Plant Mol Biol (2006) 61:917–932
DOI 10.1007/s11103-006-0059-y

The *Populus* homeobox gene *ARBORKNOX1* reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium

Andrew T. Groover · Shawn D. Mansfield · Stephen P. DiFazio · Gayle Dupper · Joseph R. Fontana · Ryan Millar · Yvonne Wang

Received: 17 October 2005 / Accepted: 29 March 2006

Abstract Secondary growth is supported by a dividing population of meristematic cells within the vascular cambium whose daughter cells are recruited to differentiate within secondary phloem and xylem tissues. We cloned a *Populus* Class 1 *KNOX* homeobox gene, *ARBORKNOX1* (*ARK1*), which is orthologous to Arabidopsis *SHOOT MERISTEMLESS* (*STM*). *ARK1* is expressed in the shoot apical meristem (SAM) and the vascular cambium, and is down-regulated in the terminally differentiated cells of leaves and secondary vascular tissues that are derived from these meristems. Transformation of *Populus* with either *ARK1* or *STM* over-expression constructs results in similar morphological phenotypes characterized by inhibition of the differentiation of leaves, internode elongation, and secondary vascular cell types in stems. Microarray analysis showed that 41% of genes up-regulated in the stems of *ARK1* over-expressing plants encode proteins involved in extracellular matrix synthesis or modification, including proteins involved in cell identity and signaling, cell adhesion, or cell differentiation. These gene expression differences are reflected in alterations of cell wall biochemistry and lignin composition in *ARK1* over-expressing plants. Our results suggest that *ARK1* has a complex mode of action that may include regulating cell fates through modification of the extracellular matrix. Our findings support the hypothesis that the SAM and vascular cambium are regulated by overlapping genetic programs.

Keywords Wood formation · Forestry · Stem cells · Extracellular matrix · Cell wall · Lignin

This work was supported by the USDA Forest Service and USDA NRI Grant 2003-00664 to AG, and a grants from the U.S. Department of Energy, Office of Science, Biological and Environmental Research Carbon Sequestration Program to AG and SD.

Electronic Supplementary Material Supplementary material is available for this article at http://www.dx.doi.org/10.1007/s11103-006-0059-y

A. T. Groover (✉) · G. Dupper · J. R. Fontana · R. Millar · Y. Wang
Institute of Forest Genetics, Pacific Southwest Research Station, USDA Forest Service, 1100 West Chiles Rd., Davis, CA 95616, USA
e-mail: agroover@fs.fed.us

S. D. Mansfield
Department of Wood Science, University of British Columbia, Vancouver, BC, Canada

S. P. DiFazio
Oakridge National Laboratory, Oak Ridge, TN, USA

Introduction

The vascular cambium is a meristematic cell population that supports secondary growth in plant stems (Larson 1994). The cambium is derived from procambial cells present during primary growth, which are in turn derived from the shoot apical meristem (SAM). Cambial daughter cells are recruited to either secondary phloem (bark) or secondary xylem (wood), where they differentiate into functional cell types. Over time, the process of secondary growth results in the radial thickening of stems and the production of larger body architectures, most dramatically illustrated by forest trees.

