Carbohydrate-Active Enzymes Involved in the Secondary Cell Wall Biogenesis in Hybrid Aspen[1]

Henrik Aspeborg, Jarmo Schrader, Pedro M. Coutinho, Mark Stam, Åsa Kallas, Soraya Djerbi, Peter Nilsson, Stuart Denman[2], Bahram Amini, Fredrik Sterky, Emma Master, Göran Sandberg, Ewa Mellerowicz, Björn Sundberg, Bernard Henrissat, and Tuula T. Teeri*

Royal Institute of Technology, Department of Biotechnology, AlbaNova University Center, SE–10691 Stockholm, Sweden (H.A., A.K., S. Djerbi, P.N., S. Denman, B.A., F.S., E. Master, T.T.T.); Umea Plant Science Center, Department of Forest Genetics and Plant Physiology, SE–90183 Umea, Sweden (J.S., G.S., E. Mellerowicz, B.S.); and Architecture et Fonction des Macromolécules Biologiques, Unité Mixte de Recherche 6098, Centre National de la Recherche Scientifique and Universities of Aix-Marseille I and II, 13402 Marseille cedex 20, France (P.M.C., M.S., B.H.)

Wood formation is a fundamental biological process with significant economic interest. While lignin biosynthesis is currently relatively well understood, the pathways leading to the synthesis of the key structural carbohydrates in wood fibers remain obscure. We have used a functional genomics approach to identify enzymes involved in carbohydrate biosynthesis and remodeling during xylem development in the hybrid aspen *Populus tremula* × *tremuloides*. Microarrays containing cDNA clones from different tissue-specific libraries were hybridized with probes obtained from narrow tissue sections prepared by cryosectioning of the developing xylem. Bioinformatic analyses using the sensitive tools developed for carbohydrate-active enzymes allowed the identification of 25 xylem-specific glycosyltransferases belonging to the Carbohydrate-Active EnZYme families GT2, GT8, GT14, GT31, GT43, GT47, and GT61 and nine glycosidases (or transglycosidases) belonging to the Carbohydrate-Active EnZYme families GH9, GH10, GH16, GH17, GH19, GH28, GH35, and GH51. While no genes encoding either polysaccharide lyases or carbohydrate esterases were found among the secondary wall-specific genes, one putative *O*-acetyltransferase was identified. These wood-specific enzyme genes constitute a valuable resource for future development of engineered fibers with improved performance in different applications.

Wood constitutes the oldest source of renewable energy for mankind and an important raw material for construction materials, textiles, pulp, and paper, as well as many other products. Recent scientific and technological advances offer new possibilities to expand the use of plant fibers toward novel types of products, such as biocomposite materials. In particular, the advent of biotechnology for short-rotation forestry and advances in enzyme technology will allow innovative fiber engineering to alter the structure, composition, and properties of the raw material. Wood formation is thus a fundamental biological process with significant economic interest.

Wood tissue develops by terminal differentiation of the vascular cambium. Xylogenesis progresses through several stages, including cell division, cell expansion, secondary cell wall formation, and programmed cell death. Although xylem fibers constitute the main body of wood biomass, other cell types contribute substantial heterogeneity. Poplar wood is typically composed of 33% (v/v) vessel elements, 53% to 55% fibers, 11% to 14% ray parenchyma, and less than 1% axial parenchyma (Mellerowicz et al., 2001). Each cell type has a characteristic structure and chemical composition and a different rate of differentiation. The resulting wood properties depend on the size, shape, and arrangement of different cell types in the xylem, as well as the structure, composition, and morphology of the secondary cell walls. Poplar wood consists of about 50% cellulose, 30% hemicellulose, and 20% lignin (Balatinecz et al., 2001). The major hemicellulose in aspens is *O*-acetyl-(4-*O*-methylglu-curono)-xylan, but a smaller portion of *O*-acetylated glucomannan is also present (Gustavsson et al., 2001; Teleman et al., 2003). While the overall process of lignin biosynthesis has been studied intensively (for review, see Boerjan et al., 2003), this is not the case for the biosynthesis of the cell wall carbohydrate polymers.

An analysis of the entries in the Carbohydrate-Active EnZYme database (CAZy; http://afmb.cnrs-mrs.fr/CAZY/) reveals that the number of genes encoding glycosyltransferases (GTs) and glycoside hydrolases (GHs) is far greater in plants than in any other organisms sequenced so far (Coutinho et al., 2003b). This apparently reflects the importance of these

enzymes in the biosynthesis and remodeling of the cell wall carbohydrate polymers during different stages of plant life. However, the majority of the corresponding genes belong to large gene families encoding enzymes with diverse substrate and product specificities. Therefore, their precise roles and functions are difficult to resolve based on the gene sequences alone (Henrissat et al., 2001; Coutinho et al., 2003a). Plant GTs are typically membrane bound, which makes their isolation and biochemical characterization difficult. Even other wood-specific enzymes present considerable challenges because of the inherent resilience of the source material. Consequently, the biosynthesis and enzymatic remodeling of the secondary cell wall carbohydrate polymers have remained poorly understood.

Functional genomics offers the means to identify enzymes and other proteins involved in developmental processes, including wood formation. The Swedish Populus Genome Project has previously established a database of 102,019 expressed sequence tag (EST) sequences from 19 different cDNA libraries from the European aspen (*Populus tremula*), the hybrid aspen (*Populus tremula* × *tremuloides* T89), and the black cottonwood (*Populus trichocarpa*; http://www.populus.db.umu.se; Sterky et al., 1998, 2004; Bhalerao et al., 2003). We have previously carried out expression profiling during wood formation using a microarray of 2,995 cDNA clones from a broad cambial region library. cDNA targets were prepared by tangential cryosectioning of tissue sections representing different stages of developing wood, including phloem (Ph), meristematic cells (A), early expansion (B), late expansion (C), secondary cell wall formation (D), and programmed cell death (E; Hertzberg et al., 2001). A larger poplar microarray with 13,490 cDNA fragments from 7 different, tissue-specific poplar cDNA libraries has been constructed to identify genes involved in leaf development (Andersson et al., 2004) and cellulose biosynthesis (Djerbi et al., 2004). The estimated redundancy on the 13.5 K microarray is 28%, indicating that approximately 10,000 unique genes are represented on the array. In this article, we have mined the old as well as previously unpublished poplar microarray data to identify xylem-specific genes encoding carbohydrate-active enzymes (CAZymes).

RESULTS AND DISCUSSION

Expression Profiling over the Developing Xylem

The 13.5 K microarray was hybridized using targets prepared from zones A to E during xylem development, as described earlier (Hertzberg et al., 2001). In order to retrieve (1) a broad group of candidate genes expressed over the developing cambium and (2) a narrower group of genes showing a relatively higher level of expression specifically in zone D, the expression data obtained were subjected to two stages of filtering. Filter I was tuned to select clones that had at

least 2-fold higher expression in zone D compared to the average level of expression in all the zones (A–E). The first filtering step yielded a collection of 1,141 clones. These clones were subjected to a second filtering step that required, in addition, a higher signal in zone D than in neighboring zones C and E. The resulting 450 clones were grouped into 9 clusters based on their expression profiles across the entire wood-forming zone (A–E) using self-organizing maps (Fig. 1). The results obtained were validated in silico by comparative analysis with data obtained by independent experiments by Hertzberg et al. (2001) and Djerbi et al. (2004; Tables I and II). Because of the redundancy of the 13.5 K microarray, some of the genes are represented by more than one EST clone. With few exceptions, these redundant clones show similar transcript profiles, and they fall into the same clusters.

Annotation of the identified ESTs or full-length cDNAs was carried out with biocomputing tools routinely used to assign novel protein sequences to the CAZy database (http://afmb.cnrs-mrs.fr/CAZY; Coutinho and Henrissat, 1999). To avoid common misannotations introduced by BLAST-based automated annotation, each gene was compared against a library of individual catalytic and ancillary modules, so that family assignment was not impaired by the frequent modularity of CAZymes. Among the 1,140 xylem-specific cDNA clones that passed the first filtering, 34 individual genes turned out to encode putative CAZymes, and 23 out of these also passed the second, more stringent filtering (Table I). Thus, genes encoding 11 CAZYmes passed filter I but did not fulfill the requirements of filter II (Table II). In contrast with the large number of GT and GH-encoding genes that were found, no genes encoding polysaccharide lyases or carbohydrate esterases were identified. In Table III, the Populus secondary cell wall-related CAZymes hereby identified are listed together with their closest hits in Arabidopsis (*Arabidopsis thaliana*) and with genes reported to have high expression in the secondary walls in Arabidopsis (Oh et al., 2003), Zinnia (*Zinnia elegans*; Demura et al., 2002), loblolly pine (*Pinus taeda*; Whetten et al., 2001; Lorenz and Dean, 2002), black locust (*Robinia pseudoacacia*; Yang et al., 2003), and cotton (*Gossypium hirsutum*; Pear et al., 1996; Zhao et al., 2001; Li et al., 2002; Ji et al., 2003). Although all of the experiments on these species are not strictly comparable to ours, coinciding expression in secondary walls serves as an additional quality control.

Classification of CAZymes Highly Expressed during Secondary Cell Wall Formation

Glycosyl Transferases

Family GT2. This is the largest GT family, with almost 3,200 members in the CAZy database, which was updated on October 1, 2004. Although few members of this family have been biochemically characterized, many different donor/acceptor specificities
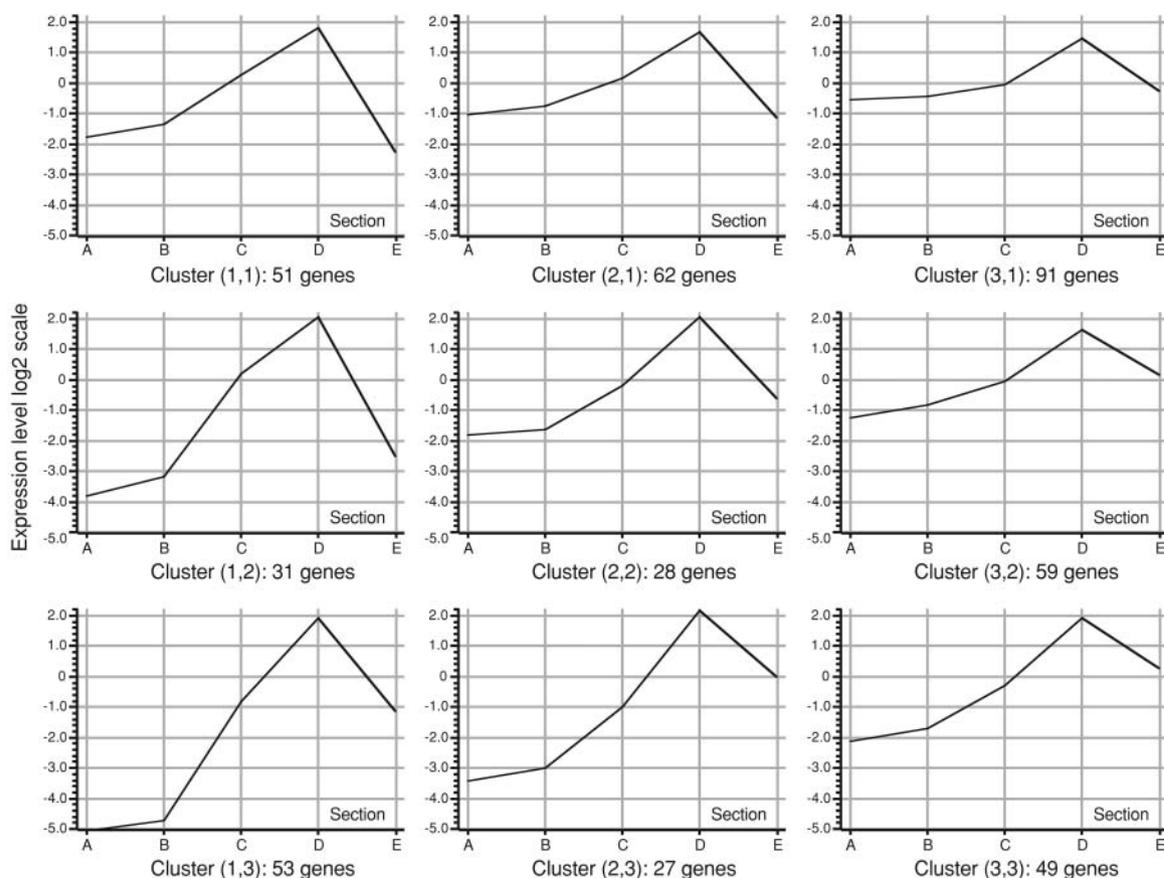
Figure 1. Clustering using self-organizing maps of 450 genes. Genes with less than two valid data points in sections C to E or with a signal less than 2-fold higher in D than in A to C and E were filtered out.

have been identified, all forming glycosidic linkages of stereochemistry opposite to that of the starting nucleotide sugar. This large family contains both single-addition and multiple-addition (or processive) GTs. Arabidopsis apparently contains two single-addition GT2 GTs (presumably dolichol-phosphate β-mannosyltransferase and dolichol-phosphate β-glucosyltransferase) and 40 processive GTs, which include cellulose synthases (CesA).

