai* transgenics accumulated salicin (5.27-fold) and threonic acid (1.95-fold) in roots. Another explanation of the increased root growth was suggested by the changes in root metabolic profiles which indicate an increased rate of respiration in roots of *gai* relative to *rgl1* transgenic *Populus*. Depletion of monosaccharides in roots, coupled with the accumulation of citric acid and amine-rich amino acids, such as Asn and Arg, could be due to increased respiratory consumption of monosaccharides to generate Krebs cycle organic acids that are required for amino acid synthesis and root growth. Transgenic regulation of the GA pathway is providing new insights into the adaptive mechanisms of trees, as well as new avenues for biomass modification that can be used to ultimately increase carbon sequestration.

LITERATURE CITED

Alder,M.N., Dames,S., Gaudet,J., and Mango,S.E. 2003. Gene silencing in *Caenorhabditis elegans* by transitive RNA interference. RNA-A Publication of the RNA Society 9: 25-32.

Badger,J.H. and Olsen,G.J. 1999. CRITICA: coding region identification tool invoking comparative analysis. Mol. Biol. Evol. 16: 512-524.

Brummell, D.A., Balint-Kurti, P.J., Harpster, M.H., Palys, J.M., Oeller, P.W. and Gutterson, N. (2003) Inverted repeat of a heterologous 3 '-untranslated region for high-efficiency, high-throughput gene silencing. *Plant J.* 33: 793-800.

Burge,C.B. and Karlin,S. 1998. Finding the genes in genomic DNA. Current Opinion in Structural Biology 8: 346-354.

Burr K, Harper R and Linacre A. (2001) One-Step Isolation of Plant DNA Suitable for PCR amplification. *Plant Mol Biol Rep* 19: 367–371.

Cutler, S and McCourt P. 2005. Dude, where's my phenotype? Dealing with redundancy in signaling networks. *Plant Phys.* 138: 558-559.

Delcher, A.L., Harmon, D., Kasif, S., White, O. and Salzberg, S.L. (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* 27: 4636-4641.

Dilworth, E and Frey JE (2000) A rapid method for high throughput DNA extraction from plant material for PCR amplification. *Plant Mol Biol Rep* 18: 61-64.

Franck A, Guilley H, Jonard G, Richards K and Hirth L. 1980. Nucleotide sequence of cauliflower mosaic virus DNA. *Cell* 21:285-294.

Gaffney, T, Friedrich L, Vernooij B, Negrotto D, Nye G, Uknes S, Ward E, Kessmann H, and Ryals J. 1993. Requirement of salicylic acid for the induction of systemic acquired resistance. *Science* 261: 754-756.

Gleave, AP (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol Biol* 20:1203-7.

Hagen, G. and Guilfoyle, T. (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Mol Biol.* 49, 373-385.

Hamann, T., Benkova, E., Baurle, I., Kientz, M. and Jurgens, G. (2002) The Arabidopsis BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. *Genes Dev.* 16:1610-1615.

Hamilton,A. and Baulcombe,D. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286: 950-952.

Harper, R.M., Stowe-Evans, E.L., Luesse, D.R., Muto, H., Tatematsu, K., Watahiki, M.K., Yamamoto, K. and Liscum, E. (2000) The NPH4 locus encodes the auxin response factor

ARF7, a conditional regulator of differential growth in aerial Arabidopsis tissue. *Plant Cell* 12:757-770.

Helliwell,C.A., Wesley,S.V., Wielopolska,A.J., and Waterhouse,P.M. 2002. High-throughput vectors for efficient gene silencing in plants. Functional Plant Biology 29: 1217-1225.

Koncz, C., H. De Greve, D. Andre, F. Deboeck, M. Van Montagu, and J. Schell. (1983) The opine synthase genes carried by Ti plasmids contain all signals necessary for expression in plants. *EMBO J.* 2(9):1597-1604.

Ma, C., S.H. Strauss and R.Meilan. 2004. *Agrobacterium*-mediated transformation of the genome-sequenced *Populus* clone, 'Nisqually-1' (*Populus trichocarpa*). *Plant Molec Biol Reporter* 22:1-9.