Fundamentally, for both the cambium and the SAM, reservoirs of meristematic cells are maintained to support indeterminate growth. Meristematic cell identity and maintenance in the SAM requires the expression of well-conserved Class I *knotted1*-like Homeobox (*KNOX*) genes. The function of *KNOX* genes in regulating cell fates was first recognized through characterization of dominant gain-of-function *Knotted1* (*Kn1*) mutations in maize (Vollbrecht et al. 1991). *Kn1* phenotypes result from the mis-expression of the normally meristem-specific *kn1* in the leaves, which results in groups of leaf cells that divide rather than adopt terminally differentiated cell fates. Similarly, the Arabidopsis Class I *KNOX* homeobox gene *SHOOT MERISTEMLESS* (*STM*) is expressed in the central and peripheral zones of the SAM, and is down-regulated in cells destined for terminal cell fates in organ primordia (Long et al. 1996). Loss-of-function Arabidopsis *stm* mutants form a SAM but fail to maintain it during embryogenesis (Long and Barton 1998), resulting in seedlings with cotyledons but no true leaves or shoot (Long et al. 1996). The over-expression of *STM* in Arabidopsis generally promotes meristematic cell identity and inhibits cell differentiation. The effects are not absolute, however; while leaf development is limited, leaves do form and display normal phyllotaxy and ectopic meristems do not form (Williams 1998; Gallois et al. 2002; Lenhard et al. 2002). Weaker over-expression phenotypes display leaf lobing, and foreshortening of the petiole (Lenhard et al. 2002). The homeobox gene *WUSCHEL* (*WUS*) is required to specify meristematic cell fate in the central zone of the SAM (Laux et al. 1996; Mayer et al. 1998). When both *STM* and *WUS* are mis-expressed, ectopic meristems can organize and are maintained (Laux et al. 1996; Mayer et al. 1998), demonstrating that STM function is modified through direct or indirect interaction with other proteins with partially overlapping expression patterns. One interpretation of these results is that *WUS* acts primarily to specify meristematic cell fate in the central zone, while *STM* functions to maintain the meristem by inhibiting differentiation throughout the central and peripheral zones of the SAM. STM likely regulates transcription of target genes as a heterodimer with BELL-like partners (e.g. Byrne et al. 2003). Although the genes regulated by STM are currently undefined, STM may directly repress the expression of the GA-biosynthesis gene *AtGA20ox1* (Hay et al. 2002) as has been shown biochemically for related KNOX genes in other species (Sakamoto et al., 2001; Chen et al. 2004).

There is clear evidence for overlap in the genetic mechanisms regulating the SAM and the vascular cambium (Groover 2005). Recent microarray analyses of gene expression during secondary growth found that key genes regulating the SAM are also expressed in the cambium region, or have paralogs that are expressed there (Ko and Han 2004; Schrader et al. 2004). For example, the Class III Homeodomain-leucine zipper and KANADI transcription factors responsible for adaxial–abaxial patterning in the SAM are also expressed in the cambium region of *Populus* (Schrader et al. 2004). *WUS*, on the other hand, appears to be uniquely expressed in the SAM, but other *WUS*-like genes are expressed in the cambium region (Schrader et al. 2004). Characterization of the function of genes co-expressed in the cambium and SAM during secondary growth would provide new insights into the regulation of the cambium and the evolution of secondary growth.

Advances in molecular genetics and genomics provide new approaches for the study of secondary growth and cambium function in higher plants, including forest trees. The genus *Populus* includes species of economic and ecological importance, and has been developed as a model for molecular genetics in angiosperm trees. Recombinant DNA-based strategies can create dominant phenotypes assayed in primary transformants, obviating the need for lengthy rounds of sexual reproduction that typically limit developmental genetic approaches in forest trees. *In planta*, gene discovery has been achieved through insertion of activation tagging (Busov et al. 2003) or gene trap constructs (Groover et al. 2004) that create dominant phenotypes. Importantly, the *Populus* genome has recently been sequenced (http://www.genome.jgi-psf.org) and should greatly aid the study of secondary growth.

STM has been well-characterized as a major regulator of the SAM through detailed molecular genetic studies (Gallois et al. 2002; Lenhard et al. 2002; Long et al. 1996). We cloned a *Populus* ortholog of *STM*, and show that the gene plays similar roles in regulating both the SAM and the vascular cambium. We also present evidence that the action of *STM* may be more complex than suggested by morphological analysis alone, and includes a role in regulating genes involved in the synthesis or modification of the extracellular matrix.

Materials and methods

Plant cultivation and transformation

Hybrid aspen clone INRA 717-IB4 (*Populus tremula* × *P. alba*) was used for all experiments and gene

cloning. Plants were in vitro propagated and transformed using the protocol of Han et al. (2000).

Nucleic acid isolations, amplifications, and analysis

Populus genomic DNA was isolated with Qiagen DNeasy, while RNA was isolated using Qiagen RNeasy following the manufacturer's protocols. All PCR was performed using an MJ Research thermocycler with heated lid and calculated temperature options engaged. RNA was reverse transcribed using an anchored poly-T primer 5′-$(T)_{25}$ N_{-1}N3′ (where N = A,C,G, or T; N_{-1} = A,G, or C) using SuperscriptII reverse transcriptase (Invitrogen) following the manufacture's protocols. For the initial identification of the *ARBORKNOX1* transcript, degenerate primers were designed to anneal in the conserved *KNOX* and homeodomain regions of Arabidopsis *SHOOT MERISTEMLESS* and related gene sequences available in Genbank. Primers for degenerate PCR were popSTM-A 5′ATG GST CAT CCT YAC TAY CAC MRH CTC TTG 3′ and popSTM-B 5′TCY TGC TTS AGR CTV CCY AAR TAW CC 3′ (where S = C,G; Y = C,T; M = A,C; R = A,G; H = A,C,T; V = A,C,G; W = A,T).