Nine different *CesA* genes have been identified so far in the EST libraries of PopulusDB (Sterky et al., 1998; Djerbi et al., 2004; Table IV). Based on the different microarray analyses (Hertzberg et al., 2001) and complementary experiments by real-time PCR (Djerbi et al., 2004), four *CesA* genes (*PttCesA1*, *PttCesA3-1*, *PttCesA3-2*, and *PttCesA9*) were found to be particularly highly expressed during xylogenesis. The full-length secondary cell wall-associated *CesA* genes from hybrid aspen cluster in the same phylogenetic groups as the xylem-specific *CesA* genes previously identified in other poplars, rice (*Oryza sativa*), barley (*Hordeum vulgare*), and Arabidopsis (Table IV), thus providing additional support for their involvement in secondary cell wall biosynthesis.

Processive GT2 GTs in plants have been assigned to different subfamilies, and the more distantly related

cellulose synthase-like (CSL) proteins apparently synthesize noncellulosic polysaccharides (Richmond and Somerville, 2001). This has been demonstrated for the guar $\beta(1\rightarrow4)$-mannan synthase in one CSL subfamily (Dhugga et al., 2004). One *CSL* gene, *PttGT2A*, was found to be highly expressed in the wood-forming tissues of poplar (Table I). Since the molecular mechanism of catalysis appears to be conserved within the GT families (Campbell et al., 1997; Coutinho et al., 2003a) and since CSL proteins are clearly related to GTs synthesizing β-linked polysaccharides such as cellulose and mannan, PttGT2a also is likely to be involved in the synthesis of a β-linked polysaccharide, such as glucomannan or *O*-acetyl-(4-*O*-methylglucurono)-xylan, for example.

The closest homolog of *PttGT*2a is the *AtCSLA9* from Arabidopsis. A T-DNA insertion mutation of *AtCslA9* renders the mutant (*rat4*) resistant to Agrobacterium transformation. No differences were observed in linkage structure of insoluble noncellulosic polysaccharides because of the defective gene (Nam et al., 1999; Zhu et al., 2003). However, elevated levels of Gal were detected in the mutant plants, as would be expected if AtCSL9A was a bona fide mannan synthase. Since the *PttGT2A* transcript is abundant at the onset of the secondary cell wall formation,

Table I. *Strictly D-specific genes encoding CAZymes: a selection of genes that passed both filter I and filter II*

Ratios are presented in log$_2$ scale, and log$_2$ ratios >1 are in bold.

Gene[a]	Sequence[a]	EST[a]	13.5 K Array[a]					3 K Array[b]	Xy versus Ph[c]
			A	B	C	D	E		
PttCesA1	FL	A064P24U	−6.41	−5.79	−0.30	**1.26**	−0.58	**Yes**	**5.66**
PttCesA1	FL	G109P08Y	−4.79	−4.41	−0.18	**1.79**	−0.70	NA	**5.30**
PttCesA3-1	FL	A040P09U	−4.77	−4.33	0.54	**1.75**	−1.64	**Yes**	**5.04**
PttCesA3-2	FL	G097P01U	−6.02	−5.45	−0.20	**1.81**	−0.32	NA	**5.27**
PttCesA3-2	FL	G103P13Y	−5.88	−5.59	−0.32	**1.34**	−0.29	NA	**5.44**
PttCesA9	Partial	G072P74Y	−6.03	−5.46	−0.08	**2.62**	−1.07	NA	**6.11**
PttCesA9	Partial	G072P58Y	−2.03	−1.70	0.03	**1.48**	0.33	NA	**4.47**
PttGT2A	Partial	B005P05U	−4.77	−4.77	−0.59	**2.06**	0.01	**Yes**	**5.89**
PttGT8B	FL	A065P67U	−5.91	−5.54	−1.32	**2.24**	−1.99	**Yes**	**6.51**
PttGT8B	FL	G069P14Y	−5.89	−5.11	−1.28	**2.35**	−1.90	NA	**6.65**
PttGT8C	FL	A073P75U	−4.74	−4.90	0.33	**2.04**	−1.61	**Yes**	**5.66**
PttGT8D	Partial	B004P07U	−6.69	−6.26	−1.65	**1.98**	−0.62	**Yes**	**6.25**
PttGT8D	Partial	G069P39Y	−4.41	−4.45	−1.48	**2.30**	−0.58	NA	**5.93**
PttGT8E	FL	G079P18Y	−1.39	−1.11	0.23	**1.98**	−2.19	NA	**2.71**
PttGT8E	FL	UB22CPE03	−0.81	−1.10	0.21	**1.95**	−2.51	NA	**3.19**
PttGT8F	Partial	G069P06Y	−2.39	−2.07	−0.74	**2.36**	−0.47	NA	**4.98**
PttGT31A	Partial	G133P32Y	−2.14	−1.97	0.86	**2.02**	−2.03	NA	3.10
PttGT43A	FL	B002P16U	−4.31	−4.25	0.33	**1.79**	−3.59	**Yes**	**4.69**
PttGT43B	FL	A027P19U	−5.05	−4.10	0.37	**1.94**	−2.32	**Yes**	**4.76**
PttGT47A	FL	A004P36U	−5.16	−5.24	−0.67	**1.25**	−2.03	**Yes**	**5.06**
PttGT47A	FL	A005P12U	−4.58	−4.75	−0.66	**1.60**	−2.05	**Yes**	**4.89**
PttGT47A	FL	A012P11U	−4.71	−4.71	−0.79	**1.61**	−1.89	**Yes**	**5.26**
PttGT47A	FL	G106P73.3pR	−4.06	−3.90	−0.57	**2.02**	−1.96	NA	**4.36**
PttGT47A	FL	G072P66.3pR	−3.36	−2.84	−0.38	**2.29**	−1.42	NA	**1.33**
PttGT47C	FL	A045P41U	−4.28	−4.03	−0.42	**1.95**	−0.20	**Yes**	**5.29**
PttGT47C	FL	A018P14U	−2.92	−2.80	−0.28	**1.95**	0.18	**Yes**	**5.12**
PttGT47C	FL	G099P90Y	−1.49	−0.95	0.23	**1.66**	−0.31	NA	**4.51**
PttGT47C	FL	I042P36P	−1.52	−1.82	0.06	**2.39**	0.15	NA	**4.18**
PttGT47D	Partial	G124P64Y	−1.10	−0.86	−0.04	**1.88**	−1.16	NA	**2.56**
PttCel9A	FL	A064P06U	−2.35	−1.94	0.02	**1.49**	0.15	**Yes**	**2.69**
PttXyn10A	FL	A020P21U	−2.53	−1.76	−0.49	**2.00**	0.60	**Yes**	**3.43**
PttGH17A	Partial	G083P33Y	−0.48	−0.65	0.10	**1.55**	−0.26	NA	−1.87
PttGH19A	FL	A044P10U	−6.73	−6.42	−1.15	0.95	−0.37	**Yes**	**6.60**
PttGH19A	FL	G066P41.5pR	−6.27	−5.75	−1.16	**1.46**	−0.74	NA	**5.93**
PttGH19A	FL	G063P44.3pR	−3.89	−3.49	−1.13	**1.41**	−0.63	NA	**4.64**
PttGH28A	Partial	A060P67U	−1.85	−1.87	−0.04	**1.55**	0.43	**Yes**	**3.52**
PttBGal35A	Partial	A066P74U	0.02	−0.20	−0.42	**1.71**	−1.58	**Yes**	**1.19**
PttGH51A	Partial	F025P61Y	−1.77	−1.80	−1.55	**2.62**	−0.73	NA	**3.49**

[a]Several genes are represented by more than one EST on the arrays. FL, Full-length sequence.   [b]Data published in Hertzberg et al. (2001). Yes/no, Gene was/was not up-regulated in the D zone; NA, not available because the EST was not represented on the array.   [c]Expression data for rows 1 to 7 have been published in Djerbi et al. (2004). Xy, Xylem; Ph, phloem.

PttGT2A might be either a mannan or a xylan synthase.

Family GT8. The 13.5 K *Populus* microarray contains 13 genes assigned to family GT8, and 8 of these (PttGT8A–8H) exhibit a high level of expression in zone D (Tables I and II). All known members of this family catalyze the formation of α-glycosidic linkages. GT8 contains enzymes with many different known substrate specificities (e.g. α-glucosyltransferase, α-galactosyltransferase, glycogenin, and galactinol synthase) as well as enzymes of unknown donor/acceptor specificity. Phylogenetic analysis of all plant GT8 sequences (Fig. 2; Lao et al., 2003) reveals several distinct subfamilies. Only one subfamily contains enzymes with known activities (galactinol synthase),

but no poplar sequence was found in this subfamily. PttGT8A to PttGT8C belong to another subfamily (Fig. 2), which contains sequences with distant similarity to animal glycogenins. Glycogenin is an autocatalytic, self-glucosylating protein that serves as the primer for glycogen synthesis. It has been proposed that a homologous enzyme in this subfamily (GenPept CAD89094) is involved in the initiation of starch biosynthesis in *Zea mays* (Chatterjee and Burrell, 2003). The Arabidopsis ortholog At3g18660 has a putative transit peptide for plastid localization, consistent with a role in starch synthesis. However, *PttGT8A* is highly expressed during the D and E zones, and both *PttGT8B* and *PttGT8C* cluster together with the secondary cell wall-specific CesAs (data not shown).

Table II. *Genes encoding CAZymes highly expressed during secondary cell wall formation: a selection of genes that passed filter I but did not fulfill the requirements of filter II*

Ratios are presented in $\log_2$ scale, and $\log_2$ ratios >1 are in bold.