Mandel, MA, Gustafson-Brown C, Savidge B, Yanofsky MF (1992) Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. *Nature* 360:273-7.

Nordstrom, A, Tarkowski P, Tarkowska D, Norbaek R, Astot C, Dolezal K and Sandberg G. (2004) Auxin regulation of cytokinin biosynthesis in Arabidopsis thaliana: a factor of potential importance for auxin-cytokinin-regulated development. *Proc Natl Acad Sci U S A.* 101(21):8039-44.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., Lee, Y., White, J., Cheung, F., Parvizi, B., Tsai, J., and Quackenbush, J. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19: 651-652.

Rutherford,K., Parkhill,J., Crook,J., Horsnell,T., Rice,P., Rajandream,M.A., and Barrell,B. 2000. Artemis: sequence visualization and annotation. Bioinformatics 16: 944-945.

Sakai, H., Honma, T., Aoyama, T., Sato, S., Kato, T., Tabata, S., and Oka, A. 2001. ARR1, A transcription factor for genes immediately responsive to cytokinins. Science 294:1519-1521.

Sessions, A, Nemhauser JL, McColl A, Roe JL, Feldmann KA and Zambryski PC (1997) ETTIN patterns the Arabidopsis floral meristem and reproductive organs. *Development* 124:4481-91.

Skoog, F. and Miller, C.O. (1957) *Symp. Soc. Exp. Biol.* 11:118–131.

Storey, JD and Tibshirani R. (2003) Statistical significance for genome-wide studies. *Proc. Natl. Acad. Sci. U. S. A.* 100: 9440-9445.

Tian, Q., Uhlir, N.J. and Reed, J.W. (2002) Arabidopsis SHY2/IAA3 inhibits auxin-regulated gene expression. *Plant Cell* 14:301-319.

Tuskan, G.A., S.P. DiFazio, U. Hellsten, S. Jansson, S. Rombauts, N. Putnam, L. Sterck, J. Bohlmann, J. Schein, R.R. Bhalerao, R.P. Bhalerao, D. Blaudez, W. Boerjan, A. Brun, A. Brunner, V. Busov, M. Campbell, J. Carlson, M. Chalot, J. Chapman, G. Chen, D. Cooper, P.M. Coutinho, J. Couturier, S.F. Covert, R. Cunningham, J. Davis, S. Degroeve, C. dePamphilis, J. Detter, B. Dirks, I. Dubchak, S. Duplessis, J. Ehlting, B. Ellis, K.

Gendler, D. Goodstein, M. Gribskov, I. Grigoriev, A. Groover, L. Gunter, B. Hamberger, B. Heinze, Y. Helariutta, B. Henrissat, D. Holligan, N. Islam-Faridi, M. Jones-Rhoades, R. Jorgensen, C. Joshi, J. Kangasjärvi, J. Karlsson, C. Kelleher, R. Kirkpatrick, M. Kirst, A. Kohler, U. Kalluri, F. Larimer, J. Leebens-Mack, J.C. Leplé, A. Déjardin, G. Pilate, P. Locascio, S. Lucas, F. Martin, B. Montanini, C. Napoli, D.R. Nelson, C.D. Nelson, K.M. Nieminen, O. Nilsson, G. Peter, R. Philippe, A. Poliakov, S. Ralph, P. Richardson, C. Rinaldi, K. Ritland, P. Rouzé, D. Ryaboy, A. Salamov, J. Schrader, B. Segerman, F. Sterky, C. Souza, C. Tsai, P. Unneberg, K. Wall, S. Wessler, G. Yang, T. Yin, C. Douglas, G. Sandberg, Y. Van de Peer & D. Rokhsar (2006) The genome of western black cottonwood, *Populus trichocarpa* (Torr. & Gray ex Brayshaw). *Science* (In review).

Tymowska-Lalanne, Z, Kreis M. 1998. Expression of the Arabidopsis thaliana invertase gene family. *Planta* 207:259-265.

Vancanneyt, G., R. Schmidt, A. O'Connor-Sanchez, and M. Rocha-Sosa. (1990) Construction of an intron-containing marker gene: Splicing of the intron intransgenic plants and its use in monitoring early events in *Agrobacterium*-mediated plant transformation. *Mol. Gen. Genet.* 220:245-250.

