Tree Genetics & Genomes (2012) 8:775–788
DOI 10.1007/s11295-011-0462-7

ORIGINAL PAPER

Correlations of expression of cell wall biosynthesis genes with variation in biomass composition in shrub willow (*Salix* spp.) biomass crops

Michelle J. Serapiglia · Kimberly D. Cameron · Arthur J. Stipanovic · Lawrence B. Smart

Received: 13 July 2011 / Revised: 5 December 2011 / Accepted: 12 December 2011 / Published online: 30 December 2011

Abstract We have measured significant genetically determined variation in biomass composition among breeding populations of shrub willow, a biomass feedstock crop. This project was aimed to ask whether patterns of cell wall gene expression can be correlated with genetic variation in biomass composition at harvest, in order to develop assays of early differences in gene expression as indicators of harvestable biomass chemical composition and potentially reduce the time of selection for new willow genotypes. Previous studies have demonstrated that manipulation of expression of cell wall biosynthetic genes results in altered biomass chemical composition. We analyzed genes encoding enzymes involved in lignin biosynthesis and carbohydrate active enzymes selected based on their functional characterization and conservation in *Populus trichocarpa* and *Arabidopsis thaliana*. Fragments of 20 genes were cloned from young stem cDNA of *Salix sachalinensis* and *Salix miyabeana*. Expression profiling in willow stem apical tissue and developing stem tissue was performed for each isolated gene using probe-based quantitative real-time PCR. Two willow parental genotypes and six progeny within a hybrid family were selected for analysis, and significant differences in expression among the individuals and between tissue types were observed for most of the genes. Significant correlations between patterns of gene expression and variation in the biomass chemical composition of those genotypes provide insight into the genetic regulation of lignocellulosic deposition in this important bioenergy crop and could be utilized as a tool for early selection of new genotypes.

Keywords Caffeic acid *O*-methyltransferase · Caffeoyl CoA *O*-methyltransferase · Cellulose synthase · Cinnamyl alcohol dehydrogenase · Glycosyltransferase · Wood composition

Communicated by A. Brunner

Electronic supplementary material The online version of this article (doi:10.1007/s11295-011-0462-7) contains supplementary material, which is available to authorized users.

M. J. Serapiglia · K. D. Cameron · L. B. Smart (✉)
Department of Horticulture, Cornell University,
New York State Agricultural Experiment Station,
630 West North Street,
Geneva, NY 14456, USA
e-mail: lbs33@cornell.edu

A. J. Stipanovic
Department of Chemistry,
State University of New York College
of Environmental Science and Forestry,
1 Forestry Drive,
Syracuse, NY 13210, USA

Introduction

For cellulosic feedstock crops, biomass chemistry is largely a reflection of cell wall composition, which is determined by the pathways associated with the biosynthesis of cellulose, hemicellulose, pectin, and lignin. Since the cell wall is an essential component for plant development, defense, and structural integrity, it has evolved to be highly recalcitrant to most biochemical deconstruction processes necessary for obtaining the simple component molecules that can be efficiently converted to biofuels and bioproducts (Himmel et al. 2007). Whether conversion processes involve biochemical means, such as fermentation into alcohols, or thermochemical conversion into heat and power, feedstock chemistry will play a role in the efficiency of the conversion process. Biochemical conversion methods exploit the polysaccharide component of the biomass, therefore increasing sugar content and yield by reducing cross-linkages with lignin that will reduce pretreatment costs. On the contrary, lignin content in biomass can be

highly beneficial in the thermochemical conversion process by adding to the heat value of the biomass and contributing to the quality of products produced through gasification or pyrolysis. The development and genetic improvement of biomass chemical composition of perennial woody bioenergy crops, such as those in the family Salicaceae, poplar (*Populus* spp.), and fast-growing shrub willow (*Salix* spp.), will be expedited by a thorough understanding of the genes involved in cell wall biosynthesis and the molecular basis for their regulation leading to variations in biomass chemical composition. In order to identify candidate genes and develop molecular tools for early selection in breeding of new varieties of shrub willow with biomass chemical composition optimized for conversion, it is important to understand the molecular regulation of cell wall biosynthesis, including key biosynthetic genes.

Over 1,600 genes encoding carbohydrate active enzymes (CAZymes, Coutinho et al. 2003) and more than 34 genes involved in phenylpropanoid metabolism and lignin biosynthesis have been identified in the *Populus trichocarpa* (Torr. & Gray) genome, highlighting the complexity of cell wall biosynthesis and the concomitant annual development of secondary xylem in perennial species (Boerjan et al. 2003; Geisler-Lee et al. 2006; Tuskan et al. 2006). Phenylpropenoid metabolism and monolignol biosynthesis have been well studied in both *Populus* ssp. and *Arabidopsis thaliana* (Campbell and Sederoff 1996; Boerjan et al. 2003; Vanholme et al. 2008). In addition, a large effort has been devoted to altering lignin content in woody species, such as poplar, through genetic engineering (Dinus et al. 2001; Boerjan 2005). Reverse genetics and the use of mutants in *A. thaliana* have been very successful in elucidating the pathways involved in the biosynthesis of the cell wall and the key CAZymes involved with polysaccharide synthesis (Joshi et al. 2004; Orfila et al. 2005; Atanassov et al. 2009; Brown et al. 2009). Twenty-five xylem-specific glycosyltransferases (GT) belonging to the CAZyme families GT2, GT8, GT14, GT31, GT43, GT47, and GT61 were successfully identified using microarrays designed from tissue-specific libraries of *Populus tremula* × *Populus tremuloides* (Aspeborg et al. 2005). Genes from these families have also been identified as key genes in cell wall biosynthesis in *Populus trichocarpa* and *A. thaliana* (Aspeborg et al. 2005; Geisler-Lee et al. 2006; Dharmawardhana et al. 2010).

Since neither a method of genetic transformation and regeneration nor a sequenced genome is available for shrub willow, traditional breeding strategies for genetic improvement rely on selection for phenotypic variation in cell wall composition. This natural variation can be used as a means to identify candidate genes regulating cell wall structure and chemistry through linkage mapping, candidate gene analysis, or association mapping (González-Martínez et al. 2007; Ranjan et al. 2010; Thumma et al. 2010). However, there have been very few direct correlations of candidate gene expression profiles with variation in cell wall composition among natural or bred populations of *Populus* (Voelker et al. 2010; Wegrzyn et al. 2010).

In this study, we have focused on patterns of mRNA accumulation of structural genes that are likely to have an impact on cell wall compositional changes. The cellulose synthase (EC 2.4.1.12; CesA) gene family and a small number of gene family members associated with noncellulosic polysaccharides, including UDP-glucose dehydrogenase (EC 1.1.1.22; UDP-GD), were chosen for analysis in this study. In addition, three major lignin biosynthetic genes were chosen for analysis in this study coding for the following enzymes, caffeoyl CoA *O*-methyltransferase (EC 2.1.1.104; CCoAOMT), caffeic acid *O*-methyltransferase (EC 2.1.1.68; COMT), and cinnamyl alcohol dehydrogenase (EC 1.1.1.195; CAD). Downregulation of *CCoAOMT* has resulted in a 12% to over 50% reduction in lignin content and a reduction in the S:G ratio (Meyermans et al. 2000; Zhong et al. 2000; Baucher et al. 2003). Declines in the S:G ratio have also been observed in transgenic poplar with downregulated *COMT* expression, in addition to the indicative reddish-brown coloration from the incorporation of coniferaldehyde (Lapierre et al. 1999; Jouanin et al. 2000; Pilate et al. 2002; Baucher et al. 2003). The decline in the S:G ratio and the incorporation of coniferaldehyde have also been observed in a wide range of plant species with downregulated *CAD* expression with reduced lignin content (Baucher et al. 1996; Ralph et al. 1997; Lapierre et al. 1999; Dauwe et al. 2007). Since discrepancies still remain as to whether or not *COMT* and *CAD* suppression results in total lignin reductions, they were of particular interest in the present study.

The goal of this study was to determine whether the patterns of expression of key cell wall biosynthetic genes in developing willow shoots are correlated with variation in the biomass chemical composition of harvest age woody stems. This would allow for the early selection of new willow genotypes with differences in biomass chemical composition, based on early expression data and not composition determined from 3-year-old tissue. This was also conducted as a pilot study for future candidate gene association genetics and potential eQTL mapping, similar in approach to Hu et al. (1999), Wagner et al. (2009), and Voelker et al. (2010), where RNA expression of *4-coumarate:coenzyme A ligase* in young stem tissue was correlated to total lignin content in mature wood. Expression analysis was accomplished through probe-based quantitative real-time PCR assays of transcript abundance of specific key genes involved in cellulose, hemicellulose, and lignin biosynthesis among sibling members of an F_1 hybrid family of *Salix sachalinensis* × *Salix miyabeana*. Finally, correlations were made between gene expression and variation in biomass chemical composition assayed by high-resolution thermogravimetric analysis

(HR-TGA). Previous TGA analyses of the biomass chemical composition of selected high-yielding shrub willow varieties identified compositional differences among clones of different species and within hybrid families (Serapiglia et al. 2008, 2009). Thermogravimetric analysis has been extensively used to individually resolve the three woody biomass components, cellulose, hemicellulose, and lignin in a range of woody species and tissues, and the thermal decomposition of these three components at specific temperature ranges is widely accepted (Shafizadeh and Chin 1977; Raveendran et al. 1996; Orfao et al. 1999; Mészáros et al. 2004; Stipanovic et al. 2004; Yang et al. 2006, 2007; Wang et al. 2008; Singh et al. 2009).

