Carbohydrate-related genes and cell wall biosynthesis in vascular tissues of loblolly pine (*Pinus taeda*)

CAMPBELL J. NAIRN,1,2 DENISE M. LENNON,1 ALICIA WOOD-JONES,1 ALISON V. NAIRN3 and JEFFREY F. D. DEAN1

1 Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA 30602, USA
2 Corresponding author (jnairn@warnell.uga.edu)
3 Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA

Received June 25, 2007; accepted August 3, 2007; published online May 1, 2008

Summary  Loblolly pine (*Pinus taeda* L.), the most widely planted tree species in the United States, is an important source of wood and wood fibers for a multitude of consumer products. Wood fibers are primarily composed of secondary cell walls, and cellulose, hemicelluloses and lignin are major components of wood. Fiber morphology and cell wall composition are important determinants of wood properties. We used comparative genomics to identify putative genes for cellulose and hemicellulose synthesis in loblolly pine that are homologous to genes implicated in cell wall synthesis in angiosperms. Sequences encoding putative secondary cell wall cellulose synthase genes, cellulose synthase-like genes, a membrane-bound endoglucanase gene, a sucrose synthase gene, a UDP-glucose pyrophosphorylase gene and GDP-mannose pyrophosphorylase genes were identified in expressed sequence tag (EST) collections from loblolly pine. Full-length coding sequences were obtained from cDNA clones isolated from a library constructed from developing xylem. Phylogenetic relationships between the genes from loblolly pine and angiosperm taxa were examined and transcriptional profiling in vascular tissues was conducted by real-time quantitative, reverse transcriptase-polymerase chain reaction. The putative cell wall synthesis genes were expressed at high levels in vascular tissues and a subset was differentially regulated in xylem and phloem tissues. Inferred phylogenetic relationships and expression patterns for the genes from loblolly pine were consistent with roles in synthesis of complex carbohydrates of the cell wall. These studies suggest functional conservation of homologous wood formation genes in gymnosperm and angiosperm taxa.

Keywords: cellulose, endoglucanase, GDP-mannose pyrophosphorylase, glucomannan, hemicellulose, phloem, real-time quantitative RT-PCR, sucrose synthase, transcript profiling, UDP-glucose pyrophosphorylase, wood formation, xylem.

Introduction

Plant cell walls are important components of food, forage and textiles. Wood, a product of secondary cell wall synthesis in plants, is a widely utilized renewable resource. In addition to the myriad of traditional uses for wood-derived products, lignocellulosic biomass from woody perennials is important in carbon sequestration and for its potential as a renewable feedstock for sustainable energy technologies. Complex carbohydrates make up 70–80% of the dry mass of the plant cell wall and are an increasingly attractive source of sugars for the production of biofuels.

Cell wall biosynthesis, including wood formation, is a complex developmental process and current knowledge of the genes and regulatory mechanisms involved is limited. Identification of the molecular and cellular processes involved in cell wall biosynthesis will enable rapid advances in wood and wood fiber improvement (Merkle and Dean 2000). Complex carbohydrates are the most abundant components of woody plant cell walls. In softwoods, such as loblolly pine (*Pinus taeda* L.), cellulose and hemicelluloses make up 43–45% and 20–23% of the cell wall mass, respectively (Galbe and Zacchi 2002). Cellulose is the most abundant biopolymer on earth, and cellulose quantity, quality and deposition in secondary walls of xylem cells are major factors underlying wood properties. Analysis of wood properties in loblolly pine populations indicated wide variation in the contents of cellulose and mannan, a major softwood hemicellulose, from 28.8 to 43.5% and 4.2 to 10.1%, respectively, and quantitative trait loci (QTL) associated with these cell wall composition differences have been detected, although specific genes were not identified (Tuskan et al. 1999).

The cellulose synthesis (CesS) complex is a multimeric enzyme assemblage that contains distinct cellulose synthase (CesA) proteins encoded by different members of the CesA gene family. It is assembled as an integral membrane complex that utilizes intracellular uridine diphosphoglucose (UDP-Glc) as a substrate for the synthesis of cellulose, which is subsequently extruded through the plasma membrane (Delmer 1999, Dobbin et al. 2002). The CesA and cellulose synthase-like (Csl) genes are members of a gene superfamily (Richmond and Somerville 2000). The genome of
Arabidopsis thaliana L. contains at least 10 CesA genes and at least 29 Csl genes in six subfamilies.


In addition to CesA proteins, other components are likely to be associated with the CesS complex (Doblin et al. 2002). Mutational analysis in Arabidopsis identified a membrane-bound form of an endo-β-1,4-glucanase (EGase), KORRIGAN, that is essential for normal cellulose synthesis in primary and secondary cell walls (Nicol et al. 1998, Szymonowicz et al. 2004). The precise function of KORRIGAN EGase (KOR) remains elusive.

Sucrose synthase (SuSy) and UDP-glucose pyrophosphorylase (UTP-glucose-1-phosphate uridylyltransferase) are enzymes implicated in the control of carbon partitioning to polysaccharide synthesis in plants. Uridine diphosphoglucose, the substrate for cellulose synthesis, can be produced by cleavage of sucrose, as catalyzed by SuSy, or by phosphorylation of glucose-1-phosphate, as catalyzed by UDP-glucose pyrophosphorylase (UGP). A role for sucrose synthase (SuSy) in providing the UDP-Glc substrate for cellulose synthesis is consistent with several lines of research (Winter and Huber 2000, Haigler et al. 2001, Salnikov et al. 2001, 2003). Specific genes encoding for SuSy and UGP are up-regulated during secondary cell wall synthesis in developing wood of Populus (Hertzburg et al. 2001, Meng et al. 2007).