Gene[a]	Sequence[a]	EST[a]	13.5 K Array[a]					3 K Array[b]	Xy versus Ph[c]
			A	B	C	D	E		
PttCesA2	FL	F120P33Y	−0.30	−0.82	0.70	**1.16**	0.87	NA	0.63
PttCesA2	FL	A020P56U	−0.34	−0.85	0.58	0.94	0.71	**No**	0.63
PttGT8A	Partial	A002P51.3pR	−3.44	−1.28	0.22	**1.75**	1.07	**Yes**	**3.75**
PttGT8A	Partial	A002P78U	−3.46	−1.38	−0.01	**1.42**	0.91	**Yes**	**3.53**
PttGT8G	Partial	A068P23U	−0.35	0.04	0.48	**1.08**	−1.39	**No**	**1.22**
PttGT8G	Partial	F056P93Y	−0.19	0.30	0.71	0.82	−0.63	NA	0.41
PttGT8H	Partial	G091P40Y	−1.02	−0.65	−0.04	**1.04**	1.84	NA	**2.34**
PttGT14A	FL	A002P61U	−1.89	−1.91	0.67	**1.60**	0.34	**Yes**	**3.91**
PttGT14B	FL	B005P04U	−2.75	−1.15	0.92	**1.34**	−1.64	**Yes**	**2.57**
PttGT31B	Partial	G108P23Y	−1.59	−1.10	0.62	**1.51**	0.20	NA	2.15
PttGT47B	FL	A006P34U	−1.51	−1.23	−0.26	**1.95**	−0.33	**Yes**	**3.25**
PttGT61A	Partial	G123P20Y	−0.41	0.14	0.85	**1.02**	0.29	NA	0.64
PttGT61A	Partial	A020P80U	−0.53	0.14	0.69	0.80	0.03	**No**	0.59
PttXET16L	Partial	UB27CPD07	0.00	−0.53	0.02	**1.42**	1.10	NA	−0.11
PttBGal35B	Partial	G079P88Y	−3.55	−3.72	0.69	**1.23**	0.94	NA	**3.34**
PttBGal35B	Partial	G107P47Y	−3.48	−4.20	0.75	**1.24**	1.07	NA	**3.16**

[a]Several genes are represented by more than one EST on the arrays.　　[b]Data published in Hertzberg et al. (2001). Yes/no, Gene was/was not up-regulated in the D zone; NA, not available because the EST was not represented on the array.　　[c]Expression data for rows 1 and 2 have been published in Djerbi et al. (2004). Xy, Xylem; Ph, phloem.

The corresponding enzymes are thus unlikely to contribute to starch biosynthesis in poplar. However, it cannot be excluded that glycogenin-like proteins may have a priming function for other cell wall polysaccharides.

PttGT8D and PttGT8G belong to another GT8 subfamily (Fig. 2). *PttGT8D* is coexpressed with the secondary cell wall-related *CesA* genes, showing a sharp peak of expression in the D zone, while *PttGT8G* is also up-regulated in the C zone. Enzymes in this subfamily are presumably linked to pectin synthesis (Bouton et al., 2002), although the precise donor/acceptor specificity remains to be established.

PttGT8E and PttGT8F are found in yet another subfamily, which contains enzymes with unknown donor/acceptor specificity but with a predicted role in pectin biosynthesis (Lao et al., 2003; Fig. 2). There is no pectin biosynthesis after the onset of secondary wall formation in wood fibers and vessel elements. However, living parenchyma cells make pectins, which participate in the formation of protective and isotropic layers as well as tyloses (Fujii et al., 1981; Rioux et al., 1998). These wall structures are formed following secondary wall deposition. Since the secondary wall is deposited earlier in vessel elements and contact ray cells (Mellerowicz et al., 2001), PttGT8D, PttGT8E, PttGT8F, and PttGT8G may be specifically involved in the pectin biosynthesis in the contact ray cells. PttGT8H, whose up-regulation is even more pronounced in the E zone than in the D zone, cannot be assigned to any of the subfamilies identified (Fig. 2).

Family GT14. Two xylem-specific GT14 members, *PttGT14A* and *PttGT14B*, were identified by expression profiling. The characterized members of family GT14 enzymes of animal origin catalyze the transfer of $\beta(1{\rightarrow}6)$-linked *N*-acetylglucosaminyl and β-linked xylosyl residues to proteins. So far, not a single plant family GT14 GT has been characterized. A phylogenetic analysis of this family shows that all the plant sequences cluster to a single subfamily, different from the subfamilies containing the animal $\beta(1{\rightarrow}6)$-*N*-acetylglucosaminyltransferases and the β-xylosyltransferases (data not shown). Whether this reflects taxonomy or enzyme specificity differences is not known and awaits functional characterization.

Family GT31. Two genes in this family, *PttGT31A and PttGT31B*, were highly expressed in the tissues undergoing secondary cell wall synthesis (zone D). Family GT31 is a large eukaryotic family grouping together a few characterized animal $\beta(1{\rightarrow}3)$-*N*-acetylglucosaminyltransferases, $\beta(1{\rightarrow}3)$-galactosyltransferases, and $\beta(1{\rightarrow}3)$-*N*-acetylgalactosaminyltransferases, together with a large number of GTs of unknown donor/acceptor specificity. Enzymes in this family form β-glycosidic linkages from α-linked nucleotide sugars. No plant homolog has been functionally characterized, but a high expression of one ortholog in Arabidopsis has been demonstrated in secondary xylem (Oh et al., 2003; Table III). The plant GT31 enzymes form distinct subfamilies, which may well coincide with more than one donor/acceptor specificity.

Family GT43. Family GT43 groups together only eukaryotic GTs. While functional (β-glucuronyltransferases) and structural data are reasonably abundant in the animal field, no function has been established experimentally for the plant homologs. Phylogenetic analysis divides plant GT43 sequences into two large

Table III. *Closest Arabidopsis homologs to the Populus CAZymes and published secondary wall-related expression data from other relevant plants*

Gene	Homolog Arabidopsis	Published Expression Data		Cotton
		Arabidopsis[a]	Zinnia	
PttGT2A CSL	At5g03760		Demura et al. (2002)	
PttCesA1	At5g44030		Milioni et al. (2002)	Pear et al. (1996)
PttCesA2	At4g39350			Li et al. (2002)
PttCesA3-1	At4g18780	S1and S4	Demura et al. (2002)	Pear et al. (1996)
PttCesA3-2	At4g18780	S1and S4	Demura et al. (2002)	Pear et al. (1996)
PttCesA9	At5g17420	S1		
PttGT8A	At4g33330			
PttGT8B	At3g18660		Demura et al. (2002)	
PttGT8C	At3g18660		Demura et al. (2002)	
PttGT8D	At5g54690			
PttGT8E	At1g19300			
PttGT8F	At1g19300			
PttGT8G	At3g02350			
PttGT8H	At3g62660			
PttGT14A	At3g03690			
PttGT14B	A1gt71070			
PttGT31A	At3g14960			
PttGT31B	At1g05170	S1	Demura et al. (2002)	
PttGT43A	At2g37090			Zhao et al. (2001)
PttGT43B	At2g37090	S1and S4		
PttGT47A	At1g27440			
PttGT47B	At3g45400			
PttGT47C	At2g28110		Demura et al. (2002)	
PttGT47D	At5g61840			
PttGT61A	At5g55500			
PttCel9A	At5g49720	S1	Demura et al. (2002)	
PttXyn10A	At1g58370		Demura et al. (2002)	
PttGH16L	At1g14720	S1		Ji et al. (2003)
PttGH17A	At4g14080			
PttGH19A	At3g16920			
PttGH28A	At3g42950			
PttBGal35A	At5g63800			Zhao et al. (2001)
PttBGal35B	At3g13750			
PttGH51A	At3g10740		Demura et al. (2002)	

[a]Data published by Oh et al. (2003). S1, Xylem up-regulated genes; S4, control up-regulated genes.

subfamilies (Fig. 3). The two poplar xylem-specific genes in this family (*PttGT43A* and *PttGT43B*) are very similar to each other and encode proteins in the same subfamily. The deduced protein sequence of a cotton gene, which is highly expressed during fiber development (GenPept AAQ54338), is a close homolog of PttGT43A and PttGT43B. Furthermore, a member of the GT43 family is highly expressed in Arabidopsis xylem (Oh et al., 2003; Table III). The transcriptional clustering of PttGT43B together with the CesAs (data not shown) also suggests a role during fiber synthesis in poplar.

Family GT47. Out of the 10 family GT47 genes represented on the *Populus* array, four genes (*PttGT47A–PttGT47D*) were highly expressed in zone D. Mechanistic conservation in the families suggests that enzymes in GT47 form glycosidic bonds of stereochemistry opposite to that of the nucleotide sugar donor. The phylogenetic tree representing all plant GT47 enzymes (Fig. 4) reveals five subfamilies similar to those established previously (Zhong and Ye, 2003; Li et al., 2004). Two of these subfamilies contain enzymes with known activities. One of them contains the recently identified pectin β-glucuronyltransferase NpGUT1 from *Nicotiana plumbaginifolia* (Iwai et al., 2002). The *nolac-H18* mutant affecting the *NpGUT1* gene leads to a weak intercellular attachment of callus, and analysis of the pectin structure of the *nolac-H18* mutant suggests that the enzyme transfers GlcUA to pectin rhamnogalacturanan II (Iwai et al., 2002). PttGT47A and PttGT47D belong to this subfamily, suggesting that they may be involved in the transfer of glucuronyl side chains to the poplar secondary wall 4-*O*-methylglucuronoxylan. Alternatively, the enzymes might synthesize rhamnogalacturonan II in ray cells that are forming the protective layer (Fujii et al., 1981). Subfamily A contains an Arabidopsis xyloglucan β-galactosyltransferase (MUR3; Madson et al., 2003),

Table IV. *Percentage identities/positives of the protein sequences of CesA proteins identified in poplar, Arabidopsis, barley, and rice*

Gene sequences were retrieved from the Stanford Web site (http://cellwall.stanford.edu). Boldface indicates genes involved in secondary cell wall synthesis (data from Taylor et al., 2000, 2003; Joshi, 2003; Tanaka et al., 2003; Burton et al., 2004; Djerbi et al., 2004). At, Arabidopsis; Ptr, *P. tremuloides*; Hv, barley; Os, rice.

