



Novel gene expression profiles define the metabolic and physiological processes characteristic of wood and its extractive formation in a hardwood tree species, *Robinia pseudoacacia*

Jaemo Yang¹, Sunchung Park¹, D. Pascal Kamdem¹, Daniel E. Keathley¹, Ernest Retzel², Charlie Paule², Vivek Kapur³ and Kyung-Hwan Han^{1,*}

¹Department of Forestry, 126 Natural Resources, Michigan State University, East Lansing, MI 48824-1222, USA (*author for correspondence; e-mail hanky@msu.edu); ²Academic Computing and Bioinformatics, Academic Health Center, University of Minnesota, 650 Children's Rehabilitation Center, UMHC Box 43, 426 Church St. S.E., Minneapolis, MN 55455-0312, USA; ³Veterinary PathoBiology, Advanced Genetic Analysis Center, 1971 Commonwealth Avenue University of Minnesota, St. Paul, MN 55108, USA

Received 17 October 2002; accepted in revised form 9 April 2003

Key words: EST, extractives, genomics, heartwood, microarray, wood extract, wood formation

Abstract

Wood is of critical importance to humans as a primary feedstock for biofuel, fiber, solid wood products, and various natural compounds including pharmaceuticals. The trunk wood of most tree species has two distinctly different regions: sapwood and heartwood. In addition to the major constituents, wood contains extraneous chemicals that can be removed by extraction with various solvents. The composition and the content of the extractives vary depending on such factors as, species, growth conditions, and time of year when the tree is cut. Despite the great commercial and keen scientific interest, little is known about the tree-specific biology of the formation of heartwood and its extractives. In order to gain insight on the molecular regulations of heartwood and its extractive formation, we carried out global examination of gene expression profiles across the trunk wood of black locust (*Robinia pseudoacacia* L.) trees. Of the 2,915 expressed sequenced tags (ESTs) that were generated and analyzed in the current study, 55.3% showed no match to known sequences. Cluster analysis of the ESTs identified a total of 2278 unigene sets, which were used to construct cDNA microarrays. Microarray hybridization analyses were then performed to survey the changes in gene expression profiles of trunk wood. The gene expression profiles of wood formation differ according to the region of trunk wood sampled, with highly expressed genes defining the metabolic and physiological processes characteristic of each region. For example, the gene encoding sugar transport had the highest expression in the sapwood, while the structural genes for flavonoid biosynthesis were up-regulated in the sapwood-heartwood transition zone. This analysis also established the expression patterns of 341 previously unknown genes.

Introduction

Wood is a unique renewable material produced by trees using solar energy through a highly ordered developmental process involving cell division/expansion, secondary cell wall synthesis/deposition, lignification, programmed cell death, and heartwood formation (Fukuda, 1996). As a result of radial growth and differentiation, the trunk wood of

many tree species has two distinctly different regions: sapwood and heartwood. Sapwood is the outermost portion of the xylem tissue and contains living cells, whereas the heartwood is defined as the 'dead' central core of the woody axis and only provides passive support to the tree. Sapwood (young xylem) has three important functions: to conduct sap (water, solutes, and gases) from the roots to all parts of the tree; to provide structural support for the entire tree; and to

serve as a reservoir for water, energy, minerals, and solutes. On average, about 10% of the cells in the sapwood are alive (Kozłowski and Pallardy, 1997). The living ray cells in sapwood serve as the source of raw materials for secondary substances. The ray parenchyma may also serve as communication channels radially from the cambium through the sapwood, while axial parenchyma functions largely as a storage tissue.

As sapwood is gradually converted to inactive heartwood, the wood parenchyma cells undergo numerous metabolic changes and produce large quantities of heartwood extractives such as phenolic compounds, lignin, and aromatic substances that accumulate in the vessels (Magel, 2000). During that process, one annual ring is converted to heartwood each year (Mauseth, 1998). The reserve materials in the parenchyma cells of the sapwood are used for wood formation and the synthesis of heartwood extractives, such as condensed tannins, terpenes, flavonoids, lignans, stilbenes, and tropolones (Burtin *et al.*, 1998; Hillinger *et al.*, 1996a and b; Hillis, 1987; Magel *et al.*, 1994 and 2000). The formation of heartwood is accompanied by a variety of alterations in metabolic conditions such as senescence. Although the events of senescence have been studied at the molecular level during leaf senescence (Miller *et al.*, 1999; Wingler *et al.*, 1998), seed germination (Cercos *et al.*, 1999), and nodule development (Matamoros *et al.*, 1999); the cell maturation and death events occurring during heartwood formation have been difficult to study because of the location and timing of the events. The presence of heartwood is a major determining factor for wood quality and influences the way in which specific woods are utilized. Various wood properties, such as dimensional stability, durability, pulpability, colors and hues, scents and beauty, are affected by extractives. Furthermore, stem wood sequesters large amounts of atmospheric CO₂ into a much slower turnover pool and consequently accounts for the largest proportion (20–40%) of total ecosystem aboveground carbon in closed forests (Saxe *et al.*, 1998). Therefore, understanding the regulation of wood formation is of great commercial and keen scientific interest.

In recent years, a genomics approach has been successfully used to examine global gene expression patterns in developing xylem tissues of pine (Allona *et al.*, 1998; Lorenz and Dean, 2002) and poplar (Sterky *et al.*, 1998; Hertzberg *et al.*, 2001). Although the information derived from those studies undoubtedly provided a powerful means for studying

the molecular mechanisms of this important differentiation pathway, it is still insufficient to account for the complete process of wood formation. So far, there has been no report on global examination of gene expression profiles inside trunk wood of mature trees. In order to gain some insight into the transcriptional hierarchy of heartwood and its extractive formation, we examined global gene expression profiles across the stems of 10-year-old *Robinia pseudoacacia* trees by sampling bark, sapwood, and sapwood-to-heartwood transition zone tissues. This report describes the first comprehensive look at global gene expression profiles in trunk wood and provides expression data for many genes of unknown function.

Materials and Methods

RNA isolation and cDNA library construction

Three cDNA libraries of bark/cambial region (BCS), sapwood region (SWS), and transition zone (TZS) of trunk-disk of 10-year-old black locusts (*Robinia pseudoacacia* L.) harvested in early summer (July 27, designated 'S') and one cDNA library (TZF) from a black locust harvested at late fall (November 27, designated 'F') were constructed using the SMART cDNA library construction kit (λ TriplEx2 vector system, Clontech, Palo Alto, CA). Mature trees (20 cm DBH) were felled using a chain saw and made into 25 cm-long logs. The logs were immediately placed on ice and brought back to a wood shop, where thin cookies (ca. 1 cm thick) were made using a table saw. The cookies were immediately submerged in RNA extraction buffer (20 mM EDTA, pH 8.0; 50 mM Tris-HCl, pH 8.0; 0.2% SDS; 10 mM 2-mercaptoethanol). While submerged, the trunk wood sections (bark/cambial region, sapwood, and transition zone) were carved out by using chisel and hammer, and washed with RNase Away solution (Invitrogen, Carlsbad, CA). The isolated sections (ca. 1 cm³ cubicles) were frozen in liquid nitrogen and stored at –80°C until needed. For RNA isolation, the frozen samples were first ground in a blender and then further ground to fine powder using mortar and pestle. The ground sample were first passed through the shredder column of DNeasy Maxi kit, and then subjected to total RNA isolation using the RNeasy Maxi kit (Qiagen, Hilden, Germany) and cleaned up by Qiagen RNeasy Mini kit.

Nucleotide sequencing

The cDNAs that were directionally cloned were randomly picked and sequenced to generate ESTs. The sequencing was carried out at the Center for Computational Genomics and Bioinformatics at the University of Minnesota and at the Genomics Technology Supporting Facility at Michigan State University. The sequencing results are posted at our website (<http://web.ahc.umn.edu/biodata/blacklocust/>).

Sequencing analysis

Raw sequence files were produced from the trace files using the Phred trace-processing program followed by the Phrap base-calling program with a quality threshold of 8–10 (Ewing *et al.*, 1998). Sequence artifacts were trimmed by the removal of leading and trailing vector sequences in the raw sequence. To obtain the best subsequence where the 'N' value is 4% or less of the total number of bases, the number of unknown or 'N' bases in a sequence of trimming, leading and trailing high-N sections was determined. Sequence similarity analyses were completed using a number of database searches. Databases include GenBank, National Center for Biotechnology Information (NCBI) GenPept (Benson *et al.*, 2000), Protein Information Resource (PIR) (Barker *et al.*, 2000), Swiss Institute of Bioinformatics SWISS-PROT (Bairoch *et al.*, 2000), TrEMBL (Bairoch *et al.*, 2000) and the National Biomedical Research Foundation NRL3D (Barker *et al.*, 2000). The EST sequences were deposited in the dbEST of the GenBank database.

Contig analysis

In order to generate the unigene sets and contigs of sequencing data, EST data was analyzed using DNA similarity algorithms and the assembly program, Phred/Phrap/Consed (Green and Ewing, 1996). Phrap ('phragment assembly program', or 'phil's revised assembly program') was used for assembling shotgun DNA sequence data, constructing a contig sequence as a mosaic of the highest quality parts of reads (rather than a consensus) and providing extensive information about assembly (including quality values for contig sequence) to assist troubleshooting. Phrap was used in conjunction with the base calls and base quality values produced by the basecaller, Phred; and with the sequence editor/assembly viewer, Consed. Cross-match was based on a 'banded' version of SWAT, an

efficient implementation of the Smith-Waterman algorithm for comparing any two sets of (long or short) DNA sequences. To calculate estimates for comparison among different libraries, we used a statistical method described by Audic and Claverie (1997). This method was developed to calculate statistical differences from different numbers of different libraries, and now it is being used for 'digital gene expression profiles'. Based on the method, we calculated significantly different numbers within the same contig among the libraries, and we applied it to estimate statistical differences within a category as described in Kirst *et al.*, (2002).

PCR amplification of the insert cDNA and microarray printing

λ TriplEx2 vector sequences flanking the insert (5'-AAGCAGTGGTATCAACGCAGAGT-3' and 5'-ATTCTAGAGGCCGAGGCCGCCGACATG-3') were used to amplify selected EST clones using polymerase chain reaction (PCR). The PCR products were precipitated in ethanol and re-suspended in 3 \times SSC (1 \times SSC is 0.15 M NaCl and 0.015 M sodium citrate). They were checked for quality by using gel electrophoresis to observe the concentration and multiple bands. PCR products of 2592 clones were arrayed from 384-well microtiter plates, and DNA was spotted on superaldehyde (Telechem, Sunnydale, CA) glass slides at a high density using an Omnigridded robot (Gene Machines, San Carlos, CA) and 16 ArrayIt chipmaker 2 pins (Telechem). Slides were washed and blocked according to the manufacturer's protocol.

