



RESEARCH PAPER

Microarray gene expression profiling of developmental transitions in Sitka spruce (*Picea sitchensis*) apical shoots

Michael Friedmann^{1,2}, Steven G. Ralph¹, Dana Aeschliman¹, Jun Zhuang³, Kermit Ritland³, Brian E. Ellis¹, Joerg Bohmann^{1,2,3,*} and Carl J. Douglas^{2,*}

¹ Michael Smith Laboratories, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

² Department of Botany, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

³ Department of Forest Sciences, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

Received 17 July 2006; Revised 20 September 2006; Accepted 16 October 2006

Abstract

The apical shoot drives the yearly new stem growth of conifer trees, is the primary site for the establishment of chemical and physical defences, and is important in establishing subsequent perennial growth. This organ presents an interesting developmental system, with growth and development progressing from a meristematic tip through development of a primary vascular system, to a base with fully differentiated and lignified secondary xylem on the inside and bark tissue with constitutive defence structures such as resin, polyphenolic phloem parenchyma cells, and sclereids on the outside. A spruce (*Picea* spp.) microarray containing approximately 16.7K unique cDNAs was used to study transcript profiles that characterize the developmental transition in apical shoots of Sitka spruce (*Picea sitchensis*) from their vegetative tips to their woody bases. Along with genes involved in cell-wall modification and lignin biosynthesis, a number of differentially regulated genes encoding protein kinases and transcription factors with base-preferred expression patterns were identified, which could play roles in the formation of woody tissues inside the apical shoot, as well as in regulating other developmental transitions associated with organ maturation. Preferential expression of known conifer defence genes, genes encoding defence-related proteins, and genes encoding regulatory proteins was observed at the apical shoot tip and in the green bark tissues at the apical shoot base, suggesting a com-

mitment to constitutive defence in the apical shoot that is co-ordinated with rapid development of secondary xylem.

Key words: Conifer ESTs, lipid transfer protein, resin duct, secondary cell wall, terpenoid secondary metabolism, white pine weevil (*Pissodes strobi*), xylem.

Introduction

Sitka spruce (*Picea sitchensis*) is native to the west coast of North America, and Sitka spruce trees originating from coastal British Columbia (Canada) are an important timber species in planted forests in Scotland and other parts of Europe. Compared with other conifers, Sitka spruce has a fast growth rate, it can attain very large sizes, and its fibre and wood are of very high quality (Burns, 1990; Petersen *et al.*, 1997). However, planting of Sitka spruce in Canada is now severely restricted by white pine weevil (*Pissodes strobi*) infestations (King *et al.*, 1997).

In the last few years, a large-scale conifer genomics programme has been developed that has facilitated studies on the defence and resistance mechanisms of Sitka spruce against white pine weevil (Miller *et al.*, 2005; Ralph *et al.*, 2006b) as well as studies on secondary xylem development in conifers. The shoot apical leader is an important organ for the study of both processes.

A key feature of conifer defence is the development of cellular and anatomical structures for chemical defence in

* To whom correspondence should be addressed. E-mail: bohlmann@interchange.ubc.ca or cdouglas@interchange.ubc.ca

Authors contributions: MF conducted experiments, analysed data and helped with manuscript preparation; SGR provided materials, helped with data analysis and manuscript preparation; DA, JZ, and KR helped with data analysis and data interpretation; BEE helped with data interpretation; JB conceived the study, designed experiments, played a lead role in data interpretation and manuscript preparation; CJD helped direct the study and played a lead role in data interpretation and manuscript preparation. JB, CJD, BEE, and KR secured funding for this study.

the apical shoot. These anatomical structures include primary resin ducts filled with terpenoid oleoresin that remain active in the bark and xylem for long periods of time, phloem polyphenolic parenchyma (PP) cells which accumulate vacuolar phenolics, and lignified sclereids (Franceschi *et al.*, 2005). While much is known about the genes and enzymes of terpenoid biosynthesis in conifer defence (Keeling and Bohlmann, 2006), as well as their inducible expression (Miller *et al.*, 2005), relatively little is known about their levels of constitutive expression relative to the early development of resin ducts in the apical shoot. Also, very little is known about expression profiles of genes potentially associated with the early development of PP cells or sclereids in the spruce apical shoot.

Wood formation is a characteristic feature of tree growth and development that requires the development of a vascular cambium and subsequent formation of specialized xylogenetic tissue at the apex of woody plants that drives secondary growth (Mellerowicz *et al.*, 2001; Plomion *et al.*, 2001; Savidge, 2001). As cells differentiate during secondary growth, and make the transition from cambial derivatives to mature tracheary elements, a progression of cell differentiation processing towards the inside of the stem is established. This progression, which involves cell elongation, formation of a cellulose-rich secondary cell wall, lignin deposition, and programmed cell death, is under tight developmental control. Transcription factors and other regulatory proteins are therefore likely to play critical roles in modulating the expression of appropriate suites of genes required for cell expansion, secondary wall formation, and lignin deposition (Anterola *et al.*, 2002; Newman *et al.*, 2004; Schrader *et al.*, 2004; Ehltling *et al.*, 2005).

Many recent studies, using genomic tools, have addressed the genetic regulation of tree growth, wood formation, and fibre quality. For instance, gene expression profiles have been described for different stages of developing secondary xylem in poplar (Hertzberg *et al.*, 2001; Schrader *et al.*, 2004), heartwood development in black locust (JM Yang *et al.*, 2003, 2004), differentiating secondary xylem and compression wood formation in pine (SH Yang *et al.*, 2004), *in vitro* differentiation of xylem vessel elements in *Zinnia* (Demura *et al.*, 2002), secondary xylem development in *Eucalyptus* (Paux *et al.*, 2004), and interfascicular fibre formation in *Arabidopsis* (Oh *et al.*, 2003; Ehltling *et al.*, 2005). Genes involved in the synthesis of lignin and cellulose, the main components of secondary cell walls, have been characterized in the genome of *Arabidopsis*, and in trees such as poplar, aspen, and pine (Anterola *et al.*, 2002; Doblin *et al.*, 2002; Boerjan *et al.*, 2003; Gardiner *et al.*, 2003; Joshi, 2003; Raes *et al.*, 2003; Djerbi *et al.*, 2004). Nevertheless, the contributions of most genes preferentially expressed in woody tissues to the control of wood formation and wood quality remain undefined. In particular, the transition from

primary to secondary growth at the apex of trees has not been studied extensively.

The apical shoot of conifer trees drives the new annual stem growth, is the primary site for the establishment of long-lasting chemical and anatomical defences, and lays the foundation for subsequent perennial and secondary stem growth. This organ presents an interesting developmental system, with growth and development progressing from a meristematic tip that controls primary growth, through development of the primary vascular system, to a woody base with fully differentiated and lignified secondary xylem on the inside and a complex bark tissue with constitutive resin ducts, polyphenolic phloem parenchyma cells, and sclereids on the outside. As far as is known, the patterns of gene expression associated with development of conifer apical shoots have not been described in the literature. Using a newly developed 16.7K element spruce cDNA microarray (S Ralph, J Bohlmann, unpublished data), the transcripts associated with developmental transition from primary to secondary growth and the parallel development of constitutive defence systems in apical shoots of Sitka spruce, from their green shoot tips to their woody bases, were profiled. This work revealed sets of structural and regulatory genes, which are likely to be involved in these processes, and apparent conservation of many genes with those in angiosperm systems.

Materials and methods

Plant material

Seedlings of Sitka spruce [*Picea sitchensis* (Bong.) Carriere] clone FB3-425 were derived from somatic embryo cultures and were generously provided by Dr David Ellis (CellFor, Victoria, BC, Canada). Maintenance of seedlings and growth conditions were as previously described (Miller *et al.*, 2005; Ralph *et al.*, 2006b). Two-year-old Sitka spruce trees were grown outdoors in 2 gallon pots containing a peat:fine bark:pumice mix (2:1:1), balanced to a target pH of 5.5–6.5 with limestone and dolomite. During the summer months the trees were watered daily and received a 15–15–15 Cal-Mag fertilizer at 120 ppm (Scotts, Marysville, OH, USA) every other day. During the winter months, the trees were watered very infrequently, and received no fertilizer. For the third year of growth, the trees were transferred to 6 gallon pots.

In early September 2003, before the onset of dormancy, the apical shoot representing the current year's growth on 2-year-old trees showed an obvious progression between primary vascular growth at the tip and secondary xylem growth at the base. Similar-sized apical shoots (13–18 cm in length) were harvested from three trees and hand-sectioned at 2 cm intervals for histological characterization of the developmental stages along the apical shoot. The extent of lignification was visualized in hand sections by UV autofluorescence using a fluorescent microscope (DMR, Leica, Wetzlar, Germany) with excitation filters at 340–380 nm and emission at 450+ nm. Images were processed using the Openlab software (version 4.0.2; Improvion, MA, USA). Tissues were then collected from 20 trees for cDNA microarray profiling. As shown in Fig. 1A, the very top 0.5 cm of the apical shoot was discarded, and a 2 cm section just below it was excised, its needles were removed, and the

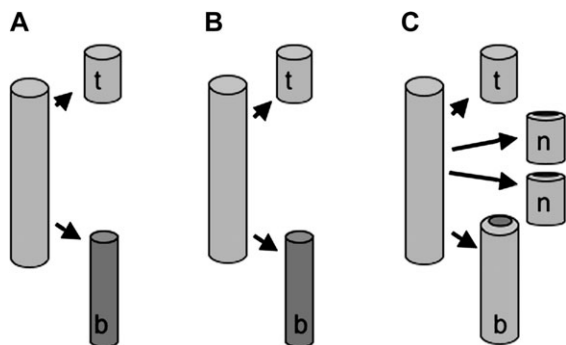


Fig. 1. Experimental design and sampling of Sitka spruce apical shoot stem sections. Apical shoots were sectioned into a tip segments (t) and a base woody stem segment (b) with the bark removed for a set of 20 trees in 2003 (A) and 10 trees in 2004 (B). An additional set of 10 trees was used in 2004 (C) and harvested for apical shoot tips (t) and woody base segments including bark (b), and two intermediate segments (n).

stem section was immediately frozen in liquid nitrogen. The collected sections were pooled. Likewise, a 3 cm section was collected from the base of the apical shoot, c. 3 cm above the junction with the previous year's growth. Needles were removed, bark was peeled off, and the remaining woody stem containing secondary xylem was immediately frozen in liquid nitrogen.

In September 2004, apical shoot tissues were collected in a similar manner from the same clonal group of trees (now 3 years old) (Fig. 1B, C). Tip and base tissues were collected from 10 trees as before (Fig. 1B). Additional 2–3-cm-long segments were taken from another 10 trees at similar positions at the tip and base of the apical shoot, as well as at two approximately central positions separated by 1–2 cm, depending on the overall length of the particular apical shoot. These intermediate segments were harvested and frozen in liquid nitrogen without separating bark and xylem (Fig. 1C). Small (0.5 cm) segments were also collected into fixative (4% formaldehyde; Canemco, Ontario, Canada) in 50 mM PIPES buffer (pH 7.2). After passage through a dehydration series in ethanol (30%, 50%, 70%, 80%, 95%, and three times 100%), these were embedded in LR White (Canemco) in a progression series with 100% ethanol (1:3, 1:1, 3:1, and 100% resin), under vacuum to remove bubbles. The embedded blocks were trimmed, and sectioned into 2 μ m sections using a Leica microtome (model Ultracut T). Sections were visualized under light microscopy, using a Zeiss model AxioPlan 2 microscope, and UV autofluorescence observed using a mercury UV source. Micrographs were captured directly into the Northern Eclipse program (EmPix Imaging Inc., Ontario, Canada).

RNA isolation

RNA was isolated from frozen tissues following the method of Chang *et al.* (1993) with modifications. Approximately 1 g of tissue was ground to a powder in liquid N₂, and extracted with 14 ml extraction buffer (100 mM TRIS-HCl, 25 mM EDTA, 2 M NaCl, 2% CTAB, 2% PVP, 0.5 μ g ml⁻¹ spermidine, 2% β -mercaptoethanol, pH 8, pre-warmed to 65 °C), vortex mixed, and an equal volume of chloroform:isoamyl alcohol (24:1 v/v) was added and the samples centrifuged at 3000 *g* (Sorvall RT-7 tabletop) for 20 min. The upper phase was re-extracted with 15 ml of chloroform:isoamyl alcohol and centrifuged. RNA was precipitated from the combined aqueous phases overnight at 4 °C by adding a one-quarter volume 10 M LiCl. RNA was pelleted by centrifugation for 30 min at 3000 *g* as above and resuspended in 1–2 ml SSTE buffer (1 M NaCl, 0.5% SDS, 10 mM TRIS-HCl, 1 mM EDTA, pH 8). RNA samples were extracted with chloroform:isoamyl alcohol, and RNA precipitated by the addition of one-tenth volume 5 M NaCl and 2

volumes EtOH at –20 °C overnight. After pelleting by centrifugation at 4 °C, RNA was rinsed with 70% EtOH, and resuspended in DEPC-treated water. RNA concentration was determined spectrophotometrically (Pharmacia Ultrospec 3000), and an aliquot visualized by agarose gel electrophoresis. RNA quality was assessed by first-strand cDNA synthesis using MMLV reverse transcriptase (Gibco-BRL, Gaithersburg, USA) with the addition of [³²P]dGTP, as described (Ralph *et al.*, 2006b).

Microarray hybridizations

The 16.7K spruce microarray contains 16 700 cDNA elements in addition to negative controls, and spots corresponding to positive spikes and orientation markers. This array shares many features with a previously described spruce 9.7K cDNA microarray (Ralph *et al.*, 2006a) and will be fully described elsewhere (S Ralph and J Bohlmann, unpublished data). First strand synthesis reactions incorporating a dendrimer-trapping oligonucleotide were conducted as described in the Dendrimer 350 kit (Genisphere) with minor modifications as described by Ralph *et al.* (2006b). Microarray pretreatments, hybridizations, and washes were carried out according to the Genisphere kit instructions with modifications as described by Ralph *et al.* (2006b). For these experiments, the hybridization mix was prepared using the Dendrimer 350 kit SDS-based hybridization buffer, complemented with LNA blocker included in the kit (3 μ l per reaction), salmon sperm DNA (2 μ g per reaction, Sigma), Cy5-labelled GFP (orientation marker), and the cDNA samples for a total volume of 45 μ l. Hybridizations were carried out at 60 °C for 16 h, and the second dendrimer hybridizations were carried out at 60 °C for 3 h and processed as described by Ralph *et al.* (2006b). For comparisons between apical shoot tip and base tissue collected in 2003, hybridizations were repeated 10 times with pooled tissue, with an equal number of dye swaps. For the 2004 tip versus base biological replicate, six hybridizations with dye swaps were performed using pooled tissue. For comparison of apical shoot tip to base in the 2004 samples without removal of the bark (Fig. 1C), six hybridizations with dye swaps were carried out. Hybridized slides were scanned on a Scan Array Express (Perkin Elmer, Foster City, CA, USA). For slides from tissues harvested in 2003, laser settings were at 90%, the Cy3 channel was scanned for most slides at a photomultiplier tube (PMT) setting of 69, and the Cy5 at a setting of 80. For slides from tissues harvested in 2004, settings were Cy3 PMT 68 and Cy5, 78. Scan intensities were comparable between each set of slides for a given hybridization.

