

# A transcriptional roadmap to wood formation

Magnus Hertzberg\*<sup>†</sup>, Henrik Aspeborg\*<sup>†</sup>, Jarmo Schrader\*, Anders Andersson<sup>‡</sup>, Rikard Erlandsson<sup>‡</sup>, Kristina Blomqvist<sup>‡</sup>, Rupali Bhalerao<sup>§</sup>, Mathias Uhlén<sup>‡</sup>, Tuula T. Teeri<sup>‡</sup>, Joakim Lundeberg<sup>‡</sup>, Björn Sundberg\*, Peter Nilsson\*<sup>¶</sup>, and Göran Sandberg\*<sup>¶</sup>

\*Umeå Plant Science Center, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, 901 83 Umeå, Sweden;

<sup>†</sup>Department of Biotechnology, Royal Institute of Technology, 100 44 Stockholm, Sweden; and <sup>§</sup>Umeå Plant Science Center, Department of Plant Physiology, Umeå University, 901 87 Umeå, Sweden

Edited by Bernard Phinney, University of California, Los Angeles, CA, and approved October 17, 2001 (received for review June 12, 2001)

**The large vascular meristem of poplar trees with its highly organized secondary xylem enables the boundaries between different developmental zones to be easily distinguished. This property of wood-forming tissues allowed us to determine a unique tissue-specific transcript profile for a well defined developmental gradient. RNA was prepared from different developmental stages of xylogenesis for DNA microarray analysis by using a hybrid aspen unigene set consisting of 2,995 expressed sequence tags. The analysis revealed that the genes encoding lignin and cellulose biosynthetic enzymes, as well as a number of transcription factors and other potential regulators of xylogenesis, are under strict developmental stage-specific transcriptional regulation.**

**T**ranscript profiling has the potential to reveal transcriptional hierarchy during development for thousands of genes, as well as providing expression data for many genes of unknown function (1, 2). This is especially true when expression patterns can be obtained for well defined tissues at specific developmental stages. However, this is technically demanding and so far there are no reports demonstrating tissue-specific analysis on cell types from a single developmental sequence. We have studied the developing secondary xylem of poplar, which is highly organized with easily recognized and distinct boundaries between the different developmental stages. Wood formation is initiated in the vascular cambium. Cambial derivatives develop into xylem cells through the processes of division, expansion, secondary wall formation, lignification, and finally, programmed cell death. The large physical size of the vascular meristem in trees offers a unique possibility to obtain samples from defined developmental stages by tangential cryo sectioning (3). To determine the steady-state mRNA levels at specific stages during the ontogeny of wood formation in *Populus tremula* × *Populus tremuloides* (hybrid aspen) we sampled 30- $\mu$ m-thick sections through the wood development region and subsequently analyzed the samples by using a spotted cDNA-microarray (4) consisting of 2,995 unique ESTs from hybrid aspen. Our study provides a unique global examination of gene expression patterns that encompasses a developmental gradient within a multicellular organism.

## Materials and Methods

The Unigene set was selected from the expressed sequence tags (ESTs) presented in ref. 5, using cluster analysis. ESTs were transformed into *Escherichia coli* by using TSS competent cells (6), plasmids were prepared by using 96-well Multiscreen FB plates (Millipore), inserts were PCR amplified by using vector-specific primers, and PCR products were purified on Multiscreen PCR filter plates (Millipore) and spotted in duplicate onto CMT GAPS slides (Corning) by using the GMS 417 Arrayer (Affymetrix, Santa Clara, CA) as described (7). All PCR products were checked on ethidium bromide-stained agarose gels. Nine clones giving double PCR bands were excluded from the analysis.

A subset of 2,085 of the 2,995 PCR products in the Unigene set were resequenced from the 5' end. The identity was confirmed for 93% of the clones, and the remaining 7% (139 clones) were corrected accordingly. New sequences for 28 clones were deposited in GenBank (accession nos. BI784488–BI784515).

Three different annotations are presented for the clones on the array. The MENDEL EST annotations (<http://www.mendel.ac.uk/>) for the hybrid aspen ESTs analyzed in this study were downloaded, and annotations with a Relative Confidence Value (RCV) higher than 0.3–0.4 are considered to be of good quality ([http://www.mendel.ac.uk/pag7\\_3.htm](http://www.mendel.ac.uk/pag7_3.htm); ref. 8). Clones that were not confirmed in the resequencing were re-searched against the MENDEL database. Secondly, Swiss Prot/Trembl was searched by using Washington BLASTX, and the description line from the best hit as well as the BLASTX score are reported. Finally, we searched the full set of *Arabidopsis* proteins (arabi\_all.proteins\_v211200.tfa, MIPS), and from this search we report the accession number of the best hit and the BLASTX score.

Poplar trees (*P. tremula* × *P. tremuloides*, clone T89) were grown under natural light in the green house to a height of 5 m. Tissue samples were prepared by taking tangential sections through the cambial region of the stem (3). Tissue samples were collected from five positions to cover the developmental sequence: (A) meristematic cells, (B) early expansion, (C) late expansion, (D) secondary wall formation, and (E) late cell maturation. A phloem sample (Phl) was also included. Individual sections (30  $\mu$ m × 2 mm × 20 mm,  $\approx$ 0.5 mg) were pooled into the developmental zones indicated in Fig. 1A and B. The phloem sample was collected by scraping the inside of the bark, as described (7).