Nested RACE PCR was performed using SMART RACE technology (Clonetech) according to the manufacturer's protocol. Primers were designed based on the sequence of the degenerate PCR amplified fragment. The primary PCR was performed with primers 5′CATGAATTGGTCAAGGGCTGG ATCTTC3′ for 5′ RACE and 5′GATCTTCAC CGATGCAGCCTG TGTTAG3′ for 3′ RACE. Primary products were diluted 1:50 and a second round of PCR was performed with primers 5′GCCATG CTTTTTCTTCAGAGGG TTGAG3′ for 5′ RACE and 5′CTGGTTGTGGTG ATGGCAATGATAGG A3′ for 3′ RACE. Cycling parameters were 25 cycles (94 1 min; 5 cycles 94°C 15 s, 70°C 30 s, 72°C 3′; 25 cycles 94°C 15 s, 68°C 30 s, 72°C 3 min). The coding sequence was reamplified from cDNA using gene-specific primers and resequenced (Genbank Accession AY755413) to confirm the sequence presented here does not represent a chimera of RACE products.

Alignment of ARK1 and STM proteins was performed with Bio Edit (Hall 1999) using default settings. Phylogenetic analysis was performed for the full amino acid sequence of each protein, using the default settings for the neighbor joining method within the Bio Edit program (Hall 1999).

Whole mount in situ hybridization and histology

For in situ hybridization, probe synthesis and hybridizations were modified from protocols developed for Arabidopsis (Long et al. 1996; Long and Barton 1998). Stems of tissue culture grown *Populus* were sectioned to 100 μm by affixing approximately 5 mm lengths of stem to a metal block with Super Glue, submerging in water, and vibratome sectioning. Each section was immediately removed from the water bath using a small paintbrush and placed in ice-cold FAA (10% Formalin, 5% acetic acid, 50% ethanol) for approximately 4 h. Sections were washed twice in 30% EtOH for 5 min, twice in distilled water, and once in 50% formamide at room temperature prior to hybridization.

Digoxigenin-labeled probes were in vitro transcribed from plasmid pBlue054 for sense and antisense probes of *ARK1*, and plasmid pBlue064 for *pop50S* as described previously (Long et al. 1996; Long and Barton 1998), treated with RQ1 DNAse (Promega) for 10 min at 37°C; the DNAse was heat-killed by incubation at 65°C for 10 min, and the RNA probes were precipitated with 1/10 volume 3 M NaAc (pH 5.2) and 2 volumes of EtOH. Probes were resuspended in 50% formamide.

Hybridizations were performed overnight at 55°C in PCR tubes in a total volume of 200 μl, using the previously described hybridization buffer (Long et al. 1996; Long and Barton 1998). Sections were subjected to a series of washes (0.2% SSC 10 min room temperature; two changes of 0.2% SSC 55°C; NTE 30 min room temperature; NTE 20 μg/ml RNAse A 15 min 37°C; two changes of NTE 5 min room temperature; 0.2% SSC 5 min room temperature; two changes of 0.2% SSC 20 min 55°C; PBS 5 min room temperature; Boehringer blocking solution 45 min room temperature; BSA block 1 h room temperature). Sections were incubated with 1:1250 dilution of alkaline phosphatase conjugated anti-digoxigenin antibody (Roche) in BSA block overnight at 4°C. Sections were then washed with four changes of BSA block for 15 min each at room temperature, washed twice in Buffer C for 15 min each at room temperature. Buffer C (100 mM Tris pH 9.7, 100 mM NaCl, 50 mM $MgCl_2$ containing 8 ul/ml NBT/BCIP (Roche) was added and color development monitored over the course of several hours.