Image processing and data analysis

Scanned images were processed using Imagene (v. 5.0) extraction software (BioDiscovery, Marina Del Rey, USA) as described by Ralph *et al.* (2006b). For all analyses, the median pixel intensities for each spot were used. Further analyses were performed excluding control elements using customized scripts for R and Bioconductor (R Development Core Team, www.r-project.org). For background correction, the mean of the lowest 10% of spot foreground intensities from a particular subgrid was taken as the background for that subgrid, and subtracted from the foreground intensity of each spot in the same subgrid. Each experiment consisted of comparing apical shoot base with apical shoot tip. A paired *t*-test on the background corrected, normalized set of intensity differences for each unique gene or cDNA in each experiment was performed. Normalization for each array independently was achieved by passing the two sets of intensities from the two channels of the arrays to the variance stabilizing transformation (Huber *et al.*, 2002) function in the Bioconductor package of the same name. The difference between signal intensities derived from shoot base and shoot tip was calculated and fold-change differentials were calculated by exponentiating the mean of the differences (Huber *et al.*, 2002). Statistical robustness

of differential intensities from each gene being assessed was based on comparison of the paired *t*-statistics obtained from the set of differences shoot base–shoot tip to the tabulated values of *t*-statistics (i.e. parametric *P*-values) and by choosing limiting values for fold-difference (>2-fold difference with a *P*-value <0.01 in the 2003A experiment).

Validation of differentially expressed genes by real-time (RT)-PCR

Appropriate reference genes for quantitative RT-PCR, WS00912_N13 (hypothetical protein) and WS0109_C05 (peroxisomal targeting signal receptor), were identified by screening the microarray data for cDNAs whose signals across all arrays remained apparently unchanged (fold-difference ratios between 0.99 and 1.01), and also displayed very low standard deviations for technical replicates. Presence of these transcripts was confirmed by RT-PCR to be consistently stable across the samples profiled (Table 1). Details of quantitative RT-PCR analysis of spruce transcripts have been described previously (Ralph *et al.*, 2006b). In brief, total RNA (18 µg) from the 2004B and 2004C samples was treated with DNase I (Invitrogen, Carlsbad, CA, USA). Absence of DNA in the treated RNA (10 ng) was confirmed by PCR using primers for WS00912_N13. Next, RNA (three reactions of 4 µg each per sample) was reverse transcribed and cDNA (10 ng) was analysed by PCR in a total volume of 20 µl, in the presence of 10 µl DyNAmo SYBR Green Mastermix (FinnZymes, Finland) and 0.3 µM each of a forward and a reverse primer (see Supplementary Table S5 at *JXB* online). Primers of 20–24 nucleotides in length were designed to amplify a gene-specific 160–200 bp fragment of the target cDNA (usually in the 3'-untranslated region) with a *T_m* of 60 °C. Reactions were carried out in triplicate in an MJR Opticon2 RT-PCR machine with an initial step of 15 min at 95 °C, followed by 40 cycles of 10 s at 94 °C, 30 s at 60 °C, and 30 s at 72 °C. Each cycle

was followed by a data-acquisition step. After the last cycle and a final 10 min extension at 72 °C, a melting curve was measured from 65 °C to 95 °C, with readings every 0.2 °C and holding for 1 s. Data analysis of the results was carried out by first adjusting threshold cycles manually, and subtracting the baseline fluorescence from cycles 3–9, and then determining individual efficiency values for each well using the LinReg program (Raemakers *et al.*, 2003). The Δ CT values between woody stem and tip were determined for each gene and for the two references, and fold change in expression determined from the Δ CT ratio between them (Pfaffl, 2001). When comparing expression along the apical shoot without removing bark, the fold-difference ratios were compared with the level of expression at the tip. In this case, the apical shoot was separated into four approximately equal segments along its length, usually with 1.5–2 cm between them, covering the apical shoot from the tip to the base.

Phylogenetic analysis

Twelve spruce full-length cDNAs for putative lipid transfer proteins (LTPs) represented on the array were identified by the presence of seven characteristic conserved Cys residues. Alignment of their deduced amino acid sequences to an *Arabidopsis* LTP gene (At3g13870) was carried out using DIALIGN (http://www.infobiogen.fr/services/analyseq/cgi-bin/dialign2_in.pl). The aligned sequences were edited manually, and a section covering a highly divergent N-terminal portion of the proteins was excluded from the analysis, so that the alignment covered a conserved region spanning most of the protein (*c.* 97 amino acids), starting at a conserved Glu residue five amino acids upstream of the first conserved Cys residue that forms part of the characteristic signature of the LTP family (Arondel *et al.*, 2000). A phylogeny was reconstructed using the PHYML maximum likelihood method (PHYML, <http://www.lirmm.fr/~guindon/phyml/body.html>) based on a default NJ tree, JTT amino acid evolution model, and generation of 100 bootstraps.

Table 1. Real-time RT-PCR validation of differentially expressed candidate genes identified by microarray expression profiling: fold-change values from RT-PCR data are compared with fold-change value from expression profiling

Clone ID	Best <i>Arabidopsis</i> BLASTX match ^a	Putative function	FC 2003A ^a	FC 2004B ^a	FC RT-PCR ^b	SD RT-PCR ^c
Reference cDNAs						
WS00912_N13	At1g14990	Hypothetical protein	1.01	0.98	1.22	0.432
WS0109_C05	At1g29260	Peroxisomal targeting signal receptor	1.05	0.95	1.35	0.368
Protein kinases						
WS0031_P03	At5g24430	CDPK	13.86	5.14	3.20	1.262
WS0031_K17	At3g46290	Receptor kinase	3.61	4.18	9.51	6.220
WS0039_K14	At1g48480	Receptor kinase	4.34	4.07	8.16	4.867
IS0012_O08	At1g73500	MAPKK	0.31	0.14	0.28	0.061
IS0014_A12	At2g31880	Receptor kinase	0.08	0.05	0.04	0.011
Reference cDNAs						
WS00912_N13	At1g14990	Hypothetical protein	1.01	0.98	0.79	0.174
WS0109_C05	At1g29260	Peroxisomal targeting signal receptor	1.05	0.95	1.18	0.290
Transcription factors						
WS00712_A21	At5g16600	myb	10.25	8.06	18.89	2.672
WS0083_F20	At2g44745	WRKY	8.43	10.55	7.35	1.568
WS0083_O11	At5g60450	ARF-4	4.39	4.24	7.03	1.734
WS00915_B21	At5g06710	Hd ZipIII	3.92	6.26	12.54	1.986
WS0092_N10	At1g52150	Hd ZipIII	2.90	2.35	5.46	1.319
WS0091_B15	At1g63650	bHLH	0.21	0.09	0.06	0.014
XET						
WS0087_M10	At5g13870	XET	16.42	29.00	33.75	7.527

^a Data from Tables 3 and 4 and Supplementary Table S1 (except reference genes).

^b The mean fold-change values derived from three replicate RT-PCR assays.

^c Standard deviation for fold-change values to the left.

Results

Anatomy of Sitka spruce apical shoots

In order to define appropriate developmental stages to be compared by gene expression profiling, spruce apical shoots were first hand-sectioned at 1 cm intervals from the top to the base, and different segments examined by light and fluorescence microscopy (Fig. 2). Sections from apical shoots harvested in 2003 and 2004 showed a rapid progression from individual vascular bundles at the top of the apical shoot, visualized by UV autofluorescence, to secondary xylem and the beginnings of a woody stem at the base (Fig. 2B, C). Micrographs of thin sections revealed additional details of the anatomy, including the presence of constitutive, axial resin ducts and PP cells in the bark (Fig. 2D, E). In the xylem, the transition to secondary growth was already detected in sections at about 2 cm below the top segment (data not shown). Extensive lignification of secondary cell walls visualized by UV autofluorescence was in the lower sections of the leader (Fig. 2C).

Expression profiling of spruce apical shoot sections

Gene expression profiling was carried out on the tip and base apical shoot sections in three separate experiments (2003A, 2004B, and 2004C; Fig. 1). In each experiment, a direct comparison was made of the gene expression profiles in the apical shoot tip, which included all tissues, and the woody stem at the base of the apical shoot by co-hybridizing Cy3- and Cy5-labelled probes derived from RNA samples from both tissues. Experiments 2003A and 2004B were identical except that the trees being sampled

were 1 year older in 2004B, and in these experiments, samples from the woody base were enriched for secondary xylem by removal for the green bark tissue. In the 2004C experiment, the same material was used as in 2004B, but the bark was not removed at the base of the apical shoot, so that gene expression data could be obtained for the apical shoot base containing all tissues (secondary xylem and bark). Functional annotation of array elements was assigned according to BLAST searches against TAIR (The Arabidopsis Information Resource; www.arabidopsis.org) *Arabidopsis* peptide set, with a BLASTX expect value (E) $<1e^{-05}$ threshold, as described by Ralph *et al.* (2006b).

The fold differential transcript abundance ['fold change' (FC)] values for all array elements for all three experiments, arranged according to preferred expression at the apical shoot base or tip, are shown in Supplementary Tables S3 and S4 at *JXB* online. Overall, the microarray data were comparable between the two years (experiments 2003A and 2004B), as shown by the year-to-year similarity of FC ratios and trends for hybridization to most of the cDNA elements (Tables 2–4; see Supplementary Tables S1–S4 at *JXB* online). For the 2003A data set, a total of 3522 genes were differentially expressed over 2-fold between the two apical shoot segments (P -value <0.01), and 2652 such genes were observed in 2004B. Of these two gene sets, 930 (35% of the 2004B set) were differentially expressed over 2-fold ($P < 0.01$) in both years of the replicated experiment (2003A and 2004B). A higher percentage was found when only the most highly differentially expressed genes were considered. For example, of the 343 genes in experiment 2004B that had an FC

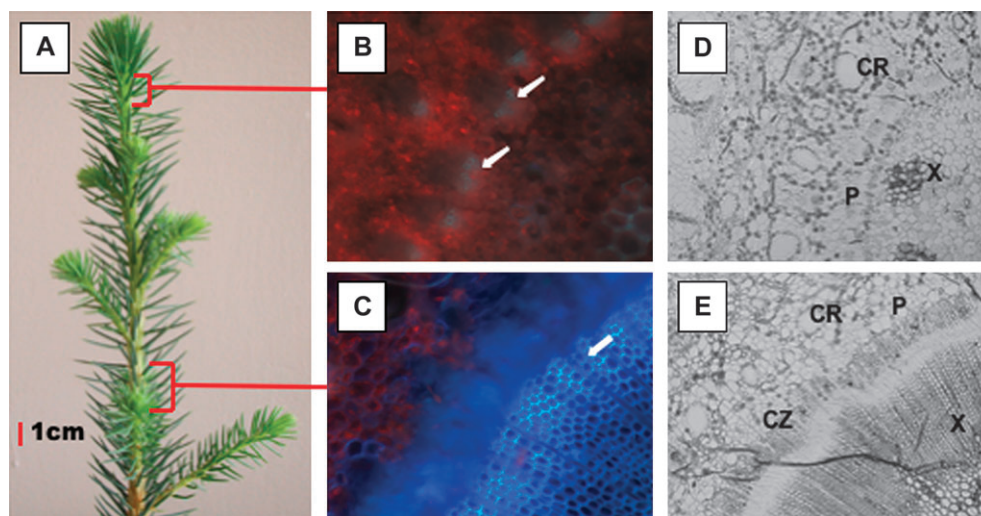


Fig. 2. Anatomy of apical shoot of 2-year-old Sitka spruce. (A) Apical shoot of Sitka spruce showing the locations where sections were taken for fluorescence and light microscopy. (B) UV autofluorescence image ($\times 100$) of a hand cross-section from the apical shoot tip. Primary vascular tissue is indicated by arrows. (C) UV autofluorescence image ($\times 100$) of a hand cross-section from the base of the apical shoot. Secondary xylem is indicated by an arrow. In (B) and (C) blue autofluorescence indicates lignin and red autofluorescence indicates chlorophyll. (D) Light microscopy ($\times 50$) of a $2\ \mu\text{m}$ cross-section of the apical shoot tip. (E) Light microscopy ($\times 50$) of a $2\ \mu\text{m}$ cross-section of the apical shoot tip. P, phloem; X, xylem, CZ, cambium zone; CR, constitutive phloem resin duct.