The transcript population was amplified, labeled, and hybridized as described (7). Sample and reference were spiked with a mixture of eight PCR-amplified human cDNAs, which were used for normalization. The control sample was an equal mixture of samples A–E. Two to four repetitive hybridizations were performed for each sample (three times for A–D; four times for E; and two times for Phl). The slides were scanned with a GMS 418 Scanner (Affymetrix). The data were analyzed with QUANTARAY 2 software (GSI Lumonics, Ottawa, ON, Canada). Using the adaptive quantification option, spots that looked visually bad (high background, bad spot morphology, dust present) were manually flagged as bad. The data were then passed through a quality filter by using EXCEL (Microsoft), where spots with weak signals and high background were excluded from further analysis, according to the following: A spot was flagged as good if the signal divided by the standard deviation of the background was higher than 2 and the signal minus background higher than 150 in both channels, or if the signal divided by the standard deviation of the background was higher than 4 and the signal

This paper was submitted directly (Track II) to the PNAS office.

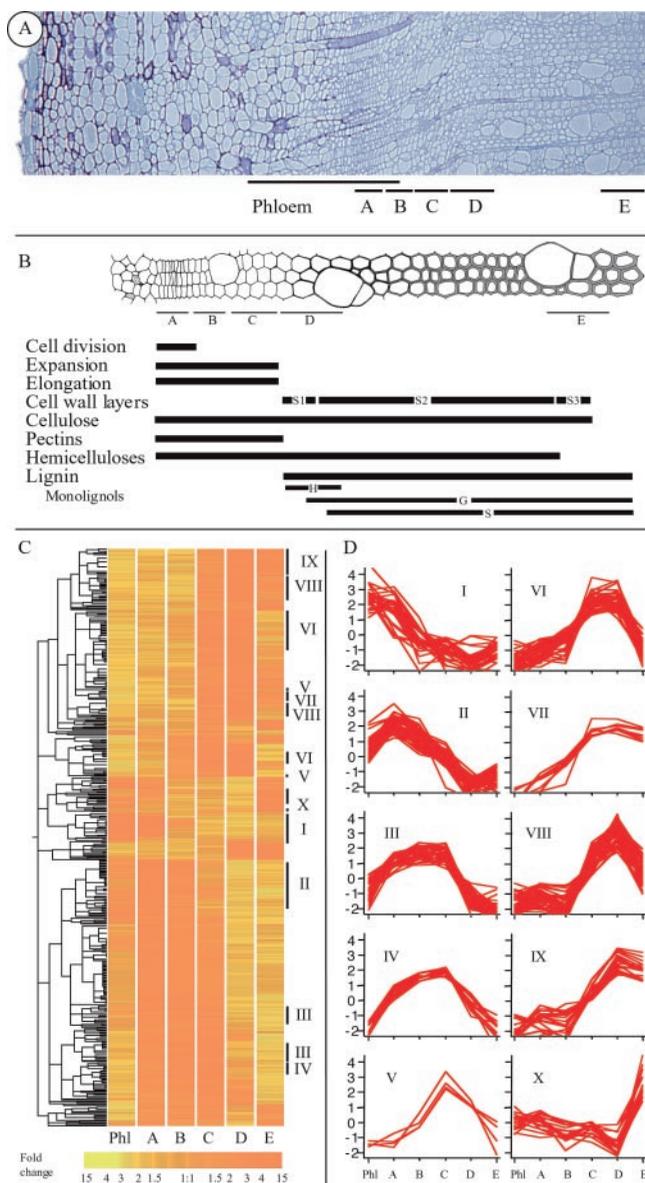
Abbreviation: EST, expressed sequence tag.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. BI784488–BI784515).

<sup>¶</sup>M.H. and H.A. contributed equally to this work.

<sup>†</sup>To whom reprint requests should be addressed. Goran.Sandberg@genfys.su.se (for G.S.) or nipe@biochem.kth.se (for P.N.).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



**Fig. 1.** (A) Cross section of a hybrid aspen stem stained with Toluidine blue. Black bars indicate the location of the sampled tissues. The phloem sample was included to give a low-resolution picture of the gene expression in the other tissue derived from the cambium. (B) Schematic representation of different cell-types and stages during vascular development. Bars depict timing and extent of the different developmental stages and the appearance of the major cell wall components. (C) Hierarchical cluster analysis of 1,791 selected genes with differential expression in the sampled tissues. The color scale at the bottom depicts fold change between samples. (D) Groups of genes with different differential expression patterns, expression ratios in  $\log_2$  scale. The samples are indicated at the bottom of the figure.

minus background higher than 150 in one channel. The filtered data were analyzed in GENESPRING 3.2.12 (Silicon Genetics, Redwood City, CA). All hybridizations were individually normalized to a median ratio of 1 for all of the ESTs. The hybridizations from each sample were subsequently normalized as a group so that the mean of the expression ratios for the human cDNA controls equaled 1. All of the data are presented as  $\log_2$  transformed ratios. The “Make Tree” function with “Change Correlation” statistics in GENESPRING was used to perform the hierarchical clustering on 539 genes that showed at

least an 8-fold differential expression over the developmental process (max expression ratio divided by min expression ratio larger than 8; Fig. 1C). A number of expression profiles were extracted from the cluster analysis and are shown in Fig. 1D.