Experimental procedures

Source material and tissue collection

A total of eight genotypes (two parental individuals and six hybrid progeny) were chosen from the family designated as 9970 (*S. sachalinensis* × *S. miyabeana*). Five of the progeny were chosen at random, but genotype 9970-036 was selected for analysis, because it is a commercial cultivar, 'Canastota,' that is being grown as a biofuels feedstock. For biomass chemical compositional analysis, willow stem biomass samples from the two parental genotypes and the six hybrid progeny genotypes were harvested in January 2006 from the Tully Genetics Field Station (Tully, NY) after the third growing season after first coppice. Three replicate plants for each genotype were selected for analysis, with the exception of 9970-036, of which only one replicate was available. All samples were collected as follows: 15-cm sections including bark were cut from the base, middle, and top of one representative canopy stem. These stem sections were dried to constant weight at 65°C and then ground in a Wiley mill with a 20-mesh screen. The ground material from the three sections of each stem was pooled and homogenized to provide a representative sample for the composition of bulk biomass at harvest age.

For molecular analysis of the family, the eight genotypes (two parents and six offspring) from the 9970 family were planted in May 2007 at the SUNY-ESF Tully Genetics Field Station in Tully, NY in a randomized complete block design with four blocks and four plants per plot. Shoot apex and young stem tissue were collected from the top 12 cm of the stem for gene expression studies in May of 2008 during the beginning of the growing season (following the first coppice). Three stems from one plant from each replicated block were sampled for a total of three biological replicates per genotype. The apical shoot portion consisted of the shoot apex, leaf primordia, young folded leaves, and stem tissue down to the first unfolded leaf as a representation of primary cell wall development (Fig. 1). The young developing stem tissue consisted of the first seven excised internodes below the first unfolded leaf (Fig. 1). Preliminary histochemical staining of stem cross-sections with phloroglucinol to HCl has indicated that secondary cell walls and vascular cambium develop and secondary growth occurs in the seven internodes excised for gene expression profiling (Supplemental Fig. S1). Both tissue types were immediately frozen in liquid nitrogen before grinding for RNA extractions.

Biomass chemical compositional analysis

All willow samples were analyzed using a Thermogravimetric Analyzer 2950 (TA Instruments, New Castle, DE) with the TA Universal Analysis 2000 software according to Serapiglia et al. (2009). The method used was "high-resolution dynamic" with a heating rate of 20°C min^{-1}, a final temperature of 600°C, a resolution of 4.0, and a sensitivity value of 1.0. The electrobalance was purged with nitrogen at a flow rate of 44 mL min^{-1}, and the furnace was purged with compressed air with a flow rate of 66 mL min^{-1}. For each analysis, 10 mg of dry tissue was used. The percent dry weight for each stem biomass component, hemicellulose, cellulose, and lignin was determined using the method developed by Serapiglia et al. (2009). The initial dry weight was corrected for moisture content. Each biological replicate was analyzed with three instrumental replicates. HR-TGA was selected as an analytical tool in this study since it

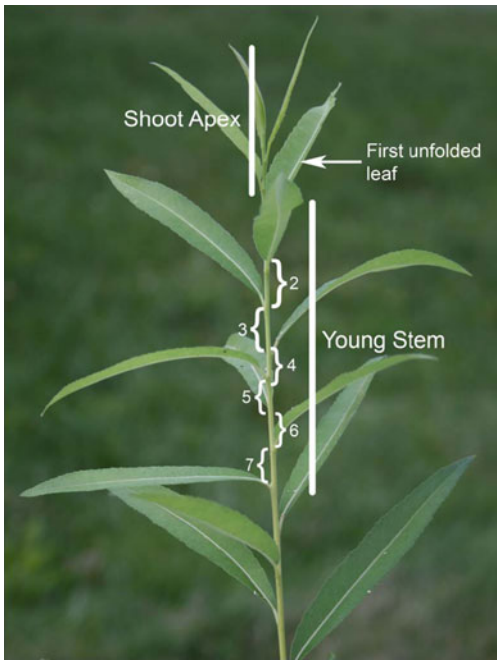

Fig. 1 Shoot of *S. sachalinensis* × *S. miyabeana* 9970-036 indicating locations of shoot apex tissue and young stem tissue collected for RNA analysis

represents a "high throughput" technique (1–2 h/sample; 16 samples per day) that is very repeatable and is sensitive to compositional differences in woody samples.

RNA extractions and cDNA synthesis

RNA was extracted from the shoot apex and young stem tissue using a modified hot borate extraction method (Wan and Wilkins 1994). The synthesis of first-strand cDNA was performed using the Superscript® III First Strand Synthesis System (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol using random hexamers for amplification. A total of 8 μg of DNase treated RNA was used for each reaction.

Cloning of partial cDNAs

Degenerate primers were designed for the amplification of gene fragments for the following genes: *COMT1*, *COMT2*, *CCoAOMT1*, *CCoAOMT2*, *CAD4*, *UDP-GD*, and representatives of specific GT families *CesA*, *GT47A*, *GT47B*, *GT47C*, and *GT8* (Supplemental Table S1). All PCR reactions were performed using a iCycler iQ™ thermocycler (Bio-Rad Laboratories, Hercules, CA). The reaction conditions for all gene fragment amplifications except *CesA* were as follows: 5 min at 95°C; 35 cycles of 45 s at 95°C, 45 s at melting temperature indicated in Supplemental Table S1, and 60 s at 72°C; 7 min at 72°C; and ending with a 4°C hold. Amplifications of *CesA* gene fragments were performed using a reaction protocol taken from Liang and Joshi (2004) and were as follows: 10 min at 95°C; 2 cycles of 60 s at 95°C, 60 s at 41°C, and 2 min at 72°C; 35 cycles of 60 s at 95°C, 90 s at 55°C, and 2 min at 72°C; 7 min at 72°C; and ending with a 4°C hold.

All PCR fragments were cloned into the pCR 2.1-TOPO® vector and were used to transform One Shot® TOP10F' chemically competent *Escherichia coli* according to the manufacturer's protocol (Invitrogen, Carlsbad, CA). Plasmid DNA was isolated from screened colonies using alkaline lysis followed by silica guanidine thiocyanate miniprep purification (Carter and Milton 1993). Plasmids from approximately 20 colonies for each gene amplified from both *S. sachalinensis* and *S. miyabeana* cDNA were sequenced to identify highly expressed genes, except for plasmids with *CesA* inserts, for which plasmids from over 100 colonies were sequenced to identify multiple *CesA* genes.

Sequence analysis

All partial cDNAs for each gene of interest were analyzed with BioEdit 7.0 (Hall 1999). BLASTN searches were performed at the NCBI server (http://ncbi.nlm.nih.gov/BLAST/) (Zhang et al. 2000) and in both version 1 (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html) and version 2 (http://www.phytozome.net/poplar.php) of the *P. trichocarpa* genome. ClustalW multiple alignments and percent identity scores were performed with BioEdit 7.0 (Hall 1999). Annotation of all lignin biosynthetic genes studied was based on the *P. trichocarpa* version 1 and version 2 and homology to annotated *A. thaliana* genes. Annotation for all glycosyltransferase genes was based on homology to genes identified in Aspeborg et al. (2005) and Geisler-Lee et al. (2006). Additional annotation for all cellulose synthase genes was performed based on *P. trichocarpa* genome assembly gene model homology to sequences in Kumar et al. (2009). All partial genes isolated were submitted to the NCBI server, and accession numbers are provided in Supplemental Table S2.

Expression analysis by probe-based real-time quantitative PCR

Primers and dual-labeled probes were designed according to corresponding sequences identified in *S. sachalinensis* and *S. miyabeana*. PCR amplification was done in a volume of 50 μL and contained 1 μg of either cDNA template or the plasmid standard in 1× IQ Supermix (Bio-Rad Laboratories), and corresponding probe and primer pairs with concentrations indicated in Supplemental Table S3. Expression assays were optimized for probe and primer concentrations and for all assays standard curves were fit to a linear model with R^2 values above 0.90, and PCR efficiencies ranged from 90% to 120%. Amplification specificity of the primers was verified by melt–curve analysis. *Actin* was used as the reference gene (Brunner et al. 2004). All real-time PCR reactions were performed using an iCycler iQ™ (Bio-Rad Laboratories). After the collection of well factors, the reaction protocol was as follows: 3 min at 95°C, and 45 cycles of 10 s at 95°C and 60 s at 60°C. Quantification of gene expression was done with the standard curve method based on the mean of three experimental replicates. Starting quantities of each gene of interest were normalized relative to the starting quantity of *Actin* by dividing the starting quantity of the gene by that of *Actin*. The starting quantity values normalized to *Actin* expression were further normalized to values of 'SX61' expression in the apex tissue in a similar manner.