Each glass slide contained two replications of the entire array, each of which consisted of 16 subarrays with 12 rows and 14 columns. Negative control genes, B-cell receptor protein genes, including genes such as Myosin heavy chain gene, Myosin regulation light chain2 and insulin-like growth factor gene, were printed on the top, middle and bottom of each array.

Preparation of labeled cDNA probes

Total RNA (1 μ g from each sample) was reverse-transcribed and amplified using the SMAT system (Clontech). To reduce nonspecific PCR amplification, cDNAs were amplified with the fewest cycles. Two micrograms of cDNA were labeled by the incorporation of either Cy5 or Cy3-dCTP (Amersham-Pharmacia, Piscataway, NJ) during oligo-dT-primed primer extension in the presence of Klenow DNA polymerase (Promega, Madison, WI) as described by

Schaffer *et al.* (1999). The labeled probes were purified using the QiaQuick PCR cleanup kit (Qiagen). The probe samples were denatured by placing them in a 100°C water bath for 3 min, left at room temperature for 30 min, and then used for hybridization.

To minimize the inherent variability of the microarray assay (Lee *et al.*, 2000) and to ensure the reliability of the results, at least two microarray slides (four replicates) were used to analyze the transcript expression of each sample pair. The first slide was probed with cDNAs labeled with Cy3 and Cy5 deoxy CTP. To probe the arrays, cDNAs were synthesized and amplified from bark/cambial region, sapwood, or transition zone and labeled by Klenow-mediated incorporation of Cy3-dCTP or Cy5-dCTP, respectively. By using independent RNA preparations, the second slide was hybridized by cDNAs reverse labeled with Cy3 and Cy5 dCTP from each sample pair to overcome potential artifacts caused by the dye-related differences in labeling efficiency, different laser settings, and nonlinearity of photomultiplier tubes in the scanner. Thus, at least two, and sometimes three or four, independent RNA preparations were made for each biological sample and were used to prepare labeled probes. The hybridization signal from each of the replicate ESTs were averaged and used for analysis.

Hybridization and washing of the DNA microarray

The labeled probes with either Cy3 or Cy5 fluorescent dye were hybridized to a microarray slide in a total volume of 30 μ L of hybridization buffer (3.4 \times SSC, 0.32% SDS, and 5 μ g of yeast tRNA) for 16 h at 65°C. The slide was then washed at room temperature in 1 \times SSC, 0.1% SDS for 10 min, in 1 \times SSC for 10 min, and in 0.01 \times SSC for 10 min. The slide was centrifuged dry and scanned with a 428 Array scanner (Affymetrix, Palo Alto, CA). Each microarray experiment was repeated twice.

Microarray data analysis

The data were analyzed with GenePix Pro3.0 (Axon Instruments Inc., Union City, CA). The scanned data were normalized by using the Global Normalization method (Hihara *et al.*, 2001), in which the image data between Cy3 and Cy5 channels are normalized by adjusting the total signal intensities of two images and the bad spots are removed. The unreliable spots were removed by the following screening. Spots containing clones that had poor amplification or multiple bands, as well as those that were

flagged due to a false intensity caused by dust or background on the array, were removed. Spots with <65% of the spot intensity at >1.5-fold that of the background in both channels were ignored (see Stanford Microarray Database Web site, <http://genome-www5.stanford.edu/MicroArray/SMD/>). Clones in one sample that had an average induction greater than 2-fold in another were determined as up-regulated. Comparison of the arrays was achieved using Microsoft Excel and Microsoft Access database. For cluster analysis, Cluster and Treeview software were used (Eisen *et al.*, 1998; available at <http://genome-www4.stanford.edu/MicroArray/SMD/restech.html>).

Antisense northern blot analysis

We conducted 'Antisense Northern blot analysis' which requires only minute amounts of RNA. Conventional Northern blot analysis protocol requires a large amount of RNA from our inner wood samples, typically transition zone, which are hard to isolate, making those protocols difficult to perform. An alternative was to use antisense RNA (aRNA) amplification method that has been successfully used in other microarray analyses (Wang *et al.*, 2000; Baugh *et al.*, 2001; Dent *et al.*, 2001; for protocol, see Patrick Brown's web site; http://cmgm.stanford.edu/pbrown/protocols/ampprotocol_3.html). Antisense RNA was generated as described previously (Wang *et al.*, 2000). Briefly, aRNA was amplified using Message AmpTM aRNA kit (Ambion, Austin, TX). We began by synthesizing first stand cDNAs from the total RNAs of seedling, bark/cambial region, sapwood, or transition zone, and the first cDNAs were used as templates for the synthesis of second cDNAs. Finally, aRNAs from the second cDNAs were generated by *in vitro* transcription (amplification). About 200 ng of aRNAs were separated in a formamide agarose gel, and transferred onto the nylon membrane using the capillary transfer method. The membrane was then hybridized with an isotope-labeled probe. The signal was exposed and detected on an X-ray film.

Results

Trunk Wood cDNA Library Construction and Sequencing

Trunk wood of mature trees can be divided into three main parts: bark/cambial region, sapwood, and

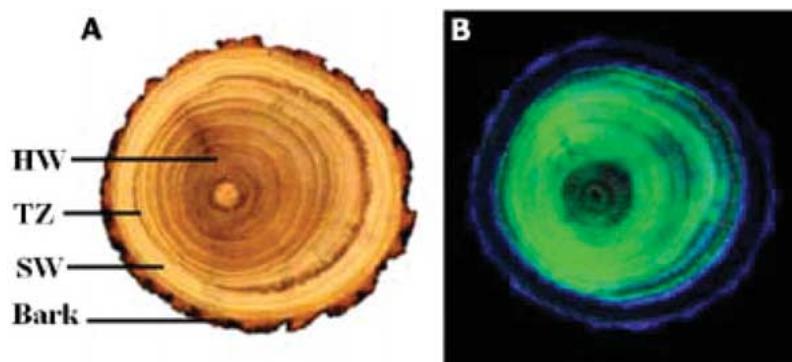


Figure 1. Cross-section of a stem from a mature *Robinia pseudoacacia* tree showing the different wood tissues. (A) Cross section under daylight. HW, Heartwood; SW, sapwood; TZ, transition zone. (B) Cross section under UV light. Bright fluorescence in the middle is due to the flavonoids.

heartwood (Figure 1). While sapwood contains living ray cells, the cells in heartwood are dead and filled with extractives, which produce intense and bright fluorescence under UV light. In order to analyze the gene expression patterns in different sections of the trunk wood, we constructed four cDNA libraries from the bark/cambial region (BCS), sapwood (SWS), and transition zone of ten-year-old black locusts. The transition zone samples were collected both in the summer and fall (TZS and TZF). Due to the large amounts of polysaccharides and phenolic compounds present in the inner wood tissues, it was not possible to obtain a large enough quantity of pure mRNA for conventional cDNA library construction. Nonetheless, we were able to construct high quality ($> 5 \times 10^5$ pfu) phagemid libraries using the PCR-based cDNA library construction kit (Clontech). The phagmids from each library were converted into plasmids by mass excision in *E. coli*. Over 3,600 individual clones were randomly selected from all four libraries and sequenced using a 5' vector sequencing primer provided in the kit. After trimming vector sequences, clones containing high ambiguous calls (high 'N' percent on a DNA sequence) were removed. Finally, a total of 2915 were chosen for further analyses and sequencing: 895 clones from the BCS library, 999 clones from the SWS library, 880 clones from the TZS library and 141 clones from the TZF library. An average length of 448 bases was obtained and used for contig analysis and database searches.

EST Analysis

The high redundancy of the mRNA in a tissue is approximately reflected in the abundance of its corresponding cDNA in non-normalized libraries. The ran-

dom sequencing of cDNAs yields information about the ESTs (Adams *et al.*, 1993). Sequence similarities were found by searching various available databases including: GenBank, National Center for Biotechnology Information (NCBI) GenPept, Protein Information Resource (PIR), Swiss Institute of Bioinformatics SWISS-PROT, TrEMBL and the National Biomedical Research Foundation NRL3D (see our website; <http://web.ahc.umn.edu/biodata/blacklocust/>). Sequence similarities identified by the BLAST program were considered statistically significant with a Poisson P value of 10^{-5} . The 1304 ESTs (44.7%) of the total 2915 ESTs matched previously sequenced genes. The 909 ESTs of the 1304 ESTs had significant homology to previously identified genes. The annotations of genes with similarities to an EST were used to assign a putative identification to our EST. The 909 ESTs with similarity to known genes were classified into 13 putative functional categories (Bevan *et al.*, 1998, Covitz *et al.*, 1998), which are listed in Table 1. Our libraries have a significantly high number of no hit clones, especially in the SWS library (80.6%). As a matter of fact, the portion of no hit clones in other libraries was also high, about 50%. Allona *et al.* (1998) suggested that the length and quality of cDNA sequences are correlated with the ability to identify similar sequences in public databases. Recent analysis of *Pinus taeda* cDNA clones with longer than 1,000 bases-read revealed that about 95% of the pine genes had homologous sequences in *Arabidopsis* genome (Sederoff *et al.*, 2002). However, the average sequenced length of a SWS clone is not entirely different from the total sequenced length (sequence length distribution data is found at <http://web.ahc.umn.edu/biodata/blacklocust/>). Many

long, high-quality sequences show neither strong nor marginal similarity to sequences in the database, and some no hit clones also show high redundancy (Table 2). Therefore, these no-hit clones may represent novel plant genes, reflecting the uniqueness of our samples.

The proportions of ESTs in each functional category differed for the EST libraries from the three trunk zones characterized in this study (Figure 2). These data, which were calculated by excluding no-hit clones in all libraries, show a common overall trend, but specific groups of genes were more or less highly represented in specific zones. Notable are the higher representation of secondary and hormone metabolism genes in the TZS library (9.1% versus 1.4% and 0.8% in the other libraries), cell wall and structural metabolism genes in the SWS library (3.6% versus 0.6% and 0.9% in the other libraries), and genes associated with membrane transport in the BCS library (6.1% versus 1.2% and 2.3% in the other libraries).