The Csl genes are considered likely candidates for encoding enzymes that polymerize β-linked non-cellulosic polysaccharides (Richmond and Somerville 2000). Identification of a CslA gene from guar (Cyamopsis tetragonoloba (L.) Taub) encoding a mannan synthase was the first evidence that CslA genes were involved in the synthesis of non-cellulosic polysaccharides (Dhugga et al. 2004). Characterization of CslA genes from bryophytes, gymnosperms and angiosperms has demonstrated that they encode enzymes with mannan and glucosmannan synthase activities (Liepmann et al. 2005, Suzuki et al. 2006, Liepmann et al. 2007).

GDP-Mannose (GDP-Man) is a substrate for glucosmannan hemicellulose synthesis catalyzed by enzymes encoded by the CslA genes. GDP-Mannose is essential for N-glycosylation in eukaryotes, synthesis of ascorbate in plants, and is an intermediate in the synthesis of GDP-fucose (Reiter and Vanzin 2001). GDP-Mannose pyrophosphorylase (GMP) catalyzes the formation of GDP-Man from mannose-1-phosphate, and an Arabidopsis mutant of GMP, cyt1, exhibits severe cell wall defects (Lukowitz et al. 2001). Characterization of the cyt1 mutant also suggested a role for N-glycosylation and GMP in cellulose synthesis (Lukowitz et al. 2001).

Comparative genomics studies of loblolly pine EST collections and more extensively developed resources for other plant species can accelerate gene discovery in loblolly pine. Loblolly pine is the most widely planted tree species in the United States with more than a billion seedlings planted annually (Whetten et al. 2001). In addition, almost 500 million loblolly pine seedlings are planted annually on other continents (Schultz 1999). Loblolly pine is largely undomesticated and genetic improvement is likely to produce significant rapid gains in tree productivity (Merkle and Dean 2000).

The genome of loblolly pine contains three CesA genes that are highly expressed in developing xylem (Nairn and Haselkorn 2005). Each of these genes is orthologous to one of the three secondary cell wall CesA genes of angiosperms indicating that evolution of functional specificity in the CesA gene family preceded the divergence of gymnosperm and angiosperm lineages. Two CslA genes from loblolly pine are homologous to CslA genes from angiosperm taxa, and encode enzymes with glucosmannan synthase activity (Liepmann et al. 2007). Collectively, studies in Arabidopsis, rice, Populus and pine suggest that functional roles of orthologous CesA and CslA genes are conserved in cell wall synthesis of vascular tissues in herbaceous and woody dicotyledons, monocotyledons and gymnosperms.

We examined genes from loblolly pine that are homologs of angiosperm genes implicated in cellulose and hemicellulose synthesis. Targeted genes included those encoding for secondary cell wall cellulose synthases (CesA genes), glucosmannan synthases (CslA genes), a KORRIGAN endoglucanase, sucrose synthase, UDP-glucose pyrophosphorylase and GDP-mannose pyrophosphorylases. We examined phylogenetic relationships and expression in xylem and phloem tissues.

Materials and methods
Gene identification, cloning and sequence analysis
Full-length coding regions for three putative secondary cell wall CesA genes and two CslA genes from loblolly pine were previously obtained (Nairn and Haselkorn 2005, Liepmann et al. 2007). Full-length Arabidopsis peptide sequences for the KORRIGAN endoglucanase, sucrose synthases, UDP-glucose pyrophosphorylase and GDP-mannose pyrophosphorylases were used to search the translated loblolly pine EST collection. The “expected threshold” for queries using BLAST (Basic Local Alignment Search Tool) was set at 10−10 for each target gene, and identified loblolly pine EST sequences were downloaded from the NCBI web site (http://www.ncbi.nlm.nih.gov). Contigs were assembled using GCG (Wisconsin Genetics Group). Primers were designed based on the 3′ ends of contig sequences and paired with vector-anchored 5′ primers to amplify full-length coding regions from a full-length enriched cDNA library constructed from developing xylem of loblolly pine. A mixture of 80% AmpliTaq Gold (Applied Biosystems, Foster City, CA) and 20% Pfu polymerase (Stratagene, La Jolla, CA) was used for RT-PCR amplification.
of cDNA. Resulting amplicons were cloned into TA vectors (Promega, Madison, WI and Invitrogen, Carlsbad, CA) and sequences for full-length coding regions were obtained by primer walking. Sequencing reactions were carried out using BigDye Terminator v3.1 reactions and analyzed using a 3730XL Genetic Analyzer (Applied Biosystems). Sequence assembly was accomplished with Sequencer (GeneCodes, Ann Arbor, MI) or GCG, and peptide sequences were derived using programs contained in the GCG suite.