The expression profiles obtained from the array experiments were validated by dot blotting the amplified cDNA populations on a nylon membrane and hybridizing it with six different ESTs by using radioactively labeled probes and stringent conditions. Additionally, amplified cDNA was prepared from a series of sections from a different tree, blotted onto a nylon membrane, and probed with 17 selected ESTs as described above.

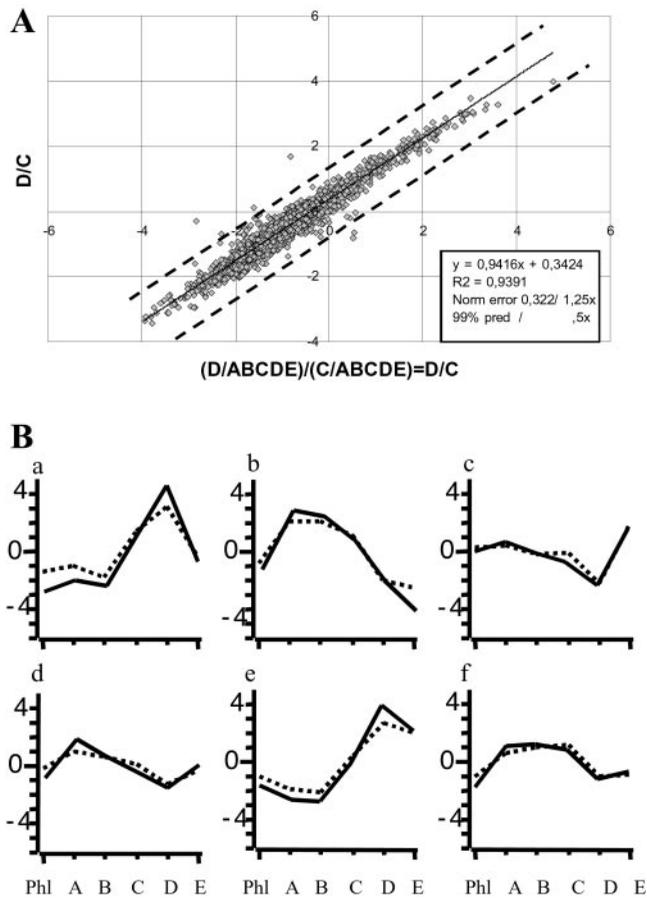
## Results and Discussion

In this report tissue samples were collected from six positions to cover the developmental sequence of wood formation (Fig. 1A and B). The transcript population from these samples was amplified and hybridized against a hybrid aspen cDNA microarray. Sequential gene expression patterns in our sample series were obtained by hybridizing amplified cDNA from each of the samples against a pool of all amplified samples. By comparing these results with results from hybridizing adjacent samples against each other, and based on earlier experiments, the limit for significant expression ratio was calculated to be  $>1.8$ -fold (see Fig. 2A; ref. 7). The expression profiles obtained from the array experiments were validated in control hybridizations on membranes, demonstrating that the entire protocol used in this paper is fully comparable with data from an alternative hybridization technique (Fig. 2B). As an additional verification, amplified cDNA was prepared from tissue sections from a different tree; membrane-based hybridizations using selected ESTs in all cases confirmed the expression patterns found in the array experiment presented in this paper. We resequenced the spotted clones to confirm clone identity—6.7% of the clones were found to be incorrect, indicating the importance of resequencing before data analysis. Resequenced clones are indicated in Table 1, which is published as supporting information on the PNAS web site, [www.pnas.org](http://www.pnas.org). The poplar microarray, combined with high-resolution sampling of cell layers in defined developmental stages, allowed us to visualize the expression pattern over the entire developmental process for 2,874 of the arrayed genes (see Table 1 for the full, annotated dataset with standard deviations).

A 4-fold differential expression over the region (max ratio/min ratio  $>4$ ) was shown in 1,246 of the genes, whereas 386 of the genes seemed to be constitutively expressed over all of the six samples (max ratio/min ratio  $<2$ ). Using hierarchical clustering of the expression profiles, 539 genes showing 8-fold differential expression over the region were clustered—this revealed gene classes with developmentally regulated expression patterns (Fig. 1C and D, I–X).

Wood formation is initiated in the vascular cambium through cell division and subsequent differentiation. Sequences expressed primarily in the cambial meristem (Fig. 1A, zone A) and in the zone of early cell expansion (Fig. 1A, zone B, and D, II) represent candidate genes involved in cell cycling, cell expansion, tip growth of fibers, and biosynthesis of the primary cell wall.