Gene	P. tremula × tremuloides (Ptt)/P. trichocarpa								
	CesA1	CesA2	CesA3-1	CesA3-2	CesA4	CesA5[a]	CesA7[a]	CesA8[a]	CesA9[a]
AtCesA1	62/74	62/74	61/71	61/72	61/73	68/78	68/80	83/89	62/72
AtCesA2	59/70	74/81	57/71	56/69	77/83	62/74	63/76	61/73	61/72
AtCesA3	63/74	62/77	61/72	60/71	62/76	81/87	79/86	68/78	65/74
AtCesA4	**72/79**	58/70	64/75	62/73	57/69	61/71	62/73	61/72	61/70
AtCesA5	60/72	73/80	56/69	57/70	75/80	63/75	63/75	61/73	60/71
AtCesA6	61/73	76/84	56/69	57/70	79/85	63/75	63/75	61/73	60/71
AtCesA7	62/73	62/74	61/72	61/72	61/73	65/74	66/76	63/74	**81/86**
AtCesA8	64/74	57/70	**80/87**	**79/87**	56/69	59/70	60/72	60/71	60/71
AtCesA9	59/72	76/84	56/69	56/70	77/84	64/74	62/76	61/73	60/71
AtCesA10	60/73	60/73	60/72	59/71	60/72	66/77	67/79	79/85	60/71
PtrCesA1	63/74	56/69	**91/93**	**92/92**	55/69	57/69	59/71	59/70	59/71
PtrCesA2	62/71	62/73	60/72	58/70	61/72	65/75	66/76	63/73	**88/89**
PtrCesA3	**85/85**	57/69	63/74	60/70	57/67	60/71	63/73	61/72	62/71
PtrCesA4	61/73	63/76	60/71	60/72	61/74	68/78	67/78	94/94	63/74
PtrCesA5	60/71	62/74	58/69	57/69	63/74	90/90	80/85	67/77	64/72
PtrCesA6	61/72	66/76	60/72	58/69	66/76	66/76	66/77	65/76	63/73
PtrCesA7	60/70	76/82	55/67	55/68	88/88	65/77	65/78	62/74	62/73
HvCesA1	61/71	62/75	58/70	58/70	61/73	75/83	76/84	67/77	63/72
HvCesA2[b]	60/71	65/75	58/71	56/68	65/75	65/75	64/76	63/75	62/72
HvCesA3[b]	61/72	61/74	59/71	58/70	61/73	76/82	77/84	67/76	64/73
HvCesA4[b]	**71/76**	61/72	67/76	64/71	60/72	63/73	64/74	66/74	64/72
HvCesA5/7[b]	68/79	65/79	**78/85**	**78/86**	65/79	67/79	67/79	67/78	65/76
HvCesA6	62/74	62/74	61/72	60/71	62/73	67/77	68/79	80/86	62/72
HvCesA8[b]	68/81	72/81	71/82	68/80	71/81	78/85	78/84	74/83	**82/85**
OsCesA1	62/73	61/72	61/71	59/70	61/72	68/78	69/80	81/87	62/72
OsCesA2	62/73	63/76	59/70	59/71	63/75	78/84	78/84	68/78	65/74
OsCesA3	62/73	68/79	58/70	58/70	67/78	66/76	65/76	66/77	64/74
OsCesA4	62/73	67/81	**74/81**	**74/81**	66/80	60/72	61/73	60/71	59/70
OsCesA5	62/73	68/78	58/70	58/70	68/78	66/76	65/76	65/77	64/74
OsCesA6	61/72	66/76	58/71	57/70	66/76	66/76	65/78	64/76	62/73
OsCesA7	**70/76**	57/69	62/72	60/69	56/66	60/70	62/73	62/71	62/71
OsCesA8	61/71	63/76	59/70	58/70	62/73	78/84	78/84	68/77	64/73
OsCesA9[b]	71/83	77/86	71/84	70/83	77/85	74/84	75/85	71/82	**71/82**

[a]Full-length sequences retrieved from the genome sequence of *P. trichocarpa*.   [b]Partial sequences.

but the poplar ortholog of *mur3* was not represented on these microarrays. *PttGT47B* and *PttGT47C* belong to the uncharacterized subfamilies E and C, respectively (Fig. 4).

Family GT61A. Only one candidate in this family (*PttGT61A*) has been identified in the Populus EST database. This gene is transcribed at a high level during xylogenesis over zones B to D (Table II). Family GT61 is a eukaryotic family with diverse plant, animal, and fungal proteins. In contrast with most GT families, only plant-derived members of this family have been characterized. The phylogenetic tree of plant GT61 enzymes reveals three subfamilies (Fig. 5). PttGT61A belongs to a subfamily that contains $\beta(1\rightarrow2)$-xylosyl-transferases from Arabidopsis, rice, and Physcomitrel-la. These enzymes catalyze the transfer of Xyl from UDP-Xyl to the core β-linked Man of *N*-linked oligo-saccharides of glycoproteins (Strasser et al., 2000).

Consequently, PttGT61A probably has no direct role in the synthesis of structural carbohydrates, but its presence among the secondary wall-specific genes reinforces the observation that correct protein glyco-sylation is required for efficient cellulose synthesis (Lukowitz et al., 2001; Williamson et al., 2002).

Glycosidases and Transglycosylases

Family GH9. Only one member of this family, *PttCel9A*, was up-regulated during secondary cell wall formation in Populus. The family GH9 enzymes are typically endoglucanases acting on $\beta(1\rightarrow4)$-glucan polymers (cellulose). Genes encoding GH9 enzymes have been found in both prokaryotic and eukaryotic organisms with 25 putative family GH9 members identified in the Arabidopsis genome. PttCel9A contains an N-terminal cytoplasmic tail

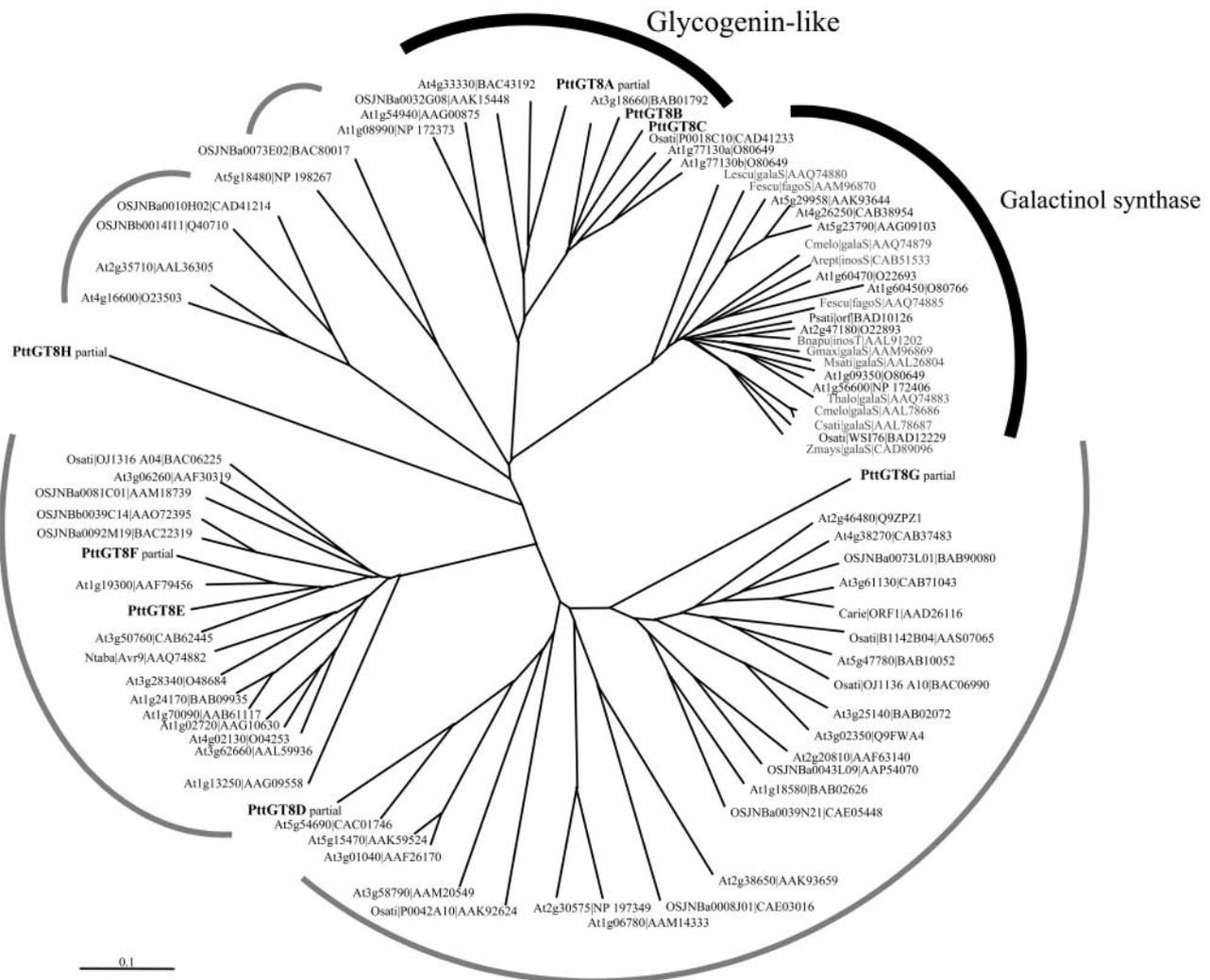**Figure 2.** Phylogenetic analysis of plant GTs from family GT8. Profile-aligned Populus sequences are in bold, and partial sequences are indicated. Subfamilies with at least one characterized enzyme are outlined by a thick black line and those without defined activities by a thin gray line.

followed by a membrane anchor and a C-terminal catalytic domain (Rudsander et al., 2003). The deduced protein sequence of *PttCel9A* shares 82% sequence identity with the Arabidopsis enzyme KOR1, which was identified by the cellulose-deficient dwarf mutant *Korrigan* (Nicol et al., 1998). KOR1 protein is abundant during intensive cellulose biosynthesis, and the corresponding gene is highly expressed during secondary cell wall synthesis (Szyjanowicz et al., 2004). KOR1 has been proposed to release nascent cellulose chains from a putative sitosterol-glucoside primer (Peng et al., 2002) or to contribute in cellulose microfibril assembly or editing (Molhoj et al., 2002). Recombinant variants of Cel9A from *Brassica napus* (Molhoj et al., 2001) and hybrid aspen (Master et al., 2004) have been expressed in *Pichia pastoris* without the cytoplasmic and membrane-spanning domains. The purified proteins hydrolyzed low-substituted, soluble $\beta(1\rightarrow4)$-glucan but not hemicelluloses. The specific activity of the recombinant

PttCel9A was orders of magnitude lower than that of corresponding microbial enzymes (Master et al., 2004), suggesting that efficient cellulose degradation is not the primary role of the enzyme.