Table 2. Selected genes preferentially expressed in apical shoot woody stem bases, with EST clone ID, annotation, and fold-change values in three experiments

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	<i>E</i> -value	Putative function	Experiment					
				FC 2003A ^c	<i>P</i> -value ^d	FC 2004B	<i>P</i> -value	FC 2004C	<i>P</i> -value
Cell-wall related and/or carbohydrate metabolism									
WS0087_M10	At5g13870	1e ⁻⁴⁶	Xyloglucan endotransferase (XET)	29.0	<0.001	4.3	0.002	29.0	<0.001
WS0035_A01	At5g15470	1e ⁻⁶⁶	Glucosyl transferase family 8	18.0	<0.001	1.3	0.614	18.0	<0.001
WS0037_E06	At1g74910	6e ⁻⁴⁶	ADP-glucose pyrophosphorylase	3.811	0.022	1.3	0.410	3.811	0.022
WS0073_D24	At4g32410	5e ⁻⁵⁹	Cellulose synthase catalytic subunit	10.9	0.003	3.6	0.061	10.9	0.003
WS0038_F16	At5g15630	8e ⁻³²	COBL4:COBRA-LIKE4	3.6	0.020	1.4	0.398	3.6	0.020
WS0038_G03	At5g60490	4e ⁻¹⁵	Fasciclin-like arabinogalactan-protein	6.6	0.003	1.8	0.067	6.6	0.003
WS0042_H15	At4g02280	9e ⁻⁴³	Putative sucrose synthase	6.8	0.001	2.3	0.105	6.8	0.001
WS0038_B19	At1g02640	2e ⁻¹³	Beta-xylosidase	6.9	0.011	2.4	0.020	6.9	0.011
WS0104_N22	At2g47550	7e ⁻¹⁰	Putative pectinesterase	14.0	0.001	6.2	<0.001	14.0	0.001
WS0015_D11	At5g49190	1e ⁻⁹⁷	Sucrose synthase	2.5	0.223	3.1	0.005	2.5	0.223
WS0032_M17	At3g06510	2e ⁻¹²	Beta-glucosidase	8.3	0.006	1.9	0.084	8.3	0.006
WS00812_N07	At5g17420	2e ⁻⁹⁰	Cellulose synthase catalytic subunit	5.8	<0.001	2.0	<0.001	5.8	<0.001
WS0104_P14	At5g53370	4e ⁻⁶	Pectinesterase	3.9	0.002	1.7	<0.001	3.9	0.002
WS00813_I07	At2g36870	9e ⁻⁶	Xyloglucan endotransferase (XET)	5.3	0.001	1.9	0.005	5.3	0.001
WS0037_D21	At1g22370	8e ⁻²⁵	UDP-glucose glucosyltransferase	10.9	0.002	2.6	0.023	10.9	0.002
WS00913_L08	At1g19170	3e ⁻²⁹	Polygalacturonase-like protein	2.6	0.050	1.5	0.103	2.6	0.050
WS0032_I18	At3g16920	2e ⁻²⁷	Glycoside hydrolase family 19 similar to class I chitinase	2.4	0.035	1.5	0.139	2.4	0.035
WS0042_J15	At5g49720	6e ⁻⁴⁹	Endo-1,4-beta-glucanase KORRIGAN-like	1.9	0.010	0.8	0.033	1.9	0.010
Lignin biosynthesis and phenylpropanoids									
WS0038_D12	At5g01190	6e ⁻³³	Laccase	22.1	<0.001	9.2	0.001	22.1	<0.001
WS0038_H05	At5g05390	3e ⁻⁵⁵	Laccase	8.2	0.009	2.0	0.079	8.2	0.009
WS0034_K23	At5g05340	2e ⁻²⁵	Peroxidase	5.3	0.035	3.2	0.086	5.3	0.035
WS0089_H07	At1g75280	4e ⁻³³	Isoflavone reductase	13.7	<0.001	5.2	0.006	13.7	<0.001
WS0042_I14	At2g30490	3e ⁻³¹	Cinnamate-4-hydroxylase (C4H)	5.0	0.011	1.6	0.119	5.0	0.011
WS0038_I15	At5g60020	5e ⁻⁷⁴	Laccase	14.9	0.002	2.3	0.154	14.9	0.002
WS0039_D14	At5g54160	1e ⁻⁵¹	Caffeic acid <i>O</i> -methyltransferase (COMT)	3.1	0.050	1.2	0.621	3.1	0.050
WS00813_M20	At2g38080	2e ⁻⁶⁵	Laccase	10.2	0.004	1.1	0.838	10.2	0.004
WS0094_C01	At1g15950	8e ⁻⁵⁵	Cinnamoyl CoA reductase (CCR)	11.1	0.010	0.5	0.205	11.1	0.010
WS0037_D03	At1g17010	1e ⁻³¹	Oxidoreductase, flavonoid biosynthesis	4.4	0.006	1.9	0.101	4.4	0.006
WS0071_D22	At3g10340	8e ⁻⁸⁰	Phenylammonia lyase (PAL)	4.3	0.053	1.6	0.341	4.3	0.053
WS0033_E05	At3g21240	8e ⁻²⁵	4-Coumarate:CoA ligase 2 (4CL2)	6.1	0.034	1.3	0.286	6.1	0.034
WS00112_D12	At2g21940	3e ⁻⁹	Shikimate kinase precursor	1.9	0.046	1.2	0.367	1.9	0.046
WS01013_J06	At3g05950	7e ⁻¹⁷	Germin-like protein	5.5	0.018	1.6	0.008	5.5	0.018
WS0038_E11	At4g01850	1e ⁻³⁰	<i>S</i> -adenosylmethionine synthase 2 (SAM)	5.5	0.019	0.8	0.644	5.5	0.019
WS0039_P08	No Hits		<i>p</i> -Coumarate 3 hydrogenase (C3H) identical to pine AY064170	3.6	0.024	2.0	0.213	3.6	0.024
WS00913_L11	At5g48930	2e ⁻²⁰	Hydroxycinnamoyl-CoA:shikimate/quininate hydroxycinnamoyltransferase (HCT)	3.0	0.149	2.2	0.005	3.0	0.149
WS01010_M10	At1g65060	8e ⁻⁷⁶	4-Coumarate:CoA ligase 3 (4CL3)	1.6	0.329	1.3	0.148	1.6	0.329
WS00712_J18	At3g29200	9e ⁻⁵³	Chorismate mutase	3.5	0.038	1.1	0.851	3.5	0.038

Table 2. (Continued)

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	<i>E</i> -value	Putative function	Experiment						
				FC 2003A ^c	<i>P</i> -value ^d	FC 2004B	<i>P</i> -value	FC 2004C	<i>P</i> -value	
WS0022_L10	At4g34050	4e ⁻⁵⁶	Caffeoyl-CoA <i>O</i> -methyltransferase (CCoOMT)	3.7	0.009	1.3	0.15	3.7	0.009	
WS0037_G16	At4g34230	2e ⁻²⁷	Cinnamyl alcohol dehydrogenase (CAD)	3.4	<0.001	0.7	0.411	3.4	<0.001	
Pathogen and disease resistance										
WS0038_N18	At5g40020	1e ⁻¹⁵	Pathogenesis response protein thaumatin domain family	25.1	<0.001	5.9	0.006	25.1	<0.001	
WS01011_J07	At1g64160	5e ⁻²⁶	Dirigent protein, spruce PDIR19	7.8	0.009	1.3	0.348	7.8	0.009	
WS00913_G16	At1g64160	2e ⁻¹⁸	Dirigent protein, spruce PDIR8	7.0	0.001	1.6	0.004	7.0	0.001	
Transport proteins										
WS0031_F12	At4g35100	7e ⁻⁵⁵	Plasma membrane intrinsic protein (SIMIP)	20.1	<0.001	2.6	<0.001	20.1	<0.001	
WS0032_G18	At4g21910	8e ⁻¹⁸	MATE efflux family	10.2	0.002	5.3	0.002	10.2	0.002	
WS0033_A08	At3g53720	1e ⁻⁷	Putative Na ⁺ /H ⁺ antiporter	27.6	<0.001	8.4	0.001	27.6	<0.001	
WS0035_A05	At1g47670	1e ⁻³⁰	Lysine and histidine specific transporter	2.2	0.003	2.1	0.004	2.2	0.003	
WS0042_C02	At1g22540	6e ⁻³⁵	Oligopeptide transporter	3.8	<0.001	2.7	0.006	3.8	<0.001	
WS0102_G17	At3g11900	4e ⁻²³	Putative amino acid transporter protein	3.8	<0.001	2.3	0.038	3.8	<0.001	
WS0105_C01	At3g28345	2e ⁻⁵³	ABC transporter family, class III	3.3	0.019	1.5	0.051	3.3	0.019	
Lipid transport										
WS0033_C20	At2g18370	6e ⁻¹¹	Putative lipid transfer protein	8.1	<0.001	2.4	0.006	8.1	<0.001	
WS0044_K03	At2g18370	6e ⁻¹⁹	Putative lipid transfer protein	7.2	0.001	2.1	0.103	7.2	0.001	
Protein metabolism										
WS0032_D05	At5g51750	9e ⁻²²	Subtilisin-like serine protease	5.6	0.001	2.7	<0.001	5.6	0.001	
WS0031_I07	At4g21326	3e ⁻¹⁸	Subtilisin-like serine protease	2.9	0.024	2.3	0.008	2.9	0.024	
WS0021_D05	At2g25710	2e ⁻²⁴	Biotin holocarboxylase synthetase	6.4	0.001	1.1	0.668	6.4	0.001	
WS00713_O03	At5g50260	4e ⁻¹⁶	Cysteine proteinase	2.6	0.038	1.4	0.008	2.6	0.038	
WS0041_E22	At2g37190	2e ⁻¹⁹	Ribosomal protein L12	2.1	0.029	1.2	0.595	2.1	0.029	
WS0052_P07	At1g11910	3e ⁻²⁰	Putative aspartic proteinase	1.5	0.120	2.8	0.044	1.5	0.120	
Metabolism										
WS0071_K02	At5g64460	2e ⁻²⁵	Similar to phosphoglycerate/ <i>bis</i> phosphoglycerate mutase	5.6	0.010	0.8	0.440	5.6	0.010	
WS0039_G06	At1g69230	e ⁻²⁰	Spiral1-like2	3.9	0.029	1.2	0.691	3.9	0.029	
WS00812_I14	At4g13660	e ⁻³⁶	Similar to pinorexinol-lariciresinol reductase	6.8	0.002	2.0	0.003	6.8	0.002	
WS0033_O18	At4g36250	5e ⁻⁶²	Aldehyde dehydrogenase	20.9	<0.001	3.0	<0.001	20.9	<0.001	
WS01013_N08	At2g39980	3e ⁻²⁴	Putative anthocyanin 5-aromatic acyltransferase	16.3	0.001	4.9	0.002	16.3	0.001	
WS0037_D17	At2g36690	4e ⁻¹⁴	Putative gibberellin beta-hydroxylase, oxygenase	2.0	0.012	0.7	0.069	2.0	0.012	
WS0038_L24	At1g52760	3e ⁻²¹	Similar to monoglyceride lipase	3.9	0.014	1.6	0.04	3.9	0.014	
WS0054_O20	At5g19530	9e ⁻³⁰	Spermine synthase	4.8	0.005	1.9	0.164	4.8	0.005	
WS0074_F01	At4g39660	8e ⁻³²	Alanine-glyoxylate aminotransferase	4.8	<0.001	2.2	0.026	4.8	<0.001	
WS0061_C15	At5g17920	2e ⁻⁴⁴	5-Methyltetrahydropteroyl triglutamate-homocysteine <i>S</i> -methyltransferase	3.6	0.072	0.7	0.480	3.6	0.072	
WS01010_N02	At3g21420	6e ⁻²⁶	Hypothetical iron/ascorbate oxidoreductase family	8.0	0.001	5.3	0.002	8.0	0.001	
WS00113_H24	At3g57380	2e ⁻⁴⁰	Putative hypothetical protein	9.7	0.006	1.3	0.285	9.7	0.006	
WS00713_N19	At5g61840	7e ⁻¹⁴	Putative protein, cell growth, exostosin	13.7	<0.001	1.9	0.333	13.7	<0.001	
WS0082_K05	At5g24580	8e ⁻¹¹	Copper-binding family protein	7.5	0.001	1.2	0.413	7.5	0.001	

Table 2. (Continued)

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	<i>E</i> -value	Putative function	Experiment					
				FC 2003A ^c	<i>P</i> -value ^d	FC 2004B	<i>P</i> -value	FC 2004C	<i>P</i> -value
WS0039_E15	At1g72230	5e ⁻²¹	Blue copper protein	12.8	0.003	1.5	0.301	12.8	0.003
WS0107_E20	At4g21270	3e ⁻⁹	Kinesin-related protein katA	5.1	0.004	1.5	0.002	5.1	0.004
WS0034_K08	At4g20890	4e ⁻¹²	Tubulin beta-9 chain	2.8	0.009	1.9	0.085	2.8	0.009
WS0036_G14	At4g14960	4e ⁻⁸⁸	Tubulin alpha-6 chain (TUA6)	3.7	0.005	1.4	0.378	3.7	0.005
WS0041_N15	At4g27270	6e ⁻⁷⁰	1,4-Benzoquinone reductase	5.4	0.046	1.9	0.366	5.4	0.046
			Unknown function						
WS0038_P09			No match	4.7	0.030	3.5	0.002	4.7	0.030
WS0031_H24			No match	5.1	<0.001	2.8	0.021	5.1	<0.001
WS0035_E10			No match	9.7	0.008	2.3	0.065	9.7	0.008
WS0018_G13			No match	8.7	0.016	2.8	0.086	8.7	0.016
WS0031_G18	At3g54820	1e ⁻⁴	No match, weakly similar to aquaporin	8.8	0.007	3.8	0.003	8.8	0.007
WS0084_I19			No match	33.2	0.001	34.5	<0.001	33.2	0.001
WS0035_G02			No match	16.4	0.002	5.3	0.001	16.4	0.002
WS0046_O03			No match	18.3	<0.001	5.3	0.001	18.3	<0.001
WS0073_A13	At3g51590	1e ⁻³	No match, weakly similar to stigma/stylar cysteine-rich	9.4	0.001	2.4	0.001	9.4	0.001
WS0047_O02			No match	10.6	0.001	1.4	0.312	10.6	0.001

^a Array element by EST designation; EST sequences have been deposited in GenBank and are available at NCBI (<http://www.ncbi.nlm.nih.gov/>).

^b Best match to *Arabidopsis* genes was determined by TBLASTX searches of the TAIR *Arabidopsis* database (<http://www.arabidopsis.org/>) using the EST sequence.

^c Fold change in expression in base segments relative to tip segments in experiment 2003A, 2003B, or 2003C, as described in the text.

^d *P*-value for fold change in the experiment to the left.

value of 5-fold or greater, 311 (90%) displayed an FC value >2 in the replicate experiment, 2003A.

Interestingly, by contrast to the 2652 differentially expressed genes detected in the 2004B experiment that excluded bark from the apical shoot base, in experiment 2004C, where bark tissue was included (Fig. 1C), the number of genes differentially expressed between tip and base was reduced to only 610, due at least partially to the exclusion of genes encoding defence-related proteins expressed both at the shoot tip and in the bark tissue of the apical shoot base. Many trends in differential expression of genes involved in wood formation in common with experiments 2003A and 2004B were still apparent, especially for the most highly differentially expressed genes in the 2003A and 2004B datasets.

Patterns of differential gene regulation were generally comparable between experiments 2003A and 2004B and the specificity of expression patterns within gene families was also reproducible. For instance, there are 39 cDNAs from the xyloglucan endo-glycosyl transferase (*XET*) gene family represented on the array, and they group into 27 different apparent gene family members, based on >90% sequence identity. Only two cDNAs (WS00813_I07 and WS0087_M10) representing two distinct gene family members showed strong preferential expression in the leader stem base in the 2003A experiment (FC >5; see Supplementary Table S1 at *JXB* online) and this specificity in

XET gene expression was found again in the subsequent year (2004B), and in the bark-included experiment (experiment 2004C; see Supplementary Table S1 at *JXB* online). Likewise, 15 cDNAs on the array were annotated as encoding members of the arabinogalactan family of proteins (AGP), and these fall into six groups, based on at least 90% sequence identity at the nucleotide level (see Supplementary Table S2 at *JXB* online). Of these, a single group containing three cDNAs very similar in sequence showed marked preferential expression in the xylem-enriched apical shoot base in both the 2003A and 2004B experiments, with a similar trend in the 2004C experiment in which bark was included in apical shoot base samples (see Supplementary Table S2 at *JXB* online). These data show that reproducible differential expression of genes, and of individual members of multi-gene families, can be readily detected when comparing the bases of apical shoots with the tips.