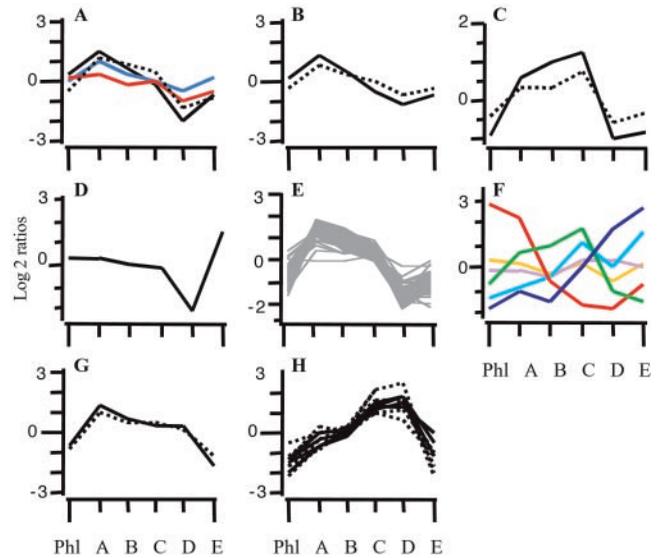
Cell division activity is controlled by the cell-cycle machinery where the activation of cyclin-dependent kinases by association with phase-specific cyclins controls the entry of the cell into different stages of the cell cycle (9). The poplar microarray contains EST homologues for a number of genes involved in cell-cycle control, like cyclins A, B, and H, *CDC2*, and *CKS1* (Fig. 3A and B), all showing similar patterns with a maximum expression in the cambial meristem (zone A). This coexpression is consistent with an interaction between these genes in cell-cycle regulation as it is seen in *Arabidopsis* and other organisms (10). The fact that cyclin-dependent kinases and cyclins are still expressed at significant levels during cell expansion (zones B and C) could be due to a role in maintaining the totipotency of these



**Fig. 2.** Quality control. (A) Variability between hybridizations. The mean of expression ratios from hybridization D/(ABCDE) were divided by the expression ratios from C/(ABCDE), and plotted against the mean of expression ratios D/C (data from hybridizing sample D against sample C), in log<sub>2</sub> scale. Based on these data, we estimate that a 1.5-fold difference in expression would be significant to 99% confidence. A small error component (1.2-fold) occurring from the normalization procedure, observed as the crossing of the regression line and the x axis at -0.3, increases the limit for significant expression ratios to >1.8-fold difference. A dashed line shows the 2-fold error line. (B) The expression profiles obtained from the array experiments were validated by dot blotting the cDNA populations on a nylon filter and hybridizing it with six different ESTs by using radioactively labeled probes. Control hybridizations in black lines and array data in dashed lines plotted in log<sub>2</sub> scale. Data from the control hybridizations are normalized to an average of one for each individual gene. (a) AI163298, unknown; (b) AI164300, transcription factor; (c) AI161744, *KNAP2* homologue; (d) AI166228, transcription factor; (e) AI161452, *CAD* (EC1.1.1.195); (f) AI164228, *ATHB-9*.

cells (11). This competence is lost in maturing fibers, but not in ray cells. Therefore, the expression of cell-cycle genes drops in zone D but increases again in E where the ray cells constitute a larger portion of the living cells. An alternative explanation for the expression of cyclins in zones B and C could be their involvement in other processes not directly related to cell cycle, as has been shown in the case of cyclin H (12). Such a role could also be assumed for one of the cyclin A homologues (AI164102), which shows only minor variation in expression throughout the developmental sequence.

Zones A and B would be expected to express genes that regulate cell fate and cell identity (Fig. 1B). One interesting example is the poplar homologue to the *Arabidopsis ATHB-8* and the closely related *ATHB-9* genes. *ATHB-8* is a member of the HD-zipIII class of transcription factors that is expressed in provascular cells of *Arabidopsis* where it has been proposed to



**Fig. 3.** Expression profiles. Log<sub>2</sub> transformed ratios. (A) Cyclins: *cycA*-like (black line, AI165944), *cycH*-like (dotted line, AI165594), *cycA*-like (red line, AI164102) and *cycB*-like (blue line, AI163933). (B) *cdc2* like (black line, AI163933) and *cks1*-like (dotted line, AI163057). (C) *ATHB-9* homologue (AI164228) in black and *ATHB-8* homologue (AI165328) in gray. (D) A *KNAP2* homologue of the *Knotted* class (AI161604). (E) Ribosomal proteins. (F) Six different MYB-domain transcription factors: AI163812 in red, AI164087 in dark blue, AI161768 in dark green, AI165848 in light blue, AI161482 in purple, and AI163448 in yellow. (G and H) Tubulin genes organized in two different expression clusters: Tubulin a (black lines) and Tubulin b (dotted lines). Data are presented as log<sub>2</sub> transformed ratios.

regulate vascular development (13). This hypothesis is supported by our transcript profiling data showing that the poplar *ATHB-8* orthologue was expressed in the cambial meristem and the expansion zone (Fig. 3C). The data set contains expression profiles for ~90 different transcription factors, including a *Knotted-1*-like gene, which is evenly expressed across the cambium and expansion zones, then sharply decreases when secondary cell-wall formation is initiated (Fig. 3D). *Knotted*-like genes are proposed to repress the progression into specific differentiation steps. The specific down-regulation observed in zone D might be necessary for the induction of secondary cell-wall synthesis, similar to the proposed function of the *knotted* genes in the apical meristem when lateral organs are initiated (14).