Contig analysis

We then attempted to estimate the redundancy of our EST clones on the basis of the contig analysis (Phrap assembly with the 70% identical value of pairwise sequence). The proportion of singletons was extremely high, representing a calculated level of singletons at 70% (2051 out of 2915 ESTs). Only 864 ESTs were identified in 228 contig sets. As a result of the contig analysis, we obtained a total of 2278 unigene sets that were submitted to GenBank database (accession numbers BI642054 to BI679372). The most frequently presented gene in our ESTs encodes a hypothetical protein (At2g41250; 1.3%) followed by an auxin-repressed protein (1.1%) that is expressed in dormant tissues and repressed by auxin treatment (Reddy and Poovaiah, 1990; Stafstrom *et al.*, 1998). It is abundant in all libraries except the sapwood library (Table 2). It is notable that the TZF library (2.1%) showed a similar proportion when compared to the TZS library (1.5%). The function of this gene has not yet been identified. The contig analysis identified three types of metallothionein or metallothionein-like protein genes with different expression patterns in the different zones of the trunk wood. For example, metallothionein (contig 226) is the highest in the TZS library, but metallothionein-like protein (contig 223) is only in the BCS library. Such differential expression of metallothionein genes has been reported in *Arabidopsis* plants (García-Hernández *et al.*, 1998). In addition

to providing physical support, trunk wood functions as a conduit for water, nutrient, and photosynthates transport. Aquaporin and phloem-specific protein, which are related to a membrane transport system or phloem, are abundant mainly in the bark/cambial region library. Two types of aquaporin genes that are related to water transport were found in our libraries Aquaporin 1 contig existed in the bark/cambial region library only, while aquaporin 2 contig was present in the transition zone library as well as bark/cambial region library. However, no contig for aquaporin genes were present in the sapwood library. Because the sole abundance of transcript encoding phloem-specific protein exists only in the bark/cambial region library, it is clear that our bark/cambial region sample contained phloem regions. Eight out of the 20 contigs in Table 2 were library-specific. In other words, the contigs were present only in one library and absent in the other two libraries. For example, maturase and hypothetical protein (PIR: A05191) are abundant only in the SWS library. Hypothetical protein (At3g03150) and cytochrome b5 DIF-F are highly abundant in the TZS library. Some contig sets contained only no hit clones resulting from the source zone of the samples; for example contig 224 and 216 consist of clones from two transition zone libraries, and clones of contig 213 are in the BCS library. These library specific contigs can be expected because the typical character of each library is dependent upon its location within the trunk wood.

Microarray experiments

In order to produce gene expression profiles related to wood formation and to compare gene expression patterns from different regions of the inner wood from a mature tree, we produced cDNA microarrays carrying 2,580 genes from all four libraries and conducted microarray hybridization experiments. This approach allowed us to examine the expression changes of ca. 2580 genes simultaneously and to search the expression patterns of bark/cambial region, sapwood, and transition zone through the use of one-by-one comparisons. We demonstrated that one-by-one experiments can be carried out for specific expressed profiling on continuous samples. Our approach involved the comparison of one sample with two other samples and then the generated ratios from the two different experimental sets were plotted. When all of the spots were plotted, we could determine which genes were specifically expressed in Sample A, when compared

Table 1. The functional classification of EST clones.

Category	Total	BCS	SWS	TZS	TZF
Cell division and cycle	7	4	0	3	0
Cell wall structure and metabolism	16	5	7	3	1
Chromatin and DNA metabolism	27	8	4	12	3
Cytoskeleton	10	6	3	0	1
Defense	90	46	5	35	4
Gene expression and RNA metabolism	76	38	20	16	2
Membrane transport	48	33	1	11	3
Miscellaneous	175	73	36	51	15
Primary metabolism	134	46	19	56	13
Protein synthesis and processing	190	84	19	80	7
Secondary and hormone metabolism	55	8	2	44	1
Signal transduction	75	26	11	32	6
Vesicular trafficking, protein sorting, secretion	6	2	2	2	0
Unknown, hypothetical	395	162	65	140	28
Hit clones	1304 (44.7)	541 (60.4)	194 (49.4)	485 (55.1)	84 (59.6)
No-hit clones	1611 (55.3)	354 (39.6)	805 (80.6)	395 (44.9)	57 (40.4)
Total sequenced clones	2915	895	999	880	141

The bold number within a row indicates significantly different number of ESTs in the library compared to all the others ($P < 0.01$). For example, the number of ESTs in the 'Gene expression and RNA metabolism' category was statistically higher in BCS than in SWS and TZS libraries. Likewise, the number of ESTs in the 'Secondary and hormone metabolism' category was significantly higher in TZS than in BCS and SWS.

to Sample B and Sample C. Unlike a reference or loop design, an approach of this type will serve as a one-by-one comparison for a small number of side-by-side samples in transcript profiling studies.

In our microarray experiments, the need to obtain large amounts of RNAs proved to be a challenge. Standard microarray protocols require isolating poly(A) RNA from samples, but our wood sample is too difficult to isolate RNAs. So, the cDNA amplification method was used, and we checked the reproducibility of our experiments. We also confirmed that the amplification method efficiently generates highly reproducible populations of cDNA. This method can be useful in transcript profiling studies with limited amounts of RNA. To test the reproducibility of the two different experiments, the expression ratio derived from one microarray experiment was compared to the expression ratio from the other, for all of the normalized clones (Figure 3A). The scatter plot shows that the gene expression ratio from the two experiments was remarkably similar (the coefficient of determination $R^2 = 0.96$). Furthermore, the coefficients of determination between other experiments were sufficiently high (>0.91) (data not shown). These results are similar to those presented in other papers, in which the

same or similar amplification protocol was used (Livezey *et al.*, 2000, Hertzberg *et al.*, 2001). Furthermore, using 'antisense northern blot' analysis, we confirmed the microarray data as well as the EST results. Thus, the microarray signals from the replicates in our study were highly reproducible and the conclusions derived from this analysis are considered reliable.

Differential gene expression across the stem

In order to investigate distinct differences in gene expression profiles among the three regions of the trunk wood (bark/cambial zone, sapwood, and sapwood-heartwood transition zone) during active tree growth, we compared differentially expressed genes in the three zones. The expression ratios for each zone in comparison to the other two zones were examined using regression analysis. Figure 3B shows the scatter plot and regression line of the expression ratios for the bark/cambial region versus the sapwood or transition zone libraries. The trend line of the scatter plot had a slope of 0.401, with $R^2 = 0.2956$. These values indicate a positive correlation, but with little of the variation in the data explained by the regression. On the basis of that graph, we identified the genes that were up or down regulated in BCS tissue us-

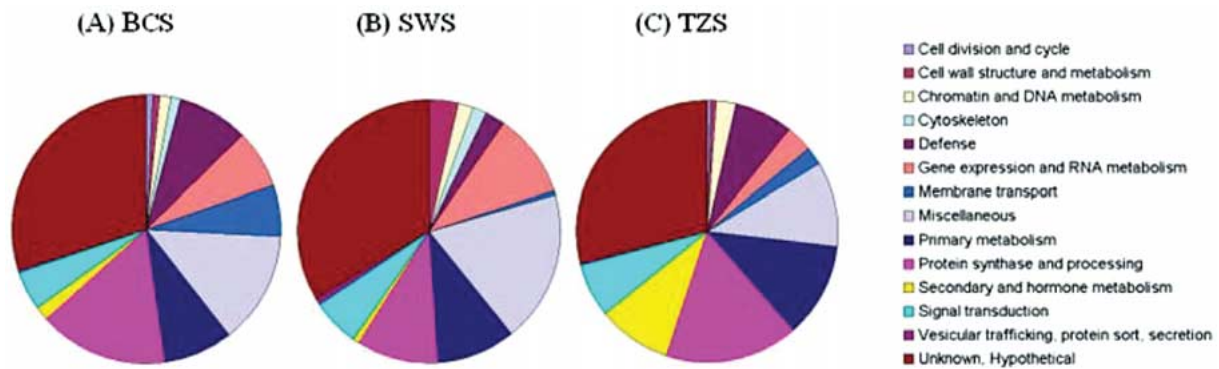


Figure 2. Functional classification of the ESTs from three libraries. Due to the high proportion of no-hit clones in the libraries, only those ESTs with significant homology with previously reported sequences were included in this classification.

ing the cutoff values of an expression ratio greater than two-fold or less than 0.5. Of all arrayed cDNAs, 292 clones were more highly expressed in the bark/cambial region than in sapwood. The expression of 174 clones in the bark/cambial region was higher than those in the transition zone. Ninety-six of the clones were considered as the BCS-specific clones (Tables 3 and 4). As expected, metallothionein-like protein (contig 223; containing all 11 BCS clones), phloem-specific protein genes (contig 217; containing all 8 BCS clones), and a no-hit clone in contig 213, which is a contig set composed of all 7 BCS clones, were highly expressed in bark/cambial region when compared to other zones. These results corroborate with those of EST redundancy analysis. In addition, genes encoding photosynthesis-related proteins, such as Photosystem II 10K proteins, were highly expressed in the bark/cambial region. In addition, 63 clones were specifically down regulated in bark/cambial region, when compared to sapwood and transition zone (Table 3). We compared each functional category to find out how many tissue-specific genes were in each category. Table 3 shows the proportion of functional categories of up or down regulated genes in bark/cambial region. Like the EST results, the proportion of the genes categorized as involving membrane transport was high relative to other functional categories in the bark/cambial region. When considering the physiological significance of the region in tree growth and development, the proteins encoded by the genes in each of these categories might be targets of special interest for biotechnological improvement of trees.

Similarly, the expression ratio of the SWS was compared to both the BCS and TZS (Figure 3c) and the number of up and down regulated genes was

calculated. Unlike the trendline of the bark/cambial region and transition zone scatter plots; the scatter plot of sapwood had a negative slope (0.401), indicating a negative correlation. Many of the clones that were highly expressed in sapwood vs. bark/cambial region were down regulated in the sapwood vs. transition zone comparison. For example, the expression ratio of the PR-10 gene in the sapwood region vs. bark/cambial region was 6.7, but the ratio in the sapwood region vs. the transition zone was 0.5. Similarly, the expression ratio of the phloem-specific protein gene in the sapwood region versus the bark/cambial region was 0.2. However, in contrast, the ratio in the sapwood region versus transition zone was 4.3, showing that the expression relationships vary for the different proteins in the three regions. In addition, the number of clones specifically up or down regulated in the sapwood region was very small (3 or 4), but when the expression in the sapwood region was compared to bark/cambial region or transition zone, the number of up or down regulated clones increased dramatically (395 up regulated for sapwood vs. bark/cambial region; 177 up regulated for sapwood vs. transition zone, and 291 down regulated for sapwood vs. bark/cambial region; 226 down regulated for sapwood vs. transition zone). This suggests that the sapwood region plays the role of a bridging zone between bark/cambial region and transitional zone. Table 5 shows the list of up-regulated genes in the sapwood compared to bark/cambial region and transition zone. In other words, only three ESTs (TZS0226, SWS0332, and SWS0562) had sapwood-specific up-regulation. Interestingly, the transcript of a sugar transport protein, which plays key roles in source-sink relationships (Lalonde *et al.*, 1999), was highly expressed in sapwood. The highly redundant