Tissue was fixed for thin sectioning in FAA at 4°C overnight, and dehydrated through an alcohol series, infiltrated with histoclear, and then infiltrated with Paraplast Xtra (Fisher) over the course of 4–7 days with twice daily changes of Paraplast. Tissue was embedded, sectioned to 8 μM, and affixed to gelatin-

subbed slides. Slides were dewaxed in histoclear and stained with Fast Green and Safrinin as described by Ruzin (1999).

Scanning electron microscopy

Tissue for scanning electron microscopy was fixed with FAA, dehydrated, subjected to critical point drying using a Tousimis Samdri 780A, and coated with gold using a Denton Desk II sputter coater. Scanning electron microscopy was performed with a Hitachi S3500N under high vacuum in depth mode with accelerating voltage of 5 kV.

Recombinant DNA constructs

The ARK1 coding sequence was PCR amplified with primers popSTM_L2 5′ACTCGAGATGGAGGGTGGTGATGG3′ and popSTM_R2 5′TGGATCCCATGCATGATCTTCAAA3′, and cloned into the TOPO TA vector (Invitrogen) to create pTAG94. The XhoI to BamHI fragment of pTAG94 was cloned into the XhoI to BamHI of pArt7 (Gleave 1992). The resulting plasmid containing the 35S promoter driving expression of the *ARK1* coding sequence followed by the nos 3′ terminator was transformed into hybrid aspen clone INRA 717-IB4 as described above.

Microarray analysis

An oligonucleotide microarray was manufactured by NimbleGen Systems, Inc. (http://www.nimblegen.com/), and will be described in detail elsewhere (A.M. Brunner, P. Dharmawardhana, and S.P. DiFazio, manuscript in preparation). In short, the array contains three distinct 60 mer probes targeting each of 55,794 nuclear and 126 mitochondrial and chloroplast gene models, and 9,995 unigenes derived primarily from *P. tremula* × *P. tremuloides* EST sequences (Sterky et al. 2004). The gene predictions were derived by the International Populus Genome Consortium from the *Populus trichocarpa* genome sequence (http://www.genome.jgi-psf.org/poplar). Gene annotations presented here represent the top scoring BLAST return for each gene's predicted protein product as query against the TAIR Arabidopsis protein database (http://www.arabidopsis.org/), or the NCBI nonredundant protein database (http://www.ncbi.nlm.nih.gov/) for gene products with no hits to TAIR.

Total RNA was isolated from entire, defoliated stems of tissue culture grown plants 2 months after subculturing. Transgenic and wild-type plants were grown under identical conditions using the maintenance medium described by Han et al. (2000). For the *35S:ARK1* plants, two biological replicates were represented by independently transformed lines (35S32-6 and 35S32-3) from which six clones each were pooled for RNA extraction. For wild type controls, two biological replicates were represented by two groups of plants, from which six clones each were pooled for RNA extraction. Total RNA was extracted by grinding stems in liquid nitrogen, and then bead-homogenizing in the RLT extraction buffer provided by the Qiagen RNA easy kit (Qiagen Inc., USA). RNA was then isolated with the Qiagen RNA easy kit following the manufacturer's protocol. Total RNA quantity and quality was assessed by absorption at 260 and 280 nm, gel electrophoresis, and the Agilent Bioanylyzer (Agilent Technologies, USA) using the manufacturer's protocol. Two independent RNA preps were used as technical replicates within each biological replicate. Biotin labeling of target RNAs, hybridization of targets to arrays, and scanning of arrays was performed by NimbleGen Corporation (http://www.nimblegen.com/).

For analysis of microarray data, data were normalized across arrays using quantile normalization, with gene calls generated by the Robust Multichip Average (RMA) protocol expressed as log2 values. Data were further analyzed in the Gene Spring microarray analysis software environment. Background hybridization levels were estimated using negative control probes which were not represented by a complementary target sequence (96 tiled probes for Ambion ArrayControl spikes derived from *Escherichia coli* sequences). The average of all three probes corresponding to each target gene was used to estimate expression levels, and all genes for which both average wild type and average *35S:ARK1* expression levels were below the estimated background hybridization were eliminated from this analysis. The remaining genes were then filtered based on two-fold difference in expression between *35S:ARK1* and wild type controls, followed by a final round of filtering requiring *t*-test value of 0.05 or less. For categorizing genes into functional categories, annotation was manually generated giving priority to primary literature, and when not available to annotation associated with the best Arabidopsis ortholog (determined by protein BLAST) at TAIR (http://www.arabidopsis.org/). Microarray data are available as Accession GSE2929 through the National Center for Biotechnology Information Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/).