Cell expansion takes place in the meristem (zone A) and in zones B and C (15). Genes with a expression across zones A–C (Fig. 1D, III) may therefore function in cell expansion. For example, high expression of pectin esterases and pectate lyases was observed in zones A–C, consistent with activity assays in poplar (16). Pectin is abundant in primary cell walls where its degree of methylation influences wall extensibility (17, 18). Many genes expressed in zones A–C are involved in general metabolism and protein synthesis, reflecting the high metabolic activity in these tissues. For example, most of the ribosomal proteins have their highest expression in these tissues (Fig. 3E). Another cluster of genes up-regulated in zones A–C has an expression shift toward the C zone (Fig. 1D, IV), and a set of genes is specifically up-regulated in zone C (Fig. 1D, V). Both classes of genes may have a role in early events of xylogenesis. For example, both classes include members of the MYB family of transcription factors (Fig. 3F) that have previously been described to be both positive and negative regulators of lignification and flavonoid biosynthesis (19). In our data set, there are six MYBs that exhibit four

distinct expression patterns, indicating roles for these MYB-like genes in different developmental processes during xylogenesis.

As soon as cell expansion is completed, the secondary cell wall is deposited in all xylem cells (zone D). The random organization of cellulose microfibrils present in the primary wall now changes to a highly organized helical structure (20). During the secondary wall thickening the interspace within the cellulose and hemicellulose network is lignified, starting in the middle lamella and progressing inwards (21). Microtubules determine the direction of cellulose microfibrils in the cell wall, thereby influencing the spatial control of cell expansion (22), as well as defining the cellulose microfibril angle in the secondary cell wall of developing wood cells (20). Among the fourteen different tubulin genes present in the poplar unigene set, ten are strongly up-regulated during late expansion and when the secondary cell wall is formed (Fig. 3 *G* and *H*).

The majority of genes involved in the biosynthesis of the secondary cell wall were predicted to be found in zones C (where the vessels initiate their secondary cell wall), D, and E (Fig. 1*D*, V–IX). A selection of enzymes related to sugar metabolism were included in the array (Fig. 4*A*). Sugars are used in plants for energy metabolism, nucleotide synthesis, and importantly, in the maintenance and synthesis of primary and secondary cell walls during wood formation. Of particular interest for wood formation are the four different members of the cellulose synthase (*CESA*) family so far identified in our EST library. Expression of two of these genes, *PttCESA1* and *PttCESA3*, was up-regulated in zones C and D with clearly reduced expression in zone E. This coordinately regulated pair of *CESA* genes may thus be specifically involved in secondary cell wall synthesis (23). Interestingly, a membrane-bound cellulase isoenzyme, homologous to KOR, an endoglucanase previously suggested to be a part of the cellulose synthesizing complex (24), was similarly up-regulated in cells forming secondary walls (zones C–E). Sucrose synthase is responsible for channeling sucrose into UDP-glucose, which is the sole precursor of cellulose. The expression of one of the sucrose synthase genes was synchronized with the *PttCESA1/PttCESA3* pair, and is thus likely to be involved in cellulose synthesis in the secondary wall. Another UDP-glucose-providing enzyme, UTP-glucose-1-phosphate uridylyltransferase, which catalyzes the formation of UDP-glucose from  $\alpha$ -D-glucose 1-phosphate, is up-regulated in zones C and D, suggesting two alternative routes to UDP-glucose for cell-wall biosynthesis.

Hemicelluloses together with pectin form a large group of heteropolysaccharides composed of D-xylose, L-arabinose, L-rhamnose, L-fucose, D-mannose, D-galactose, D-galacturonate, or D-glucose. UDP-D-glucuronate synthesis is the rate-limiting step for the biosynthesis of both hemicellulose and pectin. There is evidence of two pathways leading to UDP-D-glucuronate, inositol oxidation or the oxidation of UDP-glucose (25). The gene involved in the inositol oxidation pathway, *myo*-inositol-1-monophosphatase, was down-regulated during the secondary cell wall formation in poplar (zone C). Two UDP-glucose dehydrogenase isoenzymes in the oxidation pathway were induced in the primary walled stage, whereas a third one was up-regulated during secondary cell wall formation (zones C and D). This finding probably reflects the expression of specific isoenzymes for channeling UDP-glucose into pectin, synthesized during primary wall formation, and xylan which is a dominating component in the secondary wall. In the absence of sequence information, it was not possible to investigate the expression patterns of the rest of the xylan biosynthetic enzymes. Interestingly, one polygalacturonase isoenzyme was selectively up-regulated in zones C–E, indicating a role in pectin remodeling during secondary wall

formation. In addition, a putative phosphomannomutase was gradually up-regulated over the developing xylem, as expected because of the high mannan content of the poplar hemicelluloses (26).