Table 2. The redundancy of EST clones based on contig analysis

Annotation	Contig number	Total (%)	BCS (%)	SWS (%)	TZS (%)	TZF (%)
Hypothetical protein (At2g41250)	228	37 (1.3)	27 (3.0)	0	8 (0.9)	2 (1.4)
Auxin-repressed protein	227	33 (1.1)	17 (1.9)	0	13 (1.5)	3 (2.1)
Metallothionein	226	18 (0.6)	2 (0.2)	1 (0.1)	14 (1.6)	1 (0.7)
Maturase	225	13 (0.4)	0	9 (0.9)	4 (0.5)	0
No hit	224	11 (0.4)	0	0	10 (1.1)	1 (0.7)
Metallothionein-like protein	223	11 (0.4)	11 (1.2)	0	0	0
Hypothetical protein (At3g03150)	222	11 (0.4)	0	0	11 (1.3)	0
Cytochrome B5 DIF-F	221	10 (0.4)	0	0	9 (1.0)	1 (0.7)
Aquaporin 1	220	9 (0.3)	9 (1.0)	0	0	0
Aquaporin 2	219	8 (0.3)	4 (0.4)	0	3 (0.3)	1 (0.7)
Hypothetical protein (A05191)	218	8 (0.3)	0	8 (0.8)	0	0
Phloem-specific protein Vein1	217	8 (0.3)	8 (0.9)	0	0	0
No hit	216	7 (0.2)	0	0	7 (0.8)	0
Metallothionein Class-II	215	7 (0.2)	2 (0.2)	0	3 (0.3)	2 (1.4)
Ubiquitin 2	14	7 (0.2)	6 (0.7)	0	1 (0.1)	0
No hit	213	7 (0.2)	7 (0.8)	0	0	0
No hit	212	7 (0.2)	6 (0.7)	0	0	1 (0.7)
Integral membrane transport protein	211	6 (0.2)	2 (0.2)	0	4 (0.5)	0
Hypothetical protein (D75542)	210	6 (0.2)	5 (0.6)	0	1 (0.1)	0
Extensin	83	0	0	0	0	2 (1.4)

Letters in bold are significantly different from the numbers in the other libraries at $P < 0.05$. For example, metallothionein-like protein (contig 223) and aquaporin 1 (contig 220) are significantly more abundant in the BCS zone than in SWS and TZS. The SWS library had a significantly higher number of ESTs in contig 218 (hypothetical protein, PIRA05191) than did the BCS and TZS libraries. Contigs 222 (hypothetical protein At3g03150) and 216 (no-hit clone) are significantly more abundant in TZS than in the BCS and SWS libraries.

Table 3. Summary of up-regulated or down-regulated genes

Functional category	Up-regulation			Down-regulation		
	in BC	in SW	in TZ	in BC	in SW	in TZ
Cell division and cycle	0	0	0	0	0	0
Cell wall structure and metabolism	0	0	0	1	0	0
Chromatin and DNA metabolism	0	0	1	1	0	0
Cytoskeleton	2	0	0	0	0	2
Defense	6	0	2	2	0	6
Gene expression and RNA metabolism	4	0	0	4	0	3
Membrane transport	7	1	0	0	0	5
Miscellaneous	16	0	1	3	0	16
Primary metabolism	8	1	2	1	0	5
Protein synthase and processing	2	0	1	0	0	5
Secondary and hormone metabolism	1	0	8	3	1	0
Signal transduction	1	0	2	1	0	0
Vesicular trafficking, protein sort, secretion	1	0	0	0	0	0
Unknown, Hypothetical	22	0	20	7	3	14
No hit	26	1	37	40	0	12
Total	96	3	74	63	4	68

Table 4. Up-regulated transcripts in the bark and cambial zone.

Clone ID	GenBank accession number	Annotation	BC/SW ratio*	BC/TZ ratio*
Cytoskeleton				
CLS0035	BI677465	actin depolymerizing factor 5	3.2	3.9
CLS0977	BI678088	Structural Maintenance of Chromosomes (SMC)-like Protein	5.1	3.3
Defense				
CLS0100	BI677437	Superoxide Dismutase (Cu-Zn)	2.4	2.4
CLS0239	BI677578	Metallothionein-like protein	5.4	3.5
CLS0595	BI677862	Superoxide Dismutase (Cu-Zn) sodB	2.0	2.1
CLS0801	BI677956	Metallothionein	2.8	2.5
CLS0867	BI678223	Glyoxalase (metalloglutathione transferase superfamily)	3.2	4.1
CLS0960	BI678073	Proteinase Inhibitor	3.2	2.4
Gene expression and RNA metabolism				
CLS0340	BI677663	Transcriptional regulator, putative	2.8	2.3
CLS0439	BI677739	Zinc Finger Protein ID1	2.2	2.7
CLS0673	BI678286	Transcription Factor like Protein	4.3	4.4
SWS1456	BI679312	Zinc Finger Protein ID1	2.2	2.7
Membrane transport				
CLS0025	BI677455	Lectin precursor, Bark Agglutinin I	7.7	3.8
CLS0051	BI677477	Vacuolar V-H subunit E [Citrus limon]	2.2	2.2
CLS0052	BI677478	Aquaporin	7.5	3.8
CLS0197	BI677550	Lectin	2.1	2.9
CLS0488	BI677780	Tonoplast Intrinsic Protein, delta type	7.0	5.9
CLS0929	BI678048	Lectin-related polypeptide	4.7	3.3
CLS0950	BI678064	Lectin like protein (hypothetical)	3.3	2.4
Miscellaneous				
CLS0022	BI677452	Phloem-specific protein Vein1	5.3	10.1
CLS0068	BI677490	Chlorophyll a/b binding protein	4.6	2.5
CLS0081	BI677501	MTN5 Gene Precursor	3.5	6.7
CLS0158	BI677521	Phloem-specific protein Vein1	8.8	5.5
CLS0274	BI677608	Trypsin Inhibitor (Serine Proteinase Inhibitor)	3.1	2.9
CLS0376	BI677690	Photosystem II 10K protein	2.6	4.2
CLS0423	BI677726	Photosystem II Protein X precursor	8.1	4.9
CLS0479	BI677773	Phloem-specific protein Vein1	7.1	4.9
CLS0536	BI677820	Core protein	2.8	2.1
CLS0564	BI677840	Photosystem II 10K protein	3.0	2.1
CLS0603	BI677869	Auxin-repressed protein	3.1	3.1
CLS0805	BI677959	Photosystem I Reaction Centre Subunit VI	5.0	3.7
CLS0819	BI677971	Trypsin Inhibitor (Serine Proteinase Inhibitor)	3.5	2.6
CLS0824	BI677975	Magnesium Chelatase (chlorophyll biosynthesis)	4.1	3.3
CLS0952	BI678066	Specific Tissue Protein 2	6.3	2.7
CLS1025	BI678111	Ripening related protein	4.2	3.0

Table 4. Continued.

Clone ID	GenBank accession number	Annotation	BC/SW ratio*	BC/TZ ratio*
Primary metabolism				
CLS0017	BI677447	Copper Amine Oxidase precursor	4.4	5.4
CLS0714	BI678307	Lipid Transfer Protein	4.2	3.6
CLS0745	BI677908	Alcohol Dehydrogenase 1	6.2	3.8
CLS0813	BI677966	Alcohol Dehydrogenase 7	3.6	3.0
CLS0930	BI678049	Acetoacetyl-CoA-thiolase	2.3	2.9
CLS0956	BI678070	Blue copper protein (Plantacyanin; Phytocyanin), basic	2.5	2.1
CLS1015	BI678039	Lipid Transfer Protein	4.5	2.9
TZS0305	BI642632	Epoxide Hydrolase	4.7	2.3
Protein synthesis and processing				
CLS0811	BI677964	Ribosomal Protein S16 protein	3.7	2.3
CLS0973	BI678084	Peptidylprolyl Isomerase; Cyclophilin (Cyp)	2.4	2.0
Secondary and hormone metabolism				
CLS0373	BI677687	Monoxygenase	3.5	2.6
Signal transduction				
CLS0377	BI677691	Protein Phosphatase 2C-like	3.2	2.3
Vesicular trafficking, protein sorting, secretion				
CLS0783	BI677940	ER retention receptor Erd2	2.3	2.8
Unknown				
CLS0032	BI677462	Unknown Protein	3.3	3.1
CLS0071	BI677493	Unknown Protein	2.3	2.1
CLS0186	BI677541	Unknown Protein, hypothetical	2.6	2.8
CLS0227	BI677572	Unknown Protein	4.0	3.3
CLS0273	BI677607	Unknown Protein, hypothetical	3.8	4.8
CLS0297	BI677624	Unknown Protein, hypothetical	2.6	2.5
CLS0324	BI677649	Unknown Protein, hypothetical	2.8	2.5
CLS0333	BI677657	Unknown Protein, hypothetical	2.5	2.2
CLS0342	BI677665	Unknown Protein, hypothetical	2.9	2.8
CLS0362	BI677678	Unknown Protein, hypothetical	3.3	2.2
CLS0406	BI677712	Unknown Protein, hypothetical	3.5	3.4
CLS0493	BI677785	Unknown Protein, hypothetical	3.1	2.9
CLS0537	BI677821	Unknown Protein, hypothetical	2.5	2.3
CLS0630	BI677890	Unknown Protein, hypothetical	10.7	3.1
CLS0740	BI677903	Unknown Protein, hypothetical	4.3	2.9
CLS0776	BI677934	Unknown Protein	3.3	2.7
CLS0800	BI677955	Unknown Protein	4.1	2.9
CLS0854	BI678208	Unknown Protein, hypothetical	2.3	2.1
CLS1014	BI678038	Unknown Protein, hypothetical	2.4	2.7
CLS1037	BI678119	Unknown Protein	3.0	2.8
CLS1100	BI678177	Unknown Protein, hypothetical	3.3	2.4
CLS1114	BI678186	Unknown Protein, hypothetical	2.5	2.7

Table 4. Continued.

Clone ID	GenBank accession number	Annotation	BC/SW ratio*	BC/TZ ratio*
No hit				
CLS0036	BI677466	No hit	3.6	2.4
CLS0055	BI677481	No hit	4.4	2.1
CLS0060	BI677485	No hit (contig 213)	5.8	3.9
CLS0066	BI677489	No hit	6.3	3.2
CLS0201	BI677554	No hit	3.4	2.6
CLS0211	BI677559	No hit (contig 212)	4.3	2.4
CLS0253	BI677591	No hit	5.2	4.3
CLS0349	BI677670	No hit	9.4	12.3
CLS0366	BI677681	No hit	4.9	2.5
CLS0378	BI677692	No hit	3.8	5.5
CLS0426	BI677729	No hit	2.1	2.0
CLS0461	BI677758	No hit	2.1	4.3
CLS0568	BI677843	No hit	2.8	2.7
CLS0608	BI677873	No hit	2.1	2.0
CLS0712	BI678306	No hit	2.9	2.3
CLS0739	BI677902	No hit	4.0	3.3
CLS0827	BI678187	No hit	6.5	4.8
CLS0840	BI678197	No hit	2.8	3.6
CLS0884	BI678235	No hit	8.0	2.9
CLS0982	BI678092	No hit	2.6	2.0
CLS1048	BI678129	No hit	2.5	2.5
CLS1095	BI678170	No hit	2.4	2.0
SWS0575	BI678741	No hit	2.2	3.6
SWS0655	BI678790	No hit	2.6	2.6
TZS0560	BI642188	No hit	3.3	3.2
TZS1309	BI642865	No hit	3.1	2.5

*The ratio was estimated as average value from data points.