Walker, AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD and Gray JC. (1999) The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *Plant Cell* 11:1337-50.

Waterhouse,P.M., Graham,M.W., and Wang,M.B. 1998. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci U S A 95: 13959-13964.

ORIGINAL MILESTONES

Year 1

- Create all RNAi constructs for *AUX/IAA* and *ARF* gene families (**Completed ORNL**)
- Complete framework genetic map for pedigree (**Completed ORNL**)
- Complete metabolic profiling for 200 individuals in pedigree (**Completed ORNL**)
- Place 5 metabolites on framework map (**Completed ORNL**)
- Identify all members of *AUX/IAA* and *ARF* gene families (**Completed ORNL**)
- Develop software for oligonucleotide microarray design for *Populus* (**Completed ORNL**)
- Design and test microarray (**Completed ORNL**)
- Perform baseline phenotyping of untransformed control trees (**Completed ORNL**)
- Create all RNAi constructs for *RR* and *IVR* families (**Completed UF**)
- Perform baseline phenotyping of untransformed control trees (**Completed UF**)
- Develop and test of promoters with root-specific expression in *Populus* (**Completed OSU**)
- Develop transformation protocol for Nisqually-1 and early flowering *P. alba* (**Partially completed OSU**)
- Develop inducible promoters in *Populus* (**Completed OSU**)

Year 2

- Perform fine-mapping of five metabolites, identify candidate genes (**Completed ORNL**)
- Complete construction of microarray (**Completed ORNL**)
- Identification of candidate genes in mapped genome regions affecting carbon partitioning (**Completed ORNL**)
- Conduct expression profiling of transgenics and controls using microarrays (**On-going ORNL, UF**)
- Conduct phenotyping and metabolic profiling of all RNAi transgenics produced in year 1 (**On-going ORNL**)
- Create all constructs for root-specific overexpression and suppression (**Delayed ORNL, UF**)
- Create constructs for inducible suppression of genes that could not be regenerated with constitutive promoters (**Delayed ORNL, UF**)
- Create RNAi constructs for suppression of candidate metabolite genes (**Delayed ORNL**)
- Perform phenotyping on RNAi transgenics produced in year 1 (**On-going UF**)
- Send year 1 transgenics to UF and ORNL for further analysis (**Partially completed OSU**)
- Test root-specific RNAi (**Delayed OSU**)

Year 3

- Conduct expression profiling of transgenics using microarrays (**On-going ORNL, UF**)
- Conduct metabolic profiling of all transgenics (**On-going ORNL**)
- Perform phenotyping on all transgenics (**On-going ORNL, UF**)
- Create constructs for root-specific overexpression of candidate genes for carbon allocation and partitioning (**Delayed ORNL**)
- Produce transgenic trees using constructs supplied by ORNL and UF (**Partially completed OSU**)
- Send year 2 transgenics to UF and ORNL for further analysis (**Delayed OSU**)
- Send year 3 transgenics to UF and ORNL for further analysis (**Delayed OSU**)

PROJECT PUBLICATIONS

Busov, V., Meilan, R., Pearce, D., Rood, S., Ma, C., Tschaplinski, T., and S. Strauss. 2006. Transgenic modification of *gai* or *rgl1* causes dwarfing and alters gibberellins, root growth, and metabolite profiles in *Populus*. *Planta* (In press).

Davis, M., G.A. Tuskan, M.M. Payne and R. Meilan. 2006. Assessment of *Populus* wood chemistry following the introduction of a Bt toxin gene. *Tree Physiology* 26:557-564.

Tschaplinski, T., G.A. Tuskan, M.M. Sewell, G.M. Gebre, D.E. Todd and C.D. Pendley. 2006. Phenotypic variation and QTL identification for osmotic potential in an interspecific hybrid inbred F_2 poplar pedigree growing under contrasting environments. *Tree Physiology* 26:595-604.