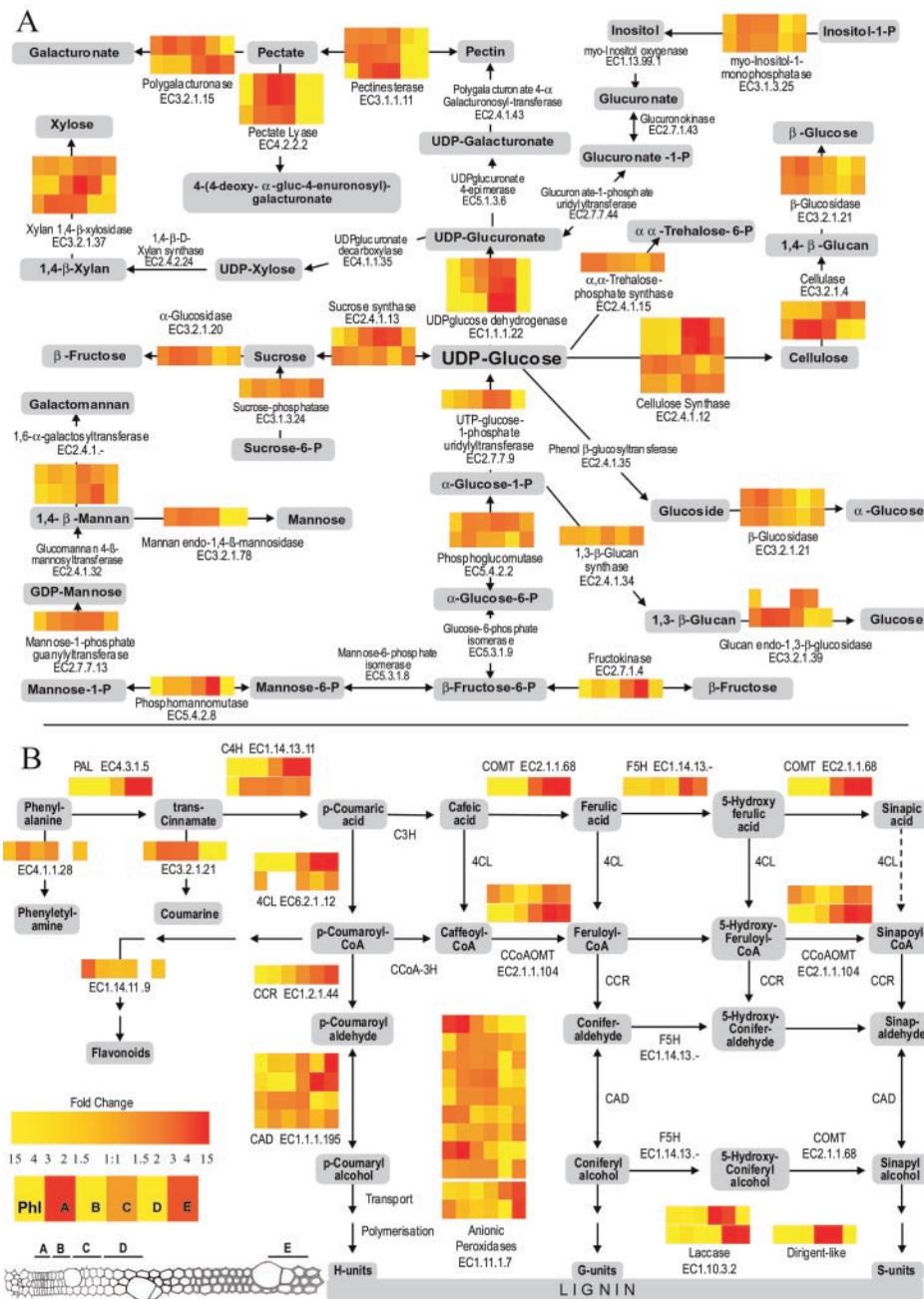
The addition of side chain residues to the polysaccharide backbone of hemicelluloses is catalyzed by glycosyltransferases. In plants, so far only two cell wall biosynthetic glycosyltransferases have been characterized (27, 28). Two putative 1,6- $\alpha$ -galactosyltransferase genes with significant similarity to the fenugreek galactomannan galactosyltransferase (27) were included on the array. The transcripts for both of these enzymes were up-regulated during the secondary cell wall formation. However, because galactomannan has not been detected in the poplar xylem (26), these genes may code for galactosyltransferases with different specificity.

The microarray contains ESTs coding for at least one isoenzyme for all cloned and published enzymes involved in lignin biosynthesis. For several of these enzymes, it was possible to examine the expression profiles of the entire gene family (Fig. 4*B*). Interestingly, completely different expression patterns were observed within several of the gene families. This is exemplified by the two 4-coumarate:CoA ligase (4CL) genes where one is up-regulated in zones C–E, and the other is expressed evenly across the cambial region. Both 4CL genes have high similarity to a previously cloned poplar 4CL gene demonstrated to be involved in lignification (29, 30). We have also analyzed four different cinnamyl alcohol dehydrogenase (CAD) genes that revealed unique expression profiles. Two of the CADs were up-regulated in zones D and E, where major lignification occurs. Both sequences encode type 2 CADs (CAD2), which have previously been shown to be involved in lignification (31). The other two CADs are type 1 sequences, and were not specifically induced in lignifying cells, in agreement with an earlier observation that type 1 CADs are not involved in lignification (32).