EST (contig 218) in the sapwood library is also highly expressed in the sapwood when compared with the bark/cambial region and transition zone (at the ratio of 2.2 and 1.7, respectively).

There were dramatic differences in gene expression patterns in the transition zone relative to the other two comparisons. As shown in Figure 3D, the gene expression ratios in the transition zone versus the other two zones showed a more definite relationship, indicating a higher correlation and lessened variability (slope 0.6999 and $R^2 = 0.5123$). Genes with low expression ratios in the transition zone versus bark/cambial region were also low at the ratio of transition zone versus sapwood. In addition, genes that were highly expressed in the transition zone compared to bark/cambial region also had a higher expression in the transition zone than in sapwood region. Of

357 genes that had a higher expression in the transition zone than in either the bark/cambial region or sapwood region, 75 genes were specifically up regulated and 68 genes were down regulated in transition zone (Tables 3 and 6). Tables 3 and 6 list the proportion of functional categories of up-regulated genes in transition zone. The results of the EST analysis and microarray results show that the proportion of secondary and hormone metabolism is relatively high (10.7%) in the transition zone. In addition, three TZS clones (one unknown and two no-hit genes) that are highly redundant in the contig sets (contig 222, 224 and 216) were also highly expressed in transition zone. These results, when combined with the EST analysis and microarray data, show that secondary metabolism-related genes are up-regulated in transition zone. No-hit clones were found in a high

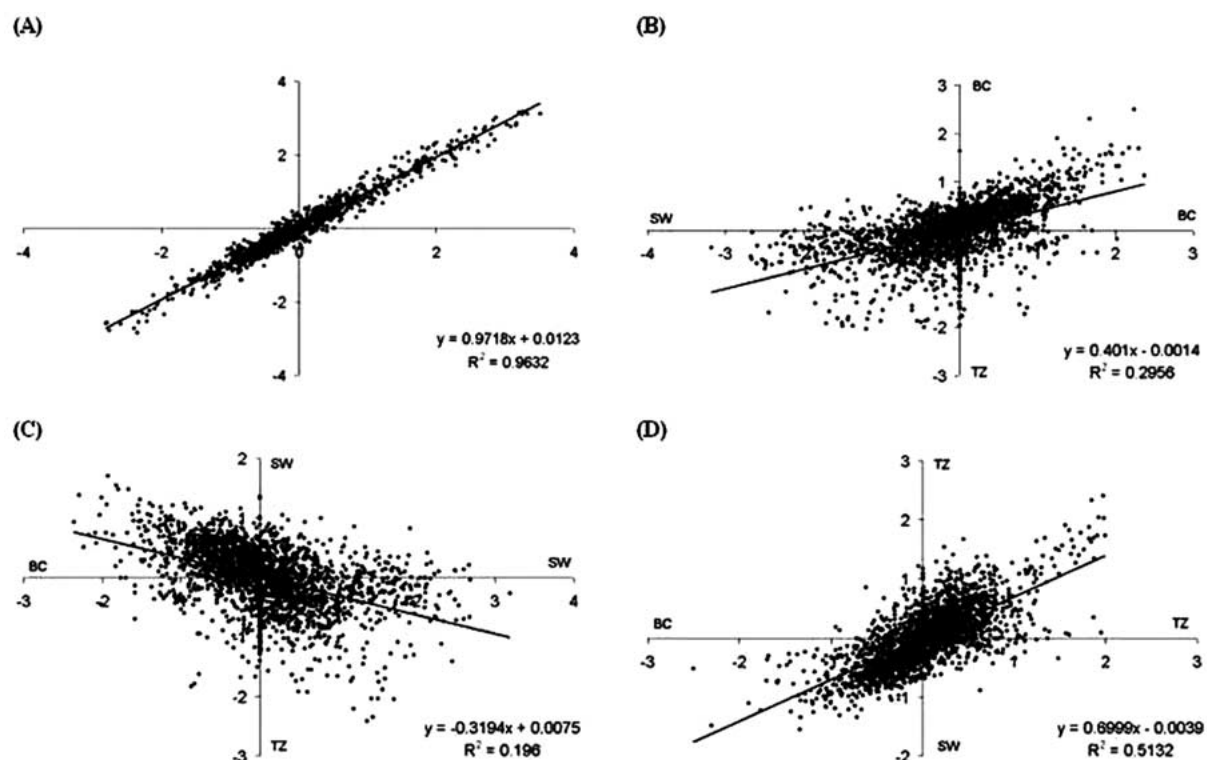


Figure 3. Scatter plots of microarray hybridization results. (A) The reproducibility of microarray experiments. Two hybridization experiments of SWS versus BCS were conducted using a dye-swap with Cy5 and Cy3 labeled probes. The natural log values of the Cy5-to-Cy3 ratios were plotted for the two replicates. (B) The BCS-specific expression pattern. The X-axis is the log-scaled ratio of gene expression for the experiment of BCS over SWS, and Y-axis is the log-scaled ratio of gene expression for the experiment of BCS over TZS. (C) The SWS-specific expression pattern. X-axis is the log-scaled ratio of SWS over BCS, and Y-axis is the log-scaled ratio of SWS over TZS. (D) The TZS-specific expression pattern. X-axis is the log-scaled ratio of TZS over BCS, and Y-axis is the log-scaled ratio of TZS over SWS. All number values of the log-scaled ratio of gene expression are average values of data points generated from each experiment. Abbreviations: BC, bark/cambial zone; SW, sapwood; TZ, transition zone.

proportion in this region as well, possibly indicating the relatively unstudied nature of this unique plant tissue zone. Interestingly, the bark/cambial region and the transition zone showed opposite gene-expression patterns, with most of the down-regulated genes in transition zone being up regulated in the bark/cambial region. For example, phloem-specific protein *VEIN1* was up regulated in bark/cambial region, but it was down-regulated in transition zone. Such contrasting gene expression is not unexpected considering the functional differences of the two regions. Transition zone is the region where many metabolic changes occur in the establishment of the inner wood. In fact, there was a diversity of genes classified into secondary and hormone metabolism as well as a remarkable increase in the proportion of genes related to secondary metabolism in the TZS library. This observation indicates that the transcript expression pattern of trans-

ition zone is highly related to secondary metabolism involving the flavonoid biosynthesis pathway.

Identification of the wood formation-associated genes

To visualize the inner-wood gene expression patterns that could potentially identify the wood formation-related genes, we performed hierarchical clustering of the arrayed genes based on microarray results. Data from the 2580 clones on our cDNA microarray were clustered from the data of two different experiments: SWS versus BCS and TZS versus BCS. Information from clustering not only allows for the identification of related expression patterns of different genes but also shows expression patterns of individual genes over different experiments. Figure 4A represents genes that show higher expression in inner wood (i.e., sapwood and transition zone) than in bark/cambial region. Many up-regulated genes

Table 5. Up-regulated transcripts in the sapwood.

Clone ID	GenBank accession number	Annotation	SW/BC ratio*	SW/TZ ratio*
Membrane transport				
TZS0226	BI642581	sugar transport protein	2.8	3.0
No hit				
SWS0332	BI678553	no hit	6.8	2.3
Primary metabolism				
SWS0562		Acetyl-CoA Carboxylase Carboxyl Transferase	2.4	3.3

*The ratio was estimated as average value from data points.

in inner wood are not characterized, but there are some interesting genes related to secondary metabolism (chalcone synthase, chalcone flavone isomerase and dihydroflavonol 4-reductase), primary metabolism (proline oxidase, cytochrome B5 DIF-F), and signal transduction (c-myc binding protein and GTP-binding protein ras-like). This observation suggests that these genes may be proximately induced in inner wood and their expression in inner wood may account for the heartwood formation in trunk wood. On the contrary, down-regulated genes are monooxygenase, alcohol dehydrogenase, methallothionein-like protein, phloem-specific protein, and so on (Figure 4B).

Confirmation of microarray results

Even though the data generated from microarray experiments are reproducible, the data should also be confirmed by northern blot analysis, western blot analysis, or RT-PCR analysis (Seki *et al.*, 2001; Perez-Amador *et al.*, 2001; Yu *et al.*, 2002). In order to compare the expression patterns of each region based on the same reference, we conducted the second microarray experiment using two probes labeled from the RNA populations of the target sample and the seedling control. Due to the nature of trunk wood samples, where only about 10% of the cells are live and the presence of wood extractives makes the isolation of RNA difficult, we were unable to obtain a large amount of mRNA for conventional northern blot analysis. Accordingly, we performed the approach of 'Antisense northern blot analysis' as well as EST results. Expression patterns in the microarray experiment were confirmed by antisense northern blot analysis (Figure 5) using Histon H 3.2 as control. Considering the data derived from ESTs, microarray, and 'antisense northern blot' analyses; the gene expression profiles reported here are considered reliable.

Discussion

In addition to the major constituents (i.e., cellulose, hemicellulose, and lignin), wood contains many extractives including tannins and other polyphenolics, coloring matter, essential oils, fats, resins, waxes, gum starch, and other secondary metabolites. These extractives make wood a veritable chemical storehouse that provides many organic compounds including biocides for biological control of insects and diseases, adhesives, biofuels, industrial oils, preservatives, pharmacologically active compounds, and rubber. Little is known about the molecular basis for such chemical diversity in heartwood. In order to gain insight on the temporal, spatial, and developmental regulation of the genes involved in the processes of heartwood and its extractive formation, we carried out global examination of gene expression profiles across the stems of mature heartwood-forming trees. It is very difficult to obtain high quality and quantity mRNA from the limited number of live inner wood cells impregnated with extractives. In this study, we characterized the transcriptional profile of black locust wood trunk through the generation of 2915 ESTs. Assembly of 2915 ESTs estimated the maximum number of unique genes represented in this set to be 2278. Because this analysis was performed on 5' end sequences that may arise from multiple non-overlapping segments of the same cDNA, the true number of unique genes may be overestimated. Of the 1304 matched ESTs that were analyzed, the largest number (30%) were uncharacterized genes, categorized as hypothetical or unknown proteins. Many of the known transcripts belonged to groups related to housekeeping genes, such as those involved in protein synthesis and processing. The low representation of genes associated with defense in the SWS library and with secondary and hormone metabolism in both the

Table 6. Up-regulated transcripts in the transition zone.