Tuskan, G.A., S.P. DiFazio, U. Hellsten, S. Jansson, S. Rombauts, N. Putnam, L. Sterck, J. Bohlmann, J. Schein, R.R. Bhalerao, R.P. Bhalerao, D. Blaudez, W. Boerjan, A. Brun, A. Brunner, V. Busov, M. Campbell, J. Carlson, M. Chalot, J. Chapman, G. Chen, D. Cooper, P.M. Coutinho, J. Couturier, S.F. Covert, R. Cunningham, J. Davis, S. Degroeve, C. dePamphilis, J. Detter, B. Dirks, I. Dubchak, S. Duplessis, J. Ehlting, B. Ellis, K. Gendler, D. Goodstein, M. Gribskov, I. Grigoriev, A. Groover, L. Gunter, B. Hamberger, B. Heinze, Y. Helariutta, B. Henrissat, D. Holligan, N. Islam-Faridi, M. Jones-Rhoades, R. Jorgensen, C. Joshi, J. Kangasjärvi, J. Karlsson, C. Kelleher, R. Kirkpatrick, M. Kirst, A. Kohler, U. Kalluri, F. Larimer, J. Leebens-Mack, J.C. Leplé, A. Déjardin, G. Pilate, P. Locascio, S. Lucas, F. Martin, B. Montanini, C. Napoli, D.R. Nelson, C.D. Nelson, K.M. Nieminen, O. Nilsson, G. Peter, R. Philippe, A. Poliakov, S. Ralph, P. Richardson, C. Rinaldi, K. Ritland, P. Rouzé, D. Ryaboy, A. Salamov, J. Schrader, B. Segerman, F. Sterky, C. Souza, C. Tsai, P. Unneberg, K. Wall, S. Wessler, G. Yang, T. Yin, C. Douglas, G. Sandberg, Y. Van de Peer & D. Rokhsar (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray ex Brayshaw). *Science* (In review).

Busov, V.B., A.M. Brunner, R. Meilan, S. Filichkin, L. Ganio, S. Gandhi, and S.H. Strauss. 2006. Genetic transformation: A powerful tool for dissection of adaptive traits in trees. *New Phytol.* (In press).

Meilan, R. and C. Ma. 2006. *Populus*. In: *Agrobacterium* Protocols, Methods in Molecular Biology, vol. 44. Ed. K. Wang, Humana Press (In review).

Filichkin, S.A., R. Meilan, V.B. Busov, C. Ma, A.M. Brunner and S.H. Strauss. 2006. Alcohol-inducible gene expression in transgenic *Populus*. *Plant Cell Reports* (In press).

Wullschleger, S.D., T.M. Yin, S.P. DiFazio, T.J. Tschaplinski, M.F. Davis, L.E. Gunter, and G.A. Tuskan. 2005. Phenotypic variation in absolute and proportional biomass allocation for two advanced-generation pedigrees of hybrid poplar (*Populus* spp.). *Can. J. For. Res.* 35:1779-1789.

Meilan, R., M. Sabatti, C. Ma. and E. Kuzminsky. 2004. An early-flowering genotype of *Populus*. *J. Plant Biol.* 47:52-56.

Ma, C., S.H. Strauss and R.Meilan. 2004. *Agrobacterium*-mediated transformation of the genome-sequenced *Populus* clone, 'Nisqually-1' (*Populus trichocarpa*). *Plant Molec Biol Reporter* 22:1-9.

Tuskan, G.A., L.E. Gunter, Z.K. Yang, T.M. Yin, M.M. Sewell and S.P. DiFazio. 2004. Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. *Can. J. For. Res.* 34:85-93.

Martin F, G.A. Tuskan, S.P. DiFazio, P. Lammers, G. Newcombe, and G.K. Podila. 2004. Symbiotic sequencing for the *Populus* mesocosm. *New Phytol.* 161(2):330-335.

Yin, T.-M., S. P. DiFazio L.E. Gunter D. Riemenschneider and G.A. Tuskan. 2004. Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. *Theor. Appl. Genet.* 109(3):451-463.

Tuskan, G.A., S.P. DiFazio and T. Teichmann. 2004. *Populus* genomics is getting popular: The impact of the *Populus* genome project on tree research. *Plant Biology* 6:2-4.