**Phylogenetic analyses**

Protein sequences for putative homologs from additional plant taxa were obtained from GenBank. Sequences were aligned using ClustalW (Thompson et al. 1994). The Gonnet 250 weight matrix was used, gap opening penalties were set at 10.0 and gap extension penalties at 0.10 for pairwise alignments. For multiple alignments, gap open penalties were set at 10.0, gap extension penalties at 0.20, and delay divergent sequences at 30%. Phylogenies were inferred from protein alignments using the Neighbor-Joining distance method and subjected to bootstrap analysis of 1000 pseudoreplicates.

**Plant materials**

Xylem and phloem tissues were harvested from four loblolly pine trees, each representing a different, but unknown genotype. The feral trees were from sites in the Piedmont region of northern Georgia. Tissue samples were collected from each tree during the growing season: April 29 (T1), June 9 (T2), July 12 (T3), and August 19 (T4). Two trees (A and B) from one site were 70–75 cm in diameter at breast height, and two trees (C and D) from a second site were 40–45 cm in diameter. Paired samples of xylem and phloem were obtained from each individual on each sampling day. A 15-cm wide by 40-cm high rectangle was scored with a circular saw to a depth of 2 cm and the bark with attached phloem was peeled away. Developing xylem was harvested from the stem by scraping the surface with a single-edged razor blade. Corresponding phloem samples were harvested by scraping the inner surface of the tissue that was peeled away with the bark. All material was immediately placed in plastic bags or tubes and flash frozen in liquid nitrogen at the site of collection. Frozen material was transferred to dry ice and transported to the laboratory where it was stored at –80 °C.

**RNA isolation and cDNA synthesis**

Total RNA was isolated from loblolly pine tissues by the method of Chang et al. (1993) with the following modifications. Solutions were treated with diethyl pyrocarbonate (DEPC), glassware was baked at 250 °C overnight, and plastic ware was treated with RNA Zap (Ambion, Austin, TX) followed by rinsing with DEPC-treated deionized H2O. Each tissue sample was ground in liquid nitrogen with a mortar and pestle. For each sample, a 3-g aliquot of ground tissue was transferred to 15 ml of grinding-lysis buffer (65 °C) and homogenized with a PowerGen125 (Fisher Scientific, Pittsburgh, PA) for 1 min. An equal volume of chloroform:isoamyl alcohol (24:1, v/v) was added to each homogenate, the mixtures were further homogenized for 1 min, and each was centrifuged at 10,000 g for 10 min. The supernatant was transferred to a clean 50-ml centrifuge tube, extracted with an equal volume of chloroform:isoamyl alcohol, transferred to a 50-ml Phase-Lock tube (Eppendorf, Westbury, NY), and centrifuged at 1500 g for 5 min. Chloroform:isoamyl alcohol extractions were repeated twice more, the supernatants transferred to glass centrifuge tubes, and RNA precipitated by the addition of 0.25 volumes of 10 M LiCl followed by incubation overnight at 4 °C. Samples were centrifuged at 10,000 g and pellets were washed with cold 2.5 M LiCl then resuspended in 4 ml of SSTE (1 M NaCl, 0.5% sodium dodecyl sulfate, 10 mM Tris-HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid). Samples were extracted four more times with equal volumes of chloroform:isoamyl alcohol, transferred to glass centrifuge tubes, and precipitated overnight at –20 °C by the addition of 0.1 volumes of 3 M sodium acetate and 2.5 volumes of ethanol. Samples were centrifuged at 10,000 g for 30 min, pellets washed with 75% cold ethanol, air-dried and resuspended in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Total RNA was treated with RNase-free DNase (Ambion) and poly A+ RNA was purified using oligo-dT magnetic beads according to the manufacturer’s protocol (Dynal, Invitrogen, Carlsbad, CA). Poly A+ enriched RNA was eluted and quantified with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE). First-strand cDNA was synthesized using 100 ng of poly A+ enriched RNA for each tissue sample in 20 µl reactions containing oligo dT and random hexamer primers with SuperScript III according to the manufacturer’s protocol (Invitrogen). Each 20 µl cDNA sample was diluted 1:20 in deionized water and used as template in subsequent real-time quantitative, reverse transcriptase-polymerase chain reaction (qRT-PCR).

**Primer design and validation**

Primer pairs were designed for each gene of interest and potential control genes, including PtAct2, using Primer3 (Rozen and Skaletsky 2000). Parameters were set for primer length at 19–21 bp, primer Tm at 60.0 ± 1.0 °C, and amplimer length of 65–75 bp. Primers (Integrated DNA Technologies, Coralville, IA) were validated with serial dilutions of loblolly pine genomic DNA. Standard curves for cycle threshold (Ct) versus log concentration were plotted and primer efficiencies were determined from the slopes. Primers with efficiencies of 100 ± 10% were selected for qRT-PCR analysis of gene expression. Dissociation curves were examined during primer validation to confirm that each primer pair generated a single product during amplification.