Representative cDNAs corresponding to genes and gene families of particular biological interest (e.g. cell wall and lignin formation, regulation, defence) were selected from the genes found to be consistently differentially expressed in experiments 2003A and 2004B and are presented in Tables 2–4. Also included in these selections are some genes that, while differentially expressed by at least 2-fold with *P*-values of ≤0.01 in the 2003A experiment, met slightly lower but still stringent criteria for differential expression in the less replicated 2004B experiment (*P*-value for

Table 3. Transcription factor and regulatory genes preferentially expressed in apical shoot woody stem bases, with EST clone ID, annotation, and fold-change values in three experiments

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	E-value	Putative function	FC 2003A ^c	P-value ^d	FC 2004B	P-value	FC 2004C	P-value
Transcriptional regulation									
WS00712_A21	At5g16600	1e ⁻²⁰	Myb family transcription factor	10.2	<0.001	8.1	0.001	2.3	0.003
WS0083_F20	At2g44745	5e ⁻⁴²	WRKY family transcription factor	8.4	<0.001	10.6	0.001	2.7	0.019
WS0083_O11	At5g60450	1e ⁻³⁶	Auxin response factor	4.4	<0.001	4.2	0.004	1.9	0.025
WS00915_B21	At5g06710	2e ⁻⁵⁵	HD-Zip transcription factor	3.9	<0.001	6.3	<0.001	1.6	0.019
WS0092_N10	At1g52150	2e ⁻⁵²	HD-Zip transcription factor	2.9	<0.001	2.4	0.055	2.7	0.074
WS00917_H19	At5g35550	8e ⁻¹⁰	Myb family transcription factor	2.5	<0.001	4.6	0.002	1.5	0.082
WS0104_K23	At5g48150	9e ⁻³⁷	GRAS family transcription factor	2.2	<0.001	1.9	0.026	1.2	0.025
WS0045_N22	At5g18830	2e ⁻³⁰	Squamosa promoter binding protein-like	2.1	<0.001	2.9	0.001	1.5	0.25
Signal transduction									
WS0033_I09	At5g20810	3e ⁻⁰⁹	Putative saur1 protein, calmodulin binding	13.9	<0.001	4.9	0.001	1.3	0.083
WS0031_P03	At5g24430	6e ⁻¹⁹	Calcium-dependent protein kinase-like	13.9	<0.001	5.1	0.014	4.4	<0.001
WS0078_M03	At3g48030	8e ⁻¹³	H2 RING finger protein (C3HC4-type)	13.3	<0.001	8.2	<0.001	2.9	0.021
WS0039_K18	At2g33380	6e ⁻⁶⁴	Calcium-binding protein RD20	12.1	<0.001	3.6	0.009	2.2	0.002
WS0082_A16	At3g07130	3e ⁻²⁵	Serine/threonine protein phosphatase	6.5	<0.001	11.8	0.003	2.5	0.106
WS0109_F16	At3g23380	8e ⁻⁷	p21-rho-binding domain-containing protein	5.4	<0.001	7.6	0.001	2.0	0.001
WS0039_K14	At1g48480	2e ⁻⁵¹	Leucine-rich repeat transmembrane protein kinase	4.3	<0.001	4.1	0.002	0.8	0.401
WS0031_K17	At5g54380	4e ⁻³⁴	Receptor protein kinase	3.6	<0.001	4.2	0.005	2.0	0.070
WS00812_M21	At3g24240	1e ⁻⁷⁶	Leucine-rich repeat transmembrane protein kinase similar to CLV1	2.9	<0.001	2.8	0.001	1.0	0.941
WS0019_P24	At3g11410	e ⁻²¹	Protein phosphatase 2C	2.6	<0.001	2.4	0.009	1.3	0.128
WS00813_I21	At2g35620	3e ⁻⁶⁷	Leucine-rich repeat transmembrane protein kinase	1.7	<0.001	2.3	0.011	1.3	0.300

^a Array element by EST designation; EST sequences have been deposited in GenBank and are available at NCBI (<http://www.ncbi.nlm.nih.gov/>).

^b Best match to *Arabidopsis* genes was determined by TBLASTX searches of the TAIR *Arabidopsis* database (<http://www.arabidopsis.org/>) using the EST sequence.

^c Fold change in expression in base segments relative to tip segments in experiment 2003A, 2003B, or 2003C, as described in the text.

^d P-value for fold change in the experiment to the left.

differential expression generally ≤ 0.05). All other differentially expressed array elements are found in the supplementary data (see Supplementary Tables S3 and S4 at *JXB* online).

Secondary cell wall synthesis, carbohydrate metabolism, and lignin biosynthesis

Genes encoding enzymes involved in secondary cell wall synthesis, carbohydrate metabolism, and lignin biosynthesis were strongly differentially expressed in secondary xylem-enriched samples at the base of apical shoots (experiments 2003A and 2004B; Table 2). These include genes

encoding cellulose synthase subunits, sucrose synthase, XET, fasciclin-like AGP, and family 8 glucosyl transferase. Also preferentially expressed at the xylem-enriched apical shoot bases were genes which collectively represent almost every step in lignin biosynthesis: phenylalanine ammonia lyase (PAL), cinnamate-4-hydroxylase (C4H), 4-coumarate CoA ligase (4CL), *p*-coumarate 3-hydroxylase (C3H), hydroxycinnamoyl-CoA:shikimate/quinate hydroxycinnamoyltransferase (HCT), cinnamoyl CoA reductase (CCR), caffeoyl-CoA *O*-methyltransferase (CCoOMT), caffeic acid *O*-methyltransferase (COMT), and cinnamyl alcohol dehydrogenase (CAD), as well as

Table 4. Selected genes preferentially expressed in apical shoot tips with EST clone ID, annotation, and fold-change values in three experiments

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	E-value	Putative function	FC 2003A ^c	P-value ^d	FC 2004B	P-value	FC 2004C	P-value
Defence responses to pests and pathogens									
WS0048_P11	At3g57240	2e ⁻³⁸	Beta-1,3-glucanase	20.4	<0.001	3.4	0.002	4.2	0.002
WS0017_B07	At2g01630	8e ⁻³⁸	Beta-1,3-glucanase	12.1	<0.001	9.3	<0.001	2.6	0.002
WS0091_O08	At4g33720	2e ⁻²⁰	Pathogenesis-related protein 1 precursor	18.2	<0.001	17.5	<0.001	4.6	0.002
WS0083_E09	At5g42500	5e ⁻⁰⁷	Disease resistance response protein-like	17.6	<0.001	10.8	0.001	1.1	0.835
WS0064_D18	At2g02120	2e ⁻¹³	Protease inhibitor II, gamma-thionin family	15.9	<0.001	21.7	<0.001	1.4	0.511
IS0014_G20	No hit	n.a.	Putative pathogenesis-related protein [<i>Picea glauca</i>]-AAF60972.2]	12.0	<0.001	13.5	<0.001	5.0	0.006
WS0108_K13	At3g54420	9e ⁻⁵⁶	Class IV chitinase	9.1	<0.001	5.3	0.002	5.3	0.002
WS0043_C01	At1g19320	4e ⁻³⁵	Pathogenesis-related protein 5 precursor	9.0	<0.001	22.7	0.001	9.7	<0.001
WS00713_O13	At1g55020	3e ⁻⁴⁹	Lipoxygenase	8.5	<0.001	3.3	0.006	1.2	0.090
WS0024_B06	At3g22400	8e ⁻¹⁹	Putative lipoxygenase	7.8	<0.001	3.6	0.011	3.8	<0.001
WS0015_C10	At4g11650	1e ⁻³⁷	Osmotin precursor, response to pathogen	6.9	<0.001	21.7	<0.001	5.0	0.004
WS00713_G19	At3g04720	2e ⁻¹²	Hevein-like protein precursor (PR-4),	6.7	<0.001	21.7	<0.001	11.8	<0.001
WS0064_G15	At2g40880	4e ⁻²⁴	Putative cysteine proteinase inhibitor B (cystatin B)	4.1	<0.001	9.0	0.001	3.0	0.018
WS0038_B23	No hit	n.a.	Spruce dirigent protein PDIR7	4.2	<0.001	6.3	0.021	1.6	0.351
WS0036_M20	At1g58170	2e ⁻²⁸	Spruce dirigent protein PDIR1	2.0	<0.001	4.9	0.004	1.2	0.346
Response to stress									
WS0076_C16	At5g14920	2e ⁻²⁶	Putative protein, response to gibberellic acid stimulus	17.0	<0.001	6.3	0.001	2.2	0.173
WS00111_I10	At5g54370	1e ⁻⁰⁹	Late embryogenesis abundant protein	10.1	<0.001	9.1	<0.001	2.3	0.004
IS0013_A20	At4g02380	e ⁻⁰⁵	Late embryogenesis abundant protein	4.8	<0.001	14.3	<0.001	1.8	0.109
WS0019_P18	At4g15910	1e ⁻⁰⁸	Drought-induced protein like	7.7	<0.001	38.5	<0.001	1.8	0.021
WS0017_L07	At1g74670	2e ⁻²³	Gibberellin-responsive protein	7.7	<0.001	4.6	0.004	3.4	0.003
WS0047_P11	At5g19880	1e ⁻⁰⁸	Peroxidase, response to ethylene stimulus	7.6	<0.001	12.5	<0.001	4.1	0.002
WS00111_H06	At3g26060	7e ⁻⁶⁸	Similar to peroxiredoxin Q, antioxidant activity	6.9	<0.001	34.5	<0.001	1.0	0.857
WS00915_G16	At3g11930	8e ⁻²⁶	Universal stress protein (USP), similar to ER6	6.3	<0.001	18.9	<0.001	1.8	0.043
WS0023_A18	At3g18080	e ⁻¹⁰²	Beta-glucosidase	8.9	<0.001	10.9	<0.001	1.6	0.005
Terpenoid metabolism									
WS0086_N12	At3g14520	6e ⁻¹¹	Abietadiene synthase [<i>Abies grandis</i>] AAK83563.1	18.7	<0.001	8.2	0.001	2.4	0.008
WS00713_H14	No hits	n.a.	Isopimaradiene synthase [<i>Picea abies</i>] AAS47690.1	18.5	<0.001	5.5	0.003	1.7	0.078
WS00911_G14	At2g18620	2e ⁻⁴⁶	Geranylgeranyl pyrophosphate synthase	14.3	<0.001	7.6	0.004	2.4	0.075
WS0063_I21	At1g61680	2e ⁻¹⁵	(-)-Pinene synthase [<i>Picea sitchensis</i>] AAP72020.1	11.2	<0.001	8.8	<0.001	1.3	0.601
WS0019_A03	At4g16740	2e ⁻⁶³	(+)-3-Carene synthase [<i>Picea abies</i>] AAO73863.1	9.8	<0.001	6.3	0.003	0.9	0.738
WS00111_A06	At4g36810	2e ⁻⁵⁷	Geranylgeranyl pyrophosphate synthase	8.3	<0.001	6.2	0.001	1.7	0.41

Table 4. (Continued)

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	E-value	Putative function	FC 2003A ^c	P-value ^d	FC 2004B	P-value	FC 2004C	P-value
WS0101_I06	At5g62790	3e ⁻⁶³	1-Deoxy-D-xylulose 5-phosphate reductoisomerase (DXR)	6.7	<0.001	3.8	0.002	2.1	0.304
WS0035_I02	At1g63970	8e ⁻⁵²	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MECPS)	6.5	<0.001	2.9	0.018	2.4	0.084
WS0092_I21	At4g02780	6e ⁻³⁵	(-)-linalool synthase [<i>Picea abies</i>]. AAS47693.1	5.2	0.001	7.8	0.003	1.2	0.718
WS0072_I23	At2g17370	2e ⁻³⁶	3-Hydroxy-3-methylglutaryl-coenzyme A reductase 2 (HMG2)	3.4	<0.001	3.0	0.001	1.1	0.794
Other secondary metabolism									
WS0041_L20	At1g24735	4e ⁻¹⁰	Caffeoyl-CoA O-methyltransferase like CCOMTL6	23.8	<0.001	8.9	0.001	3.7	0.011
IS0013_L11	At5g13930	3e ⁻⁶¹	Chalcone synthase	16.7	<0.001	13.0	0.002	1.4	0.464
WS0022_P21	At1g64930	2e ⁻²¹	Cytochrome P450 monooxygenase	15.4	<0.001	15.4	0.001	7.1	0.001
WS0018_L17	At2g02400	5e ⁻⁴³	Cinnamoyl-CoA reductase like CCRL8	14.7	<0.001	10.0	0.001	1.7	0.438
WS0063_L14	At4g22870	2e ⁻¹⁸	Anthocyanidin synthase-like	13.2	<0.001	7.4	0.008	1.4	0.517
WS00914_G22	At4g22880	1e ⁻¹³	Putative leucoanthocyanidin dioxygenase (LDOX)	12.6	<0.001	5.3	0.009	1.9	0.03
WS0047_L15	At2g37040	5e ⁻²²	Phenylalanine ammonia lyase (PAL)	9.9	<0.001	8.5	0.014	1.7	0.495
WS00916_M24	At5g08640	1e ⁻²³	Flavonol synthase (FLS)	8.5	<0.001	25.6	0.001	2.2	0.019
WS0092_M15	At4g10490	4e ⁻⁶³	Putative flavanone 3-beta-hydroxylase naringenin 3-dioxygenase	8.5	<0.001	5.8	0.001	1.4	0.37
WS0071_H13	At5g54160	3e ⁻³⁷	Quercetin 3-O-methyltransferase 1 (OMT1)	7.9	0.001	27.8	<0.001	7.0	<0.001
WS0018_F19	At5g07990	1e ⁻⁰⁷	Flavonoid 3-hydroxylase	7.7	<0.001	12.0	0.004	0.5	0.062
WS0016_M16 *	At5g07990	6e ⁻⁶²	Flavonoid 3'-monooxygenase, transparent testa 7 protein (TT7)	7.4	<0.001	17.9	0.003	0.8	0.373
WS0021_N17	At5g42800	7e ⁻³⁶	Dihydroflavonol 4-reductase	6.6	<0.001	8.2	0.001	1.2	0.450
WS00111_I09	At1g77330	3e ⁻²²	1-Aminocyclopropane-1-carboxylate oxidase	5.9	<0.001	3.0	0.002	1.2	0.671
Lipid transfer proteins									
WS0019_O12	At5g59310	8e ⁻²⁴	Putative nonspecific lipid-transfer protein, LTP4	25.6	<0.001	17.2	<0.001	2.2	0.086
WS0022_I21	At5g59320	8e ⁻¹⁶	Putative lipid transfer protein, LTP3	28.4	<0.001	17.5	<0.001	1.0	0.995
WS0018_K17	At5g59320	6e ⁻¹⁸	Putative lipid transfer protein, LTP4	26.3	<0.001	25.0	<0.001	1.5	0.295
WS0022_C17	At2g38540	5e ⁻¹⁰	Non-specific lipid transfer protein LTP1 – loblolly pine	25.2	<0.001	9.0	0.011	1.4	0.499
WS0023_D03	At5g59310	1e ⁻¹⁵	Non-specific lipid-transfer protein precursor	24.4	<0.001	17.9	<0.001	1.7	0.101
WS0016_L19	At5g59310	5e ⁻¹⁰	Putative non-specific lipid-transfer protein, LTP4	16.9	<0.001	34.5	<0.001	2.5	0.068