Statistical analysis

Statistical analyses were performed using SAS® version 9.1.2 at a critical α=0.05 (SAS Institute Inc.). A two-way ANOVA was performed using PROC GLM to analyze significant differences among biomass chemical composition and gene expression profiles among genotypes for experimental replicate and variety. Each tissue type was analyzed independently of each other. The variance components for the total data set, between and within varieties, and within

instrumental replicate were estimated with nested random effects by PROC NESTED. A two-way ANOVA was performed using PROC GLM to analyze significant effects by replicate and tissue type. When a significant interaction ($P<0.05$) was observed, Tukey's mean studentized range test was used for pairwise comparisons among genotypes. Pearson correlations were performed using PROC CORR to identify any significant correlations among gene expression patterns and cell wall composition, using the mean values for both composition and gene expression.

Results and discussion

The ability to generate and select for variation in biomass chemical composition through breeding will allow for identification of genotypes best suited for different conversion technologies. The 9970 genetic field study, consisting of the two parental genotypes, *S. sachalinensis* and *S. miyabeana* and their progeny, was established with the intention of characterizing gene expression patterns associated with cell wall biosynthesis and relating those with biomass chemical composition in different progeny genotypes. Both parental genotypes and hybrid progeny were used in the study of expression profiles of the genes identified in this research, with the result that unique expression profiles and differences in expression between the progeny and the parents were observed. Variation in gene expression could potentially be a good indicator of the underlying additive genotypic variation associated with a trait, in this case biomass chemical composition (Kirst et al. 2005; Voelker et al. 2010). For this particular interspecific family, the variation in expression observed within the family may be due to recombination within each parent, as well as the allelic, epistatic, and environmental interactions among the progeny genotypes. Based on flow cytometry analysis of nuclear DNA content, *S. sachalinensis*, *S. miyabeana*, and the progeny genotype '9970-036' have been identified as tetraploid (Cameron et al. 2010). There could be up to four alleles for each gene present within each parental genome, not including previously duplicated genes. It is not known how many alleles are contributing to the expression profiles observed in this study, nor the alleles specific to each parental genotype and how they may have been recombined within the progeny, but this study does verify that there are quantifiable differences.

Biomass chemical composition of selected members of the 9970 family

Harvest age (3-year-old) stem biomass from the two parental genotypes, *S. sachalinensis* 'SX61' and *S. miyabeana* 'SX64', and from six progeny of the *S. sachalinensis* 'SX61' × *S. miyabeana* 'SX64' family 9970 was analyzed using HR-TGA, according to Serapiglia et al. (2009). Significant differences in cellulose, hemicellulose, and lignin content were identified among all eight genotypes (Fig. 2), but between the parental genotypes, only hemicellulose content differed significantly. Genotype 9970-014 had high cellulose content and was the only progeny genotype with cellulose content that was significantly different from that of one of the parents (*S. miyabeana* 'SX64'). Genotype 9970-052 had high hemicellulose content, which was significantly different from that of both parents. Lignin content in 9970-014 was the only genotype significantly lower than that in the parents. Analysis of the total variance in the data set

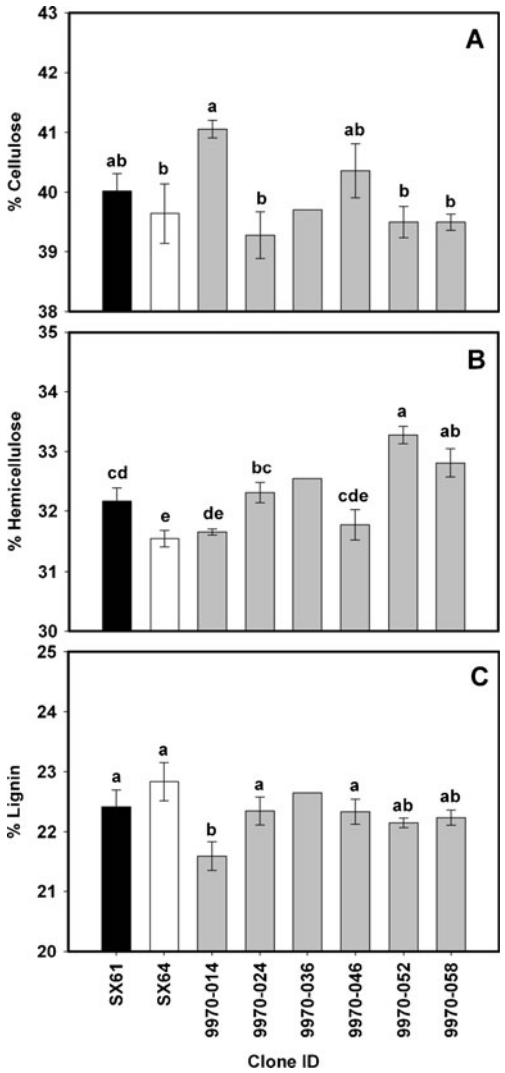

Fig. 2 Willow composition as percentage dry weight as determined by HR-TGA for *S. sachalinensis* × *S. miyabeana* 9970 family. *Bars* indicate the mean and standard error among three biological replicates ($n=3$). Parental genotypes are *S. sachalinensis* SX61 and *S. miyabeana* SX64, *black* and *white*, respectively. **a** Cellulose, **b** hemicellulose, and **c** lignin

showed that over 70% of the total variance was due to variation between the varieties for each of the three biomass components, and only a small amount of variance was due to experimental and instrumental replication.

Cloning and expression profiles of selected genes

Partial-length cDNAs from genes coding for enzymes involved with lignin, cellulose, and noncellulosic polysaccharide biosynthesis were isolated from the two parental genotypes, *S. sachalinensis* and *S. miyabeana*, for which there are no published reports of any cloning of cell wall biosynthetic genes (Supplemental Table S2). Gene fragments corresponding to cloned cDNAs were named with the prefix "Ssa" for those isolated from *S. sachalinensis* and "Smi" for those isolated from *S. miyabeana*. In this study, all cloned gene fragments aligned with conserved paralogs in *P. trichocarpa* subfamilies. When we compared polymorphisms between willow, *P. trichocarpa*, and *A. thaliana*, we observed the expected degree of sequence divergence, with gene fragments from willow being highly similar to coding regions in *P. trichocarpa*. Corresponding *A. thaliana* sequences were less similar to willow than poplar, with identity scores of 47% to 80%. Comparisons of the same gene fragments cloned from *S. sachalinensis* and *S. miyabeana* indicated that the gene sequences were over 99% identical. Single-nucleotide polymorphisms (SNPs) unique to each species could also be identified but were rare. SNPs present in similar sequences of the same gene may be indicative of allelic variation.

Lignin biosynthetic genes

Fragments of three genes involved with monolignol biosynthesis and known in *P. trichocarpa* to be strongly associated with lignin biosynthesis, *CCoAOMT1*, *COMT2*, and *CAD4*, were amplified from willow. Only a single gene fragment of *CCoAOMT* was amplified from *S. sachalinensis* and *S. miyabeana* and was highly similar to *PtrCCoAOMT1* (95% identity; POPTR_0009s10270). Similarly, only a single gene fragment of *COMT* was amplified from *S. sachalinensis* and *S. miyabeana* and was most similar to *PtrCOMT2* (91% identical; POPTR_0012s00670). A single SNP, a C/T substitution, was identified between *SsaCOMT2* and *SmiCOMT2* sequences; the T being unique to *SmiCOMT2*. For *CCoAOMT1*, no SNPS between the two species were identified. The sequences of *SsaCAD4* and *SmiCAD4* from willow were all 95% identical to *PtrCAD4* (POPTR_0009s09870), and no SNPs between *S. sachalinensis* or *S. miyabeana* were identified.

Gene expression profiles of these three genes were analyzed among the 9970 family parents and progeny. There were apparent differences in *CCoAOMT1* expression among 9970 progeny. Expression levels were higher in the shoot apex than in the young stem of SX61, SX64, and 9970-014 (Fig. 3). There were no significant differences in *COMT2* expression in either tissue among the genotypes analyzed (Fig. 3). *CAD4* gene expression in SX64 young stem tissue was significantly greater than the other seven genotypes (Fig. 3). Likewise, gene expression of *CCoAOMT1* was greatest in SX64 young stems than in the other genotypes. Higher *CAD4* gene expression was observed in the shoot apex than the young stem in three genotypes, SX64, 9970-024, and 9970-036.