**Real-time quantitative RT-PCR**

Expression profiling was conducted by qRT-PCR with triplicate reactions for each tissue sample and a gene-specific primer pair. Reactions were assembled with 5 µl cDNA template, 5 µl primer mix containing 125 nM each of gene specific forward and reverse primers, and 10 µl SYBR Green Supermix
in 96-well plates sealed with optical film (Bio-Rad Laboratories, Hercules, CA). Reactions were carried out with a MyiQ Single-Color Real-Time PCR Detection System (Bio-Rad Laboratories). The amplification protocol was 95 °C for 3 min, 40 cycles of 95 °C for 30 s, 65 °C for 45 s, 78 °C for 20 s, then 95 °C for 1 min and 55 °C for 1 min. Dissociation curves were obtained to confirm that each reaction yielded single, specific products. Expression profiling for all genes was conducted on four biological replicates with three technical replicates for each cDNA sample and gene-specific primer pair. Several genes were evaluated for use as an internal control. The PtAct2 gene exhibited the most consistent Ct values across all samples and was selected as the endogenous control gene for normalization. The logarithmic mean Ct value for each target and control gene was converted to a linear value using 2^(-Ct) (Livak and Schmittgen 2001). Converted values were normalized to the PtAct2 control gene by dividing the converted value for each target gene by the converted value of the control gene PtAct2.

Results

Gene identification and phylogenetic analysis

A putative KORRIGAN EGase cDNA, PtKor1, was obtained from the loblolly pine xylem library. The protein sequence encoded by PtKor1 (EF619968) shared 79.6% sequence similarity with the Arabidopsis KOR protein sequence (NP_199783). The EGase protein sequences from Arabidopsis, rice, Populus and Pinus radiata D. Don. (Table 1) were obtained from GenBank (Looper and others 1998, Libertini and others 2004, Bhandari and others 2006, Zhou and others 2006). Phylogenies were inferred from alignments with the loblolly pine KOR protein sequence (Figure 1). The PtKor1 sequence grouped within a monophyletic clade that was supported in 100% of 1000 bootstrap pseudoreplicates. This clade included all EGase sequences containing the putative membrane-spanning domain.

Assembly of loblolly EST sequences identified by queries with sucrose synthase (SuSy) proteins yielded two contigs containing over 30 EST sequences each, as well as several singletons. Approximately 800–1200 bp of the 5′ portions of the coding regions were not represented by EST sequences. A 3′ reverse primer was designed for each of the two-contig sequences. Approximately 800–1200 bp of the 5′ portions of the coding regions were not represented by EST sequences. A 3′ reverse primer was designed for each of the two-contig sequences. Therefore, the second containing two EST sequences from roots treated with paraquat. One clone, PtUGP1, was obtained from the xylem library. Sequence analysis of PtUGP1 (EF619969) indicated that it contained an open reading frame encoding a putative UGP of 480 amino acids. A phylogeny inferred from the alignment of UGP protein sequences contained three clades that have strong bootstrap support, and each of these contained the two sequences from individual taxa (Figure 1). Placement of the PtaUGP1 sequence was supported in only 54% of bootstrap pseudoreplicates and was not well resolved in this analysis.

Two contigs were obtained from assembly of EST sequences identified by queries of the loblolly pine database using GDP-mannose pyrophosphorylase protein sequences from Arabidopsis. The PCR amplification of the loblolly pine cDNA library with contig-specific 3′ reverse primers yielded an amplicon for each of the two contigs. The DNA sequence analysis of the two clones revealed that each contained open reading frames of 361 amino acids representing two distinct

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis thaliana</td>
<td>NP_171779</td>
</tr>
<tr>
<td></td>
<td>NP_175323</td>
</tr>
<tr>
<td></td>
<td>NP_177228</td>
</tr>
<tr>
<td></td>
<td>NP_180858</td>
</tr>
<tr>
<td></td>
<td>NP_181984</td>
</tr>
<tr>
<td></td>
<td>NP_192138</td>
</tr>
<tr>
<td></td>
<td>NP_194157</td>
</tr>
<tr>
<td></td>
<td>NP_199783</td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>P0C1U4</td>
</tr>
<tr>
<td></td>
<td>Q5NAT8</td>
</tr>
<tr>
<td></td>
<td>Q6H3F2</td>
</tr>
<tr>
<td></td>
<td>Q62Z06</td>
</tr>
<tr>
<td></td>
<td>Q84Q51</td>
</tr>
<tr>
<td>Pinus radiata</td>
<td>AAC12684</td>
</tr>
<tr>
<td></td>
<td>AAS45400</td>
</tr>
</tbody>
</table>

Table 1. Accession numbers are listed for sequences used to generate multiple alignments of endo-β-1,4-glucanase protein sequences for phylogenetic analysis. The sequences for Arabidopsis thaliana were identified in Libertini et al. (2004), for rice (Oryza sativa) in Zhou et al. (2006), for Pinus radiata in Loopstra et al. (1998), and the sequence from Populus tremuloides was identified in the GenBank database (http://www.ncbi.nlm.nih.gov/).
genes. *PtaGMP1* (EF619965) and *PtaGMP2* (EF619966) shared 80 and 90% identity between their DNA and amino acid sequences, respectively. The loblolly pine GMP protein sequences shared 74–92% with GMP sequences from other taxa examined. A phylogeny inferred from alignments of amino acid sequences grouped together GMP sequences from individual species, with the exception of the CYT1 GMP of *Arabidopsis*, which clustered with the GMP from *Populus* (Figure 1). Taxon-specific groups had strong bootstrap support, and the grouping of the *Arabidopsis* and *Populus* sequences was supported in 61% of bootstrap pseudoreplicates.

The loblolly pine GMP sequences were grouped with the *Arabidopsis-Populus* clade, but relationships were not well resolved, as reflected by bootstrap support of only 52%.