Stirling, B., Z. Yang, L. Gunter, J. Vrebolav, G. Tuskan and T. Bradshaw. 2003. Comparative sequence analysis between orthologous regions of the Arabidopsis and *Populus* genomes reveals substantial synteny and microcolinearity. *Can. J. For. Res.* 33(11):2245-2251.

PROJECT MANUSCRIPTS IN PREPARATION

Kalluri, U.C., Tschaplinski, T., Patel, M., Jawdy, S., DiFazio, S.P. and Tuskan G.A, 2006. Transcriptome and metabolome analysis of transgenic poplar down-regulated in an Aux/IAA gene expression. *In Prep.*

Aux/IAA proteins are transcriptional repressors that regulate auxin-mediated root development. Transcriptome and metabolome analysis of transgenic plant altered in Aux/IAA functioning (reported here) will reveal the underlying molecular networks/ mechanisms that result in the observed root development.

Brunner, A., DiFazio, S.P., Dharmavardhana, P., Kalluri, U.C., Gunter, L., Tuskan, G.A. 2006. Development of a *Populus* whole-genome oligoarray and assessment of applicability of cottonwood microarrays in studying aspen transcriptome. *In Prep.*

Systems-level knowledge of the transcriptional response of an organism to various internal and external signals can be revealed by the microarray technology. This paper reports the design, development, experimental testing and applications of the first whole-genome microarray for *Populus*.

Busov, V., U. Kalluri, G. Tuskan, J. Schrader, K. Nieminen, J. Kangasjarvi. 2006. Genome-wide survey and validation of genes involved in metabolism and signaling of plant growth regulators in a *Populus* tree. *In Prep.*

Aux/IAA and ARF proteins are known to regulate auxin responses and auxin mediated plant development. This paper reports the bioinformatics and expression studies of Aux/IAA and ARF gene families in *Populus* and its implications in tree growth and development.

Bocock, P., Huang, L.F., Koch, K. and Davis, J.M. 2006. The poplar acid invertase family and response to glucose signaling. *In Prep.*

Invertases can drive sink strength, which in turn can condition shifts in carbon allocation. This paper reports the identification and expression (in roots and in response to glucose signals) of all members of the poplar invertase family of enzymes.

Dervinis, C., Lawrence, S.D., Novak, N.G. and Davis, J.M. 2006. Poplar wound signaling interacts with cytokinin and reduces insect growth. *In Prep.*

This paper reports that cytokinin treatment makes poplar trees more resistant to insect herbivory. The basis of this response is observed at the gene expression level. By extension, we predict that cytokinin RR transgenics will be more resistant to insects due to shifts in carbon partitioning.

Lawrence, S.D., Dervinis, C., Novak, N. and Davis, J.M. 2006. Wound and herbivory induced genes in poplar. *In Prep.*

We collaborated with SL and NN (USDA-ARS Insect Biocontrol Lab) after microarray analysis indicated linkages between cytokinin responses and wound-induced shifts in carbon partitioning. This paper reports that insect- and wound-induced responses are similar.

Morse, A.M., Tschaplinski, T., Dervinis, C., Pijut, P., Day, W., Schmeltz, E. and Davis, J.M. 2006. Metabolic and transcript profiling in poplar reveal compensatory mechanisms in response to a salicylate hydroxylase transgene. *In Prep.*

This paper reports the development of metabolic and transcriptome profiling methods to quantify carbon partitioning shifts in transgenics that lack a whole-plant phenotype. The partitioning shifts involve a suite of quinate compounds known to impact microbial decomposition and thus soil carbon.

Quesada, T., Li, Z., Dervinis, C., Casella, G., DiFazio, S., Tuskan, G., Peter, G., Davis, J.M. and Kirst, M. 2006. Chromatin domains and the evolution of wood. *In Prep.*

This paper reports the development of genome-scale discovery and statistical analysis methods to identify the genes responsible for the properties of roots and stems. By knowing the "addresses" of all poplar genes, we identified which chromosome segments direct gene products to roots and stems, which will allow rapid identification of QTL candidates in the future.