Following synthesis, the monophenols are transported to the apoplast where they are polymerized in a controlled fashion starting in the middle lamella and cell corners (21). The dirigent protein has recently been proposed to act as an initiation site for lignin polymerization (33). Consistent with this, our dataset showed induction of a gene coding for a dirigent homologue coincident with lignification. The final polymerization of the lignin monomers has been linked to the function of peroxidases and laccases, but the specific isoforms have yet to be identified (31). The array detected two peroxidases not previously linked to lignification that were up-regulated in lignifying cells. One is similar to glutathione peroxidases, and the other has high homology to a putative peroxidase from *Arabidopsis*. We also detected two laccases expressed in the lignifying tissues that have overlapping but different expression patterns, indicating that they could have specific and different roles in the polymerization process.

Genes strongly up-regulated in zone E (Fig. 1*D*, X) included many wall-degrading enzymes required for cell-wall sculpturing through final stages of the formation of pits and pores, or genes related to late phases of fiber maturation such as lignification and programmed cell death (see Fig. 4*A* and *B*). This class (X) also contains genes specifically involved in metabolism and transport in ray cells, which, as opposed to the fibers, remain alive and maintain their metabolic activity.

Many of the poplar genes analyzed by transcript profiling are homologous to previously identified genes in other plant species. Sixty percent of the analyzed genes have high similarity to sequences in *Arabidopsis* (BLASTX scores  $>200$  and *P* values less than  $10^{-15}$ ). Nevertheless, a significant subset of sequences appear novel, with 78 of the arrayed genes showing no significant similarity to any sequences in the current databases and 142 genes showing only weak similarity (BLASTX



**Fig. 4.** Metabolic pathways related to cell-wall formation. The metabolites are presented in gray boxes; arrows represent enzymatic reactions. Colored bars next to arrows indicate relative expression ratios of the corresponding gene in the different tissue samples. (A) Selected steps in the carbohydrate metabolic pathway important for the formation of cell-wall components. Genes included (with accession nos. in parentheses) EC3.1.1.11 (A1164340, A1164970, A1165089), EC4.2.2.2 (A1163756, A1162298, A1162963), EC3.2.1.15 (A1163516, A1164358), EC3.2.1.37 (A1164515, A1162157, A1163643), EC2.4.1.15 (A1163996), EC1.1.1.22 (A1163328, A1162135, A1166238), EC2.4.1.13 (A1162073, A1163591), EC3.1.3.24 (A1165881), EC3.2.1.20 (A1163888), EC2.7.7.9 (A116198), EC2.4.1.12 (A1163338, A1164546, A1162632), EC3.2.1.4 (A1164537, A1162370), EC3.2.1.21 (A1162064, A1161473), EC5.4.2.2 (A1162727, A1165714), EC2.4.1.34 (A1164628), EC3.2.1.39 (A1162475, A1164898), EC2.7.1.4 (A1162647), EC5.4.2.8 (A1162548), EC2.7.7.13 (A1165373), EC3.2.1.78 (A1162354), EC2.4.1.- (A1161656, A1161661), and EC3.1.3.25 (A1162400, A1165680). (B) Metabolic pathway for lignin biosynthesis; included genes EC4.1.1.28 (A1161926), phenylalanine ammonia-lyase (PAL) EC4.3.1.5 (A1166034), EC3.2.1.21 (A1163643), cinnamate 4-hydroxylase (C4H) EC1.14.13.11 (A1164016, A1164359), caffeate O-methyltransferase (COMT) EC2.1.1.68 (A1162318), ferulate 5-hydroxylase (F5H) EC1.14.13.- (A1166206), 4-coumarate:CoA ligase (4CL) EC6.2.1.12 (A1163594 and A1162611), caffeoyl CoA O-methyltransferase (CCoAOMT) EC2.1.1.104 (A1165013, A1165215), EC1.14.11.9 (A1163786), cinnamoyl-CoA reductase (CCR) EC1.2.1.44 (A1163319), cinnamyl alcohol dehydrogenase (CAD) EC1.1.1.195 (A1165107, A1161452, A1164048, A1164606), Peroxidases A1163015 and A1163276, Laccase EC1.10.3.2 (A1165655 and A1165982), Dirigent-like (A1163758).

scores <70). Functional analysis of these novel genes is thus best performed in poplar, where the technology for high throughput gene function assays is currently under development. On the other hand, homology to *Arabidopsis* genes offers a unique opportunity to explore their function by screening the

insertion mutant libraries available for *Arabidopsis*. Among the 2,995 poplar sequences, ~540 show high similarity to *Arabidopsis* proteins with an unknown function (as classified by MIPS, Nature Arabidopsis Genome CD; ref. 32). Of these, 211 genes are differentially expressed during xylem development

in poplar (see Table 2, which is published as supporting information on the PNAS web site, [www.pnas.org](http://www.pnas.org)). Fig. 5, which is published as supporting information on the PNAS web site, shows these expression patterns, divided into groups by using K-means clustering (GENESPRING). Combining the cell-specific transcript profiling in poplar with phenotypic analysis of the corresponding *Arabidopsis* mutants will provide the

means for a rapid functional analysis of these unknown genes and their specific involvement in wood formation.

We thank O. Nilsson, M. Sievertzon, and R. P. Bhalerao for valuable comments and discussions. This work was supported by The Swedish Research Council, the Foundation for Strategic Research, the European Union Popwood Project, and in part by the Corning Incorporated Science Products Division, through its CMT-Partners program.