Table 4. (Continued)

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	<i>E</i> -value	Putative function	FC 2003A ^c	<i>P</i> -value ^d	FC 2004B	<i>P</i> -value	FC 2004C	<i>P</i> -value
WS00111_N24	At5g59310	6e ⁻²⁴	Putative non-specific lipid-transfer protein, LTP4	14.9	<0.001	38.5	<0.001	2.9	0.013
WS0022_P23	At5g57800	8e ⁻⁴³	Lipid transfer protein; glossy1 homologue, CER1 protein	9.6	<0.001	8.0	<0.001	1.9	<0.001
Transcriptional regulation									
WS0072_F05	At3g61460	1e ⁻¹¹	RING finger protein, response to brassinosteroid stimulus	7.9	<0.001	28.6	<0.001	15.9	0.001
WS0094_A03	At5g43840	2e ⁻⁰⁸	Heat shock transcription factor-like protein	7.0	<0.001	27.8	<0.001	1.0	0.982
WS0011_C05	At4g00730	7e ⁻¹⁶	HD-GLABRA2 family homeodomain protein	5.0	<0.001	15.6	<0.001	1.4	0.051
WS0091_B15	At1g63650	7e ⁻¹⁷	ANL2 bHLH transcription factor EGL3	4.9	<0.001	11.0	0.002	1.6	0.082
WS0082_I17	At1g29860	1e ⁻²⁷	WRKY family transcription factor	4.6	<0.001	2.1	0.070	12.7	<0.001
WS0017_K19	At1g32640	4e ⁻⁶²	bHLH MYC-like protein MYC2	4.3	<0.001	2.9	0.005	0.8	0.545
WS00919_O17	At2g21900	3e ⁻⁰⁶	WRKY family transcription factor	4.3	<0.001	1.5	0.120	0.9	0.424
WS0023_D19	At3g07650	1e ⁻²⁵	CONSTANS-like zinc finger protein COL9 (B-box type)	4.2	<0.001	5.4	0.003	1.0	0.897
Signalling									
WS0076_C16	At5g14920	2e ⁻²⁶	Gibberellin-regulated family protein, similar to cell wall-plasma membrane linker protein	17.0	<0.001	6.3	0.001	2.2	0.173
IS0014_A12	At2g31880	4e ⁻⁴⁸	Putative receptor-like protein kinase	12.5	<0.001	20.4	<0.001	3.6	0.002
WS0011_K11	At3g48090	3e ⁻⁰⁷	Disease resistance protein EDS1	5.1	<0.001	3.5	0.016	1.0	0.962
IS0014_E17	At5g49760	1e ⁻¹⁴	Receptor protein kinase-like protein	4.7	<0.001	17.9	0.001	1.0	0.984
WS00913_F03	At1g69990	3e ⁻²⁷	Receptor-like protein kinase	4.3	<0.001	5.9	0.001	3.9	0.005
WS00715_J19	At5g48740	1e ⁻²⁶	Receptor protein kinase-like protein	4.0	<0.001	2.1	0.005	1.2	0.647
IS0012_O08	At1g73500	3e ⁻⁶³	MAPKK9	3.2	<0.001	7.0	0.001	1.5	0.086
WS0023_A18	At3g18080	1e ⁻¹⁰²	Glycosyl hydrolase family 1	8.9	<0.001	10.9	<0.001	1.6	0.005
WS0021_J24	At1g15380	4e ⁻²⁸	Lactoglutathione lyase family, glyoxalase I family.	7.2	<0.001	2.6	0.028	2.0	0.063
WS0052_E19	At5g65730	6e ⁻⁰⁶	Xyloglucan endotransferase	6.8	<0.001	2.7	0.070	0.5	0.001
WS00112_O10	At1g26770	2e ⁻⁵⁵	Expansin	6.8	<0.001	3.6	<0.001	2.6	0.018
WS0064_M21	At2g43840	3e ⁻³⁷	Putative glucosyltransferase	6.0	<0.001	13.9	<0.001	0.9	0.634
WS0072_C14	At4g37800	1e ⁻³²	Xyloglucan endotransferase	5.9	<0.001	14.9	0.001	2.9	0.002
WS0018_C19	At1g09420	1e ⁻³²	Putative glucose-6-phosphate dehydrogenase	5.2	<0.001	10.3	<0.001	1.4	0.506
Carbon fixation and photosynthesis									
WS00112_J05	At1g44575	1e ⁻⁴³	Photosystem II 22 kDa protein	10.0	<0.001	11.9	0.003	1.3	0.303
WS0012_I03	At1g20340	5e ⁻¹⁵	Plastocyanin	8.3	<0.001	5.7	0.011	2.2	0.344

Table 4. (Continued)

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	E-value	Putative function	FC 2003A ^c	P-value ^d	FC 2004B	P-value	FC 2004C	P-value
WS0012_E01	At3g14930	3e ⁻⁴⁷	Uroporphyrinogen decarboxylase	7.6	<0.001	9.6	0.010	0.8	0.703
WS00712_D11	At4g14690	1e ⁻⁴⁷	Expressed protein, chlorophyll binding	6.3	<0.001	6.8	0.002	1.2	0.470
WS0023_C07	At2g20260	4e ⁻³²	Putative photosystem I reaction centre subunit IV	6.0	<0.001	3.6	0.020	1.2	0.560
WS0017_F05	At3g22840	9e ⁻²⁷	Chlorophyll A-B binding protein, early light-induced (ELIP)	5.1	<0.001	7.4	0.002	1.2	0.573
WS0024_E14	At5g38430	2e ⁻⁴⁰	Ribulose biphosphate carboxylase small chain 1b precursor	5.1	<0.001	7.9	0.001	0.8	0.506
			Metabolism						
WS0014_N06	At4g28780	7e ⁻²⁸	GDSL-motif lipase/hydrolase	15.5	<0.001	20.8	<0.001	10.9	0.002
WS00910_N15	At2g04570	2e ⁻⁴⁶	Putative GDSL-motif lipase/hydrolase similar to APG proteins	14.1	<0.001	37.0	<0.001	0.9	0.676
WS00110_C02	At5g33370	4e ⁻³⁴	GDSL-motif lipase/hydrolase	13.7	<0.001	83.3	<0.001	4.6	<0.001
WS00916_G05	At1g73340	2e ⁻³⁰	Cytochrome P450 family	11.9	<0.001	6.5	0.008	4.4	0.017
WS0056_J05	At3g52500	3e ⁻²⁹	Aspartyl protease, cell wall	11.2	<0.001	5.2	0.003	1.0	0.935
WS0041_I14	At2g36580	5e ⁻⁵⁹	Putative pyruvate kinase	8.4	<0.001	6.2	<0.001	1.5	0.292
WS0012_E01	At3g14930		Uroporphyrinogen decarboxylase	7.6	<0.001	9.6	0.010	0.8	0.703
WS0035_I02	At1g63970	8e ⁻⁵²	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	6.5	<0.001	2.9	0.018	2.4	0.084
IS0013_E03	At1g08080	6e ⁻⁴⁰	Storage protein, carbonic anhydrase	6.5	<0.001	3.2	0.035	4.8	0.009
WS0024_L04	At5g59480	4e ⁻²⁶	Haloacid dehalogenase-like hydrolase	6.1	<0.001	5.0	0.001	1.9	0.125
WS00811_B15	At5g62420	1e ⁻⁴⁰	Aldose reductase-like protein	5.6	<0.001	4.5	0.001	1.5	0.006
WS0045_G20	At5g23870	7e ⁻¹⁸	Pectinacetyl esterase	5.2	<0.001	3.5	0.003	6.6	0.001
			Miscellaneous						
WS0045_F23	At5g42830	2e ⁻⁰⁹	Similar to anthranilate N-hydroxycinnamoyl/benzoyltransferase	11.1	<0.001	10.4	<0.001	1.5	0.197
WS0012_J09	At5g05340	2e ⁻⁴⁴	Peroxidase	10.9	<0.001	8.2	0.001	2.0	0.112
WS0053_F21	At4g31840	3e ⁻²⁴	Plastocyanin-like domain-containing protein	10.8	<0.001	2.9	0.001	0.5	0.175
WS0047_L11	At2g25060	2e ⁻¹²	Plastocyanin-like domain-containing protein	10.0	<0.001	2.6	0.015	0.9	0.875
WS00910_N01	At5g06720	2e ⁻¹⁶	Peroxidase	9.8	<0.001	5.1	0.005	1.1	0.533
WS0075_O24	At3g26060	9e ⁻²⁷	Putative peroxiredoxin similar to peroxiredoxin Q	7.9	<0.001	19.6	<0.001	1.0	0.846
IS0014_M04	At1g10360	7e ⁻¹⁷	Putative glutathione S-transferase TSI-1	6.1	<0.001	9.6	<0.001	1.5	0.18
WS0044_N18	At3g27200	3e ⁻⁰⁹	Plastocyanin-like domain-containing protein	5.5	<0.001	6.2	0.002	1.9	0.015
WS0021_L21	At5g48810	1e ⁻⁰⁹	Cytochrome b ₅	5.3	<0.001	17.2	<0.001	1.1	0.734
			Unknown function						
WS0018_P03			No match	31.4	<0.001	47.6	<0.001	1.9	0.049
WS00713_C21			No match	30.4	<0.001	34.5	<0.001	1.3	0.501
WS0015_G05			No match	26.6	<0.001	8.3	0.004	2.3	0.171
WS0022_P17			No match	21.5	<0.001	47.6	0.001	2.7	0.041
WS0013_B02			No match	21.1	<0.001	9.9	0.002	1.0	0.991

Table 4. (Continued)

Clone ID ^a	Best <i>Arabidopsis</i> BLASTX match ^b	E-value	Putative function	FC 2003A ^c	P-value ^d	FC 2004B	P-value	FC 2004C	P-value
IS0011_E04			No match	19.9	<0.001	17.9	<0.001	2.2	0.053
WS0093_N18			No match	19.6	<0.001	8.0	0.022	1.2	0.174
WS0024_I20			No match	18.7	<0.001	13.7	0.001	3.0	0.023
WS0108_K23			No match	18.1	<0.001	9.4	0.009	3.1	0.027
WS0044_C24			No match	18.0	<0.001	8.9	0.003	1.0	0.988
WS0023_D23			No match	16.9	<0.001	7.4	0.001	1.4	0.118
WS0018_J23			No match	15.7	<0.001	41.7	<0.001	3.1	<0.001

^a Array element by EST designation; EST sequences have been deposited in GenBank and are available at NCBI (<http://www.ncbi.nlm.nih.gov/>).

^b Best match to *Arabidopsis* genes was determined by TBLASTX searches of the TAIR *Arabidopsis* database (<http://www.arabidopsis.org/>) using the EST sequence.

^c Fold change in expression in tip segments relative to base segments in experiment 2003A, 2003B, or 2003C, as described in the text.

^d P-value for fold change in the experiment to the left.

members of the shikimate pathway that provides precursors for lignin and other phenolic pathways. Other genes whose functions could be related to cellular processes involved in secondary xylem differentiation, secondary wall formation, and lignin deposition, such as proteases, laccases, and aquaporins were also more highly expressed in secondary xylem-enriched tissue at the stem base (experiments 2003A and 2004B; Table 2). This group also includes a number of apparent spruce orthologues of *Arabidopsis* genes that have been shown to be co-expressed during secondary cell wall formation in *Arabidopsis*, such as the fasciclin-like AGP gene mentioned above, a germin-like protein, a chitinase-like protein (family 19), and a COBRA-like 4 protein. Additional differentially expressed genes representative of other functional categories are also shown in Table 2.

Regulatory proteins and transcription factors at the shoot base

Transcripts with consistent up-regulation in the xylem-enriched woody stem base (experiments 2003A and 2004B) included genes encoding potential signalling proteins and transcription factors (Table 3). Of 136 cDNAs for transcription factors on the array (annotated by similarity to *Arabidopsis*; <http://arabidopsis.med.ohio-state.edu/AtTFDB>; see Supplementary Tables S3 and S4 at *JXB* online), eight were differentially expressed (Table 3). A *Myb* gene family member (WS00712_A21) with marked similarity to *AtMyb43* showed >10-fold higher expression in the xylem-enriched apical shoot base, which is the highest differential expression ratio observed for any transcription factor on the array. Another *Myb* gene family member (WS00917_H19) highly up-regulated at the shoot base is most closely related to *AtMyb123* (*TT2*), a regulator of proanthocyanidin accumulation in *Arabidopsis* (Nesi et al., 2001). Other transcription factors preferentially expressed in the xylem-enriched apical shoot base showed similarity to two different members of the *Arabidopsis* HD-Zip family (WS0092-N10/ATHB-15;

WS00915_B21/HAT14), as well as to members of the WRKY, GRAS, and squamosa promoter binding protein-like families (Table 3). Transcripts hybridizing with array element WS0083_O11, whose sequence is similar to the *Arabidopsis* auxin response factor (ARF) genes *ARF4* and *ARF5*, were also preferentially expressed at the xylem-enriched base.

There are at least 81 cDNAs in the array with annotations related to protein kinases, and several of these were differentially expressed in the xylem-enriched apical shoot base (see Supplementary Tables S3 and S4 at *JXB* online). For both the shoot tip and woody base, elements annotated as receptor kinases were among the most prominent differentially expressed kinase-related genes, and six of these showed consistent, preferential expression at the xylem-enriched shoot base (Table 3). These include putative leucine-rich repeat (LRR) receptor-like kinases (RLKs) (WS00812_M21, WS0039_K14, WS00813_I21), which are similar to *Arabidopsis* RLKs of unknown functions (Table 3). WS00812_M21 is most closely related to the CLV1 family of LRR-RLKs, and shares a stretch of *c.* 90 highly conserved amino acids with the *Arabidopsis* CLV1-like gene, At3g24240. Preferential expression of other genes with putative kinase and phosphatase functions in the xylem-enriched tissue at the shoot base suggests an active role for protein phosphorylation associated with secondary xylem development (Table 3). Intriguingly, one of the genes (WS0031_P03) most highly specific to the xylem-enriched apical shoot base, resembles the *Arabidopsis* Ca-dependent protein kinase CDPK4 (At5g24430) and a Ca-CaM-dependent protein kinase from maize (MCK2). Another highly apical shoot base-specific gene (WS0039_K18) is similar to the *Arabidopsis* Ca²⁺-binding protein, RD20 (At2g33380).