We included a loblolly pine gene for a phenylalanine ammonia-lyase (PAL) that is involved in lignification of xylem during secondary cell wall formation. A PAL sequence, U39792, was identified in the GenBank database for loblolly pine. The encoded protein sequence was used to query the translated loblolly pine EST database. Sequences corresponding to the PAL gene were highly represented in the EST database and assembled into a single contig. A cDNA clone, *PtaPAL1*, was
obtained from the xylem library that contained a full-length coding region.

Expression profiling by qRT-PCR

Expression profiling by qRT-PCR was employed to investigate candidate genes involved in loblolly pine wood formation. Several genes were evaluated for use as internal controls for normalization of expression data, including two ribosomal protein L4 genes (PtaRPL4a and PtaRPL4b), a proteasome subunit gene (PtaRpt3), and several actin and tubulin genes. One of the actin genes, PtaAct2, was expressed most consistently across all samples, and so was selected for inclusion in all of the qRT-PCR experiments.

Three CesA genes, PtaCesA1, PtaCesA2 and PtaCesA3, from loblolly pine are expressed in developing xylem (Nairn and Haselkorn 2005). Analysis by qRT-PCR indicated that expression of PtaCesA1, PtaCesA2 and PtaCesA3 in developing xylem was 10- to > 100-fold higher than in the adjacent phloem tissues (Figure 2). Relative differences in transcript levels between xylem and phloem were consistently observed in four individuals from two size classes growing at two locations. Expression of PtaCesA1 and PtaCesA3 was similar within dates, usually within an order of magnitude, among the four sample trees. Differences in phloem expression were observed between trees on some dates. Expression of PtaCesA2 was higher than that of PtaCesA1 and PtaCesA3.

Expression of PtaKor1 was similar to that observed for the PtaCesA genes and was generally higher in xylem than in phloem (Figure 2). Transcript abundance was either similar or up to 4.4-fold higher in xylem than in phloem. In one sample pair (Tree B, T2), expression in phloem was slightly higher than in xylem.

The PtaSuSy1 gene exhibited levels and patterns of expression similar to those of the secondary cell wall CesA genes (Figure 2). Expression of PtaSuSy1 was higher in xylem samples than in corresponding phloem samples with the exception of the T2 sample from tree (A). Expression in xylem relative to phloem ranged from about 5- to 30-fold across all trees and sampling dates.

Expression of the UDP-glucose pyrophosphorylase (UGP) gene, PtaUGP1, was relatively high in vascular tissues, similar to that of the loblolly pine CesA, Kor and SuSy genes (Figure 2). The UGP transcript abundance was higher in phloem than in corresponding xylem samples. Difference in expression averaged across all trees and sampling times was about 2-fold higher in phloem than in corresponding xylem samples. However, relative expression between the two tissue types was not consistent for all individuals or sampling times. Several of

![Figure 2. Transcript analysis of genes involved in cellulose synthesis from xylem and phloem of four pine trees throughout the growing season. Gene-specific primer pairs for PtaCesA1 (CesA1), PtaCesA2 (CesA2), PtaCesA3 (CesA3), PtaKor1 (Kor1), PtaSuSy1 (SuSy1) and PtaUGP1 (UGP1) were used to determine relative transcript abundance from xylem (solid bars) and phloem (open bars) from four pine trees (Panels A–D) on four dates during the 2006 growing season: April (T1), June (T2), July (T3) and August (T4). Transcript abundance data are grouped by the gene assayed. Transcript analysis of indicated genes in mRNA isolated from xylem and phloem tissues were performed as follows: mean Ct values from triplicate samples obtained for each gene were converted to linear values and normalized to the endogenous control gene PtaAct2. The relative transcript abundance for genes assayed is plotted on a log scale.](image)
the comparisons showed less than 2-fold differences, and in one sample, expression was slightly higher in xylem than in phloem. Tree D exhibited consistently higher expression in phloem than in xylem on all dates, ranging from about 2-fold to 4-fold.

Two cellulose synthase-like (Csl) genes, \textit{PtaCslA1} and \textit{PtaCslA2}, containing full-length coding regions were previously obtained from the loblolly pine xylem cDNA library. The gene \textit{PtaCslA1} was shown to encode an enzyme with mannan and glucomannan synthase activity (Leipman et al. 2007). Expression profiling by qRT-PCR indicated that the genes had different, but overlapping expression patterns in xylem and phloem tissues (Figure 3). Expression of \textit{PtaCslA1} were more similar in xylem and phloem tissues than that of \textit{PtaCslA2}. Transcription of \textit{PtaCslA1} was similar to or up to 2.5-fold higher in xylem than in phloem. The \textit{PtaCslA2} gene exhibited 4-fold to 30-fold greater expression in xylem than in corresponding phloem samples.

The \textit{PtaGMP1} and \textit{PtaGMP2} genes were expressed in xylem and phloem tissues of loblolly pine (Figure 3). Transcription of these genes differed between tissues, with \textit{PtaGMP2} expression being 5- to 10-fold higher than that of \textit{PtaGMP1}. Expression of both genes was greater in xylem than in phloem, except on two dates (T2 and T4) in Tree B. All trees exhibited 2- to 4.5-fold greater expression of \textit{PtaGMP1} in xylem than in corresponding phloem samples at T1 (late April), but patterns of expression were not consistent between individuals on dates T2 through T4.