Clone ID	GenBank accession number	Annotation	TZ/BC ratio*	TZ/SW ratio*
Chromatin and DNA metabolism				
TZS0924	BI642324	SAP1 Protein	3.1	3.2
Defense				
TZS0124	BI642508	PR-10 Protein	5.3	3.6
TZS0357	BI642069	PR-10 Protein	3.2	2.1
Miscellaneous				
TZS1380	BI642911	NAM (no apical meristem)-like protein	2.1	2.7
Primary metabolism				
TZS0097	BI642647	Proline Oxidase	2.9	2.4
TZS0519	BI642157	Cytochrome B5 DIF-F	6.3	5.1
Protein synthesis and processing				
TZS0133	BI642514	Eukaryotic translation initiation factor eIF4F chain p28	4.8	6.8
Secondary and hormone metabolism				
TZS0021	BI642439	Chalcone-Flavone Isomerase	4.4	4.5
TZS0108	BI642499	Naringenin 3-Dioxygenase	6.5	5.5
TZS0312	BI642638	Flavonoid 3',5'-Hydroxylase	4.6	4.8
TZS0424	BI642100	Chalcone Synthase	5.0	4.2
TZS0751	BI642292	Flavonoid 3'-hydroxylase	3.4	2.6
TZS0854	BI642412	Chalcone Flavone Isomerase	2.6	2.3
TZS0870	BI642387	Dihydroflavonol 4-reductase	7.4	5.7
TZS0962	BI642432	Chalcone Reductase	5.5	4.1
Signal transduction				
TZF0100	BI642916	Serine/Threonine-specific protein kinase [<i>Arabidopsis thaliana</i>]	2.2	2.0
TZS0163	BI642537	GTP-binding Protein, ras-like	5.2	4.7
Unknown				
CLS1022	BI678108	Unknown Protein	3.4	3.1
SWS0040	BI678344	Unknown Protein, hypothetical	3.5	5.2
TZF0001	BI642988	Unknown protein	2.2	2.9
TZF0010	BI642996	Unknown protein	2.1	3.0
TZF0020	BI642927	Unknown protein putative protein [<i>Arabidopsis thaliana</i>]	2.3	2.8
TZF0137	BI643021	Unknown protein [<i>Arabidopsis thaliana</i>]	2.6	2.3
TZF0160	BI643042	Unknown protein F17L21.12 [<i>Arabidopsis thaliana</i>]	2.2	3.1
TZS0046	BI642463	Unknown Protein (contig 222)	4.9	6.2
TZS0295	BI642621	Unknown Protein, hypothetical	3.1	3.8
TZS0445	BI642109	Unknown Protein, hypothetical	6.5	3.9
TZS0547	BI642176	Unknown Protein, hypothetical	3.2	3.1
TZS0826	BI642402	Unknown Protein, hypothetical	2.0	2.1
TZS0828	BI642404	Unknown Protein, hypothetical	2.7	2.0
TZS0984	BI642764	Unknown protein [<i>Arabidopsis thaliana</i>]	5.1	5.3
TZS1012	BI642649	Unknown Protein, hypothetical	6.7	5.7
TZS1095	BI642715	Unknown Protein, hypothetical	5.6	6.0
TZS1161	BI642738	Unknown Protein, hypothetical	2.1	2.3
TZS1232	BI642802	Unknown Protein, hypothetical	3.2	3.2
TZS1287	BI642850	Unknown Protein, hypothetical	2.4	2.6
TZS1375	BI642907	Unknown Protein, hypothetical	2.5	2.6

Table 6. Continued.

Clone ID	GenBank accession number	Annotation	TZ/BC ratio*	TZ/SW ratio*
No hit				
SWS0033	BI678339	No hit	2.2	3.2
SWS0663	BI679126	No hit	2.6	2.2
SWS0732	BI678837	No hit	2.2	2.1
SWS0757	BI678855	No hit	2.6	2.0
SWS0820	BI679148	No hit	3.3	3.3
SWS0855	BI679178	No hit	3.5	2.7
SWS0931	BI678931	No hit	2.7	3.6
SWS0983	BI679030	No hit	3.5	2.2
SWS1086	BI679045	No hit	2.2	2.6
SWS1123	BI679072	No hit	2.0	4.0
TZF0086	BI642977	No hit	2.2	2.3
TZF0109	BI643000	No hit	2.1	2.9
TZS0038	BI642455	No hit (contig 224)	4.3	4.6
TZS0145	BI642532	No hit	6.2	5.4
TZS0161	BI642544	No hit	5.8	4.5
TZS0278	BI642608	No hit	3.7	3.3
TZS0362	BI642070	No hit	3.3	3.4
TZS0388	BI642081	No hit	2.1	2.5
TZS0392	BI642084	No hit	6.2	6.6
TZS0437	BI642106	No hit	3.9	2.6
TZS0440	BI642107	No hit	7.2	7.6
TZS0550	BI642179	No hit	6.8	7.7
TZS0617	BI642214	No hit (contig 216)	7.2	11.2
TZS0660	BI642240	No hit	2.2	2.1
TZS0837	BI642379	No hit	4.1	3.9
TZS0903	BI642337	No hit	3.4	2.9
TZS0922	BI642430	No hit	2.1	2.1
TZS0944	BI642356	No hit	6.3	10.4
TZS0953	BI642359	No hit	4.2	3.4
TZS0973	BI642756	No hit	3.6	3.0
TZS0986	BI642766	No hit	3.7	4.1
TZS1067	BI642697	No hit	2.2	3.3
TZS1112	BI642719	No hit	2.2	2.5
TZS1226	BI642796	No hit	3.8	5.4
TZS1268	BI642831	No hit	2.9	3.0
TZS1331	BI642873	No hit	3.1	3.0
TZS1334	BI642879	No hit	2.0	3.4

*The ratio was estimated as average value from data points.

BCS and SWS libraries are also notable considering the physiological importance of the regions. Thus, the libraries from the different trunk regions have distinctive characteristics based on the position of the sample in the wood. They show a common overall trend of the proportion of genes in the various categories, but with differential expression in some categories that

highlights the unique metabolic characteristics of each zone.

The results of our contig analysis show that functional categorization also presents the specified function of each tissue region. A direct proportion comparison of functional categories within each library could not be completed reliably, due to the high percent of

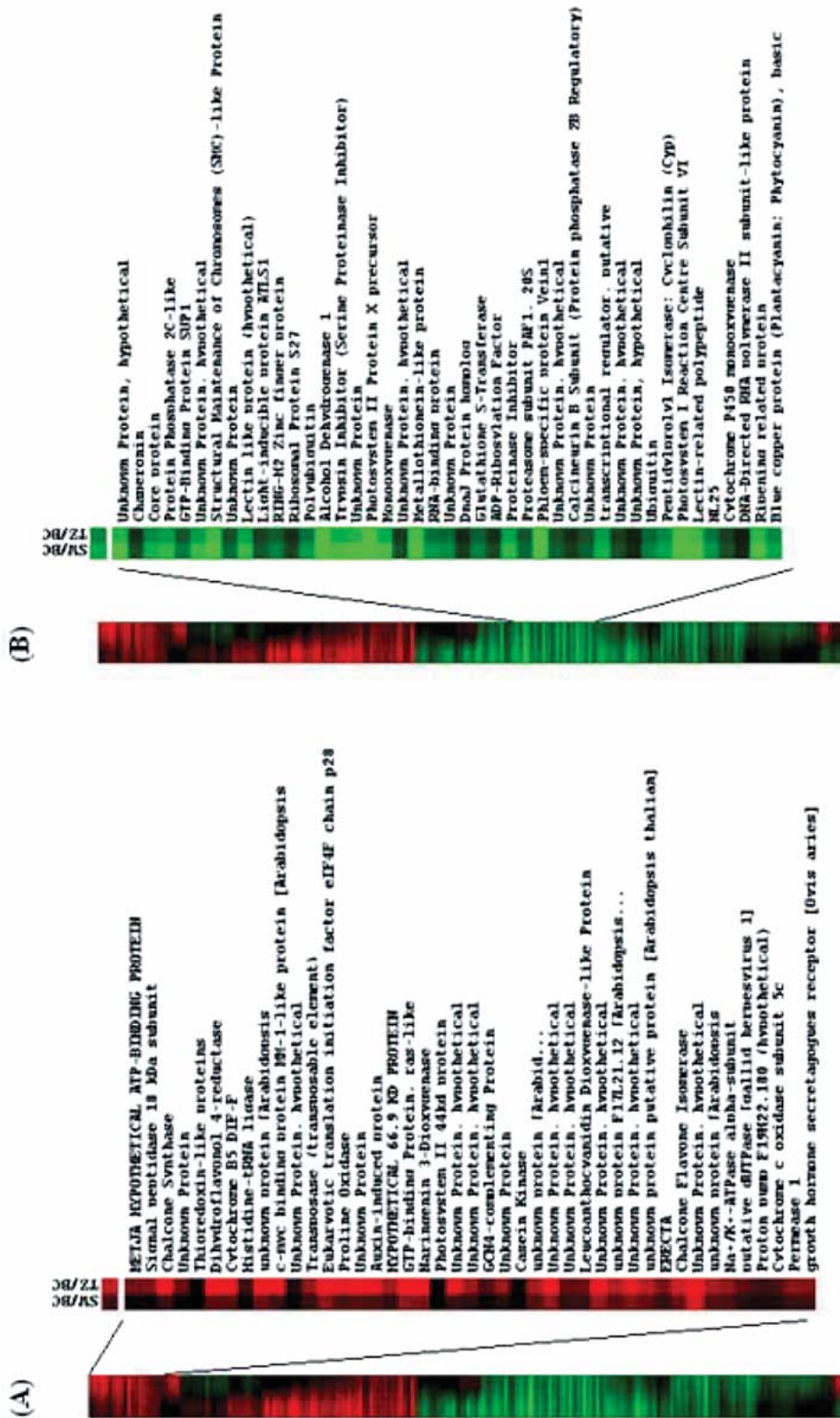


Figure 4. Cluster analysis of expression ratios from SWS or TZS vs. BCS. (A) The subcluster of highly expressed genes in inner wood (the sapwood and transition zone) from total hierarchical clustering display. (B) The subcluster of high expression genes in the bark/cambial zone from total hierarchical clustering display. Only hit EST clones were included for cluster analysis. (CZ); bark/cambial zone, SW; sapwood, and TZ; transition zone.