Family GH10. One putative Populus xylanase gene, *PttXyn10*, was highly up-regulated in zone D. Most known family GH10 enzymes are endo-$\beta(1\rightarrow4)$-xylanases found in many different microbes and plants. The Arabidopsis genome contains 12 family GH10 genes (http://afmb.cnrs-mrs.fr/CAZY/). In addition to a catalytic module, other functional modules are commonly found in family 10 xylanases. PttXYN10 is similar to AtXyn1, which is predominantly expressed in vascular bundles (Suzuki et al., 2002). Full-length sequencing reveals three consecutive family CBM22 carbohydrate-binding modules at the N terminus of the predicted protein sequence. It seems thus that the *PttXYN10* gene encodes an endoxylanase, which may be involved in xylan hydrolysis or remodeling during xylogenesis.
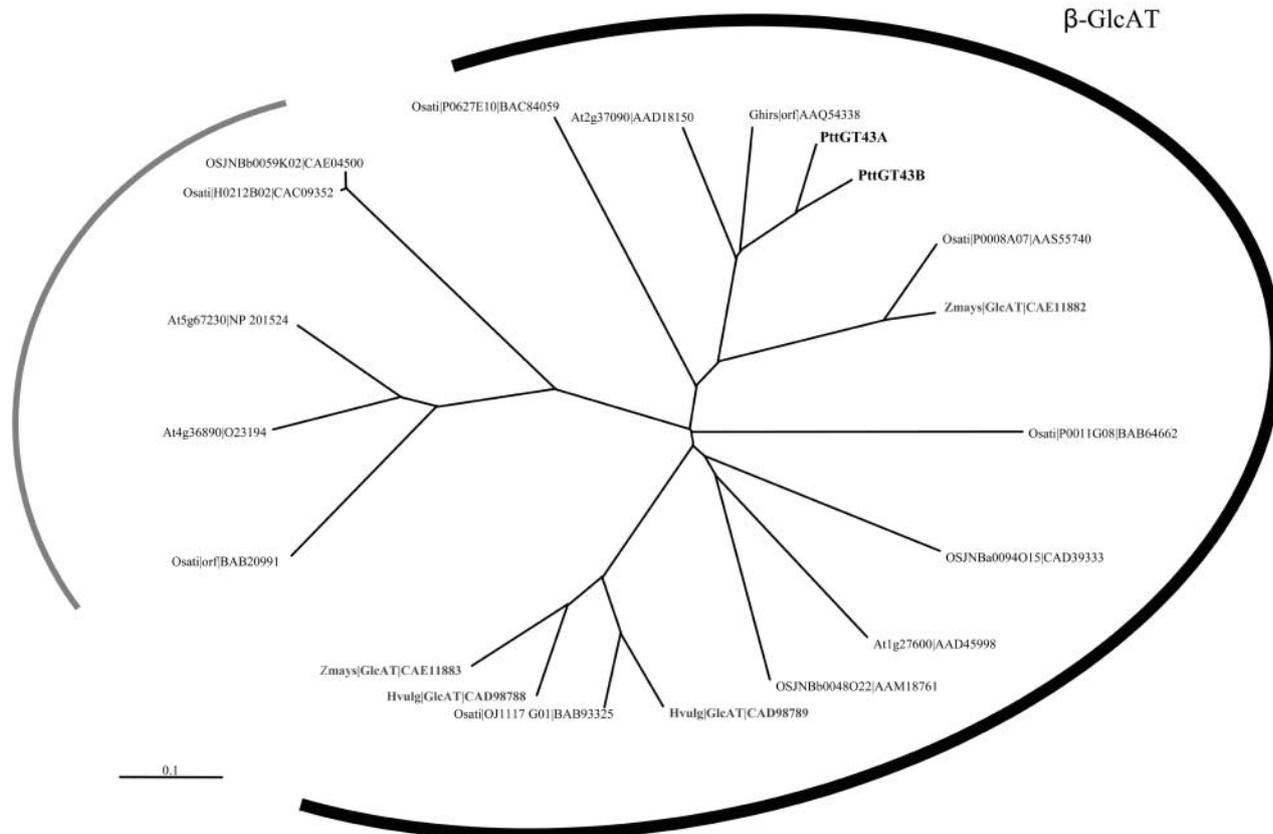
Figure 3. Phylogenetic analysis of plant GTs from family GT43. Legend as in Figure 2.

Family GH16. Close to 30 different family GH16 genes have been identified in the PopulusDB, and 16 of them were found in wood-forming tissues (N. Nishikubo and E. Mellerowicz, unpublished data). This family of enzymes groups together prokaryotic and eukaryotic enzymes of different substrate specificities, but all of the plant members characterized so far act exclusively on xyloglucan (Rose et al., 2002). Some plant family GH16 enzymes, including the hybrid aspen PttXET16A (Å. Kallas, S. Denman, K. Piens, H. Henriksson, P. Johansson, T.A. Jones, H. Brumer, and T.T. Teeri, unpublished data), are strict xyloglucan endotransglycosylases (XET; E.C. 2.4.1.207), whereas others have varying degrees of xyloglucan endohydrolase (XH; E.C. 3.2.1.151) activity (Rose et al., 2002). Only one member of GH16, *PttXET16L*, exhibited a high level of transcription during secondary cell wall formation in poplar. *PttXET16L* is most similar to *At1g14720* (*XTH28*, also known as *XTR2*). Constitutive expression of *XTR2* has been reported during various conditions in Arabidopsis (Xu et al., 1995; Akamatsu et al., 1999), with specific up-regulation in the developing secondary xylem (Oh et al., 2003).

Although xyloglucan is not supposed to be an abundant component of secondary walls, XET activity has been demonstrated recently in xylem and phloem fibers at an early stage of secondary cell wall formation (Bourquin et al., 2002). Xyloglucan incorporation into primary walls of fibers is probably completed at this stage, but xyloglucan degradation has been detected in xylogenic Zinnia cultures during secondary wall deposition (Ohdaira et al., 2002). *PttXET16L* belongs to the subfamily 3 of *XTH* genes, which also contains the putative hydrolytic xyloglucan endotransglucosylases/hydrolases (XTH) from Nasturtium seeds (Rose et al., 2002). It is thus possible, although not yet experimentally demonstrated, that these enzymes contribute to xyloglucan degradation by endohydrolytic activity.

Family GH17. Only one member from the family GH17, *PttGH17*, was found among the secondary cell wall-associated genes of hybrid aspen. This family contains enzymes able to degrade mainly $\beta(1\rightarrow3)$- or $\beta(1\rightarrow3)(1\rightarrow4)$-glucans. They have been well characterized in plants, both at the biochemical and at the three-dimensional structural level. The D-specific cDNA fragment (*PttGT17A*) shows high similarity with a $\beta(1\rightarrow3)$-glucanase (SgGN) from willow (*Salix gilgiana* Seemen; Futamura et al., 2000). Both *PttGT17A* and *SgGN* show high similarity with anther-specific $\beta(1\rightarrow3)$-glucanases from Arabidopsis and Brassica (Hird et al., 1993). Some $\beta(1\rightarrow3)$-glucanases have a glycosylphosphatidylinositol anchor, but the Arabidopsis homologs to the poplar *PttGH17* lack this feature (Borner et al., 2002). About one-half of the 49 GH17 family members contained in the Arabidopsis genome are appended to an auxiliary module of unknown
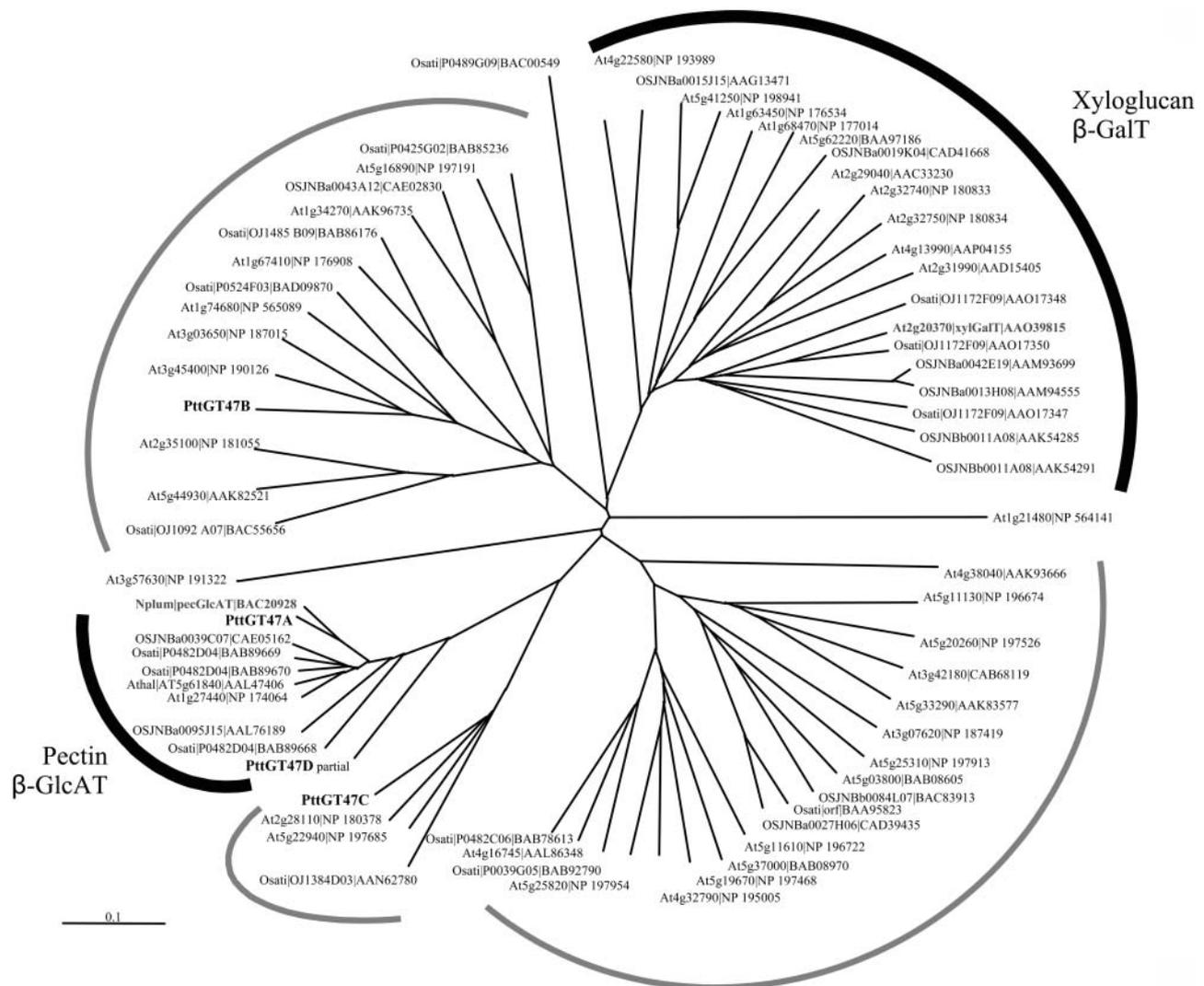
Figure 4. Phylogenetic analysis of plant GTs from family GT47. Legend as in Figure 2.

function denoted X8 (Henrissat and Davies, 2000). However, the D-specific poplar *PttGT17A* gene does not appear to include an X8 module. Recently, it has been established that callose [$\beta(1 \rightarrow 3)$-glucan] occurs as a regular feature in the exoplasmic zone of cotton fiber secondary walls, not just as a result of a stress response (Salnikov et al., 2003). However, these expression data indicate higher expression of *PttGH17A* in phloem than xylem (Table I). Since the sieve areas of phloem sieve tubes are particularly rich in callose, PttGH17 may indeed have a role in callose degradation.

Family GH19. One family GH19 gene, *PttGH19A*, was picked in the screen for genes highly expressed during poplar secondary cell wall synthesis. The only function established so far for members in this family is the hydrolysis of chitin, and plant chitinases are a component of the defense system against fungal infection. However, xylem is unlikely to be the site of primary attack by fungi. The molecular mechanism and the three-dimensional structure of several GH19

chitinases have been determined (Hart et al., 1995; Hahn et al., 2000). Two Glu perform catalysis, one acting as a general acid and the other as a base to activate a water molecule for the single-step displacement reaction. Close homologs to *PttGH19A* have been identified in other plants, two from Arabidopsis (At3g16920 and At1g05850) and cotton (GenPept AAP80801 and AAP80801) as well as one from rice (GenPept BAC81645), maize (*Zea mays*; GenPept AAQ31297), and pea (*Pisum sativum*; GenPept BAC81645). Interestingly, in all of them, including PttGH19A, the catalytic acid seems to be missing.