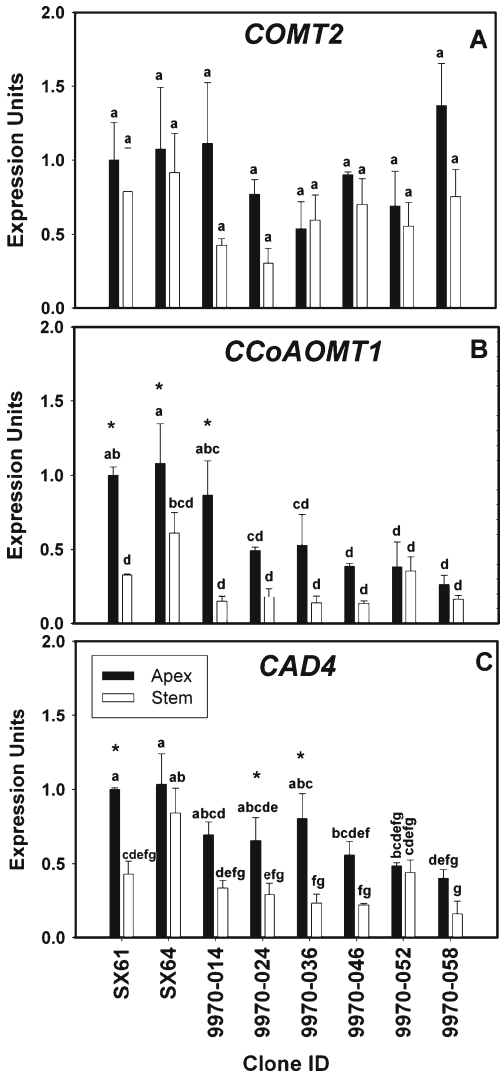

Fig. 3 Quantitative RT-PCR of lignin biosynthetic genes from shoot apex and young stem cDNA. *Bars* indicate the mean and standard error of expression among biological replicates (n=3). *Letters* above each bar indicate the results of mean separation test among genotypes. *Asterisks* indicate significant differences between tissues. A lack of asterisks indicates no significant differences were identified. Actin was used as the constitutive control. **a** Caffeic acid *O*-methyltransferase (*COMT2*), **b** caffeoyl CoA *O*-methyltransferase (*CCoAOMT1*), and **c** cinnamyl alcohol dehydrogenase (*CAD4*)

In our study, most genotypes had similar gene expression for *COMT2*, *CAD4*, and *CCoAOMT1* between the two tissue types. Since lignin deposition is necessary for normal xylem development and this process coincides with secondary cell wall development in primary xylem, it can be concluded that both tissue types analyzed in this study captured both developing and differentiating xylem. This is supported by the gene expression data. These results are comparable to studies in *P. trichocarpa* that have shown *PtrCCoAOMT1*, *PtrCOMT2*, and *PtrCAD4* to be abundantly expressed only in differentiating xylem (Tsai et al. 2006; Barakat et al. 2009; Shi et al. 2010).

Glycosyltransferases and the synthesis of noncellulosic polysaccharides

The *P. trichocarpa* genome, queried with *SsaUDP-GD* and *SmiUDP-GD*, retrieved two gene matches from BLAST searches (POPTR_0004s11760 and POPTR_0017s12760). It is possible that this is a duplicated gene, as a result of the genome duplication or a separate gene family member. Analysis of the *UDP-GD* gene family in *A. thaliana* has identified four *UDP-GD* genes (Klinghammer and Tenhaken 2007), and the size of this gene family in *P. trichocarpa* has yet to be determined. The probe and primers designed for gene expression studies were designed only for those that aligned to the *P. trichocarpa* sequence on LG IV. Expression profiling of *UDP-GD* in this study revealed significant differences among the members of the 9970 family (Fig. 4). Gene expression of *UDP-GD* was similar in both parental genotypes for both tissue types, but the expression in all of the progeny in the shoot apex and the young stem was lower than in the parents (Fig. 4). Expression was significantly greater in the young stem compared to that in the shoot apex for most individuals. These results support UDP-GDs involvement in cell wall development and complement the work done by Johansson et al. (2002) that showed high expression in developing xylem and young leaves of one *UDP-GD* gene in *P. tremula* × *P. tremuloides.*

Two members of the GT8 gene family were identified in *S. sachalinensis* and *S. miyabeana* that aligned well to *PttGT8B* (ortholog of POPTR_0005s06280) and *PttGT8C* (ortholog of POPTR_0007s04030). These two genes are

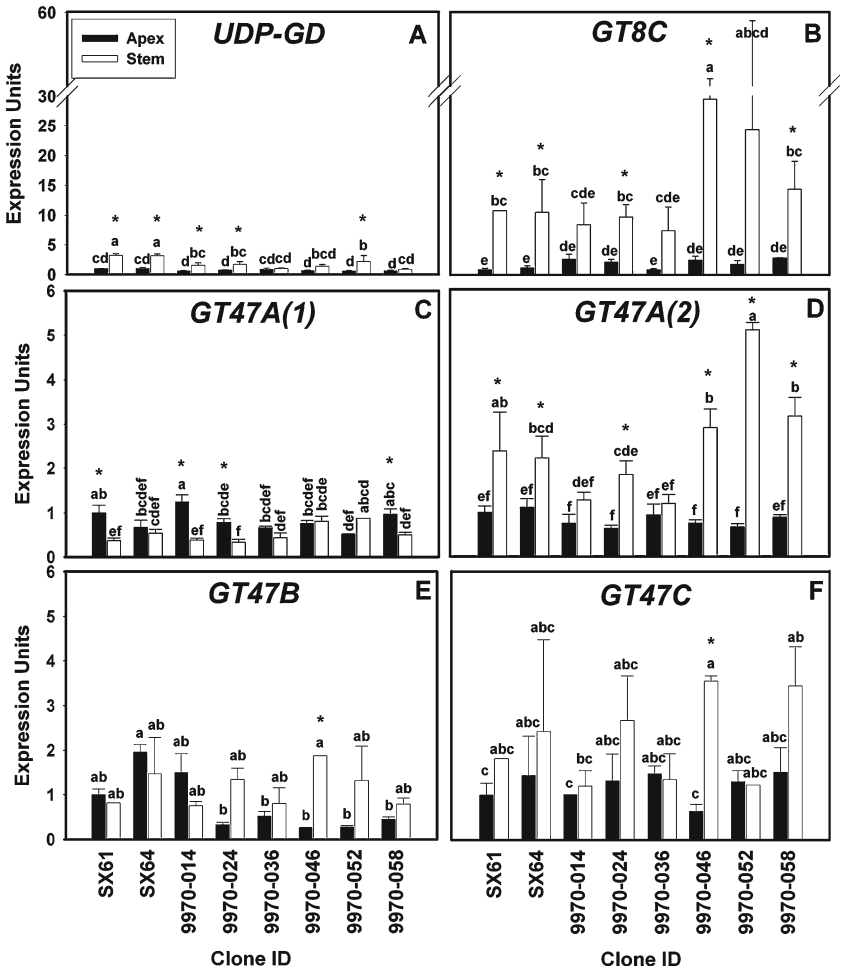

Fig. 4 Quantitative RT-PCR from shoot apex and stem cDNA. *Bars* indicate the mean and standard error of expression among biological replicates (n=3). *Letters* above each bar indicate the results of mean separation test among genotypes. *Asterisks* indicate significant differences between tissues. A lack of asterisks indicates no significant differences were identified. Actin was used as the constitutive control. **a** UDP-glucose dehydrogenase (*UDP-GD*), **b** glycosyltransferase family 8 (*GT8C*), **c** *GT47A1*, **d** *GT47A2*, **e** *GT47B*, and **f** *GT47C*

approximately 90% identical to each other. Primers and probes were designed specifically to amplify only the *GT8C* gene. The expression profile observed for *GT8C* displayed variation in expression among the members of the 9970 family and much greater expression in the young stem tissue than in the shoot apex (Fig. 4). For the shoot apex, the parental genotypes had low expression of *GT8C* compared to most of the progeny. In the young stem, expression was relatively similar for all the genotypes except for two of the progeny, but overall expression in the young stem was 2 to 40 times greater than in the shoot apex. Similar to our study, a *GT8* gene in *Pinus sitchensis*, closely related to *GAUT14* in *A. thaliana* (a member of the GT8C subfamily), showed an 18-fold increase in expression from the apical shoot tip to the woody stem below the tip (Friedmann et al. 2007).