	(A) Microarray				(B) Antisense Northern Blot			
	SD/sd	BC/sd	SW/sd	TZ/sd	SD	BC	SW	TZ
Phloem specific protein	1	2.8	0.1	0.1				
Sugar transport protein	1	0.3	1.2	0.2				
Ser/Thr specific protein kinase	1	5.5	45	63				
WRKY-like protein	1	11	64	56				
Auxin-induced protein (TZS1033)	1	0.6	2.4	4.6				
PR-10	1	0.1	0.7	1.7				
Cytochrome b5 DIF-F	1	1.1	3.5	12				
Flavonoid 3,5-hydroxylase	1	2.8	-	23				
Naringenin 3-dioxygenase	1	8.6	10	50				
No hit (TZS0038, contig224)	1	0.5	9.2	45				
Histon H 3.2	1	0.8	1.1	1.0				

Figure 5. Confirmation of microarray Data with antisense northern blot analysis. (A) The expression ratio of the selected genes in microarray analysis with each sample (BC; bark/cambial zone, SW; sapwood, and TZ; transition zone) versus seedling (SD). Smaller than 1.0 means down-regulation in the target sample compared to the seedling control (sd). Greater than 1.0 indicates up-regulation in the target sample. (B) The expression patterns of the selected genes in seedling, bark/cambial zone, sapwood, and transition zone by antisense northern blot analysis. Each sample lane contained the equal amount of 200 ng of aRNAs.

no hit clones in the SWS. So, we excluded the portion of no-hit clones and compared the proportions of functional categories in each library. All three libraries contained similar percents of genes in their unknown function groups, as well as their housekeeping function category, which is related to protein synthesis and processing. When compared with other libraries, the proportion of membrane transport-related genes was high in the BCS, but the category of secondary and hormone metabolism membrane was highly scored in the TZS.

We predicted that the bark/cambial region library would contain many genes categorized into the cytoskeleton, vesicle trafficking, cell division and cycle functioning categories because this region includes actively growing and differentiating cells. However, even though representatives of these classes were found in the library, the proportions were not significantly different from those of other libraries. Additionally, we found that this library has a comparatively large number of genes classified into membrane transport such as: aquaporin, aquaporin-like protein PIP2, plasma membrane intrinsic protein, plasma membrane integral protein ZmPIP2-7, and delta type tonoplast intrinsic protein. Many transcripts (e.g., the genes en-

coding phloem-specific proteins) of unknown function or no hit were also highly expressed in this region. When considering the cell growth and division in this region, the proteins encoded by the genes in each of these categories are of interest for future analysis.

The sapwood sample included developing-xylem cells. Like the ESTs derived from the developing-xylem cells of pine and poplar, our sapwood library contained a higher concentration of those genes involved in cell wall synthesis, when compared to the other two libraries. However, this library possessed only a few transcripts coding for enzymes involved in the synthesis of lignin. Instead, there were clones corresponding to cell wall structural proteins including: extensin-like proteins, proline-rich proteins, leucine-rich repeat proteins, arabinoxylan arabinofuranohydrolase isoenzyme AXAH-II, and pectinacetylsterases. These results indicate that sapwood gene expression patterns more closely resemble the characteristics of inner wood gene expression than that of developing-xylem. This finding parallels the fact that sapwood is a part of inner wood and has ray cells, which remain alive and maintain their metabolic activity. Based on microarray results few genes are specifically expressed in the sapwood when com-

pared with bark/cambial region and transition zone. One possible explanation for the small number of specifically expressed genes in sapwood is that this region might play bridging roles between bark/cambial region and transition zone.

As in the EST results, genes involved in secondary metabolism were highly expressed in transition zone. Recent research has shown a specific example of this in black walnut (*Juglans nigra*), where flavonoid biosynthesis was up-regulated in the transition zone (Beritognolo *et al.*, 2002). In the July samples of *Juglans nigra*, both chalcone synthase (*CHS*) and flavonoid 3'-hydroxylase (*F3H*) reached their maximum levels of expression in the transition zone, while phenylalanine ammonia lyase (*PAL*) had increased expression in the outer sapwood. Its activity was much higher in the transition zone than the sapwood of black locust (Magel and Hübner, 1997). In the current study, its expression was increased only in the sapwood (8-fold), but no change in the transition zone. This may reflect the difference in the genotypes used in the studies, sampling time, age, or environmental conditions where the trees were grown. This suggests that the basis for the heartwood chemical diversity among different trees may be at the transcriptional level. Our study showed that *CHS* was up-regulated 5-fold in the transition zone as compared to the bark/cambium region, 4-fold compared to the sapwood, and the level of *F3H* was increased about 3-fold in the transition zone. These results corroborate with the gene expression patterns observed in black walnut (Beritognolo *et al.*, 2002). The expression of dihydroflavonol 4-reductase (*DFR*) could not be detected by northern blot, only by RT-PCR in the trunk of walnut, and showed no significant change across the stem sections. However, it was up-regulated 7- and 6-fold in the transition zone of black locust as compared to the bark/cambium region and the sapwood, respectively. This differential expression of the structural genes may explain the difference in the heartwood chemical profiles of the two species. The high expression of cytochrome B5 in the transition zone may be required for flavonoid biosynthesis during heartwood formation. This gene is known to enhance the activity of flavonoid 3',5'-hydroxylase, which catalyzes the 3',5'-hydroxylation of dihydroflavonols, the precursors of purple anthocyanins (de Vetten, *et al.*, 1999). We found that the expression level of *B5 DIF-F* (accession number B1642157) was up-regulated 6- and 5-fold in the transition compared to the bark/cambium region and the sapwood, respectively. Accordingly,

the expression of flavonoid 3',5'-hydroxylase gene was dramatically increased in the transition zone. The high level expression of the genes related to flavonoid biosynthesis in the transition zone fits well with the darkening of the heartwood that is mediated by these genes. The pathogenesis-related class 10 (PR-10), which is a protein related to defense, was also highly expressed in the transition zone, as well as the sapwood region. WRKY, which is a family of plant-specific zinc-finger-type transcriptional factors, was expressed in trunk wood at low levels. Eulgem *et al.* (1999) reported that the promoter of PR-10 gene is regulated by WRKY proteins as an early defense response in parsley. In addition, Ser/Thr specific protein kinase genes and auxin-induced protein genes were also highly expressed in the transition zone.

The stems of growing trees are characterized by low oxygen and high carbon dioxide (CO₂) concentrations (Carrodus and Triffett, 1975). Under such conditions, Magel (2000) suggested that the products of the accelerated oxidative pentose-phosphate pathway might be used for the synthesis of phenolics. Our microarray has two pentose-phosphate pathway genes, transaldolase and fructose biphosphate aldolase. Both of those genes were up-regulated in the sapwood. The plant hormone ethylene has been suggested as an important regulator of heartwood formation (Nilsson *et al.*, 2002). This view is further supported by the fact that ethylene production was greater in the transition zone than in the outer sapwood (Nelson, 1978) and it stimulates the activity of important enzymes for polyphenol biosynthesis (Roberts and Miller, 1983; Ingermarsson, 1995). Neither of the ethylene receptor gene (accession number B1677627) and the ethylene-responsive small GTP-binding protein (accession number B1677741) on our microarray showed changes in their expression levels across the stem.

One important question in the study of heartwood formation is the source of the carbon-skeletons used for the biosynthesis of heartwood extractives. The evidence gathered so far support the hypothesis that the substrates are derived from imported carbohydrates (Magel and Hübner, 1997; Hauch and Magel, 1998; Magel, 2000). The microarray analysis showed that the gene encoding sucrose transporter (accession number B1642581) was up-regulated 2.8-times in the sapwood, suggesting increased activity of sucrose transporter in the region. Carbohydrates have been shown to be distributed across trunk wood (Magel *et al.*, 1994; Ugglä *et al.*, 2001), and there are many re-

ports confirming that sugar transporters play a role in the cell-to-cell and long-distance distribution of sugars throughout the plant (for review, Williams *et al.*, 2000). The high expression of sugar transport protein genes in the sapwood region may indicate that carbohydrates from source tissues are transported to inner wood by the sugar transport proteins. During summer months, starch is accumulated in the sapwood and its accumulation correlates with enhanced sucrose synthase (SuSy) activities in the sapwood (Magel, 2000). Especially, the activity of SuSy increases dramatically at the sapwood-heartwood transition zone, which may lead to enhanced degradation of sucrose in the region where the synthesis and accumulation of phenolic heartwood extractives occur (Magel *et al.*, 1994; Magel and Huber, 1997; Magel, 2000). The activity of SuSy has been proposed as a measure for sink strength in tissues with extensive synthesis of phenolic compounds (Magel, 2000). In this case, our data add additional evidence for the view that heartwood extractives are synthesized at the transition zone using imported carbohydrates, not translocated via the phloem and the wood rays to the heartwood (Steward, 1966).

In summary, we report the first comprehensive study of gene expression profiles deep inside trunk wood of a hardwood tree. Our ESTs present unique gene sets that are expressed in uncharted plant tissues. Using DNA microarray analysis, we have identified genes associated with inner wood formation and profiled gene expression patterns in trunk wood. These genes will assist future investigations to unravel the molecular mechanisms regulating the formation of inner wood. Resolving the dilemma of achieving greater environmental protection of forest ecosystems while meeting the increasing demand for forest utilization necessitates gaining a fundamental understanding of the biochemical processes involved in tree growth and development. The findings described here will provide a platform for such attempts.

Acknowledgements

We thank Dr. Jeff Landgrof for his help with microarray printing, Paul Bloese for help with tree sampling, and Matias Kirst for help with the statistic analysis. Merylyn Ruthig, Jin Kim, and Dawn Overmeyer provided technical assistance. This project is supported by an USDA NRI Grant No. 2001-35504-10040 and USDA Special Grant No. 98-34158-5995.