Gene expression differences estimated by microarray analysis were re-evaluated using real time PCR. Gene-specific oligonucleotide PCR primers were

designed to target genes showing differential expression in the microarray comparison of *35S:ARK1* and wild type plants. Primers with Tm of >59°C were designed to produce a product 200–500 bp. Amplification mixture (25 μl) consisted of 40 mM M Tris HCl pH8.4, 100 mM KCl, 6 mM MgCl2, 20 nM Fluorescein, 8% Glycerol, 1X BSA, 0.4X SYBR Green I (Molecular Probes), 1.6 mM dATP, dTTP, dCTP and dGTP, 12 μM of each primer, and 0.6 U of Takara Mirus Hot-Start Taq (Fisher Scientific), and cDNA corresponding to 2.5 μg total RNA. Reactions were run on an iCycler iQ real time PCR machine (BioRad). Cycling parameters were as follows: 3 min at 95°C, 50 cycles of 30 s at 95°C, 30 s at 59°C, and 1 min at 72°C, 1 min at 95°C, and 1 min at 59°C. This was followed by a melting curve program starting from 59 to 99°C with a 10s hold at each 0.5°C. Fluorescence data were collected at the 59°C step and during the melting curve program. A seven step 5-fold cDNA dilution series starting at 1/15 was queried with 50S ribosomal RNA primers to produce a standard curve.

Cell wall chemistry analysis

The chemical composition of stems was determined according to a modified micro-Klason analysis (Huntley et al. 2003). In brief, ~0.5 g of freeze-dried *Populus* stems were ground to pass a 40-mesh screen using a Wiley mill, soxhlet-extracted with acetone for 6 h, digested with 72% H_2SO_4 for 2 h, and then hydrolyzed in 4% H_2SO_4 for 1 h at 121°C. The total weight of non-hydrolyzed, extracted components was determined gravimetrically (acid-insoluble lignin), while the filtrate was analyzed for acid-soluble lignin by absorbance at 205 nm according to TAPPI Useful Method UM250. The carbohydrate composition of the treatment filtrates, expressed in the anhydro-form, was determined by high performance anion-exchange chromatography on a CarboPac PA-1 column using a Dionex High Pressure Liquid Chromatography (HPLC) system (Dionex, Sunnyvale, CA, USA) equipped with a pulsed amperometric detector. Carbohydrates were expressed as a percentage (mg/100 mg dry tissue) of the dry weight of the sample. Thioacidolysis on the extracted, ground stems was employed to determine the compositional ratio of H, G, and S-type lignin monomers, and was conducted according to published methods (Lapierre et al. 1999), with the volumes scaled to accommodate 20 mg of starting material.

Soluble carbohydrates (glucose, fructose and sucrose) were extracted from ground freeze-dried plant material using methanol–chloroform–water (12:5:3) overnight at −20°C. The sample was centrifuged and the supernatant removed. The remaining pellet was washed twice with fresh methanol–chloroform–water (12:5:3) and all fractions pooled. A volume of 5 ml of water was added to the combined supernatant and centrifuged to facilitate phase separation. The aqueous fraction was removed to a round bottom flask and rotary evaporated to dryness. The sample was resuspended in 3 ml of distilled water and analyzed using anion exchange HPLC (Dionex, Sunnyvale, CA, USA) on a DX-600 equipped with a Carbopac PA20 column and an electrochemical detector.

Upon request, all novel materials described in this publication will be made available in a timely manner for non-commercial research purposes.