In Arabidopsis, the mutation *ELP1* defective in the At1g05850 gene (also denoted *AtCTL*) caused ectopic deposition of lignin and aberrant shapes of cells with incomplete cell walls in the pith of inflorescence stems. Ethylene production also was enhanced in this mutant. High expression of the *AtCTL* gene was detected in seedlings, stems, and flowers during normal plant growth and development. The expression of this gene
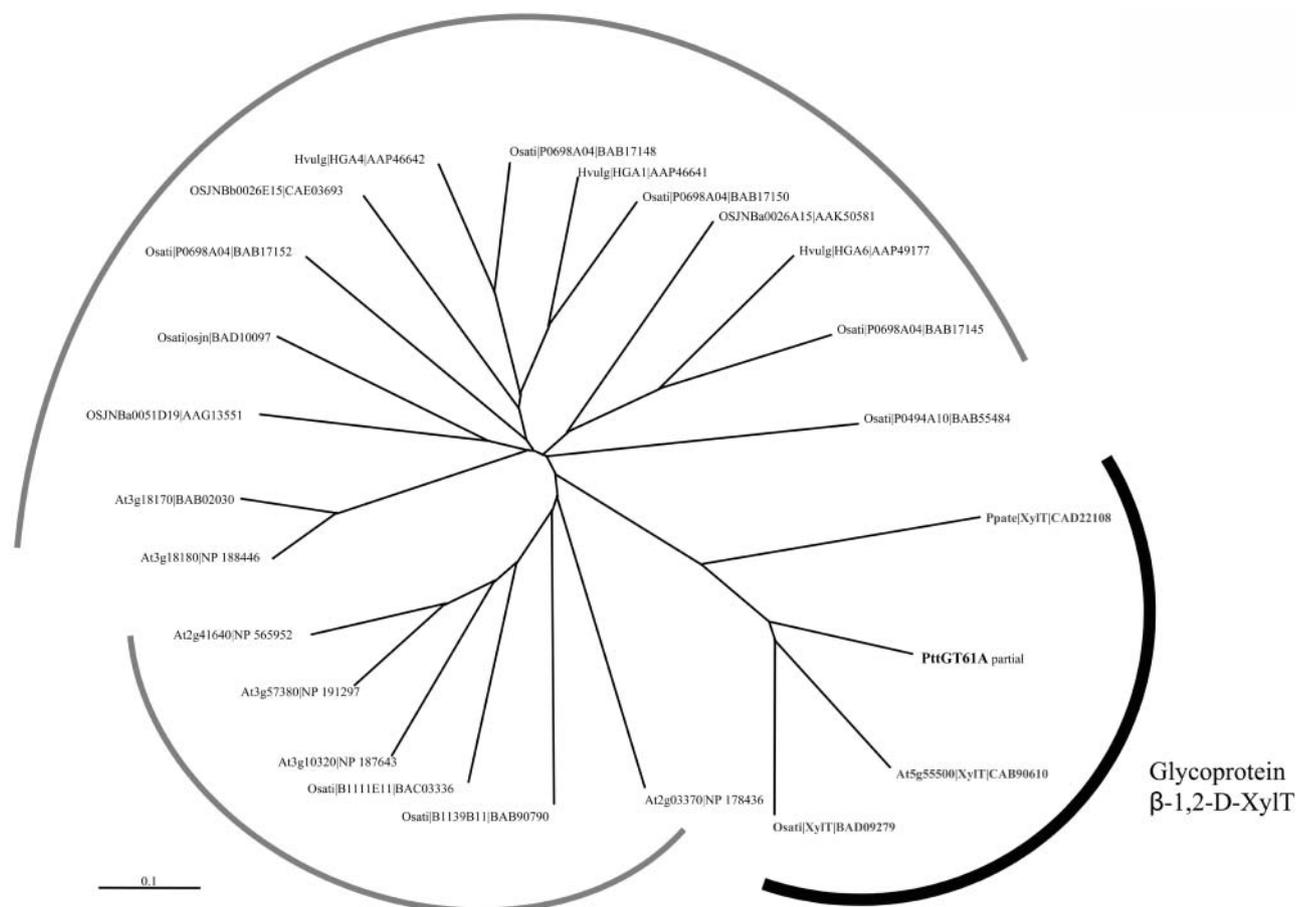
Figure 5. Phylogenetic analysis of plant GTs from family GT61. Legend as in Figure 2.

was not influenced by wounding or induction by salicylic acid, pectin fragments, or ethylene (Zhong et al., 2002). Another Arabidopsis mutant, *POM1*, is apparently identical to *ELP1*. In a dendrogram obtained by hierarchical cluster analysis of Arabidopsis lines based on their Fourier transform infrared spectra, the two *pom1* alleles clustered together with other cell wall mutants with reduced cellulose content in the walls (Mouille et al., 2003; Pilling and Hofte, 2003). Similar to PttGH19A, a chitinase-like gene, also lacking the apparent catalytic acid, is preferentially expressed in fibers during secondary wall deposition in cotton (Zhang et al., 2004). Different hypotheses can be offered to explain the apparent lack of an essential catalytic residue in these plant proteins. It seems most likely these proteins have acquired another function that does not depend on chitinase or any enzymatic activity. The acquisition of novel function by plant glycosidase scaffolds has been discussed by Coutinho et al. (2003b), who have noted that there is no obvious relationship between the ancestral and the novel, acquired function. Another possibility is that this subset of the GH19 proteins includes enzymes relying on an externally recruited acid/base, as described for an *S*-glycosidase, myrosinase (Burmeister et al., 2000), or one that uses substrate assistance from a ligand-

donated acid. However, further functional analyses in vivo and in vitro are required to uncover the specific function of these xylem-specific proteins.

Family GH28. Family GH28 contains polygalacturonases (PGs), which are among the most important enzymes associated with pectin degradation and cell wall modification in the late phases of fruit ripening, in organ abscission, in pod dehiscence, and in pollen maturation (Dal Degan et al., 2001). Among several PG-like genes of family GH28 in the EST database PopulusDB, only one (*PttGH28A*) exhibited a high level of expression in zone D. Other examples of putative secondary cell wall-associated endo-PGs have been identified in differentiating tracheary elements of Zinnia (Demura et al., 2002; Nakashima et al., 2004) and in differentiating vascular tissue in tomato (*Lycopersicon esculentum*; Sitrit et al., 1999). It has been suggested that PGs may participate in the primary wall degradation between secondary thickenings of primary xylem (Ohdaira et al., 2002) and in the regions where future perforations develop in secondary xylem vessel elements (for review, see Mellerowicz et al., 2001). Pits, on the other hand, retain a thin layer of their primary wall (pit membrane) that probably undergoes some form of wall modification before cell autolysis. Pit membranes contain high levels of pectin

(Zwieniecki et al., 2001), which could act as a substrate for the xylem-specific PGs.

Family GH35. Two D-specific genes (*PttBGal35A* and *PttBGal35B*) in family GH35 were also identified. This family is one of the few isofunctional families in the CAZy database, as it contains only β-galactosidases from plants, animals, fungi, bacteria, and archaea. It is thus likely that all plant homologs are β-galactosidases. In plants, these enzymes form a multigene family with 18 copies in Arabidopsis, and the characterized members act on various substrates, including arabinogalactans, galactolipids, and pectin to release Gal (Smith and Gross, 2000; Esteban et al., 2003). A particular feature of plant GT35 members is that about one-half of them carry a C-terminal extension distantly related to animal rhamnose-binding lectins, suggesting that this extension is a carbohydrate-binding module. Out of the two genes that we have identified, *PttBGal35A* and *PttBGal35B*, only *PttBGal35B* encodes this C-terminal putative binding domain. β-Galactosidases might play a role in pit membrane modification and hydrolysis of primary walls in the perforations, as discussed earlier, by degrading galactan side chains of rhamnogalacturonan I, or xyloglucan. Alternatively, they could act on arabinogalactan antenna of the arabinogalactan proteins (AGPs).

Family GH51. Only one secondary cell wall-related gene was found in family GH51 (*PttGH51*). Family GH51 is isofunctional and all characterized members, bacterial and eukaryotic, have α-L-arabinofuranosidase activity. α-L-Arabinofuranosidases are accessory enzymes releasing Ara from substrates such as arabinan, arabinoxylan, gum arabic, and arabinogalactan (Saha, 2000). PttGH51A is very similar to α-L-arabinofuranosidases from apple (*Malus domestica*), tomato (LeARF1), Japanese pear (*Pyrus pyrifolia*), and Arabidopsis (AtASD1 and AtASD2).

Since Ara is rare in poplar wood, a likely function for α-L-arabinofuranosidase during the secondary cell wall formation is modification of AGPs. Another possibility is that arabinofuranosidase participates in the pectin degradation in the defined cell wall areas, such as developing perforations or pit membranes. A similar expression pattern of other pectin-degrading enzymes, such as β-galactosidases and PGs, suggests degradation of homogalacturonan and rhamnogalacturonan I during the secondary wall deposition. In support of the latter hypothesis are the data on sugars released to the medium by a xylogenetic Zinnia culture at the stage of secondary cell wall formation (Stacey et al., 1995; Ohdaira et al., 2002). Linkage analysis indicated a degradation of homogalacturonan and side chains of rhamnogalacturonan I (arabinan and galactan) rather than a degradation of AGP antenna.

Other Secondary Cell Wall-Specific Enzymes and Proteins

Interestingly, some of the enzymes known to be involved in xylogenesis, such as Suc synthases in family GT4 (Babb and Haigler, 2001), were not secondary wall-specific in poplar. In our earlier work (Hertzberg et al., 2001), one Suc synthase was up-regulated in both zones C and D, while this dataset only showed up-regulation in zone C.

Hemicelluloses, such as xyloglucan (Pauly and Scheller, 2000), glucuronoxylan (Timell, 1967), glucomannan (Teleman et al., 2003), and pectins (Pauly and Scheller, 2000), are often *O*-acetylated. It has been shown that acetyl-CoA is the donor substrate for *O*-acetylation of cell wall polysaccharides (Pauly and Scheller, 2000). Recent results show a positive correlation between xyloglucan *O*-acetylation and fucosylation, suggesting that a suitable substrate for at least one Arabidopsis *O*-acetyltransferase is fucosylated xyloglucan (Perrin et al., 2003). Cas1p is a membrane protein required for the *O*-acetylation of glucuronoxylomannan in the fungal pathogen *Cryptococcus neoformans*, and four orthologs were found in Arabidopsis (Janbon et al., 2001). By performing a TBLASTN search with the Cas1p sequence against the PopulusDB, a gene coding for a putative *O*-acetyltransferase was identified. The transcript of the Populus putative *O*-acetyltransferase is highly expressed during zone D, which is consistent with a role in *O*-acetylation of the secondary wall polysaccharides glucomannan and glucuronoxylan. Nevertheless, functional characterization of this gene and its products is required to verify or discard its hypothesized role in *O*-acetylation.

Another interesting class of proteins is those associated with cortical microtubules, which have been suggested to facilitate the alignment of cellulose microfibrils as they are deposited in the cell wall (Baskin, 2001). Notably, among 14 different tubulin genes present in the cambial cDNA library of poplar, 10 were very highly expressed during late expansion (C) and secondary wall biosynthesis (D; data not shown).

The rest of the xylem-specific genes encoded many proteins of unknown function (data not shown). Some of these genes encode proteins with putative transmembrane domains or glycosylphosphatidylinositol anchors, and it is possible that new GTs remain to be identified in this group. Another abundant group of unknown genes seems to encode proteins containing domains involved in DNA-DNA or DNA-protein interactions and signaling events.

In summary, this functional genomic approach allowed the identification of 25 GTs in 7 different GT families and 9 GHs from 8 different GH families in poplar xylem. Even though the precise substrate specificity currently can be predicted only for a few of these enzymes, the assignments to CAZy families performed here help to narrow future functional studies to a limited number of enzymes and polysaccharides. Finally, the remaining collection of xylem-specific genes of entirely unknown function offers a rich source of new targets for comprehensive studies of the secondary cell wall biosynthesis in plants.