Sequences from the GT47 family amplified from *S. sachalinensis* and *S. miyabeana* were similar to two members from the GT47A subfamily in *P. trichocarpa*, indicated as *GT47A(1)* (POPTR_0001s12940) and *GT47A(2)* (POPTR_0012s11100); one member from the GT47B subfamily (POPTR_0013s06370); and one member from the GT47C subfamily (POPTR_0009s01200). Expression profiling of *GT47C* in the 9970 willow family did not show any significant differences among genotypes. Only 9970-046 showed a significant increase in expression in the young stem compared to the shoot apex. Significant differences in expression among the genotypes were observed for *GT47A(1)* and *GT47B* in the shoot apex tissue and for *GT47A(1)* and *GT47A(2)* in the young stem tissue (Fig. 4). *GT47A(2)* expression was greater in young stem tissue compared to shoot apex tissue, but this was not observed for *GT47A(1)*. Expression levels in the parental genotypes were significantly different for *GT47B* in the shoot apex tissue with SX64 and 9970-014 having the greatest expression. Our analysis supports other studies that have shown that *GT47A(2)* is associated with secondary cell wall growth (Brown et al. 2009). The *irx10* gene identified in the xylan deficient mutant, *IRREGULAR XYLEM10* in *A. thaliana*, a member of the GT47A subfamily, is critical for xylan development and normal secondary cell wall biosynthesis (Brown et al. 2009). In contrast, *GT47A(1)* expression profiles suggest a different role unrelated to secondary cell wall formation.

Cellulose synthase

Gene fragments representing ten cellulose synthase genes were identified in *S. sachalinensis* or *S. miyabeana* and are numbered according to Kumar et al. (2009). The PCR primers used for amplification of cellulose synthase were designed from the highly conserved subdomains flanking the second hypervariable region (HVR, Liang and Joshi 2004). HVRs, also named class-specific regions, are highly variable among the *CesA* paralogs within the same species, but are highly conserved among the *CesA* orthologs from different species. All amplified regions and the genes they represented were aligned to the orthologous *CesA* genes from *P. trichocarpa*: *PtiCesA1*, *PtiCesA3*, *PtiCesA4*, *PtiCesA6*, *PtiCesA7*, and *PtiCesA8* (Table 1). They also aligned with the corresponding orthologs in *A. thaliana*. Two distinct

Table 1 Paralogs and orthologs of cellulose synthase genes identified in *S. sachalinensis*, *S. miyabeana*, *A. thaliana*, *E. grandis*, and *P. trichocarpa*

S. sachalinensis, *S. miyabeana*	*A. thaliana*	*E. grandis* (Ranik and Myburg 2006)	*P. trichocarpa* (Joshi et al. 2004)	*P. trichocarpa* (Kumar et al. 2009)	*P. trichocarpa* locus	*E* value *Populus*/*Salix*
SmiCesA1A	AtCesA1(AF027172)	EgCesA5 (DQ014509)	PtrCesA4B	PtiCesA1A	POPTR_0006s26810	$3e^{-166}$
SsaCesA1B	AtCesA1(AF027172)	EgCesA5 (DQ014509)	PtrCesA4A	PtiCesA1B	POPTR_0018s01540	$7e^{-178}$
SsaCesA3B	AtCesA3(NM_120599)	EgCesA4 (DQ014508)	PtrCesA8B	PtiCesA3B	POPTR_0016s05520	$7e^{-147}$
SsaCesA4	AtCesA4(AF458083)	EgCesA2 (DQ014506)	PtrCesA3A	PtiCesA4	POPTR_0002s25970	0
SmiCesA6A	AtCesA5(NM_121024)	EgCesA6 (DQ014510)	PtrCesA7A	PtiCesA6A	POPTR_0005s08970	$2e^{-163}$
SsaCesA6F	AtCesA5(NM_121024)	EgCesA6 (DQ014510)	PtrCesA6B	PtiCesA6F	POPTR_0005s02770	$1e^{-159}$
SsaCesA7A	AtCesA7(AF088917)	EgCesA3 (DQ014507)	PtrCesA2B	PtiCesA7A	POPTR_0006s19580	$1e^{-173}$
SsaCesA7B	AtCesA7(AF088917)	EgCesA3 (DQ014507)	PtrCesA2	PtiCesA7B	POPTR_0018s11290	$3e^{-160}$
SsaCesA8A	AtCesA8(AF267742)	EgCesA1 (DQ014505)	PtrCesA1B	PtiCesA8A	POPTR_0011s07040	$1e^{-164}$
SsaCesA8B	AtCesA8(AF267742)	EgCesA1 (DQ014505)	PtrCesA1A	PtiCesA8B	POPTR_0004s05830	$4e^{-161}$

genes from each of the subgroups *CesA1*, *CesA6*, *CesA7*, and *CesA8* were identified in *S. sachalinensis* or *S. miyabeana*. Only one isoform was identified for each of the genes *CesA3* and *CesA4*. Since the sequences within each subgroup are highly similar with very few SNPs, it was not possible to design primer and probes to distinguish between the two genes in each subgroup.

For the six *CesA* subgroups studied in willow expression levels were not significantly different between the parental genotypes in both tissues (Fig. 5). The highest level of *CesA7* and *CesA8* expression in the shoot apex was observed for 9970-014. In the shoot apex tissue *CesA1*, *CesA4*, and *CesA6* expression levels were significantly different among the progeny genotypes, while *CesA3* was not. The expression of *CesA1*, *CesA3*, and *CesA4* was greater in the shoot apex compared to the young stem for the progeny in the 9970 family and their parental genotypes. The expression of *CesA1* in the shoot apex was negligible in 9970-014 and 9970-024. In the majority of the progeny, *CesA4* and *CesA6* expression levels were greater than in the parental genotypes. In young stems, *CesA4* expression was greater in 9970-052 than in the other progeny, while gene expression of CesA6 in young stems of 9970-052 was greater than 9970-014, 9970-024, and 9970-036. There were no significant differences in expression in any of the genotypes for either *CesA1* or *CesA3* in the young stem.

Corresponding *CesA* genes identified in *S. sachalinensis* and *S. miyabeana* were determined by sequence identity to *A. thaliana*, *P. trichocarpa*, and *Eucalyptus grandis* (Table 1, Joshi et al. 2004; Ranik and Myburg 2006; Kumar et al. 2009). Phylogenetic comparisons of *A. thaliana* and *P. trichocarpa* show that *PtiCesA4*, *PtiCesA7*, and *PtiCesA8* are orthologous to the secondary cell wall genes in *A. thaliana*, *AtCesA4*, *AtCesA7*, and *AtCesA8*, respectively (Joshi et al. 2004). In addition, *PtiCesA1*, *PtiCes2*, *PtiCesA3*, and *PtiCesA6* are orthologous to the primary cell wall genes in *A. thaliana*, *AtCesA1*, *AtCesA2*, *AtCesA3*, and *AtCesA6*, respectively. Functional characterization of CSCs through immunoprecipitation assays using developing xylem tissue from *Populus deltoides* × *P. trichocarpa* indicated the presence of two types of CSCs (Song et al. 2010). Similar to *A. thaliana*, secondary cell wall formation involves a "type I" CSC made up of the proteins, Pd×tCesA4, Pd×tCesA7A and B, and Pd×tCesA8A and B.