Results

Cloning of ARBORKNOX1, a *Populus* SHOOT MERISTEMLESS ortholog

A putative *SHOOT MERISTEMLESS* (*STM*) ortholog was cloned from hybrid aspen clone INRA 717-IB4 (*Populus tremula* x *P. alba*) and named *ARBORKNOX1* (*ARK1*). The full length transcript (Genebank Accession AY755413) includes 1,086 bp of protein coding sequence, 214 bp of 3′ UTR including poly adenylation, and 721 bp of 5′ UTR (see Materials and methods). Sequence and phylogenetic analysis suggests that *ARK1* is orthologous to *STM*. As shown in Fig. 1A, the ARK1 protein contains conserved domains found in STM and related proteins, including KNOX, ELK and HOMEODOMAINs. ARK1 and STM proteins are 66% identical and 81% similar at the amino acid level (Fig. 1A). Phylogenetic analysis (see Materials and methods) of ARK1 (from *P. tremula* x *P. alba*), all KNOX homeodomain proteins predicted to be encoded by the *P. trichocarpa* genome sequence, and all Arabidopsis KNOX homeodomain proteins confirms that *ARK1* is a Class 1 *KNOX* gene, and that the STM protein sequence is more similar to ARK1 than to any other Arabidopsis homeodomain protein (Fig. 1B). The *P. trichocarpa* genome has two *STM* paralogs, fgenesh4_pm.C_scaffold_166000014 and gw1.XI.1499.1 (Fig. 1B). The currently annotated gw1.XI.1499.1 gene model is truncated at the 5′ end (not shown).

ARBORKNOX1 is expressed in the shoot apical meristem and vascular cambium

ARK1 expression was monitored by RT-PCR in shoot apices, expanded leaves, stem nodes, and internodes

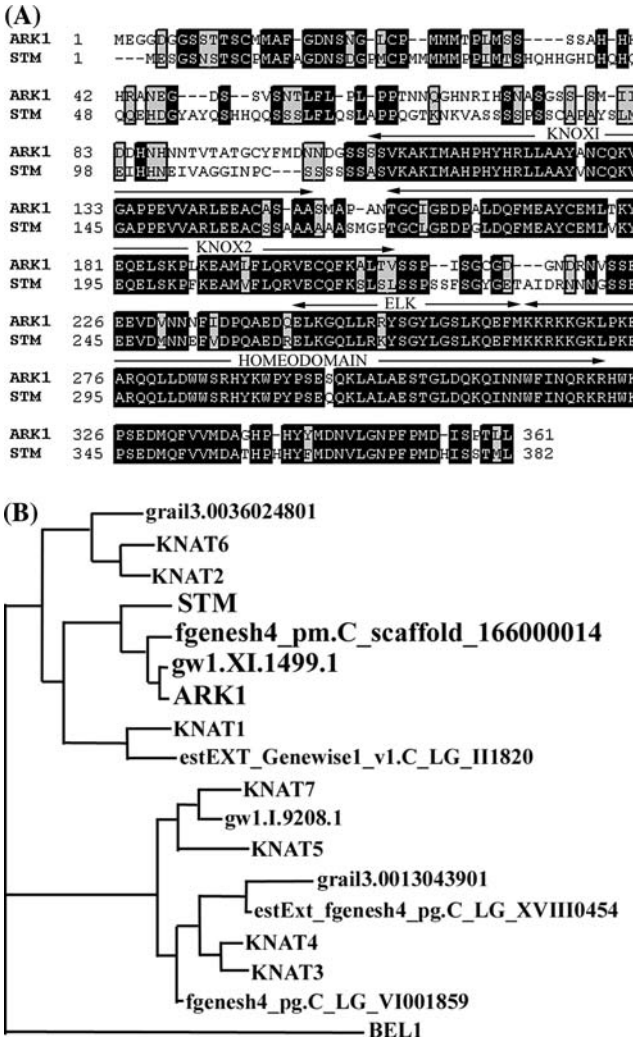

Fig. 1 Sequence of ARK1 and relationship to STM, and to other Arabidopsis and *Populus trichocarpa* KNOX proteins. (**A**) Amino acid sequence alignment of *Populus* ARK1 and Arabidopsis STM. Black and gray boxes indicate identical and similar amino acids, respectively. (**B**). Un-rooted phylogenetic tree of all Arabidopsis KNOX gene protein products, all *Populus trichocarpa* KNOX proteins, and ARK1 from *P. tremula* × *P. alba*. ARK1 is a member of the Class 1 KNOX family, which is comprised of *STM*, *KNAT1*, *KNAT2*, and *KNAT6* in Arabidopsis. *STM* is more similar to *ARK1* than to any other Arabidopsis gene