MATERIALS AND METHODS

Microarray Analysis

Construction of the Populus 13.5 K microarray containing cDNA fragments, printed in duplicate, from seven tissue-specific libraries has been described by Andersson et al. (2004). Complementary DNA for microarray hybridization was obtained by reamplifying material from a study by Hertzberg et al. (2001), consisting of five tissue sections over the developing xylem, i.e. meristematic cells (A), early expansion (B), late expansion (C), secondary cell wall formation (D), and programmed cell death (E). Because there is no obvious reference tissue section in developing xylem, a mixture of equal amounts of the individual samples (A + B + C + D + E) was used as a common reference. This ensures that genes expressed in only one tissue will be represented in the reference, thereby decreasing the risk of losing such genes in the quality filter process. However, the transcript profiles of genes that are highly expressed in one sample will be truncated because that sample comprises one-fifth of the reference.

All amplification and hybridization steps were carried out as described previously (Hertzberg et al., 2001), except that each Cy3-labeled sample was hybridized twice (A–C) or three times (D and E) against a Cy5-labeled pool of all samples. The slides were scanned at 5-μm resolution using ScanArray 4000/ScanArray Lite scanners (Perkin-Elmer, Boston). The images were quantified and analyzed using GenePix Pro 5.0 (Axon Instruments, Union City, CA). Spot intensity was computed as a median foreground minus median background signal. A quality filter was applied where the following conditions classified a spot as bad: (1) The spot was manually flagged as bad; (2) the spot intensity (background subtracted) was lower than 2 times a global background estimate (calculated as the median of mean background pixel intensities) in one or both channels; and (3) more than 30% of the pixels were saturated in both channels. If one or more of these criteria were met, the spot was flagged and excluded from further analysis. The data were normalized to a median of 1 for the ratios of all genes by using the intensity-dependent locally weighted linear regression (Lowess) procedure as implemented in GeneSpring 6.1 (Silicon Genetics, Redwood City, CA). Intensities of within-slide replicates were averaged before calculating ratios. The complete dataset from this hybridization has been published elsewhere (Schrader et al., 2004).

The secondary cell wall-specific genes were selected for this study based on the following criteria: at least two valid data points in each of the sections, C to E, and a signal at least 2-fold higher in D than in the common reference channel (a pool of tissues A–E). A second, more stringent selection required the signal in D to be >2-fold higher than in any of the other samples. Nine clusters were generated for the more stringent selection using self-organizing maps in GeneSpring 6.1 (Silicon Genetics). Hybridization of the 2 × 13.5 K microarray with targets prepared from the developing xylem and cambium/phloem tissues has been described in detail by Djerbi et al. (2004).

Full-Length Cloning

The sequences available for bioinformatic analysis consisted of ESTs, contig sequences (Sterky et al., 2004; http://www.populus.db.umu.se), and full-length cDNA sequences, as indicated in Tables I and II. For each identified CAZyme, the longest available sequence was used for the detailed modular annotation. For full-length sequencing of the cDNA clones, the insert of the cDNA clone corresponding to the EST sequence was sequenced on ABI 3700 sequencers (Applied Biosystems, Foster City, CA) by a primer-walking strategy. In most cases, the complete open reading frame could be revealed, but some gene sequences remained partial. Full-length sequences for a subset of genes were obtained by using the FirstChoice RLM-RACE kit (Ambion, Austin, TX). Initial cambial and xylem RNA was extracted by the RNeasy plant mini kit (Qiagen, Valencia, CA). The 5′ sequences were cloned in silico with the previously partial sequences to obtain full-length cDNA sequences. For a selection of genes, the full-length cDNA was cloned from cDNA synthesized from hybrid aspen cambial RNA by using an oligoT primer followed by PCR amplification with gene-specific primers. The products from the PCR reactions were cloned into a pGEM-T easy vector (Promega, Madison, WI) and sequenced.

Sequence Analysis and Data Storage

All Populus ESTs, contig sequences, and full-length sequences were compared against an annotated sequence library derived from the CAZy database (29,442 entries as of October 1, 2004). The CAZy database classification assigns nonoverlapping segments of each protein sequence to one of the different families of enzymes (presently including GHs and transglycosylases, GTs, polysaccharides lyases, and carbohydrate esterases). Because CAZymes are frequently modular, we have undertaken a systematic analysis and assignment of the modular structure of these enzymes (Coutinho and Henrissat, 1999). More than 56,000 modules arranged in 360 families have already been defined. Sequence segments corresponding to these modules are compiled into a library that is used for updating the CAZy database with new sequences by BLAST-based analyses (Altschul et al., 1997). The assignment of the ESTs reported here was performed following exactly the same procedure as that used for updating CAZy. For low levels of sequence identity, assignments were verified by the presence of known family signatures. Occasionally, hydrophobic cluster analysis (Callebaut et al., 1997) was used to help refine the boundaries of distantly related modules. Because CAZymes often display variable modular structures, only matches to catalytic modules were retained.

Phylogenetic Analyses

An analysis of the evolutionary relationships between plant CAZymes was performed in a three-step procedure. First, the catalytic domains from full sequences of plant origin belonging to relevant CAZy families were excised and aligned using Muscle 3.0 (Edgar, 2004). Then this initial alignment served as a profile for the alignment of all EST-derived sequences and other relevant incomplete sequences using the profile mode of ClustalW 1.83 (Thompson et al., 1994). Finally, phylogenetic trees were calculated from the resulting alignments from ClustalW using 1,000 bootstrap steps. The display of the trees was made using Treeview 1.6.6 (Page, 1996).

LITERATURE CITED

Akamatsu T, Hanzawa Y, Ohtake Y, Takahashi T, Nishitani K, Komeda Y (1999) Expression of endoxyloglucan transferase genes in acaulis mutants of Arabidopsis. Plant Physiol **121:** 715–721

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389–3402

Andersson A, Keskitalo J, Sjodin A, Bhalerao R, Sterky F, Wissel K, Tandre K, Aspeborg H, Moyle R, Ohmiya Y, et al (2004) A transcriptional timetable of autumn senescence. Genome Biol **5:** R24

Babb VM, Haigler CH (2001) Sucrose phosphate synthase activity rises in correlation with high-rate cellulose synthesis in three heterotrophic systems. Plant Physiol **127:** 1234–1242

Balatinecz JJ, Kretschmann DE, Leclerc A (2001) Achievements in the utilization of poplar wood—guideposts for the future. For Chron **77:** 265–269

Baskin TI (2001) On the alignment of cellulose microfibrils by cortical microtubules: a review and a model. Protoplasma **215:** 150–171

Bhalerao R, Keskitalo J, Sterky F, Erlandsson R, Bjorkbacka H, Birve SJ, Karlsson J, Gardestrom P, Gustafsson P, Lundeberg J, et al (2003) Gene expression in autumn leaves. Plant Physiol **131:** 430–442

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

Borner GHH, Sherrier DJ, Stevens TJ, Arkin IT, Dupree P (2002) Prediction of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A genomic analysis. Plant Physiol 129: 486–499

Bourquin V, Nishikubo N, Abe H, Brumer H, Denman S, Eklund M, Christiernin M, Teeri TT, Sundberg B, Mellerowicz EJ (2002) Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues. Plant Cell 14: 3073–3088

Bouton S, Leboeuf E, Mouille G, Leydecker MT, Talbotec J, Granier F, Lahaye M, Hofte H, Truong HN (2002) Quasimodo1 encodes a putative membrane-bound glycosyltransferase required for normal pectin synthesis and cell adhesion in Arabidopsis. Plant Cell 14: 2577–2590

Burmeister WP, Cottaz S, Rollin P, Vasella A, Hnerissat B (2000) High resolution x-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. J Biol Chem 275: 39385–39393

Burton RA, Shirley NJ, King BJ, Harvey AJ, Fincher GB (2004) The CesA gene family of barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes. Plant Physiol 134: 224–236

Callebaut I, Labesse G, Durand P, Poupon A, Canard L, Chomilier J, Henrissat B, Mornon JP (1997) Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. Cell Mol Life Sci 53: 621–645

Campbell JA, Davies GJ, Bulone V, Henrissat B (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem J 326: 929–939

Chatterjee M, Burrell MM (2003) Starch modification in transgenic plants using glycogenin-like genes from Arabidopsis and other plants. International PCT Application No. WO03014365

Coutinho PM, Deleury E, Davies GJ, Henrissat B (2003a) An evolving hierarchical family classification for glycosyltransferases. J Mol Biol 328: 307–317

Coutinho PM, Henrissat B (1999) Carbohydrate-active enzymes: an integrated database approach. In HJ Gibert, G Davies, B Henrissat, B Svensson, eds, Recent Advances in Carbohydrate Bioengineering. The Royal Society of Chemistry, Cambridge, pp 3–12

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

Dal Degan F, Child R, Svendsen I, Ulvskov P (2001) The cleavable N-terminal domain of plant endopolygalacturonases from clade B may be involved in a regulated secretion mechanism. J Biol Chem 276: 35297–35304

Demura T, Tashiro G, Horiguchi G, Kishimoto N, Kubo M, Matsuoka N, Minami A, Nagata-Hiwatashi M, Nakamura K, Okamura Y, et al (2002) Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells. Proc Natl Acad Sci USA 99: 15794–15799

Dhugga KS, Barreiro R, Whitten B, Stecca K, Hazebroek J, Randhawa GS, Dolan M, Kinney AJ, Tomes D, Nichols S, et al (2004) Guar seed beta-mannan synthase is a member of the cellulose synthase super gene family. Science 303: 363–366

Djerbi S, Aspeborg H, Nilsson P, Sundberg B, Mellerowicz EJ, Blomqvist K, Teeri TT (2004) Identification and expression analysis of genes encoding putative cellulose synthases (CesA) in the hybrid aspen, Populus tremula (L.) × P. tremuloides (Michx.). Cellulose 11: 301–312

Edgar R (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5: 113

Esteban R, Dopico B, Munoz FJ, Romo S, Martin I, Labrador E (2003) Cloning of a Cicer arietinum beta-galactosidase with pectin-degrading function. Plant Cell Physiol 44: 718–725

Fujii T, Harada H, Saiki H (1981) Ultrastructure of ''amorphous layer'' in xylem parenchyma cell wall of angiosperm species. Mokuzai Gakkaishi 27: 149–156

Futamura N, Mori H, Kouchi H, Shinohara K (2000) Male flower-specific expression of genes for polygalacturonase, pectin methylesterase and beta-1,3-glucanase in a dioecious willow (Salix gilgiana Seemen). Plant Cell Physiol 41: 16–26

Gustavsson M, Bengtsson M, Gatenholm P, Glasser W, Teleman A, Dahlman O (2001) Isolation, characterisation and material properties of 4-O-methylglucuronoxylan from aspen. In Biorelated Polymers:

Sustainable Polymer Science and Technology. Combined Proceedings of the International Conference on Biopolymer Technology, 1st (Coimbra, Portugal) and 2nd (Naples, Italy), pp 41–52

Hahn M, Hennig M, Schlesier B, Hohne W (2000) Structure of jack bean chitinase. Acta Crystallogr D Biol Crystallogr 56: 1096–1099

Hart PJ, Pfluger HD, Monzingo AF, Hollis T, Robertus JD (1995) The refined crystal-structure of an endochitinase from Hordeum vulgare L. seeds at 1.8 A resolution. J Mol Biol 248: 402–413

Henrissat B, Coutinho PM, Davies GJ (2001) A census of carbohydrate-active enzymes in the genome of Arabidopsis thaliana. Plant Mol Biol 47: 55–72

Henrissat B, Davies GJ (2000) Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. Plant Physiol 124: 1515–1519

Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlen M, Teeri TT, Lundeberg J, et al (2001) A transcriptional roadmap to wood formation. Proc Natl Acad Sci USA 98: 14732–14737