Effect of xylem enrichment on analysis of apical shoot base differential gene expression

In experiments 2003A and 2004B (Fig. 1), bark tissue was removed from the base of the apical shoot in order to

enrich the sample for xylem tissue. Consequently, it was possible to detect transcripts that were preferentially present in xylem compared with green apical shoot tips, but whose signals would otherwise have been diluted by the presence of transcripts associated with green bark tissue. A number of highly differentially expressed genes in the xylem-enriched tissue at the base of the apical shoot, such as cellulose synthase, XET, aquaporins, and laccases also retained this specificity even when the bark and phloem were included in the leader base samples in experiment 2004C (Tables 2, 3), although the FC ratios were correspondingly lower. Other genes with lower FC values for base versus tip in experiments 2004A and 2004B had non-significant FC values when bark was retained in experiment 2004C. Overall, most of the differentially regulated genes with highest expression at the apical shoot base (Tables 2, 3) showed an expression pattern that was strikingly enhanced in the xylem-enriched base samples. These include genes represented by cDNAs without significant sequence similarity to *Arabidopsis* genes (Table 2), suggesting that they may play conifer-specific roles in secondary xylem differentiation and wood formation.

Genes preferentially expressed in apical shoot tips

A large number of differentially expressed transcripts in experiments 2003A and 2004B showed preferential expression in the green apical shoot tip relative to the xylem-enriched base. Prevalent within this subset were genes generally associated with defence responses against pests or pathogens, responses to stress, and known conifer defence genes of terpenoid (oleoresin) and phenolic secondary metabolism. Representative examples of defence genes predominantly expressed in the green shoot tip relative to xylem-enriched shoot base are shown in Table 4. When gene expression was examined in tips versus intact bases including green bark tissues (experiment 2004C; Fig. 2) about 25% of these defence- and stress-related genes retained significant differential expression in tips compared with the full apical shoot base (experiment 2004C; Table 4), supporting a model according to which the apical shoot tip commits significantly enhanced gene expression to the early formation of a constitutive defence barrier. Spruce genes annotated as PR-proteins showed particularly high ratios (Table 4; see Supplementary Table S4 at *JXB* online). In addition, of the 34 chitinase elements represented on the array, most showed tip-specific transcript hybridizations (see Supplementary Table S4 at *JXB* online) similar to element WS0108_K13 (Table 4). Similarly, most of the highly tip-specific β -1,3 glucanase transcripts detected on the array also showed patterns of preferential expression in apical shoot tip relative to the intact apical shoot base in experiment 2004C (Table 4). A group of protease inhibitor (PI) transcripts were also preferentially detected in the apical shoot tip relative to the xylem-enriched shoot base

(e.g. element WS0064_D18; Table 4; see Supplementary Table S4 at *JXB* online). Many of these (15 cDNA array elements) share similarity to the protein encoded by the *Arabidopsis PI II* gene, At2g02120, a protease inhibitor II containing a gamma-thionin domain. However, unlike the β -1,3 glucanase and chitinase gene family members, these spruce PI genes appeared to be expressed at similar levels in the tips and the intact bases of leader stems in experiment 2004C (Table 4), suggesting that they may contribute to constitutive defence both in the apical shoot tip and the young apical shoot bark.

Genes for terpenoid and phenolic secondary metabolism and constitutive defence

Several genes of the methylerythritol phosphate (MEP) pathway leading to the formation of the isopentenyl diphosphate and dimethylallyl diphosphate precursors for mono- and diterpenoid oleoresin defence, as well as other terpenoids, were preferentially expressed in the green apical shoot tips when compared with the xylem-enriched apical shoot base. Genes corresponding to cDNAs for 1-deoxy-D-xylulose-5-phosphate (DXR, At5g62790) and 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MECPS, At1g63970) showed the highest shoot tip-preferred expression in this group of enzymes (Table 4). Similarly, spruce terpene synthase genes, specifically involved in spruce oleoresin defence and represented on the array by several previously characterized cDNAs (Martin *et al.*, 2004), showed higher expression in apical shoot tips relative to the xylem-enriched base (5–20-fold), corresponding with the phloem and shoot tip location of developing constitutive resin ducts. When the apical shoot tip was compared with the intact apical shoot base, including bark and phloem tissues (experiment 2004C; Fig. 2), genes encoding enzymes in the terpenoid pathway were similarly expressed in both sections. These results support association of constitutive terpenoid biosynthesis with constitutive resin ducts that are initiated very early in the development of the green apical shoot tip and remain active in mature phloem.

In phenolic secondary metabolism, several genes encoding proteins related to *Arabidopsis* flavonoid biosynthetic enzymes were also strongly differentially expressed in apical shoot tips relative to xylem-enriched bases, as were other phenylpropanoid pathway-related genes (Table 4). Most of these genes remained abundantly expressed in bark tissues at the apical shoot base (see data on experiment 2004C; Table 4). At least 28 cDNAs on the array have been annotated as dirigent or dirigent-like proteins (Ralph *et al.*, 2006a), and almost all of these elements showed preferential expression in the apical shoot tip relative to the xylem-enriched leader base in experiments 2003A and 2004B (see Table 4 for examples; see Supplementary Table S4 at *JXB* online). However, some were preferentially expressed in xylem-enriched tissue at

the leader base (experiments 2003A and 2004B; Table 2), in particular those with similarity to At1g64160 (*DIR5*).

Genes encoding putative regulatory proteins with preferential shoot tip expression

Several genes with potential regulatory functions were differentially expressed in the apical shoot tip relative to the xylem-enriched base (experiments 2003A and 2004B; Table 4). For example, elements WS0011_C05 and WS0091_B15 encode homeodomain and bHLH transcription factors related to the *Arabidopsis* genes *ANL* and *EGL3*, respectively, which are regulators of anthocyanin biosynthesis in *Arabidopsis* (Kubo *et al.*, 1999; Broun, 2005). At least four WRKY transcription factor family members showed differential expression in tip samples. Of these, element WS0082_I17 retained a strong tip preferred expression even in experiment 2004C (Table 4). Several genes potentially involved in protein kinase signalling also showed differential expression in the apical shoot tip relative to the xylem-enriched base and, among these, two RLKs (IS0014_A12, WS00913_F03; Table 4) appeared strongly tip-preferred.

Lipid transfer proteins

Among the most prevalent transcripts showing particularly high FC ratios between the apical shoot tip and the xylem-enriched shoot base were those hybridizing with cDNA array elements annotated as lipid transfer proteins (LTPs). There are at least 35 LTP-like cDNAs on the spruce array that show similarity to this gene family in *Arabidopsis*, and these fall into 10 groups. Within a group, each *LTP* shares over 90% EST nucleotide sequence similarity with other members. A large fraction of the *LTP* array elements showed very high differential expression in tips relative to the xylem-enriched tissue at the shoot base (10–30× FC, experiments 2003A and 2004B; Table 4), and most of these share the greatest similarity with *Arabidopsis LTP4* or *LTP3* genes. However, they did not retain a significant differential expression in tips relative to the intact leader base including the green bark tissue (Table 4; experiment 2004C). On the other hand, a small number of *LTP* genes were highly differentially expressed in the xylem-enriched leader base (e.g. cDNAs WS0033_C20 and WS0044_K03; Table 2). Sequence alignment and phylogenetic reconstruction of 12 spruce full-length *LTP* sequences revealed that *LTP* genes showing preferential expression in xylem-enriched leader base samples grouped into a distinct phylogenetic clade (Fig. 4).

Validation of candidate gene expression by RT-PCR

Candidate genes for expression validation by quantitative RT-PCR were chosen from those genes encoding putative signalling proteins and transcription factors with apparently preferential expression either at the apical shoot tip

or base. The *XET* gene whose expression was specific to xylem-enriched tissue at the apical shoot base was also included in this analysis. The corresponding cDNAs were fully sequenced, which allowed the design of sequence-specific PCR primer pairs. RT-PCR assays performed on the same 2004 RNA samples validated microarray results for all the selected cDNAs, not only in specificity of expression, but also in the magnitude (Table 1). Expression of a subset of this group of genes was also monitored along the length of the spruce leader by RT-PCR. For this analysis, four segments were taken from the tip to the base of the leader, without removing the bark (Fig. 1C). Expression relative to the tip changed gradually along the leader for some genes (Fig. 3), while for others there was a sharp transition in expression between the tip and the segment immediately below it, where secondary growth was already evident.

Discussion

Apical shoots of Sitka spruce provide a useful system to study gene expression associated with early events of stem development in conifers. In the present study, it was possible to identify genes that displayed consistent patterns of differential expression over the course of apical shoot development, including those involved in secondary xylem formation and the developmental transition to the woody apical shoot base, and in constitutive defence in the shoot tip and young bark tissues. Also, differentially expressed regulatory proteins were identified that provide new candidates for control with secondary xylem development and constitutive defence structures and secondary metabolite production.

Secondary cell wall development

The spruce apical shoot begins to form secondary xylem immediately below the apical shoot tip, characterized by extensive development of lignified secondary cell walls (Fig. 2). Prominent among the cell secondary cell wall-related genes preferentially expressed in the xylem tissue present at the apical shoot base (Table 2) were members of the cellulose synthase (*CesA*) superfamily (Doblin *et al.*, 2002). In *Arabidopsis*, distinct *CesA* subunits are associated with secondary and primary cell wall formation (Doblin *et al.*, 2002; Zhong *et al.*, 2003). The two spruce *CesA* genes (WS0073_D24, WS00812_N07) that stood out in their differential expression in the xylem-enriched apical shoot base are very similar to the loblolly pine (*Pinus taeda*) *CesA3* gene, proposed to be the pine orthologue of the *Arabidopsis* secondary cell wall-associated *CesA7* gene (Nairn and Haselkorn, 2005). This is consistent with the suggestion that distinct *CesA* isoforms with functions in primary and secondary wall biosynthesis evolved both in conifers and angiosperms (Nairn and

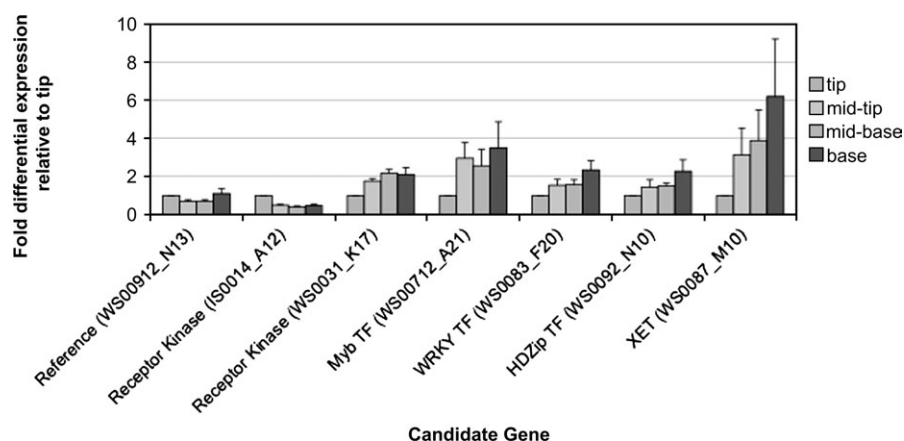


Fig. 3. Expression analysis of candidate genes along the leader stem analysed by quantitative RT-PCR. Expression was analysed progressing from the apical shoot tip to the base and was assessed in four segments, as shown in Fig. 1C, without removing the bark tissue. RT-PCR was performed with the same RNA utilized for microarray gene expression profiling in experiment 2004C. Data for the RT-PCR were analysed as the expression ratio relative to the tip, and is shown as the mean of three replicated assays. Standard deviations of the means are shown as error bars. Included are the ratios observed for a control reference cDNA element (WS00912_N13; hypothetical protein) which did not show differential expression in the microarray experiments. Candidate genes chosen for RT-PCR validation were a receptor kinase (IS0014_A12) with high expression at the tip relative to the base by microarray analysis, and a receptor kinase (WS0031_K17), Myb transcription factor (WS00712_A21), WRKY transcription factor (WS0083_F20), HDZip transcription factor (WS0092_N10), and XET (WS0087_M10) with high expression at the base relative to the tip by microarray analysis.

Haselkorn, 2005). As well, recent studies have identified, *in silico*, *Arabidopsis* genes whose expression is highly correlated with *CesA* gene expression during secondary wall formation (Brown *et al.*, 2005; Ehling *et al.*, 2005; Persson *et al.*, 2005). It is interesting that a set of genes represented on the spruce cDNA microarray with similarity to these *Arabidopsis* genes is also up-regulated in the xylem-enriched tissue, together with the *CesA* genes. These spruce genes are annotated as encoding COBRA-like 4 (COBL4), germin-like, chitinase-like 2 (CTL2), fasciclin-like arabinogalactan (FLA12), and KORRIGAN (KOR) proteins. This correspondence could suggest that the spruce proteins play similar roles in cellulose deposition during secondary wall formation, either as part of the cellulose synthase complex, or in subsequent modification of cellulose microfibrils, and that this process is highly conserved in angiosperms and gymnosperms. Two genes with homology to sucrose synthase that appear to be up-regulated in the apical shoot base may correspond to spruce orthologues of a sucrose synthase that delivers the UDP-Glc substrate to the *CesA* complex (Haigler *et al.*, 2001).

The present results also showed differential expression among the members of the arabinogalactan-proteins (AGP) gene family (see Supplementary Table S2 at *JXB* online). The spruce AGP gene (WS0038_G03) preferentially expressed in the xylem-enriched apical shoot base is nearly identical to an AGP abundant in loblolly pine xylem (p14A9) (Loopstra and Sederoff, 1995; Yang *et al.*, 2005), and the similarity of this gene to *Arabidopsis* FLA12 reinforces the suggestion that FLA12-like AGPs play an important role in cellulose deposition during secondary xylogenesis. However, two other Sitka spruce AGP genes, with similarities to *Arabidopsis* FLA1 and

FLA9, were more preferentially expressed in the leader tip (Tables 2, 4; see Supplementary Table S2 at *JXB* online). Similarly, of the several xyloglucan endoglycosyl transferase (*XET*) genes represented on the microarray, one (WS0087_M10) showed marked preferential expression at the apical shoot base (Tables 1, 2; see Supplementary Table S1 at *JXB* online; Fig. 3). The sequence of the corresponding cDNA is very similar to that of a loblolly pine cDNA isolated from a salt-treated root cDNA library (TC69813, pine gene index, TIGR: www.tigr.org/tigrscripts/tgi/T_index.cgi?species=pinus), and is similar (65% amino acid sequence similarity) to a cotton *XET* gene highly expressed in elongating fibres (Ji *et al.*, 2003). This result is consistent with previous studies showing differential expression of *XET* genes in cambial and expanding xylem tissues (Vander Mijnsbrugge *et al.*, 2000; Hertzberg *et al.*, 2001; Mellerowicz *et al.*, 2001; Paux *et al.*, 2004; Schrader *et al.*, 2004; SH Yang *et al.*, 2004), suggesting that the woody stem-specific spruce *XET* is involved in cell expansion during secondary xylem formation.