The \textit{PtaPAL1} gene was included in the analysis as a reference because it is expressed during secondary cell wall synthesis and lignification of developing xylem. The \textit{PtaPAL1} gene was expressed in xylem at levels similar to those of the three \textit{PtaCesA} genes, whereas expression in phloem was higher than that of the \textit{CesA} genes (Figure 3). Expression in xylem ranged from 2- to 25-fold higher than in corresponding phloem samples. In each of the four trees sampled, gene expression was highest in the xylem sample collected in July (T3).

**Discussion**

Identification and functional characterization of genes involved in growth and development of loblolly pine will greatly facilitate our ability to improve growth and wood properties. Loblolly pine EST collections enable comparative genomic research in conifers (Allona et al. 1998, Cairney et al. 2006, Lorenz et al. 2006). Such approaches can utilize progress in plant developmental studies from model species to identify genes of interest in loblolly pine and other plants of commercial importance. Computational approaches have demon-
strated the potential for using loblolly pine EST collections to identify candidate genes for functional analyses (Pavy et al. 2005).

Genes involved in carbohydrate metabolism and polysaccharide assembly are not as well characterized as genes encoding enzymes of the lignin biosynthetic pathway. We investigated carbohydrate-related genes that are involved in the synthesis of cellulose and hemicellulose as well as those implicated in carbon partitioning to cell wall biosynthesis. Previous work in our laboratory has identified CesA genes and CslA genes from loblolly pine that are homologs of those found in angiosperm genomes (Nairn and Haselkorn 2005, Liepman et al. 2007). Here we report the identification of genes encoding a loblolly pine KORRIGAN endoglucanase, a sucrose synthase, a UDP-glucose pyrophosphorylase and two GDP-mannose pyrophosphorylases. These genes are homologous to genes described in angiosperm taxa and are implicated in the synthesis of cellulosic and hemicellulosic polysaccharides.

Different groups of three CesA genes are involved in cellulosic polysaccharide assembly of the primary and secondary cell wall of angiosperms. Each of the three CesA genes from loblolly pine is orthologous to one of the three secondary cell wall CesA genes in angiosperms, and each is highly expressed in developing xylem. Phylogenies inferred from comparative sequence analysis indicated that duplication of CesA genes and gene-specific function in cellulose synthesis of the primary and secondary cell wall evolved before the divergence of gymnosperms and angiosperms (Nairn and Haselkorn 2005).

Expression patterns of PtaCesA1, PtaCesA2 and PtaCesA3 in xylem and phloem of loblolly pine were consistent with a role in cellulose synthesis of the secondary cell wall (Figure 2). Overall, PtaCesA2 was expressed at higher levels than PtaCesA1 and PtaCesA3, which exhibited similar levels of expression. Transcription of each of the CesA genes in xylem tissues undergoing high rates of secondary cell wall synthesis ranged from about 10- to more than 100-fold higher than in the corresponding phloem tissue samples. Expression levels in xylem were similar in the four trees (A-D) sampled on four dates (T1–T4), and with individual trees representing two size/age classes. The two trees in the 70–75 cm stem diameter class were located on a different site from the two trees in the 40–45 cm class. Expression in phloem samples at T4 (mid-August) ranged from about 5- to 10-fold higher in the larger size class individuals than in the smaller size class individuals for each of the three PtaCesA genes. The factor or factors influencing these differences in expression are unknown; however, they could include differences in transcriptional regulation at different sizes or ages in loblolly pine, environmental factors such as nutrient and water availability, or differences in genotype.

The KORRIGAN (KOR) gene is required for normal cellulose synthesis in the primary and secondary cell wall of angiosperm taxa and encodes a membrane-bound EGase (Brummell et al. 1997, Nicol et al. 1998). The EGases containing a membrane-spanning domain are part of a subfamily within the larger EGase gene family (Libertini et al. 2004). Characterization of Arabidopsis plants with other mutant alleles of the KOR locus demonstrated that mutations cause defects in cell plate formation, reductions in cellulose content of the primary cell wall, and severe defects in cellulose synthesis of the secondary cell wall (Szanownicz et al. 2004). However, the precise role of the KOR protein in cellulose synthesis remains unknown. Hypothetical functions include cleavage of the cellulose microfibrils to terminate synthesis of the cellulose chain, removal of noncrystalline glucan chains, or release of tensional stress presumably caused by assembly of the long-chain molecules (Szanownicz et al. 2004, Somerville 2006). No direct evidence or consensus has been reached as to whether KOR is a part of the cellulose synthesis complex containing the CesA catalytic subunits. Some investigators speculate that it is associated with the complex based on functional evidence (Somerville 2006) whereas others disagree because of the inability to co-purify the protein with the CesA catalytic subunits (Szanownicz et al. 2004).