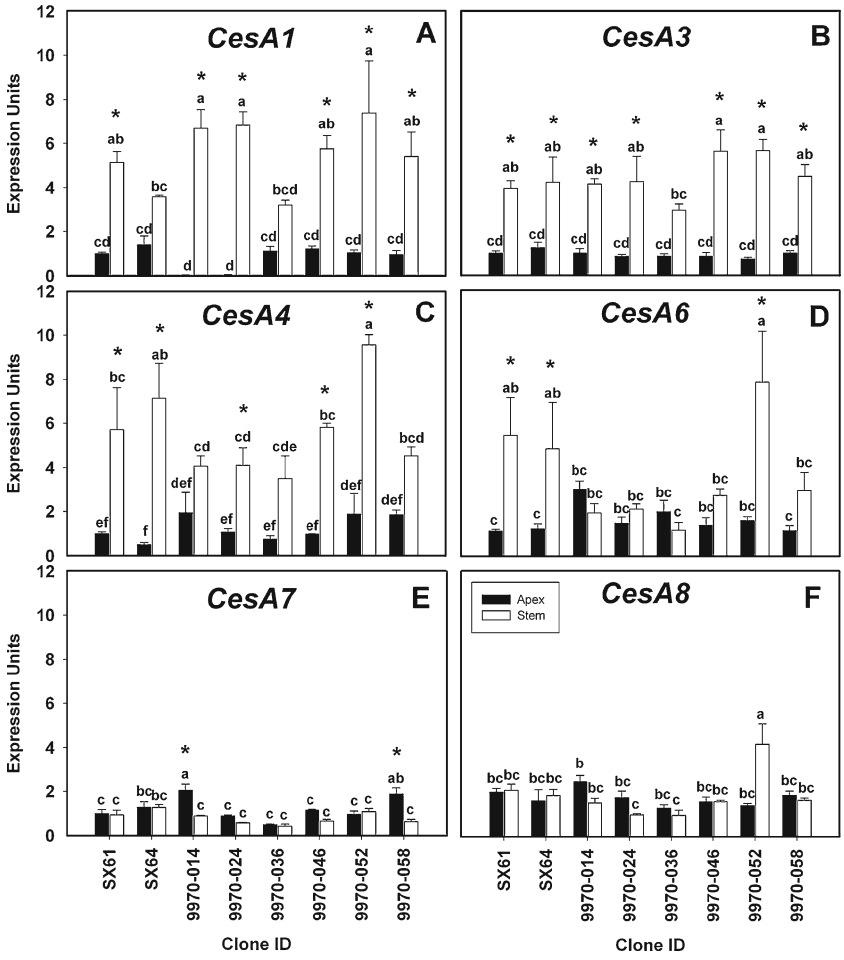

Fig. 5 Quantitative RT-PCR analysis of cellulose synthase genes from shoot apex and young stem cDNA. *Bars* indicate the mean and standard error of expression among biological replicates (n=3). *Letters* above each bar indicate the results of mean separation test among genotypes. *Asterisks* indicate significant differences between tissues. A lack of asterisks indicates no significant differences were identified. Actin was used as the constitutive control. **a** *CesA1*, **b** *CesA3*, **c** *CesA4*, **d** *CesA6*, **e** *CesA7*, and **f** *CesA8*

However, the "type II" CSC is associated with both primary and secondary cell wall formation, made up of the proteins Pd×tCesA1A and B, Pd×tCesA3A and B, and Pd×tCesA6A and B. These findings suggest the possibility that secondary cell wall cellulose microfibrils are synthesized by different types of CSCs. In our study with willow, gene expression profiles of *CesA4* and *CesA6* agree with the findings of Song et al. (2010). However, higher gene expression of *CesA1* and *CesA3* in the young stem tissue compared with the shoot tip does not support their involvement in primary cell wall development, nor do the gene expression profiles of *CesA7* or *CesA8* indicate association with secondary cell wall formation. Further expression studies will need to be done to clarify why these obvious differences in transcript abundance were observed. In spite of this, *CesA4* gene expression was strongly correlated to CesA7 (R^2=0.75) and CesA8 (R^2=0.91) expression in the stem tissue, indicating the possible relationship between the three in cell wall development.

Gene expression and correlations to biomass chemical composition

In this study, we correlated expression profiles in the shoot tip and young stem to biomass compositional data as a preliminary analysis of potential candidate genes. This method of analysis is based on the basic evolutionary principle that variations in gene expression can be directly associated with variation in the phenotype (Doebley and Lukens 1998; Carroll 2000; Stern 2000; Bergmann et al. 2003). Intraspecific variation in gene expression has been used to study changes in various phenotypic traits (Broekgaarden et al. 2010; Machabo and Cruzan 2010). Patterns of differential gene expression were correlated with variations in total polysaccharide content (percentage hemicellulose and percentage cellulose) and lignin composition of stem biomass from the two parents and six progeny (Table 2). Gene expression of *COMT2* was not included in the correlation analysis, since gene expression was not significantly different within the progeny. For this analysis, α was set at 0.1, with a threshold for correlation coefficients above 0.6 for all positive correlations or below −0.6 for negative correlations. Strong correlations (P<0.05) were observed with both *CesA4* and *UDP-GD* transcript abundance in the shoot apex with lignin (R^2=0.86 for *CesA4* and 0.75 for *UDP-GD*) and total polysaccharide (R^2=0.78 for *CesA4* and 0.68 for *UDP-GD*) content (Fig. 6).

Based on the expression analysis and correlations to biomass chemical composition, critical information was obtained about how gene expression may impact final composition and, subsequently, total biomass chemical composition. Notably, all correlations for gene expression profiles with biomass chemical composition were observed with apical shoot tissue, with the exception of *CesA1* young stem (Table 2). In willow, the apical shoot initiates stem growth and establishes the foundation for all concomitant and successive secondary stem growth and development. The development of the primary vascular system paves the way for wood formation and fully lignified mature secondary xylem. The gene expression associated with the development of the apical shoot may be vital to secondary wood growth and formation having a significant impact on final biomass chemical composition.

Both *COMT1* and *CAD4* are necessary for lignin biosynthesis, but they may or may not significantly impact the lignin content in the cell wall expressed as percent of dry matter. The lack of correlations with *COMT1* levels and lignin content observed in this study support a lack of impact on lignin biomass chemical composition, but there may be an effect on lignin chemistry, which we cannot resolve by HR-TGA. There was a weak positive trend with lignin content and *CAD4* expression, as well as a negative

Table 2 Pearson correlation coefficients for biomass composition and gene expression

Gene (tissue)	Percentage of polysaccharides	Percentage of lignin
CAD4 (apex)	−0.533	0.498
CAD4 (stem)	−0.531	0.412
CCoAOMT1 (apex)	−0.371	0.160
CCoAOMT1 (stem)	−0.511	0.500
CesA1 (apex)	−0.190	**0.685**
CesA1 (stem)	0.481	**−0.749**
CesA3 (apex)	−0.576	0.279
CesA3 (stem)	0.224	−0.274
CesA4 (apex)	**0.781**	**−0.857**
CesA4 (stem)	0.081	0.119
CesA6 (apex)	0.487	**−0.683**
CesA6 (stem)	0.135	0.089
CesA7 (apex)	0.234	**−0.638**
CesA7 (stem)	−0.143	0.034
CesA8 (apex)	0.209	**−0.696**
CesA8 (stem)	0.435	−0.163
GT8C (apex)	0.300	**−0.665**
GT8C (stem)	−0.222	0.102
GT47A(1) (apex)	0.269	**−0.646**
GT47A(1) (stem)	0.280	0.015
GT47A(2) (apex)	−0.443	**0.622**
GT47A(2) (stem)	0.231	0.001
UDP-GD (apex)	**−0.680**	**0.748**
UDP-GD (stem)	−0.385	0.331

Correlations with P values <0.1 are indicated in bold

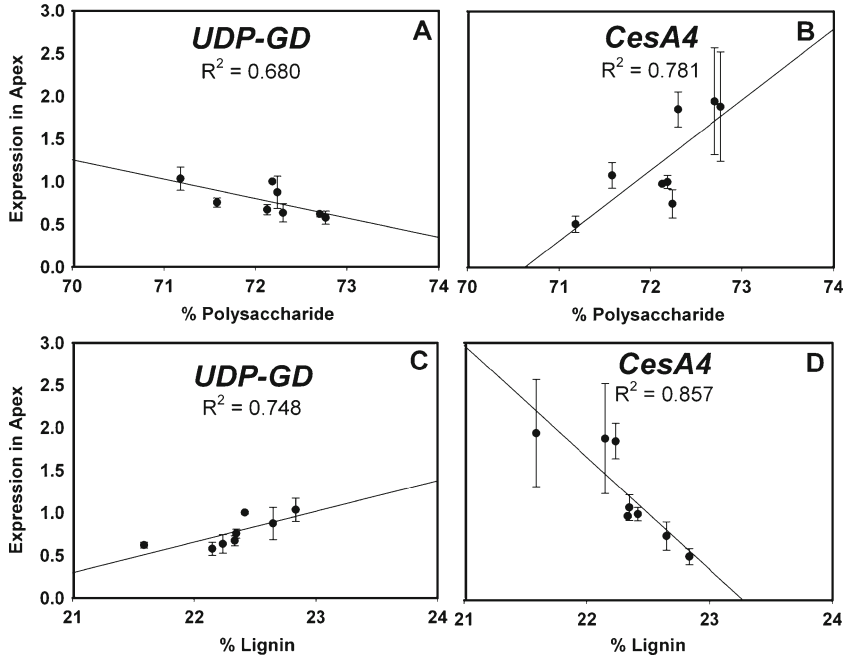

Fig. 6 Regression analysis showing correlations of UDP-glucose dehydrogenase (*UDP-GD*; **a**, **c**) and cellulose synthase 4 (*CesA4*; **b**, **d**) gene expression profiles to total polysaccharide (**a**, **b**) and lignin content (**c**, **d**)

relationship with polysaccharide content, which may indicate some impact on compositional variation. If a future goal is to select for variation in S to G composition in willow biomass, further analysis with respect to *COMT* and *CAD4* expression may be warranted. In contrast to *COMT1* and *CAD4*, downregulation of *CCoAOMT2* has led to a decrease in total lignin content and S and G units (Meyermans et al. 2000; Baucher et al. 2003; Day et al. 2009). Our study showed a weak relationship between *CCoAOMT2* gene expression and lignin content, and a negative relationship with polysaccharide content (Table 2). Future studies of gene expression could focus on other lignin biosynthetic enzymes, such as coumaroyl shikimate 3-hydroxylase and ferulate 5-hydroxylase, which may have a stronger correlation to lignin content (Chen and Dixon 2007; Elissetche et al. 2011).