- White, K. P., Rifkin, S. A., Hurban, P. & Hogness, D. S. (1999) *Science* **286**, 2197–2184.
- Aharoni, A., Keizer, L. C., Bouwmeester, H. J., Sun, Z., Alvarez-Huerta, M., Verhoeven, H. A., Blaas, J., van Houwelingen, A. M., De Vos, R. C., van der Voet, H., *et al.* (2000) *Plant Cell* **12**, 647–662.
- Ugla, C., Moritz, T., Sandberg, G. & Sundberg, B. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 9282–9286.
- Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. (1995) *Science* **270**, 467–470.
- Sterky, F., Regan, S., Karlsson, J., Hertzberg, M., Rohde, A., Holmberg, A., Amini, B., Bhalerao, R., Larsson, M., Villarroel, R., *et al.* (1998) *Proc. Natl. Acad. Sci. USA* **95**, 13330–13335.
- Chung, C. T., Niemela, S. L. & Miller, R. H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2172–2175.
- Hertzberg, M., Sievertzon, M., Aspeborg, H., Nilsson, P., Sandberg, G. & Lundberg, J. (2001) *Plant J.* **25**, 585–591.
- Lonsdale, D., Arnold, B. & Arnold, B. (1999) *Plant Mol. Biol. Rep.* **17**, 239–247.
- Reynard, G. J., Reynolds, W., Verma, R. & Deshaies, R. J. (2000) *Mol. Cell. Biol.* **20**, 5858–5864.
- Stals, H., Casteels, P., Van Montagu, M. & Inze, D. (2000) *Plant Mol. Biol.* **43**, 583–593.
- Martinez, M. C., Jørgensen, J. E., Lawton, M. A., Lamb, C. J. & Doerner, P. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7360–7364.
- Yamaguchi, M., Fabian, T., Sauter, M., Bhalerao, R. P., Schrader, J., Sandberg, G., Umeda, M. & Uchimiya, H. (2000) *Plant J.* **24**, 11–20.
- Baima, S., Nobili, F., Sessa, G., Lucchetti, S., Ruberti, I. & Morelli, G. (1995) *Development (Cambridge, U.K.)* **121**, 4171–4182.
- Scanlon, M. J. (2000) *Trends Plant Sci.* **3**, 31–36.
- Larson, P. R. (1994) in *The Vascular Cambium: Development and Structure* (Springer, Berlin).
- Micheli, F., Bordenave, M. & Richard, L. (2000) in *Cell and Molecular Biology of Wood Formation*, eds. Savidge, J. R., Barnett, A. R. & Napier, R. (BIOS Scientific Publishers, Oxford), pp. 295–304.
- Goldberg, R., Morvan, C. & Roland, J. C. (1986) *Plant Cell Physiol.* **27**, 417–429.
- McCann, M. C., Stacey, N. J., Wilson, R. & Roberts, K. (1993) *J. Cell Sci.* **106**, 1347–1356.
- Teragnone, L., Merida, A., Parr, A., Mackay, S., Cuiñez-Macia, F. A., Roberts, K. & Martin, C. (1998) *Plant Cell* **10**, 135–154.
- Funada, R. (2000) in *Plant Microtubules*, ed. Nick, P. (Springer, Berlin), pp. 51–82.
- Terashima, N., Fukushima, K., He, L. F. & Takabe, K. (1993) in *Forage Cell Wall Structure and Digestibility*, ed. Jung, H. G. (American Society of Agronomy–Crop Science Society of America–Soil Science Society of America, Madison, WI), pp. 247–270.
- Nick, P. (2000) in *Plant Microtubules*, ed. Nick, P. (Springer, Berlin), pp. 1–24.
- Fagard, M., Höfte, H. & Vernhettes, S. (2000) *Plant Physiol. Biochem.* **38**, 15–25.
- Nicol, F., His, I., Jauneau, A., Vernhettes, S., Canut, H. & Hofte, H. (1998) *EMBO J.* **17**, 5563–5576.
- Witt, H. J. (1992) *J. Plant Physiol.* **140**, 276–281.
- Mellerowicz, E. J., Baucher, M., Sundberg, B. & Boerjan, W. (2001) *Plant Mol. Biol.* **47**, 239–276.
- Edwards, M. E., Dickson, C. A., Chengappa, S., Sidebottom, C., Gidley, M. J. & Reid, J. S. (1999) *Plant J.* **19**, 691–697.
- Perrin, R. M., DeRocher, A. E., Bar-Peled, M., Zeng, W., Norambuena, L., Orellana, A., Raikhel, N. V. & Keegstra, K. (1999) *Science* **284**, 1976–1979.
- Hu, W. J., Harding, S. A., Lung, J., Popko, J. L., Ralph, J., Stokke, D. D., Tsai, C. J. & Chiang, V. L. (1999) *Nat. Biotechnol.* **17**, 808–812.
- Hu, W. J., Kawaoka, A., Tsai, C. J., Lung, J., Osakabe, K., Ebinuma, H. & Chiang, V. L. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 5407–5412.
- Boudet, A. M. (2000) *Plant Physiol. Biochem. (Paris)* **38**, 81–96.
- The Arabidopsis Genome Initiative (2000) *Nature (London)* **408**, 796–815.