Genes encoding enzymes involved in the synthesis of the monolignol units from phenylalanine have been well characterized in *Arabidopsis* (Boerjan *et al.*, 2003; Raes *et al.*, 2003; Ehling *et al.*, 2005), and many of them have also been described in pine (Anterola *et al.*, 2002) and spruce (Ralph *et al.*, 2006b). A set of spruce genes corresponding to cDNA microarray elements annotated as encoding lignin biosynthetic enzymes showed strongly up-regulated expression in the woody base of Sitka spruce apical shoots (Table 2; see Supplementary Table S3 at *JXB* online). Most of these enzymes were represented on the spruce microarray by multiple cDNAs representing putative gene family members with distinct sequences, yet

for many of the putative family members only a subset were up-regulated in the woody stem base. This set, with representatives shown in Table 2, includes all enzymes known to be required for the biosynthesis of G-lignin, (PAL, C4H, 4CL, HCT, C3H, CCoOMT, CCR, CAD) and thus represents those spruce phenylpropanoid genes most likely to encode the spruce enzymes involved in developmental lignification. Interestingly, the *Arabidopsis* genes to which members of this group of spruce genes showed highest similarity were, in most cases, those known or inferred by expression analysis to be involved in developmental lignification in that plant as well (e.g. *4CL2*, *CCOMT1*, *CAD1*, *CCR1*; Raes et al., 2003; Ehltling et al., 2005).

The role of conifer COMT, also known as AEOMT (Li et al., 1997), in monolignol biosynthesis is controversial (Anterola et al., 2002), since, according to current understanding in angiosperms (Boerjan et al., 2003), it is involved in S-lignin biosynthesis, which does not occur in conifers. However, of the 13 spruce cDNAs with similarity to COMT on the microarray, six displayed preferential expression in the woody apical shoot base (see Supplementary Table S3 at *JXB* online), and among these was WS0039_D14, the cDNA that shows the greatest similarity to the dual-function COMT (AEOMT) (Table 2) previously described in loblolly pine (Li et al., 1997). This result supports an association of AEOMT with active wood formation, while leaving unresolved the potential biosynthetic role of the bifunctional *O*-methylation potentially catalysed by this enzyme.

Two classes of oxidative enzymes believed to be involved in oxidative polymerization of monolignols are laccases and peroxidases (Boerjan et al., 2003). Laccases predominated among the genes showing a high expression in the woody portion of the Sitka spruce apical shoot (Table 2). Three of these showed a high degree of similarity to *Arabidopsis* laccases LAC4 (At2g38080), LAC12 (At5g05390), and LAC17 (At5g60020) that were found to be strongly co-expressed with cellulose synthase genes involved in secondary cell wall formation (Brown et al., 2005; Ehltling et al., 2005; Persson et al., 2005). Similarly, a spruce peroxidase with strong woody apical shoot base-preferred expression (WS0034_K23) is similar to *Arabidopsis* peroxidase P49 (At5g05340), which is also strongly up-regulated in concert with lignification in *Arabidopsis* (Ehltling et al., 2005). These results point to a conserved repertoire of oxidative enzymes involved in lignin polymerization in both conifers and angiosperms.

Several classes of proteases also show preferential expression in the woody base of Sitka spruce apical shoots, including subtilisin-like, Ser-, Cys-, and Asp-proteases (Table 2). These classes of proteases were shown to be up-regulated during programmed cell death during *trans*-differentiation of mesophyll cells into tracheary elements in *Zinnia* (Demura et al., 2002) and have been character-

ized recently in plant cell death during wood formation in poplar (Moreau et al., 2005), suggesting that they participate in this final step of spruce xylem differentiation.

Regulatory proteins in xylem development

A number of transcription factors showed consistent up-regulation at the apical shoot woody base, making them candidates for positive regulators of secondary xylem differentiation (Table 3). Given their association with wood formation and lignification in both conifers and angiosperms (Patzlaff et al., 2003; Newman et al., 2004), the two spruce *Myb* genes represented by spruce cDNAs WS00712_A21 and WS00917_H19, are of particular interest. Although WS00712_A21 is distinct from loblolly pine *PtMYB4*, reported to be a regulator of lignification (Patzlaff et al., 2003), it shows similarity to loblolly pine *PtMYB1* (81% amino acid identity at the N-terminal R2R3 DNA binding domain), a putative regulator of phenylpropanoid metabolism (Patzlaff et al., 2003). Interestingly, the spruce gene sequence is moderately similar to those of the related *Arabidopsis* *AtMyb43* and *AtMyb20* genes (81% and 55% amino acid identity, respectively, in the R2R3 DNA-binding domains) whose transcription has been correlated with secondary wall formation and lignification in developing *Arabidopsis* interfascicular fibres (Ehltling et al., 2005). This is consistent with phylogenetic analysis, which places *PtMYB1* in the same clade as *AtMyb43* and *AtMyb20* (Patzlaff et al., 2003). The second *Myb* gene WS00917_H19 shares similarity (60% amino acid identity in the R2R3 DNA-binding domain) to *Arabidopsis* *Myb* transcription factors *AtMyb123* (*TT2*) (Nesi et al., 2001) and *AtMyb75* (*PAP1*) (Borevitz et al., 2000), regulators of phenylpropanoid metabolism. The similarity of these two spruce genes to clades of *Arabidopsis* *Myb* genes associated with phenylpropanoid metabolism and secondary wall formation suggest potential functional conservation of *Myb* proteins between conifers and angiosperms.

HD-ZipIII transcription factor family members are candidates for regulators of xylogenesis in angiosperms (Baima et al., 2001; Hertzberg et al., 2001; Schrader et al., 2004; Prigge et al., 2005), and the two different spruce *HD-ZipIII* genes up-regulated in xylem-enriched tissue at the base (WS0092_N10 and WS00915_B21) may play analogous roles in spruce xylem differentiation. Other transcription factors showing strong woody stem-preferred expression include a member of the WRKY family (WS0083_F20) that has a high degree of amino acid sequence similarity (up to 81%) to the WRKY42 protein in *Arabidopsis* (At2g44745) that is involved in control of cell expansion and proanthocyanidin biosynthesis in developing seeds (Johnson et al., 2002; Garcia et al., 2005), and a spruce *auxin response factor* (*ARF*) transcription factor gene (WS0083_O11) similar to *Arabidopsis* *ARF4* and *ARF5/MONOPTEROS*. These results

suggest potential conservation of a transcription factor function regulating cell expansion and auxin-regulated secondary xylem development in conifers and angiosperms.

The presence of multiple cDNA microarray elements related to protein signalling functions, such as LRR RLKs and protein kinases, allowed potential roles for these classes of proteins to be assessed during tissue differentiation processes in the spruce leader. Four spruce LRR *RLK* genes showed a pronounced up-regulation in the woody apical shoot base (Table 3) and one of these, WS00812_M21 shows sequence similarity to the *Arabidopsis CLVI* gene, raising the possibility of conserved signalling pathways underlying key pathways of RLK-mediated cellular differentiation in conifers and angiosperms. In the present experiment, a putative Ca-dependent protein kinase gene, WS0031_P03, with similarity to *Arabidopsis CPK4* (At5g24430) was strongly up-regulated in woody stem tissue of the apical shoot. Another gene with preferential expression in woody stem tissues of the apical shoot, WS0039_K18, encodes a protein similar to the *Arabidopsis* Ca²⁺-binding protein, RD20 (At2g33380), which has been shown to be induced by ABA upon desiccation (Takahashi *et al.*, 2000). These observations suggest that Ca²⁺-signalling may also play a key role in xylem differentiation in Sitka spruce.

Terpenoid, phenolic, and other defence-related genes

The present experiment revealed genes that are preferentially expressed in non-woody tissues of the Sitka spruce apical shoot, specifically the apical shoot tip and green bark tissues at the base where anatomical structures for chemical and physical defences develop. Elevated expression of genes involved in formation of terpenoid and phenolic secondary metabolism, along with other defence-related genes such as chitinases, glucanases, and protease inhibitors in the growing apical shoot tip supports a model of an early allocation of substantial resources towards constitutive defence and protection. For example, in terpenoid defence metabolism, many of the terpene synthase gene family members on the microarray showed high expression ratios in apical shoot tips (5–20-fold) relative to the woody leader base. Similarly, almost all of the spruce *DIR* genes represented on the spruce microarray were preferentially expressed in apical shoot tip and green tissues (Table 4), consistent with previous results of preferential expression of many spruce *DIR* genes in outer stem tissues and a proposed function in conifer defence (Ralph *et al.*, 2006a). By contrast, only two *DIR* genes (WS01011_J07, WS00913_G16) were found strongly differentially expressed in the woody apical shoot base, where they could be involved in lignin deposition. Many of the defence-related genes with high expression in the apical shoot tip continue to be expressed in the developing green bark tissues of the apical shoot base, suggesting that

these defences remain active for a long time and are not confined to the non-woody shoot tip.

The apical shoot of Sitka spruce is the preferred site for feeding and oviposition by adult white pine weevil and the developing larvae of this insect, which is the single most destructive pest in regenerating Sitka spruce forests (King, 1997). In addition, a substantial body of research has shown that the same bark and phloem tissues of Sitka spruce apical shoots also respond strongly with a plethora of inducible defences in response to attack by adult white pine weevil (Miller *et al.*, 2005; Ralph *et al.*, 2006a, b). For example, in addition to the prominent up-regulation of terpenoid defences in Sitka spruce bark tissues (Miller *et al.*, 2005) and the weevil-induced up-regulation of *DIR*-proteins (Ralph *et al.*, 2006a), a number of the chitinases and glucanases identified here in the apical shoot tip and green bark (i.e. WS0017_B07; Table 4) were also induced in stem bark tissues by mechanical wounding and insect feeding (Ralph *et al.*, 2006b). Furthermore, a number of additional genes with diverse functions (chalcone synthase, glutathione *S*-transferase, AP2 transcription factor, β -glucosidase, and cytochrome *b*₅; Table 4) with apical shoot tip-preferred expression are also induced by wounding, adult weevil feeding, or by budworm caterpillar feeding (Ralph *et al.*, 2006b).

Differential expression of LTPs: possible roles in constitutive chemical defence and secondary cell wall formation

LTPs are small proteins with highly conserved Cys residues (Kader, 1996; Arondel *et al.*, 2000) that are thought to transfer lipids between membranes but which appear to have other diverse roles in plant development, signalling, and response to pathogens (Ma *et al.*, 1995; Blein *et al.*, 2002; Horvath *et al.*, 2002; Eklund and Edqvist, 2003; Gomes *et al.*, 2003; Cheng *et al.*, 2004). Most *LTP* genes represented on the spruce microarray showed very high relative expression at the apical shoot tip (15–30× higher than in xylem-enriched tissue at the apical shoot base). In fact, they comprised a significant portion of the genes with the highest tip-preferred expression (Table 4; see Supplementary Table S4 at *JXB* online). Most of these spruce LTPs show greatest similarity to *Arabidopsis LTP1*, *LTP3*, or *LTP4* genes (Arondel *et al.*, 2000). To date, mechanisms of transport and secretion of the highly lipophilic terpenoid oleoresin compounds are not known. While functions of conifer LTPs have not been characterized, some of these spruce proteins could be involved in the secretion and accumulation of terpenoid and phenolic secondary metabolites in resin ducts and PP cells, respectively, at the shoot tip. In support of a role in defence, one of the *LTP* genes, WS0019_O12, showed strong induction in spruce leader tips upon feeding by western budworm caterpillars (Ralph *et al.*, 2006b).

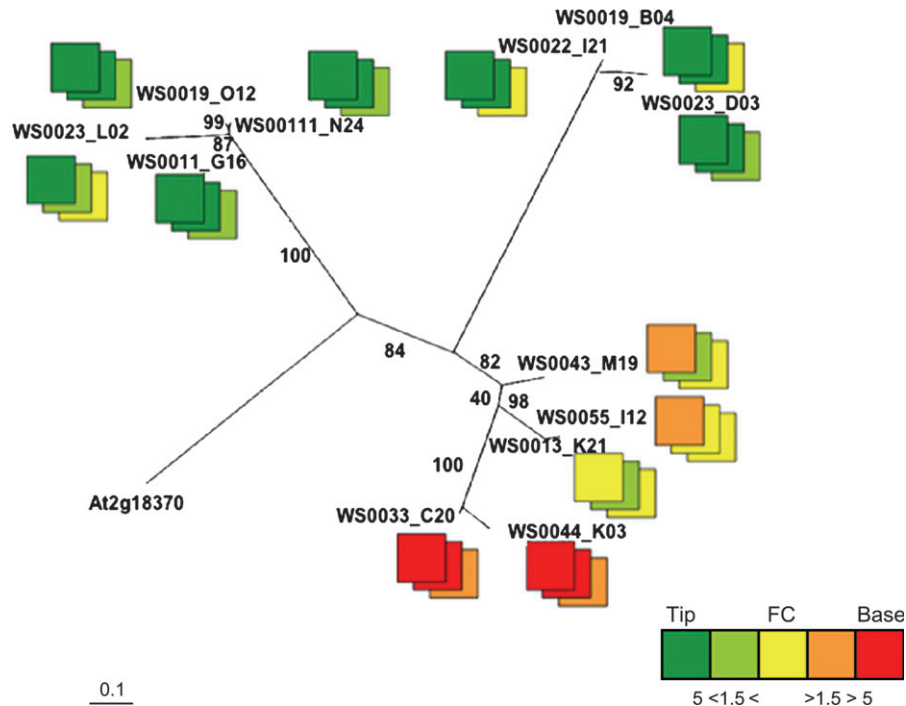


Fig. 4. Phylogeny of selected differentially expressed lipid transfer proteins (LTP). The PHYML maximum-likelihood method was applied to a multiple sequence alignment created using DIALIGN of predicted LTP amino acid sequences, as described in the Materials and methods. The unrooted tree was created in Treeview (<http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>). Shown are bootstrap values from 100 generated replications. Coloured squares indicate relative expression values (FC) for the given LTP gene in the tip or base of the apical shoot, with the top square representing data from experiment 2003A, the middle square experiment 2004B, and the bottom square experiment 2004C. Green squares indicate experiments in which the gene showed greater than a 1.5-fold higher expression in the tip relative to the base, while red squares indicate experiments in which the gene showed greater than 1.5-fold higher expression in the base relative to the tip. The degree of FC is represented by colour shading. Yellow squares indicate experiments in which the gene showed no significant difference in expression in tip versus base (FC < 1.5).