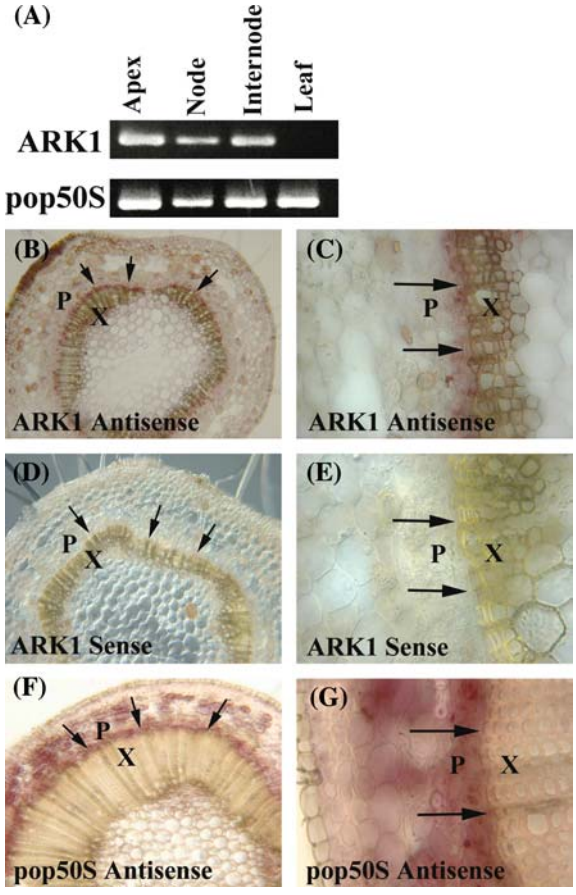

Fig. 2 *ARK1* expression pattern. (**A**) RT PCR of RNAs harvested from apices, nodes, internodes, and leaves. *ARK1* is expressed in all tissues except leaves. Positive control amplifications for a gene encoding a 50S ribosomal subunit (*pop50S*) yield products for all tissues. Internode samples did not contain axillary meristems, and were confirmed to have a vascular cambium by phloroglucinol staining of secondary xylem. RNA samples subjected to PCR amplification without prior reverse transcription did not produce products (data not shown). (**B**) Whole mount in situ hybridization of *Populus* stem cross-section with antisense *ARK1* probe showing signal in the cambial zone. (**C**) Higher magnification of *Populus* stem cross-section probed with antisense *ARK1* probe. (**D**) Whole mount in situ hybridization of *Populus* stem cross-section with negative control, sense *ARK1* probe (no signal). (**E**) Higher magnification of *Populus* stem cross-section with negative control, sense *ARK1* probe. (**F**) Whole mount in situ hybridization of *Populus* stem cross-section with antisense probe for a gene encoding a broadly expressed ribosomal 50S subunit. (**G**) Higher magnification of *Populus* stem cross-section with antisense probe for a gene encoding a broadly expressed ribosomal 50S subunit. Abbreviations: Secondary Xylem (X) and Secondary Phloem (P). Arrows indicate location of cambium region

(see Materials and methods). As shown in Fig. 2A, *ARK1* transcripts are detectable in shoot apices but not in mature leaves, similar to previous reports for *STM* which is down-regulated in differentiating cells of organ anlagen and leaves of simple-leaved dicots (Long et al. 1996). *ARK1* transcripts are also detectable in nodes and internodes of stems (Fig. 2A). While the node samples included axillary meristems that could be responsible for the detected *ARK1* expression, the internode samples were taken from the center of the internode and did not contain axillary meristems. Internode stem samples contained a continuous vascular cambium, as confirmed by phloroglucinol staining of secondary xylem (data not shown), and therefore it is apparent that *ARK1* expression is associated with both the SAM and stems undergoing secondary growth.

Subsequently, the *ARK1* expression pattern in the stem was more precisely determined using in situ hybridization. Fresh stems of tissue culture-grown plants were vibratome sectioned to 100 μm, fixed, and subjected to a whole-mount hybridization procedure using digoxigenin labeled RNA probes (see Materials and methods). An antisense probe revealed *ARK1* expression in the vascular cambium (Fig. 2B). As shown in Fig. 2C, *ARK1* expression includes both the presumptive initials and recent derivatives. This result is in keeping with the broad expression of related *Populus* Class I KNOX genes *pttKNOX1* and *pttKNOX2* in the cambium zone, as reported in a microarray survey of gene expression in the cambium zone (Figure 7, Schrader et al. 2004). A sense *ARK1* negative control probe indicated minimal background for the whole mount in situ procedure (Fig. 2D and E). Hybridization signal was detected in all living cell types for an antisense positive control probe recognizing a broadly expressed gene encoding a *Populus* 50S ribosomal protein (Fig. 2F and G).