A common observation in cell wall manipulation has been the compensatory nature of the cell wall, where a loss of lignin has been offset by an increase in the polysaccharide counterpart (Sticklen 2006; Chen and Dixon 2007). The reverse holds true as well. Increased expression levels of *UDP-GD* in transgenic alfalfa resulted in plants with decreased cell wall polysaccharide content and a concomitant increase in Klason lignin content (Samac et al. 2004). It was speculated that redirecting more UDP-glucose to UDP-GD would reduce the flux to cellulose biosynthesis and result in increased lignin content. Our study revealed a direct correlation of *UDP-GD* gene expression with lignin content and a negative correlation with polysaccharide content (Fig. 6; Table 2). This strong correlation observed in our study indicates that UDP-GD has a significant impact on cell wall development.

In addition, expression profiles of *GT8C* and most of the *CesA* genes were negatively correlated with lignin content. Only *CesA4* showed a direct correlation with cellulose content (Fig. 6). These results support the hypothesis that higher polysaccharide content results in lower lignin content and could possibly show a strong relationship with gene expression and carbon sinks. Increased *CesA* gene expression would result in increases in the carbon fraction directed to cellulose biosynthesis, concurrently reducing the amount of carbon entering the phenylpropanoid pathway for lignin biosynthesis reducing total lignin content.

The analysis of shrub willow biomass accomplished throughout this research has enhanced the understanding of candidate structural genes that could potentially impact total biomass chemical composition in this emerging feedstock crop. This study represents one of the first analyses on cell wall development in shrub willow, showing differential transcript accumulation of candidate genes among genotypes and two distinct tissue types. This study also shows that early stages of cell wall development and stem growth can be correlated to final mature wood composition and the potential use of early gene expression as a selection tool in the breeding process. This would greatly reduce the time required to select new genotypes based on compositional characteristics. Gene expression analysis could potentially be performed on very young progeny, even prior to field selection, to provide insight into the chemical composition of their future biomass. In addition, these preliminary results will allow future research in shrub willow to progress toward the identification of important genes that control wood formation and cell wall development and toward the development of marker-assisted selection. Further sequencing

of the *Salix* genome will provide tools for expanded analysis of these networks controlling biomass chemical composition. Future research with candidate gene association mapping is promising and has the potential to enhance the breeding process for elite shrub willow varieties for bioenergy.

Acknowledgments This work was funded by a grant from the USDA-CSREES McIntire-Stennis Cooperative Forestry Research Program to Dr. Larry B. Smart. Michelle Serapiglia was generously supported by the Josiah Lowe and Hugh Wilcox Scholarship Fund and the Edna Bailey Sussman Foundation. Appreciation is also expressed to Drs. William Powell, Larry Abrahamson, Tim Volk, Ed White, and Bill Winter for their support and advice as collaborators with this research and to Mark Appleby and Ken Burns for the excellent technical support. Special thanks go to Dr. Martin Goffinet for the assistance with histochemical and microscopy analysis of willow stems and to two anonymous reviewers for their insightful suggestions of improvements to this paper.

References

Aspeborg H, Schrader J, Coutinho PM, Stam M, Kallas A, Djerbi S, Nilsson P, Denman S, Amini B, Sterky F, Master E, Sandberg G, Mellerowicz E, Sundberg B, Henrissat B, Teeri TT (2005) Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen. Plant Physiol 137:983–997

Atanassov II, Pittman JK, Turner SR (2009) Elucidating the mechanisms of assembly and subunit interaction of the cellulose synthase complex of *Arabidopsis* secondary cell walls. J Biol Chem 284:3833–3841

Barakat A, Bagniewska-Zadworna A, Choi A, Plakkat U, DiLoreto D, Yellanki P, Carlson J (2009) The cinnamyl alcohol dehydrogenase gene family in *Populus*: phylogeny, organization, and expression. BMC Plant Biol 9:26

Baucher M, Chabbert B, Pilate G, Van Doorsselaere J, Tollier M, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inze D, Jouanin L, Boerjan W (1996) Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. Plant Physiol 112:1479–1490

Baucher M, Halpin C, Petit-Conil M, Boerjan W (2003) Lignin: genetic engineering and impact on pulping. Crit Rev Biochem Mol Biol 38:305–350

Bergmann S, Ihmels J, Barkai N (2003) Similarities and differences in genome-wide expression data of six organisms. PLoS Biol 2:85–93

Boerjan W (2005) Biotechnology and the domestication of forest trees. Curr Opin Biotechnol 16:159–166

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Ann Rev Plant Biol 54:519–546

Broekgaarden C, Poelman E, Voorrips R, Dicke M, Vosman B (2010) Intraspecific variation in herbivore community composition and transcriptional profiles in field-grown *Brassica oleracea* cultivars. J Exp Bot 61:807–819

Brown DM, Zhang Z, Stephens E, Dupree P, Turner SR (2009) Characterization of *IRX10* and *IRX10*-like reveals an essential role in glucuronoxylan biosynthesis in *Arabidopsis*. Plant J 57:732–746

Brunner A, Yakovlev I, Strauss S (2004) Validating internal controls for quantitative plant gene expression studies. BMC Plant Biol 4:14

Cameron KD, Kopp RF, Abrahamson LP, Volk TA, Smart LB (2010) High-yielding shrub willow (*Salix* spp.) emerges as a robust bioenergy crop suitable for growth in the Northeast, Upper Midwest, and Southern Canada. http://abstracts.aspb.org/pb2010/public/P01/P01054.html. *American Society of Plant Biologists—Plant Biology* 2010, Montreal, Canada

Campbell MM, Sederoff RR (1996) Variation in lignin content and composition (mechanisms of control and implications for the genetic improvement of plants). Plant Physiol 110:3–13

Carroll S (2000) Endless forms: the evolution of gene regulation and morphological diversity. Cell 101:577–580

Carter MJ, Milton ID (1993) An inexpensive and simple method for DNA purifications on silica particles. Nucleic Acids Res 21:1044

Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. Nat Biotechnol 25:759–761

Coutinho PM, Stam M, Blanc E, Henrissat B (2003) Why are there so many carbohydrate-active enzyme-related genes in plants? Trends Plant Sci 8:563–565

Dauwe R, Morreel K, Goeminne G, Gielen B, Rohde A, Beeumen JV, Ralph J, Boudet A-M, Kopka J, Rochange SF, Halpin C, Messens E, Boerjan W (2007) Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. Plant J 52:263–285

Day A, Neutelings G, Nolin F, Grec S, Habrant A, Cronier D, Mahar B, Rolando C, David H, Chabbert B, Hawkins S (2009) Caffeoyl coenzyme A *O*-methyltransferase down-regulation is associated with modifications in lignin and cell-wall architecture in flax secondary xylem. Plant Physiol Biochem 47:9–19

Dharmawardhana P, Brunner A, Strauss S (2010) Genome-wide transcriptome analysis of the transition from primary to secondary stem development in *Populus trichocarpa*. BMC Genomics 11:150

Dinus RJ, Payne P, Sewell MM, Chiang VL, Tuskan GA (2001) Genetic modification of short rotation poplar wood: properties for ethanol fuel and fiber productions. Crit Rev Plant Sci 20:51–69

Doebley J, Lukens L (1998) Transcriptional regulators and the evolution of plant form. Plant Cell 10:1075–1082

Elissetche JP, Valenzuela S, Garcia R, Norambuena M, Iturra C, Rodriguez J, Mendonca RT, Balocchi C (2011) Transcript abundance of enzymes involved in lignin biosynthesis of *Eucalyptus globulus* genotypes with contrasting levels of pulp yield and wood density. Tree Genet Genome 7:697–705

Friedmann M, Ralph SG, Aeschliman D, Zhuang J, Ritland K, Ellis BE, Bohlmann J, Douglas CJ (2007) Microarray gene expression profiling of developmental transitions in Sitka spruce (*Picea sitchensis*) apical shoots. J Exp Bot 58:593–614

Geisler-Lee J, Geisler M, Coutinho PM, Segerman B, Nishikubo N, Takahashi J, Aspeborg H, Djerbi S, Master E, Andersson-Gunneras S, Sundberg B, Karpinski S, Teeri TT, Kleczkowski LA, Henrissat B, Mellerowicz EJ (2006) Poplar carbohydrate-active enzymes. Gene identification and expression analysis. Plant Physiol 140:946–962

González-Martínez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB (2007) Association genetics in *Pinus taeda* L. I. wood property traits. Genetics 175:399–409

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Himmel ME, Ding S-Y, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007) Biomass recalcitrance: engineering plants and enzymes for biofuel production. Science 315:804–807

Hu W, Harding S, Lung J, Popko J, Ralph J, Stokke D, Tsai C, Chiang V (1999) Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nat Biotechnol 17:808–812