Hird DL, Worrall D, Hodge R, Smartt S, Paul W, Scott R (1993) The anther-specific protein encoded by the Brassica napus and Arabidopsis thaliana A6 gene displays similarity to beta-1,3-glucanases. Plant J 4: 1023–1033

Iwai H, Masaoka N, Ishii T, Satoh S (2002) A pectin glucuronyltransferase gene is essential for intercellular attachment in the plant meristem. Proc Natl Acad Sci USA 99: 16319–16324

Janbon G, Himmelreich U, Moyrand F, Improvisi L, Dromer F (2001) Cas1p is a membrane protein necessary for the O-acetylation of the Cryptococcus neoformans capsular polysaccharide. Mol Microbiol 42: 453–467

Ji SJ, Lu YC, Feng JX, Wei G, Li J, Shi YH, Fu Q, Liu D, Luo JC, Zhu YX (2003) Isolation and analyses of genes preferentially expressed during early cotton fiber development by subtractive PCR and cDNA array. Nucleic Acids Res 31: 2534–2543

Joshi CP (2003) Xylem-specific and tension stress-responsive expression of cellulose synthase genes from aspen trees. Appl Biochem Biotechnol 108: 17–25

Lao NT, Long D, Kiang S, Coupland G, Shoue DA, Carpita NC, Kavanagh TA (2003) Mutation of a family 8 glycosyltransferase gene alters cell wall carbohydrate composition and causes a humidity-sensitive semi-sterile dwarf phenotype in Arabidopsis. Plant Mol Biol 53: 687–701

Li CH, Zhu YQ, Meng YL, Wang JW, Xu KX, Zhang TZ, Chen XY (2002) Isolation of genes preferentially expressed in cotton fibers by cDNA filter arrays and RT-PCR. Plant Sci 163: 1113–1120

Li XM, Cordero I, Caplan J, Molhoj M, Reiter WD (2004) Molecular analysis of 10 coding regions from Arabidopsis that are homologous to the MUR3 xyloglucan galactosyltransferase. Plant Physiol 134: 940–950

Lorenz WW, Dean JFD (2002) SAGE profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (Pinus taeda). Tree Physiol 22: 301–310

Lukowitz W, Nickle TC, Meinke DW, Last RL, Conklin PL, Somerville CR (2001) Arabidopsis cyt1 mutants are deficient in a mannose-1-phosphate guanylyltransferase and point to a requirement of N-linked glycosylation for cellulose biosynthesis. Proc Natl Acad Sci USA 98: 2262–2267

Madson M, Dunand C, Li XM, Verma R, Vanzin GF, Calplan J, Shoue DA, Carpita NC, Reiter WD (2003) The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. Plant Cell 15: 1662–1670

Master ER, Rudsander UJ, Zhou W, Henriksson H, Divne C, Denman S, Wilson DB, Teeri TT (2004) Recombinant expression and enzymatic characterization of PttCel9A, a KOR homologue from Populus tremula x tremuloides. Biochemistry 43: 10080–10089

Mellerowicz EJ, Baucher M, Sundberg B, Boerjan W (2001) Unravelling cell wall formation in the woody dicot stem. Plant Mol Biol 47: 239–274

Molhoj M, Pagant SR, Hofte H (2002) Towards understanding the role of membrane-bound endo-beta-1,4-glucanases in cellulose biosynthesis. Plant Cell Physiol 43: 1399–1406

Molhoj M, Ulvskov P, Dal Degan F (2001) Characterization of a functional soluble form of a Brassica napus membrane-anchored endo-1,4-beta-glucanase heterologously expressed in Pichia pastoris. Plant Physiol 127: 674–684

Mouille G, Robin S, Lecomte M, Pagant S, Hofte H (2003) Classification and identification of Arabidopsis cell wall mutants using Fourier-Transform InfraRed (FT-IR) microspectroscopy. Plant J 35: 393–404

Nakashima J, Endo S, Fukuda H (2004) Immunocytochemical localization of polygalacturonase during tracheary element differentiation in *Zinnia elegans*. Planta 218: 729–739

Nam J, Mysore KS, Zheng C, Knue MK, Matthysse AG, Gelvin SB (1999) Identification of T-DNA tagged Arabidopsis mutants that are resistant to transformation by Agrobacterium. Mol Gen Genet 261: 429–438

Nicol F, His I, Jauneau A, Vernhettes S, Canut H, Hofte H (1998) A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in Arabidopsis. EMBO J 17: 5563–5576

Oh S, Park S, Han KH (2003) Transcriptional regulation of secondary growth in *Arabidopsis thaliana*. J Exp Bot 54: 2709–2722

Ohdaira Y, Kakegawa K, Amino S, Sugiyama M, Fukuda H (2002) Activity of cell-wall degradation associated with differentiation of isolated mesophyll cells of *Zinnia elegans* into tracheary elements. Planta 215: 177–184

Page RDM (1996) Treeview: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12: 357–358

Pauly M, Scheller HV (2000) *O*-Acetylation of plant cell wall polysaccharides: identification and partial characterization of a rhamnogalacturonan *O*-acetyl-transferase from potato suspension-cultured cells. Planta 210: 659–667

Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM (1996) Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. Proc Natl Acad Sci USA 93: 12637–12642

Peng LC, Kawagoe Y, Hogan P, Delmer D (2002) Sitosterol-beta-glucoside as primer for cellulose synthesis in plants. Science 295: 147–150

Perrin RM, Jia ZH, Wagner TA, O'Neill MA, Sarria R, York WS, Raikhel NV, Keegstra K (2003) Analysis of xyloglucan fucosylation in Arabidopsis. Plant Physiol 132: 768–778

Pilling E, Hofte H (2003) Feedback from the wall. Curr Opin Plant Biol 6: 611–616

Richmond TA, Somerville CR (2001) Integrative approaches to determining Csl function. Plant Mol Biol 47: 131–143

Rioux D, Nicole M, Simard M, Ouellette GB (1998) Immunocytochemical evidence that secretion of pectin occurs during gel (gum) and tylosis formation in trees. Phytopathology 88: 494–505

Rose JKC, Braam J, Fry SC, Nishitani K (2002) The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. Plant Cell Physiol 43: 1421–1435

Rudsander U, Denman S, Raza S, Teeri TT (2003) Molecular features of family GH9 cellulases in hybrid aspen and the filamentous fungus *Phanerochaete chrysosporium*. J Appl Glycobiol 50: 253–256

Saha BC (2000) α-L-Arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. Biotechnol Adv 18: 403–423

Salnikov VV, Grimson MJ, Seagull RW, Haigler CH (2003) Localization of sucrose synthase and callose in freeze-substituted secondary-wall-stage cotton fibers. Protoplasma 221: 175–184

Schrader J, Nilsson J, Mellerowicz E, Berglund A, Nilsson P, Hertzberg M, Sandberg G (2004) A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. Plant Cell 16: 2278–2292

Sitrit Y, Hadfield KA, Bennett AB, Bradford KJ, Downie AB (1999) Expression of a polygalacturonase associated with tomato seed germination. Plant Physiol 121: 419–428

Smith DL, Gross KC (2000) A family of at least seven beta-galactosidase genes is expressed during tomato fruit development. Plant Physiol 123: 1173–1183

Stacey NJ, Roberts K, Carpita NC, Wells B, McCann MC (1995) Dynamic changes in cell surface molecules are very early events in the differentiation of mesophyll cells from *Zinnia elegans* into tracheary elements. Plant J 8: 891–906

Sterky F, Bhalerao RR, Unneberg P, Segerman B, Nilsson P, Brunner AM, Charbonnel-Campaa L, Lindvall JJ, Tandre K, Strauss SH, et al (2004) A Populus EST resource for plant functional genomics. Proc Natl Acad Sci USA 101: 13951–13956

Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R, et al (1998) Gene discovery in the wood-forming tissues of poplar: analysis of 5,692 expressed sequence tags. Proc Natl Acad Sci USA 95: 13330–13335

Strasser R, Mucha J, Mach L, Altmann F, Wilson IBH, Glossl J, Steinkellner H (2000) Molecular cloning and functional expression of beta 1,2-xylosyltransferase cDNA from *Arabidopsis thaliana*. FEBS Lett 472: 105–108

Suzuki M, Kato A, Nagata N, Komeda Y (2002) A xylanase, AtXyn1, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the *Arabidopsis thaliana* genome. Plant Cell Physiol 43: 759–767

Szyjanowicz PMJ, McKinnon I, Taylor NG, Gardiner J, Jarvis MC, Turner SR (2004) The irregular xylem 2 mutant is an allele of korrigan that affects the secondary cell wall of *Arabidopsis thaliana*. Plant J 37: 730–740

Tanaka K, Murata K, Yamazaki M, Onosato K, Miyao A, Hirochika H (2003) Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. Plant Physiol 133: 73–83

Taylor NG, Howells RM, Huttly AK, Vickers K, Turner SR (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. Proc Natl Acad Sci USA 100: 1450–1455

Taylor NG, Laurie S, Turner SR (2000) Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in Arabidopsis. Plant Cell 12: 2529–2540

Teleman A, Nordstrom M, Tenkanen M, Jacobs A, Dahlman O (2003) Isolation and characterization of *O*-acetylated glucomannans from aspen and birch wood. Carbohydr Res 338: 525–534

Thompson J, Higgins D, Gibson T (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680

Timell TE (1967) Recent progress in the chemistry of wood hemicelluloses. Wood Sci Technol 1: 45–70

Whetten R, Sun YH, Zhang Y, Sederoff R (2001) Functional genomics and cell wall biosynthesis in loblolly pine. Plant Mol Biol 47: 275–291

Williamson RE, Burn JE, Hocart CH (2002) Towards the mechanism of cellulose synthesis. Trends Plant Sci 7: 461–467

Xu W, Purugganan MM, Polisensky DH, Antosiewicz DM, Fry SC, Braam J (1995) Arabidopsis Tch4, regulated by hormones and the environment, encodes a xyloglucan endotransglycosylase. Plant Cell 7: 1555–1567

Yang JM, Park S, Kamdem DP, Keathley DE, Retzel E, Paule C, Kapur V, Han KH (2003) Novel gene expression profiles define the metabolic and physiological processes characteristic of wood and its extractive formation in a hardwood tree species, *Robinia pseudoacacia*. Plant Mol Biol 52: 935–956

Zhang D, Hrmova M, Wan CH, Wu C, Balzen J, Cai W, Wang J, Densmore LD, Fincher GB, Zhang H, et al (2004) Members of a new group of chitinase-like genes are expressed preferentially in cotton cells with secondary walls. Plant Mol Biol 54: 353–372

Zhao GR, Liu JY, Du XM (2001) Molecular cloning and characterization of cotton cDNAs expressed in developing fiber cells. Biosci Biotechnol Biochem 65: 2789–2793

Zhong RQ, Kays SJ, Schroeder BP, Ye ZH (2002) Mutation of a chitinase-like gene causes ectopic deposition of lignin, aberrant cell shapes, and overproduction of ethylene. Plant Cell 14: 165–179

Zhong RQ, Ye ZH (2003) Unraveling the functions of glycosyltransferase family 47 in plants. Trends Plant Sci 8: 565–568

Zhu YM, Nam J, Carpita NC, Matthysse AG, Gelvin SB (2003) Agrobacterium-mediated root transformation is inhibited by mutation of an Arabidopsis cellulose synthase-like gene. Plant Physiol 133: 1000–1010

Zwieniecki MA, Melcher PJ, Holbrook NM (2001) Hydrogel control of xylem hydraulic resistance in plants. Science 291: 1059–1062