The PtaKor1 gene encodes a full-length protein that shares 80% sequence similarity with that of Arabidopsis KOR and contains the membrane-spanning region unique to the γ-EGase subfamily. A phylogeny inferred from alignments of KOR protein sequences indicates that PtaKor1 is orthologous to the KOR genes reported from Arabidopsis, rice and Populus (Figure 1). Two groups were present within the clade. The terminal clade was monophyletic and contained two sequences from each of the Arabidopsis and rice genomes, including the Arabidopsis KOR sequence. This group also included the KOR sequence from Populus tremuloides and PtaKor1 from loblolly pine. One sequence each from Arabidopsis and rice formed a separate group basal to the KOR clade. Two EGases from Pinus radiata, which lack the membrane-spanning domain, are grouped with a different clade that contains other Arabidopsis and rice EGase sequences suggesting that evolution of the gene family preceded the divergence of the gymnosperm and angiosperm lineages.

The KOR transcripts are present in elongating cotton fibers and increase in cotton fibers undergoing secondary wall cellulose synthesis (Peng et al. 2002). Expression in Arabidopsis leaves is low compared with that in developing stem segments (Szanownicz et al. 2004), and a KOR homolog in Populus is up-regulated in tension wood where cellulose synthesis is rapid (Bhandari et al. 2006). These studies suggest that KOR expression may be related to rates of cellulose synthesis in cell wall biogenesis. Studies of KOR homologs in angiosperm taxa suggest a requirement for the KOR EGase in cellulose synthesis, and expression patterns suggest a role in primary and secondary cell wall synthesis.

The PtaKor1 gene is expressed at high levels in xylem and phloem tissues of loblolly pine (Figure 2). In the majority of samples, expression was 2- to 5-fold higher in the xylem sample relative to the corresponding phloem sample. These differences were more consistently observed in the two smaller trees. The large differences observed between xylem and phloem for the PtaKor1 genes were not seen in PtaKor1 expression patterns. Different expression patterns between the PtaCesA genes and the PtaKor1 gene may be explained by specificity for cellulose synthesis in the secondary cell wall for
the former and a requirement for the latter in cellulose synthesis in primary and secondary cell walls (Nicol et al. 1998, Szyjanowicz et al. 2004). Evolutionary relationships between the PtaKor1 gene and expression patterns in loblolly pine tissues are consistent with a conserved function for KOR in cellulose synthesis during cell wall formation in gymnosperms.

The enzymes SuSy and UGP catalyze formation of UDP-Glc, the immediate substrate for cellulose synthesis. The dominant paradigm is that SuSy is the key enzyme involved in channeling UDP-Glc to cellulose synthesis by hydrolysis of sucrose. The SuSy enzyme is localized in the cytosol and is associated with the actin cytoskeleton or the plasma membrane (Amor et al. 1995, Haigler et al. 2001) or both. It has been localized at sites of cellulose synthesis in several studies (Salnikov et al. 2001, Albrecht and Mustroph 2003). The SuSy enzyme is post-translationally modified by a reversible phosphorylation mechanism, which may be involved in regulating localization of the enzyme (Winter and Huber 2000). Treatment of tomato fruit with kinase inhibitors resulted in an increase in membrane localization of SuSy, presumably due to a reduction in phosphorylation of the SuSy enzyme (Anguenot et al. 2006). Evidence that SuSy is involved in channeling UDP-Glc to sites of cell wall polysaccharide synthesis was shown from studies in transgenic Populus overexpressing a SuSy gene. In this system, labeled sucrose loaded into the phloem of the leaf was channeled to cellulose and xyloglucan in the stem (Konishi et al. 2004).

The PtaSuSy1 gene encodes an 833 amino acid protein that shares 64–79% sequence similarity with SuSy proteins from Arabidopsis. Efforts to obtain a full-length cDNA clone for a second putative SuSy gene were unsuccessful. This may be due to low representation of the transcript in xylem tissues because the corresponding contig contains only one EST from wood with the remaining sequences coming from root tissues. The phylogeny inferred from alignments of SuSy protein contains three well-supported clades containing monocot and dicot sequences (Figure 1). PtaSuSy1 groups with the clade containing AtSus2 and AtSus3. This grouping is supported in 70% of bootstrap pseudoreplicates. The loblolly pine SuSy sequence is more closely related to these and AtSus4 than to AtSus5 and AtSus6. Additional SuSy genes are present in the loblolly pine genome, as inferred by EST sequences, but no other full-length sequences are available at present. The presence of three orthologous groups in Arabidopsis and rice indicates that evolution of the SuSy gene family preceded the divergence of monocot and dicot lineages (Figure 1). PtaSuSy1 transcripts were more abundant than those of angiosperms was poorly resolved in the analysis.

The PtaUGP1 gene is expressed at similar levels in xylem and phloem (Figure 2). Patterns of expression in xylem and corresponding phloem samples were different from those for PtaSuSy1 and the PtaCesA genes. Transcription in xylem was within the range observed for PtaSuSy1 and PtaCesA genes, but expression in phloem was considerably higher and more similar to that observed for PtaKor1. PtaUGP1 was the only gene examined in this study for which expression was generally higher in phloem than in corresponding xylem samples. Differences were up to 5-fold. PtaSuSy1 transcripts were more abundant than those of PtaUGP1 and differences ranged from 2- to 7-fold higher in most tissue samples. In Populus, expression profiling has shown that specific genes encoding SuSy and UGP are up-regulated in the zone of secondary cell wall synthesis in developing wood and that SuSy transcripts are more abundant than UGP transcripts (Hertzberg et al. 2001, Meng et al. 2007). Patterns of expression for SuSy and UGP genes in developing wood of loblolly pine are generally consistent with those from Populus, although our current analysis does not resolve differences in expression between the different zones of development.