Johansson H, Sterky F, Amini B, Lundenberg J, Kleczkowski LA (2002) Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation. Biochim Biophys Acta Gene Struct Expr 1576:53–58

Joshi CP, Bhandari S, Ranjan P, Kalluri U, Liang X, Fujino T, Samuga A (2004) Genomics of cellulose biosynthesis in poplars. New Phytol 164:53–61

Jouanin L, Goujon T, de Nadai V, Martin M-T, Mila I, Vallet C, Pollet B, Yoshinaga A, Chabbert B, Petit-Conil M, Lapierre C (2000) Lignification in transgenic poplars with extremely reduced caffeic acid *O*-methyltransferase activity. Plant Physiol 123:1363–1374

Kirst M, Basten CJ, Myburg AA, Zeng Z-B, Sederoff RR (2005) Genetic architecture of transcript-level variation in differentiating xylem of a *Eucalyptus* hybrid. Genetics 169:2295–2303

Klinghammer M, Tenhaken R (2007) Genome-wide analysis of the UDP-glucose dehydrogenase gene family in *Arabidopsis*, a key enzyme for matrix polysaccharides in cell walls. J Exp Bot 58:3609–3621

Kumar M, Thammannagowda S, Bulone V, Chiang V, Han K-H, Joshi CP, Mansfield SD, Mellerowicz E, Sundberg B, Teeri T, Ellis BE (2009) An update on the nomenclature for the cellulose synthase genes in *Populus*. Trends Plant Sci 14:248–254

Lapierre C, Pollet B, Petit-Conil M, Toval G, Romero J, Pilate G, Leple J, Boerjan W, Ferret V, De Nadai V, Jouanin L (1999) Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid *O*-methyltransferase activity have an opposite impact on the efficiency of industrial Kraft pulping. Plant Physiol 119:153–164

Liang X, Joshi CP (2004) Molecular cloning of ten distinct hypervariable regions from the cellulose synthase gene superfamily in aspen trees. Tree Physiol 24:543–550

Machabo H, Cruzan M (2010) Intraspecific variation in gene expression under prolonged drought in *Piriqueta* hybrids and their parental taxa. Plant Sci 178:429–439

Mészáros E, Jakab E, Várhegyi G, Szepesváry P, Marosvölgyi B (2004) Comparative study of the thermal behavior of wood and bark of young shoots obtained from an energy plantation. J Anal Appl Pyrol 72:317–328

Meyermans H, Morreel K, Lapierre C, Pollet B, De Bruyn A, Busson R, Herdewijn P, Devreese B, Van Beeumen J, Marita JM, Ralph J, Chen C, Burggraeve B, Van Montagu M, Messens E, Boerjan W (2000) Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis. J Biol Chem 275:36899–36909

Orfao JJM, Antunes FJA, Figueiredo JL (1999) Pyrolysis kinetics of lignocellulosic materials—three independent reactions model. Fuel 78:349–358

Orfila C, Sorensen S, Harbolt J, Geshi N, Crombie H, Truong H-N, Reid J, Knox J, Scheller H (2005) *Quasimodo1* is expressed in vascular tissue of *Arabidopsis thaliana* inflorescence stems, and affects homogalacturonan and xylan biosythesis. Planta 222:613–622

Pilate G, Guiney E, Holt K, Petit-Conil M, Lapierre C, Leple J, Pollet B, Mila I, Webster E, Marstorp H, Hopkins D, Jouanin L, Boerjan W, Schuch W, Cornu D, Halpin C (2002) Field and pulping performances of transgenic trees with altered lignification. Nat Biotechnol 20:607–612

Ralph J, Mackay JJ, Hatfield R, O'Malley D, Whetten R, Sederoff RR (1997) Abnormal lignin in a loblolly pine mutant. Science 277:235–239

Ranik M, Myburg AA (2006) Six new cellulose synthase genes from *Eucalyptus* are associated with primary and secondary cell wall biosynthesis. Tree Physiol 26:545–556

Ranjan P, Yin T, Zhang X, Kalluri U, Yang X, Jawdy S, Tuskan G (2010) Bioinformatics-based identification of candidate genes from QTLs associated with cell wall traits in *Populus*. Bioenerg Res 3:172–182

Raveendran K, Ganesh A, Khilar KC (1996) Pyrolysis characteristics of biomass and biomass components. Fuel 75:987–998

Samac DA, Litterer L, Temple G, Jung H, Somers DA (2004) Expression of UDP-glucose dehydrogenase reduces cell-wall polysaccharide concentration and increases xylose content in alfalfa stems. Appl Biochem Biotech 113–116:1167–1182

SAS Institute Inc. SAS 9.1.3 Help and Documentation. Cary, NC: SAS Institute Inc., 2000–2004

Serapiglia MJ, Cameron KD, Stipanovic AJ, Smart LB (2008) High-resolution thermogravimetric analysis for rapid characterization of biomass composition and selection of shrub willow varieties. Appl Biochem Biotechnol 145:3–11

Serapiglia MJ, Cameron KD, Stipanovic AJ, Smart LB (2009) Analysis of biomass composition using high-resolution thermogravimetric analysis and percent bark content as tools for the selection of shrub willow bioenergy crop varieties. Bioenerg Res 2:1–9

Shafizadeh F, Chin PPS (1977) Thermal decomposition of wood. In: Goldstein IS (ed) Wood technology: chemical aspects, pp. 57–81. American Chemical Society Symposium Series 43

Shi R, Sun Y-H, Heber S, Sederoff R, Chiang V (2010) Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes. Plant Cell Physiol 51:144–163

Singh K, Risse M, Das KC, Worley J (2009) Determination of composition of cellulose and lignin mixtures using thermogravimetric analysis. J Energ Resour Tech 131:022201–022206

Song D, Shen J, Li L (2010) Characterization of cellulose synthase complexes in *Populus* xylem differentiation. New Phytol 187:777–790

Stern DL (2000) Perspective: evolutionary developmental biology and the problem of variation. Evolution 54:1079–1091

Sticklen M (2006) Plant genetic engineering to improve biomass characteristics for biofuels. Curr Opin Biotechnol 17:315–319

Stipanovic AJ, Goodrich J, Hennessy P (2004) High resolution thermogravimetric analysis (HR-TGA) in the compositional characterization of polysaccharides American Chemical Society symposium on "novel analytical tools in the characterization of polysaccharides". Cellulose and Renewable Materials Division, Philadelphia

Thumma BR, Southerton SG, Bell JC, Owen JV, Henery ML, Moran GF (2010) Quantitative trait locus (QTL) analysis of wood quality traits in *Eucalyptus nitens*. Tree Genet Genome 6:305–317

Tsai C-J, Harding SA, Tshaplinski TJ, Lindroth RL, Yuan Y (2006) Genome-wide analysis of the structural genes regulating defense phenylpropanoid metabolism in *Populus*. New Phytol 172:47–62

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen G-L, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, dePamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjarvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leple J-C, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai C-J, Uberbacher E, Unneberg P et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313:1596–1604

Vanholme R, Morreel K, Ralph J, Boerjan W (2008) Lignin engineering. Curr Opin Biotechnol 11:278–285

Voelker SL, Lachenbruch B, Meinzer FC, Jourdes M, Ki C, Patten AM, Davin LB, Lewis NG, Tuskan GA, Gunter L, Decker SR, Selig MJ, Sykes R, Himmel ME, Kitin P, Shevchenko O, Strauss SH (2010) Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar. Plant Physiol 154:874–886

Wagner A, Donaldson L, Kim H, Phillips IS, Flint H, Steward D, Torr K, Koch G, Schmitt U, Ralph J (2009) Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata*. Plant Physiol 149:370–383

Wan CY, Wilkins TA (1994) A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). Anal Biochem 223:7–12

Wang L, Weller CL, Jones DD, Hanna MA (2008) Contemporary issues in thermal gasification of biomass and its application to electricity and fuel production. Biomass Bioenerg 32:573–581

Wegrzyn JL, Eckert AJ, Choi M, Lee JM, Stanton BJ, Sykes R, Davis MF, Tsai CJ, Neale DB (2010) Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. New Phytol 188:515–532

Yang H, Yan R, Chen H, Zheng C, Lee DH, Liang DT (2006) In-depth investigation of biomass pyrolysis based on three major components: hemicellulose, cellulose, and lignin. Energ Fuel 20:388–393

Yang H, Yan R, Chen H, Lee DH, Zheng C (2007) Characteristics of hemicellulose, cellulose, and lignin pyrolysis. Fuel 86:1781–1788

Zhang Z, Schwartz S, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7:203–214

Zhong R, Morrison WH, Himmelsbach DS, Poole FL, Ye Z-H (2000) Essential role of caffeoyl coenzyme A *O*-methyltransferase in lignin biosynthesis in woody poplar plants. Plant Physiol 124:563–578