Other spruce LTP genes were preferentially expressed in the woody apical shoot base, and all of these have sequences distinct from those with highest expression in shoot tips. Alignment and phylogenetic analysis of LTP amino acid sequences represented by full-length cDNA sequences (Ralph *et al.*, 2006b) shows that LTPs associated primarily with the woody shoot base are located in a separate phylogenetic clade (Fig. 4). These LTPs might play roles in secondary cell wall formation, perhaps as part of the cellulose synthase complex, for which there is evidence in cotton (Ma *et al.*, 1995; Doblin *et al.*, 2002).

Conclusions

The apical shoot of Sitka spruce trees is important for initiating stem growth and secondary xylem formation and is the primary target for feeding and ovipositing white pine weevil. During the growing season, resources in the apical shoot must be allocated for growth and development, constitutive defence, and induced defences when needed, requiring co-ordinated regulation of these processes. Using large-scale transcript profiling, differentially expressed genes associated with secondary xylem differ-

entiation and constitutive defence in this developing organ have been identified. This information now serves as a basis for work on elucidating functions of specific genes or combinations of genes in spruce and other conifers. The associated sequence information available for a large spruce EST and full-length cDNA resource (Ralph *et al.*, 2006b), together with the indications of conservation in gene function and expression patterns between conifers and angiosperms, also provide fertile ground for comparative genomic studies that address the evolution of important traits such as vascular development and metabolic specialization of secondary metabolism in the plant kingdom.

Supplementary data

Supplementary data can be found at *JXB* online.

Table S1. Comparison of XET expression between years.

Table S2. Comparison of AGP expression between years.

Table S3. Base versus tip fold-change and *P*-values for all array elements on experiments 2003A, 2004B, and 2004C.

Table S4. Tip versus base fold-change and *P*-values for all array elements in experiments 2003A, 2004B, and 2004C.

Table S5. Quantitative RT-PCR Primer Sequences.

Acknowledgements

We gratefully acknowledge financial support from Genome Canada, Genome British Columbia, and the Province of British Columbia through funding provided for the Treenomix Conifer Forest Health Project, Genome Canada Competition III (grant to JB and KR) and for the Treenomix Project, Genome Canada Competition I (grant to JB, KR, BEE, and CJD); as well as financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC; Discovery Grants to JB and CJD). We thank Claire Oddy, Hesther Yueh, and Sharon Jancsik for technical assistance with array development, David Kaplan for greenhouse support, members of the UBC Bio-imaging facility for assistance with microscopy, and Dr Jürgen Ehling for discussions.

References

- Anterola AM, Jeon JH, Davin LB, Lewis NG. 2002. Transcriptional control of monolignol biosynthesis in *Pinus taeda*: factors affecting monolignol ratios and carbon allocation in phenylpropanoid metabolism. *Journal of Biological Chemistry* **277**, 18272–18280.
- Arondel V, Vergnolle C, Cantrel C, Kader JC. 2000. Lipid transfer proteins are encoded by a small multigene family in *Arabidopsis thaliana*. *Plant Science* **157**, 1–12.
- Baima S, Possenti M, Matteucci A, Wisman E, Altamura MM, Ruberti I, Morelli G. 2001. The Arabidopsis ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiology* **126**, 643–655.
- Blein JP, Coutos-Thevenot P, Marion D, Ponchet M. 2002. From elicitors to lipid-transfer proteins: a new insight in cell signalling involved in plant defence mechanisms. *Trends in Plant Science* **7**, 293–296.
- Boerjan W, Ralph J, Baucher M. 2003. Lignin biosynthesis. *Annual Review of Plant Biology* **54**, 519–546.
- Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C. 2000. Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *The Plant Cell* **12**, 2383–2394.
- Broun P. 2005. Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in Arabidopsis. *Current Opinion in Plant Biology* **8**, 272–279.
- Brown D, Zeef L, Ellis J, Goodacre R, Turner S. 2005. Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *The Plant Cell* **17**, 2281–2295.
- Burns RM, Honkala BH. 1990. *Silvics of North America. 1. Conifers*. Washington, DC: US Department of Agriculture, Forest Service.
- Chang S, Puryear J, Cairney J. 1993. Simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* **11**, 113–116.
- Cheng CS, Samuel D, Liu YJ, Shyu JC, Lai SM, Lin KF, Lyu PC. 2004. Binding mechanism of nonspecific lipid transfer proteins and their role in plant defense. *Biochemistry* **43**, 13628–13636.
- Demura T, Tashiro G, Horiguchi G, et al. 2002. Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells. *Proceedings of the National Academy of Sciences, USA* **99**, 15794–15799.
- Djerbi S, Aspeborg H, Nilsson P, Sundberg BR, Mellerowicz E, Blomqvist K, Teeri TT. 2004. Identification and expression analysis of genes encoding putative cellulose synthases (CesA) in the hybrid aspen, *Populus tremula* (L.) × *P. tremuloides* (Michx.). *Cellulose* **11**, 301–312.
- Doblin MS, Kurek I, Jacob-Wilk D, Delmer DP. 2002. Cellulose biosynthesis in plants: from genes to rosettes. *Plant and Cell Physiology* **43**, 1407–1420.
- Ehling J, Mattheus N, Aeschliman DS, et al. 2005. Global transcript profiling of primary stems from *Arabidopsis thaliana* identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. *The Plant Journal* **42**, 618–640.
- Eklund DM, Edqvist J. 2003. Localization of nonspecific lipid transfer proteins correlate with programmed cell death responses during endosperm degradation in *Euphorbia lagascae* seedlings. *Plant Physiology* **132**, 1249–1259.
- Franceschi VR, Krokene P, Christiansen E, Krekling T. 2005. Anatomical and chemical defenses of conifer bark against bark beetles and other pests. *New Phytologist* **167**, 353–375.
- Garcia D, Fitz Gerald JN, Berger F. 2005. Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in Arabidopsis. *The Plant Cell* **17**, 52–60.
- Gardiner JC, Taylor NG, Turner SR. 2003. Control of cellulose synthase complex localization in developing xylem. *The Plant Cell* **15**, 1740–1748.
- Gomes E, Sagot E, Gaillard C, Laquitaine L, Poinssot B, Sanejouand YH, Delrot S, Coutos-Thevenot P. 2003. Nonspecific lipid-transfer protein genes expression in grape (*Vitis* sp.) cells in response to fungal elicitor treatments. *Molecular Plant-Microbe Interactions* **16**, 456–464.
- Haigler CH, Ivanova-Datcheva M, Hogan PS, Salnikov VV, Hwang S, Martin K, Delmer DP. 2001. Carbon partitioning to cellulose synthesis. *Plant Molecular Biology* **47**, 29–51.
- Hertzberg M, Aspeborg H, Schrader J, et al. 2001. A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences, USA* **98**, 14732–14737.
- Horvath BM, Bachem CWB, Trindade LM, Oortwijn MEP, Visser RGF. 2002. Expression analysis of a family of nsLTP genes tissue specifically expressed throughout the plant and during potato tuber life cycle. *Plant Physiology* **129**, 1494–1506.
- Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. 2002. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18**, S96–S104.
- Ji SH, Lu YC, Feng JX, Wei G, Li J, Shi YH, Fu Q, Liu D, Lou JC, Zhu YX. 2003. Isolation and analyses of genes preferentially expressed during early cotton fiber development by subtractive PCR and cDNA array. *Nucleic Acids Research* **31**, 2534–2543.
- Johnson CS, Kolevski B, Smyth DR. 2002. TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. *The Plant Cell* **14**, 1359–1375.
- Joshi CP. 2003. Xylem-specific and tension stress-responsive expression of cellulose synthase genes from aspen trees. *Applied Biochemistry and Biotechnology* **105**, 17–25.
- Kader JC. 1996. Lipid-transfer proteins in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 627–654.
- Keeling CI, Bohlmann J. 2006. Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *New Phytologist* **170**, 657–675.
- King JN, Yanchuk A, Kiss G, Alfaro R. 1997. Genetic and phenotypic relationships between weevil resistance and height growth in spruce populations of British Columbia. *Canadian Journal of Forest Research* **27**, 732–739.

- Kubo H, Peeters AJM, Aarts MGM, Pereira A, Koornneef AM.** 1999. ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in *Arabidopsis*. *The Plant Cell* **11**, 1217–1226.
- Li LG, Popko JL, Zhang XH, Osakabe K, Tsai CJ, Joshi CP, Chiang VL.** 1997. A novel multifunctional *O*-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine. *Proceedings of the National Academy of Sciences, USA* **94**, 5461–5466.
- Loopstra C, Sederoff R.** 1995. Xylem-specific gene expression in loblolly pine. *Plant Molecular Biology* **27**, 277–291.
- Ma DP, Tan H, Si Y, Creech RG, Jenkins JN.** 1995. Differential expression of a lipid transfer protein gene in cotton fiber. *Biochimica et Biophysica Acta – Lipids and Lipid Metabolism* **1257**, 81–84.
- Martin DM, Faldt J, Bohlmann J.** 2004. Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiology* **135**, 1908–1927.
- Mellerowicz EJ, Baucher M, Sundberg B, Boerjan W.** 2001. Unravelling cell wall formation in the woody dicot stem. *Plant Molecular Biology* **47**, 239–274.
- Miller B, Madilao LL, Ralph S, Bohlmann J.** 2005. Insect-induced conifer defense. White pine weevil and methyl jasmonate induce traumatic resinosis, *de novo* formed volatile emissions, and accumulation of terpenoid synthase and putative octadecanoid pathway transcripts in Sitka spruce. *Plant Physiology* **137**, 369–382.
- Moreau C, Aksenov N, Lorenzo MG, Segerman B, Funk C, Nilsson P, Jansson S, Tuominen H.** 2005. A genomic approach to investigate developmental cell death in woody tissues of *Populus* trees. *Genome Biology* **6**, R34.
- Nairn CJ, Haselkorn T.** 2005. Three loblolly pine Cesa genes expressed in developing xylem are orthologous to secondary cell wall Cesa genes of angiosperms. *New Phytologist* **166**, 907–915.
- Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L.** 2001. The *Arabidopsis* TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *The Plant Cell* **13**, 2099–2114.
- Newman LJ, Perazza DE, Juda L, Campbell MM.** 2004. Involvement of the R2R3-MYB, AtMYB61, in the ectopic lignification and dark-photomorphogenic components of the det3 mutant phenotype. *The Plant Journal* **37**, 239–250.
- Oh S, Park S, Han KH.** 2003. Transcriptional regulation of secondary growth in *Arabidopsis thaliana*. *Journal of Experimental Botany* **54**, 2709–2722.
- Patzlaff A, McInnis S, Courtenay A, et al.** 2003. Characterisation of a pine MYB that regulates lignification. *The Plant Journal* **36**, 743–754.
- Paux E, Tamsloukht M, Ladouce N, Sivadon P, Grima-Pettenati J.** 2004. Identification of genes preferentially expressed during wood formation of *Eucalyptus*. *Plant Molecular Biology* **55**, 263–280.
- Persson S, Wei H, Milne J, Page G, Somerville C.** 2005. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proceedings of the National Academy of Sciences, USA* **102**, 8633–8638.
- Petersen E, Peterson N, Weetman G, Martin P.** 1997. *Ecology and management of Sitka spruce. Emphasis on its natural range.* Vancouver: UBC Press.
- Pfaffl M.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29**, 2003–2007.
- Plomion C, Leprovost G, Stokes A.** 2001. Wood formation in trees. *Plant Physiology* **127**, 1513–1523.
- Prigge MJ, Otsuga D, Alonso JM, Ecker JR, Drews GN, Clark SE.** 2005. Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development. *The Plant Cell* **17**, 61–76.
- Raemakers C, Ruijter JM, Depreza RHL, Moorman AFM.** 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* **339**, 62–66.
- Raes J, Rohde A, Christensen JH, Van de Peer Y, Boerjan W.** 2003. Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiology* **133**, 1051–1071.
- Ralph S, Park JY, Bohlmann J, Mansfield SD.** 2006a. Dirigent proteins in conifer defense: gene discovery, phylogeny, and differential wound- and insect-induced expression of a family of DIR and DIR-like genes in spruce (*Picea* spp.). *Plant Molecular Biology* **60**, 21–40.
- Ralph S, Yueh H, Friedmann M, et al.** 2006b. Conifer defense against insects: microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobi*) reveals large scale changes of the host transcriptome. *Plant, Cell and Environment* **29**, 1545–1570.
- Savidge RA.** 2001. Intrinsic regulation of cambial growth. *Journal of Plant Growth Regulation* **20**, 52–77.
- Schrader J, Nilsson J, Mellerowicz E, Berglund A, Nilsson P, Hertzberg M, Sandberg G.** 2004. A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. *The Plant Cell* **16**, 2278–2292.
- Takahashi S, Katagiri T, Yamaguchi-Shinozaki K, Shinozaki K.** 2000. An *Arabidopsis* gene encoding a Ca²⁺-binding protein is induced by abscisic acid during rehydration. *Plant and Cell Physiology* **41**, 898–903.
- Vander Mijnsbrugge K, Meyermans H, Van Montagu M, Bauw G, Boerjan W.** 2000. Wood formation in poplar: identification, characterization, and seasonal variation of xylem proteins. *Planta* **210**, 589–598.
- Yang JM, Kamdem DP, Keathley DE, Han KH.** 2004. Seasonal changes in gene expression at the sapwood–heartwood transition zone of black locust (*Robinia pseudoacacia*) revealed by cDNA microarray analysis. *Tree Physiology* **24**, 461–474.
- Yang JM, Park S, Kamdem DP, Keathley DE, Retzel E, Paule C, Kapur V, Han KH.** 2003. Novel gene expression profiles define the metabolic and physiological processes characteristic of wood and its extractive formation in a hardwood tree species, *Robinia pseudoacacia*. *Plant Molecular Biology* **52**, 935–956.
- Yang SH, van Zyl L, No EG, Loopstra CA.** 2004. Microarray analysis of genes preferentially expressed in differentiating xylem of loblolly pine (*Pinus taeda*). *Plant Science* **166**, 1185–1195.
- Yang S, Wang H, Sathyan P, Stasollae C, Loopstra C.** 2005. Real-time RT-PCR analysis of loblolly pine (*Pinus taeda*) arabinogalactan protein and arabinogalactan-protein-like genes. *Physiologia Plantarum* **124**, 91–106.
- Zhong RQ, Morrison WH, Freshour GD, Hahn MG, Ye ZH.** 2003. Expression of a mutant form of cellulose synthase AtCesA7 causes dominant negative effect on cellulose biosynthesis. *Plant Physiology* **132**, 786–795.