The potential for manipulating SuSy and UGP genes for modification of plant growth has recently been demonstrated (Coleman et al. 2006). Transgenic tobacco (Nicotiana tabacum) plants expressing a cotton SuSy gene or a bacterial UGP gene had a higher soluble carbohydrates content and a greater total biomass. The expression of both genes had an additive effect. The cellulose content relative to total mass showed little increase; however, the increase in biomass produced increased cellulose yields from individual plants.

Hemicelluloses, including glucomannans and xyloglucans, are major components of the plant secondary cell wall. Glucomannans are more abundant in the secondary cell wall of loblolly pine and other softwoods, whereas xyloglucans are more abundant in angiosperms. Glucomannans are undesir-
able in wood fiber for pulping, but may be beneficial for production of biofuels (Suzuki et al. 2006). CslA genes from a broad range of land plant taxa encode enzymes with mannan and glucomannan synthase activities (Liepman et al. 2007).

Two CslA genes were identified in loblolly pine. Although the genome of a pine has not been sequenced, to our knowledge there is currently no evidence for additional CslA genes in loblolly pine. *PtaCslA1* and *PtaCslA2* were expressed at high levels in xylem and transcript levels were generally similar or higher for *PtaCslA1* than for *PtaCslA2* (Figure 3). Patterns of expression for *PtaCslA2* in xylem and phloem were similar to those observed for the *PtaCesA* genes, *PtaSuSy1* and *PtaPAL1*. The smaller trees (C and D) showed greater differences between xylem and phloem expression levels than those of the larger trees (A and B). Overall, the differences in expression patterns between xylem and phloem for *PtaCslA1* and *PtaCslA2* reflected lower expression of *PtcslA2* in phloem, whereas transcription in xylem was more similar for the two *PtaCslA* genes. A previous study has shown that expression of three of the five CslA genes in *Populus* is greater in xylem than in other tissues (Suzuki et al. 2006). These observations are consistent with conserved functions for CslA genes in hemicellullose synthesis of gymnosperms and angiosperms.

The GMP enzyme catalyzes the formation of GDP-Man from mannose-1-phosphate. GDP-Mannose is a substrate for assembly of glucomanann hemicelluloses in the cell walls of woody forming tissues. Two genes from loblolly pine encoding GMP, *PtaGMP1* and *PtaGMP2*, share 80 and 90% sequence similarity at the DNA and protein levels, respectively. The derived protein sequences are 74–92% similar to GMP sequences from angiosperm taxa. Comparisons of the loblolly pine protein sequences with those from angiosperm taxa and a phylogeny inferred from protein sequence alignments indicate that the two loblolly pine genes are more closely related to each other than to genes from other taxa.

*PtaGMP1* and *PtaGMP2* transcripts are abundant in xylem and phloem tissues of loblolly pine (Figure 3). *PtaGMP2* transcripts are more highly represented than *PtaGMP1* in both tissues. *PtaGMP2* expression is 5- to 9-fold greater than that of *PtaGMP1* in xylem samples and from 4.5- to 22-fold higher in phloem samples.

Both *PtaCslA* and *PtaGMP* genes were expressed at relatively high levels in xylem and phloem tissues of loblolly pine with a pattern of elevated transcription in xylem relative to corresponding phloem samples for *PtaCslA2* and in three of the four sampled trees for *PtaGMP1* (Figure 3). These data were consistent with roles for CslA and GMP in hemicellulose synthesis in the secondary cell wall of xylem.

In addition to performing a structural function in the cell wall, mannans are implicated in a variety of essential functions for plant growth and viability. Mutations in *CslA* genes of *Arabidopsis* exhibit pleiotropic effects, including aberrant pollen tube growth (Liepman et al. 2007). A severe mutation in the CYT1 locus of *Arabidopsis* encoding a GMP, is lethal during embryonic development (Lukowicz et al. 2001). Plants with weaker mutations at this locus were deficient in N-glycosylation and had altered cell wall polysaccharide composition, including a 5-fold decrease in cellulose. *PtaCslA1* and *PtaCslA2* were represented in EST collections from developing embryos of loblolly pine. Together, these data are consistent with roles for mannan and mannan-containing polysaccharides in a variety of plant processes in gymnosperms and angiosperms, including cell wall biogenesis, storage and embryonic development.

The genome of loblolly pine contains homologs of genes from angiosperms that are important in the synthesis of plant cell wall polysaccharides suggesting functional conservation of these genes between gymnosperms and angiosperms. These observations are consistent with an analysis demonstrating that a high percentage of genes in loblolly pine are putative homologs of genes in the herbaceous angiosperm *Arabidopsis thaliana* (Kirst et al. 2003). Expression profiles of the loblolly pine genes in xylem and phloem are consistent with functions in wood formation of gymnosperms. The information obtained in our analysis validates these genes as candidates for further functional studies that will more precisely define their role in wood formation of loblolly pine and as potential targets for genetic manipulation.

Acknowledgments

We thank Michael Marsh and Jay Brown for assistance in obtaining plant materials. This work was supported by grants from the USDA Forest Service (SRS05CA11330126240) and the Georgia Research Alliance.

References