References

- Adams, M., Kerlavage, A., Fields, C., and Venter, J. 1993. 3,400 new expressed sequence tags identify diversity of transcripts in human brain. *Nat Genet* 4: 256–267.
- Allona, I., Quinn, M., Shoop, E., Swope, K., Cyr, S., Carlis, J., Riedl, J., Retzel, E., Campbell, M., Sederoff, R., and Whetten, R. 1998. Analysis of xylem formation in pine by cDNA sequencing. *Proc Natl Acad Sci USA* 95: 9693–9698.
- Audic, S. and Claverie, J.-M. 1997. The significance of digital gene expression profiles. *Genome Research* 7: 986–995.
- Bairoch, A. and Apweiler, R. 2000. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 28: 45–48.
- Barker, W., Garavelli, J., Huang, H., McGarvey, P., Orcutt, B., Srinivasarao, G., Xiao, C., Yeh, L.-S.L., Ledley, R., Janda, J., Pfeiffer, F., Mewes, H.-W., Tsugita, A. and Wu, C. 2000. The Protein Information Resource (PIR). *Nucleic Acids Res.* 28: 41–44.
- Baugh, L., Hill, A., Brown, E., and Hunter, C. 2001. Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acid Res.* 29: e29.
- Benson, D., Karsch-Mizrachi, I., Lipman, D., Ostell, J., Rapp, B., and Wheeler, D. 2000. GenBank. *Nucleic Acids Res.* 28: 15–18.
- Beritognolo, I., Magel, E., Abdel-Latif, A., Charpentier, J., Jay-Allemand, C., and Breton, C. 2002. Expression of genes encoding chalcone synthase, flavanone 3-hydroxylase, and dihydroflavonol 4-reductase correlates with flavanol accumulation during heartwood formation in *Juglans nigra* L. *Tree Physiol.* 22: 291–300.
- Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirkse, W., Van Staveren, M., Stiekema, W., Drost, L., Ridley, P., Hudson, S.A., Patel, K., Murphy, G., Piffanelli, P., Wedler, H., Wedler, E., Wambutt, R., Weitzenegger, T., Pohl, T.M., Terry, N., Gielen, J., Villarroel, R., and Chalwatzis, N. 1998. Analysis of 1.9 Mb contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* 391: 485–488.
- Burtin, P., Jay-Allemand, C., Charpentier, J., and Janin, G. 1998. Natural wood colouring process in *Juglans* sp. (*J. nigra*, *J. regia* and hybrid *J. nigra* 23 × *J. regia*) depends on native phenolic compounds accumulated in the transition zone between sapwood and heartwood. *Trees* 12: 258–264.
- Carrodus, B., and Triffett, A. 1975. Analysis of respiratory gases in woody stems by mass spectrometry. *New Phytol.* 74: 243–246.
- Cercos, M., Santamaria, S., and Carbonell, J. 1999. Cloning and characterization of TPE4A, a thiol-protease gene induced during ovary senescence and seed germination in pea. *Plant Physiol.* 119: 1341–1348.
- Covitz, P., Smith, L., and Long, S. 1998. Expressed Sequence Tags from a Root-Hair-Enriched *Medicago truncatula* cDNA library. *Plant. Physiol.* 117: 1325–1332.
- Dazzo, F., and Hubbell, H. 1975. Cross-reactive antigens and lectins as determinants of symbiotic specificity in the *Rhizobium*-clover association. *Appl. Microbiol.* 30: 1017–1033.
- de Vetten, N., ter Horst, J., van Schaik, H.-P., de Boer, A., Mol, J., and Koes, R. 1999. A cytochrome *b5* is required for full activity of flavonoid 3',5'-hydroxylase, a cytochrome P450 involved in the formation of blue flowers. *Proc Natl Acad Sci USA* 96: 778–783.
- Dent, G., O'Dell, D., and Eberwine, H. 2001. Gene expression profiling in the amygdala: An approach to examine the molecular substrates of mammalian behavior. *Physiol. Behavior* 73: 841–847.

- Eisen, M., Spellman, P., Brown, P., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95: 14863–14868.
- Eulgem, T., Rushton, P., Schmelzer, E., Hahlbrock, K., and Somssich, I. 1999. Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. *EMBO J.* 18: 4689–4699.
- Ewing, B., Hillier, L., Wendl, M., and Green, P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8: 175–185.
- Fukuda, H. (1996). Xylogenesis: initiation, progression, and cell death. *Ann Rev Plant Physiol Plant Mol Biol* 47, 299–325.
- García-Hernández, M., Murphy, A., and Taiz, L. 1998. Matallothioneins 1 and 2 have distinct but overlapping expression patterns in *Arabidopsis*. *Plant Physiol.* 118: 387–397.
- Hauch, S. and Magel, E. 1998. Extractable activities and protein content of sucrose-phosphate synthase, sucrose synthase and neutral invertase in trunk tissues of *Robinia pseudoacacia* L. are related to cambial wood production and heartwood formation. *Planta* 207: 266–274.
- Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., Bhalerao, R., Uhlén, M., Teeri, T., Lundberg, J., Sundberg, B., Nilsson, P., and Sandberg, G. 2001. A transcriptional roadmap to wood formation. *Proc. Natl. Acad. Sci. USA* 98: 14732–14737.
- Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., and Ikeuchi, M. 2001. DNA Microarray Analysis of Cyanobacterial Gene Expression during Acclimation to High Light. *Plant Cell* 13: 793–806.
- Hillinger, C., Holl, W., and Ziegler, H. 1996a. Lipids and lipolytic enzymes in the trunkwood of *Robinia pseudoacacia* L. during heartwood formation. I. Radial distribution of lipid classes. *Trees* 10: 366–375.
- Hillinger, C., Holl, W., and Ziegler, H. 1996b. Lipids and lipolytic enzymes in the trunkwood of *Robinia pseudoacacia* L. during heartwood formation. II. Radial distribution of lipases and phospholipases. *Trees* 10: 376–381.
- Hillis W. 1987. Heartwood and the exudates, Springer, Berlin, Heidelberg, New York.
- Ingemarsson, B. 1995. Ethylene effects on peroxidases and cell growth patterns in *Picea abies* hypocotyl cuttings. *Physiol. Plant.* 94: 211–218.
- Kirst, M., Johnson, A., Retzel, E., van Zyl, L., Craig, D., Li, Z.-J., Whetten, R., Baucom, C., Ulrich, E., Hubbard, Kristy, and Sederoff, R. (2002). Quantitative inference in functional genomics of loblolly pine (*Pinus taeda* L.) using ESTs and microarrays. In *Proc. the 10th IAPTC&B Congress*. (Ed) I. Vasil. Kluwer Academic Publishers, Dordrecht, Netherlands. Kozłowski, T. and Pallardy, S. (1997). *Physiology of woody plants*. San Diego, Academic Press.
- Lalonde, S., Boles, E., Hellmann, H., Barker, L., Patrick, J., Frommer, W., and Ward, J. 1999. The dual function of sugar carriers: transport and sugar sensing. *Plant Cell* 11: 707–726.
- Lee, M., Kuo, F., Whitmore, G., and Sklar, J. 2000. Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Natl. Acad. Sci. USA* 97: 9834–9839.
- Livesey, F., Furukawa, T., Steffen, M., Church, G., and Cepko, C. 2000. Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene *Crx*. *Current Biology* 10: 301–310.
- Lorenz, W. and Dean, J. 2002. SAGE profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (*Pinus taeda*). *Tree Physiology* 22: 301–310.
- Magel, E. 2000. Biochemistry and physiology of heartwood formation. In *Molecular and Cell Biology of Wood Formation*. Eds. R. Savidge, J. Barnett and R. Napier. BIOS Scientific Publishers, Oxford, pp. 363–376.
- Magel, E., Jay-Allemand, C., and Ziegler, H. 1994. Formation of heartwood substances in the stemwood of *Robinia pseudoacacia* L.: II. Distribution of nonstructural carbohydrates and wood extractives across the trunk. *Trees* 165–171.
- Magel, E. and Hübner, B. 1997. Distribution of phenylalanine ammonia lyase and chalcone synthase within trunks of *Robinia pseudoacacia* L. *Bot. Acta* 110: 314–322.
- Matamoros, M., Baird, L., Escuredo, P., Dalton, D., Minchin, F., Iturbe-Ormaetxe, I., Rubio, M., Moran, J., Gordon, A., and Becana, M. 1999. Stress-induced legume root nodule senescence. Physiological, biochemical, and structural alterations. *Plant Physiol* 121: 97–112.
- Mauseth, J. 1998. *Botany: an introduction to plant biology*. Jones and Bartlett Publishers. Sudbury, Massachusetts.
- Miller, J., Arteca, R., and Pell, E. 1999. Senescence-associated gene expression during ozone-induced leaf senescence in *Arabidopsis*. *Plant Physiol.* 120: 1015–1024.
- Nelson, N. 1978. Xylem ethylene, phenol oxidising enzymes and nitrogen and heartwood formation in walnut and cherry. *Can. J. Bot.* 56: 626–634.
- Nilsson, M., Wikman, S., and Eklund, L. 2002. Induction of discolored wood in Scots pine (*Pinus sylvestris*). *Tree Physiology* 22: 331–338.
- Perez-Amador, M., Lidder, P., Johnson, M., Landgraf, J., Wisman, E., and Green, P. 2001. New molecular phenotypes in the *dst* mutants of *Arabidopsis* revealed by DNA microarray analysis. *Plant Cell* 13: 2703–2717.
- Reddy, A., and Poovaiyah, B. 1990. Molecular cloning and sequencing of a cDNA for an auxin-repressed mRNA: correlation between fruit growth and repression of the auxin-regulated gene. *Plant Mol Biol.* 14: 127–136.
- Roberts, L. and Miller, A. 1983. Is ethylene involved in xylem of differentiation? *Plant Physiol.* 6: 1–24.
- Saxe, H., Ellsworth, D. S., and Heath, J. 1998. Tree and forest functioning in an enriched CO₂ atmosphere. *New Phytol.* 139: 395–436.
- Schaffer, R., Landgraf, J., Accerbi, M., Simon, V., Larson, M., and Wisman, E. 2001. Microarray Analysis of Diurnal and Circadian-Regulated Genes in *Arabidopsis*. *Plant Cell* 13: 113–123.
- Sederoff, R., Kirst, M., Johnson, A., Retzel, E., Whetten, R., Vasques-Kool, J. & O'Malley, D. 2002. Homology of expressed genes in loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. The 10th IAPTC&B Congress, February 23–28, 2002, Orlando, FL.
- Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y., and Shinozaki, K. 2001. Monitoring the Expression Pattern of 1300 *Arabidopsis* Genes under Drought and Cold Stresses by Using a Full-Length cDNA Microarray. *Plant Cell* 13: 61–72.
- Stafstrom, J., Ripley, B., Devitt, M., and Drake, B. 1998. Dormancy-associated gene expression in pea axillary buds. Cloning and expression of PsDRM1 and PsDRM2. *Planta* 205: 547–552.
- Sterky, F., Regan, S., Karlsson, J., Hertzberg, M., Rohde, A., Holmberg, A., Amini, B., Bhalerao, R., Larsson, M., Villarroel, R., Van Montagu, M., Sandberg, G., Olsson, O., Teeri, T. T., Boerjan, W., Gustafsson, P., Uhlen, M., Sundberg, B., and Lundberg, J. 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.
- Steward, C. 1966. Excretion and heartwood formation in living trees. *Science* 153: 1068–1074.
- Uggla, C., Magel, E., Moritz, T., and Sundberg, B. 2001. Function and dynamics of auxin and carbohydrates during early wood/late wood transition in Scots pine. *Plant Physiol.* 125: 2029–2039.
- Wang, E., Miller, L., Ohnmacht, G., Liu, E., and Marincola, F. 2000. High-fidelity mRNA amplification for gene profiling. *Nat. Biotech* 18: 457–459.
- Williams, L., Lemoine, R., and Sauer, N. 2000. Sugar transporters in higher plants – a diversity of roles and complex regulation. *Trends in Plant Science* 5: 283–290.
- Wingler, A., von Schaewen, A., Leegood, R., Lea, P., and Quick, W. 1998. Regulation of leaf senescence by cytokinin, sugars, and light. *Plant Physiol.* 116: 329–335.
- Yu, Q., He, M., Lee, N., and Liu, E. (2002). Identification of MYC-mediated death response pathways by microarray analysis. *J Biol Chem.* 277: 13059–13066.