ARK1 and STM over-expression cause similar alterations of primary growth

T-DNA constructs were assembled harboring the cauliflower mosaic virus 35S promoter driving expression of either the *Populus ARK1* or Arabidopsis *STM* protein coding sequence, and introduced into hybrid poplar clone INRA 717-IB4 using *Agrobacterium*-mediated leaf disk transformation. Leaf disks transformed with a gene trap vector that does not cause mutant phenotypes served as the transformation controls (Groover et al. 2004).

As shown in Fig. 3A, as early as 12 weeks following transformation, approximately 40% of control cultures developed well-defined shoots with leaves (Fig. 3D). In contrast, leaf disks transformed with either *35S:ARK1* or *35S:STM* produced callus with numerous nodules, but were delayed in organogenesis and failed to produce well-defined shoots or leaves. The 8% of 334 *35S:ARK1* explants (Fig. 3B, E) and 14% of 292 *35S:STM* explants (Fig. 3C, F) eventually produced elongated stem-like outgrowths that were disorganized in their growth, were frequently branching and fasciated and lacked recognizable organs. Six weeks after transfer of shoots to fresh medium, typical control seedlings were 8 cm tall, and had well-formed leaves and expanded internodes (Fig. 3G). In contrast, 7% (22 independently transformed lines) of *35S:ARK1* explants (Fig. 3H) and 4% (11 independently transformed lines) of *35S:STM* explants (Fig. 3I) produced elongated stems, but the leaves were drastically reduced to small outgrowths, and displayed extremely short internodes. Older *35S:ARK1* and *35S:STM* plants had vigorous growth, but were highly branched and often showed multiple branches emerging from a single node. Phenotypes ranged in severity but in general *35S:STM* plantlets displayed a more severe phenotype, could not be easily rooted, and failed to survive outside of culture. In contrast, *35S:ARK1* plants were more amenable to rooting, and in three lines well-formed leaves emerged that eventually could effectively be transferred to soil. Two of these lines have wild type morphology (Fig. 3J, L), while the third line (Fig. 3K, M) has highly lobed leaves, margins that are frequently defined by a vein, mesophyll rumples above the plane of the leaf blade, and drastically reduced petioles. While the severe phenotype of *35S:STM* plantlets did not allow extensive root formation, *35S:ARK1* plants did not exhibit any obvious defects in root formation, architecture or growth. After extended culture, adventitious shoots appeared from the roots of 10 of 22 *35S:ARK1* plantlets (Fig. 3N).

The highly branched and bushy appearance of older *35S:ARK* and *35S:STM* plants reflects the formation and outgrowth of ectopic meristems. Shoots from *35S:ARK1, 35S:STM*, and control plants were examined by scanning electron microscopy (Fig. 4). As shown for *35S:STM* in Fig. 4A, both *35S:STM* and *35S:ARK1* shoots have leaves that fail to fully develop or expand, do not have elongated petioles, and are lobed similar to previous reports for KNOX gene over-expression in other species (Sinha et al. 1993; Chuck et al. 1996; Long et al. 1996). Closer examination of the epidermis reveals disorganization of cells including patches of cells that do not properly contact and interlace with neighboring cells (Fig. 4B). In wild type plants, the axil contains a single axillary meristem associated with the stem that is flanked by stipules (Fig. 4C). After extended culture of *35S:ARK1* plants, ectopic meristems form (Fig. 4D). However, these meristems are not associated with the stem or axil, but rather are associated with the adaxial surface of the leaf blade. Although they also fail to produce normal leaves, the ectopic meristems are competent to produce a shoot (Fig. 4D). Ectopic meristems are also prominent on older *35S:STM* plantlets, as shown in Fig. 4E.

ARK1 and STM over-expression alters secondary growth

A striking aspect of both the *35S:ARK1* and *35S:STM* plants is their thin stems, which is consistent with defects in secondary growth. To determine the effect of *ARK1* and *STM* over-expression on secondary growth,

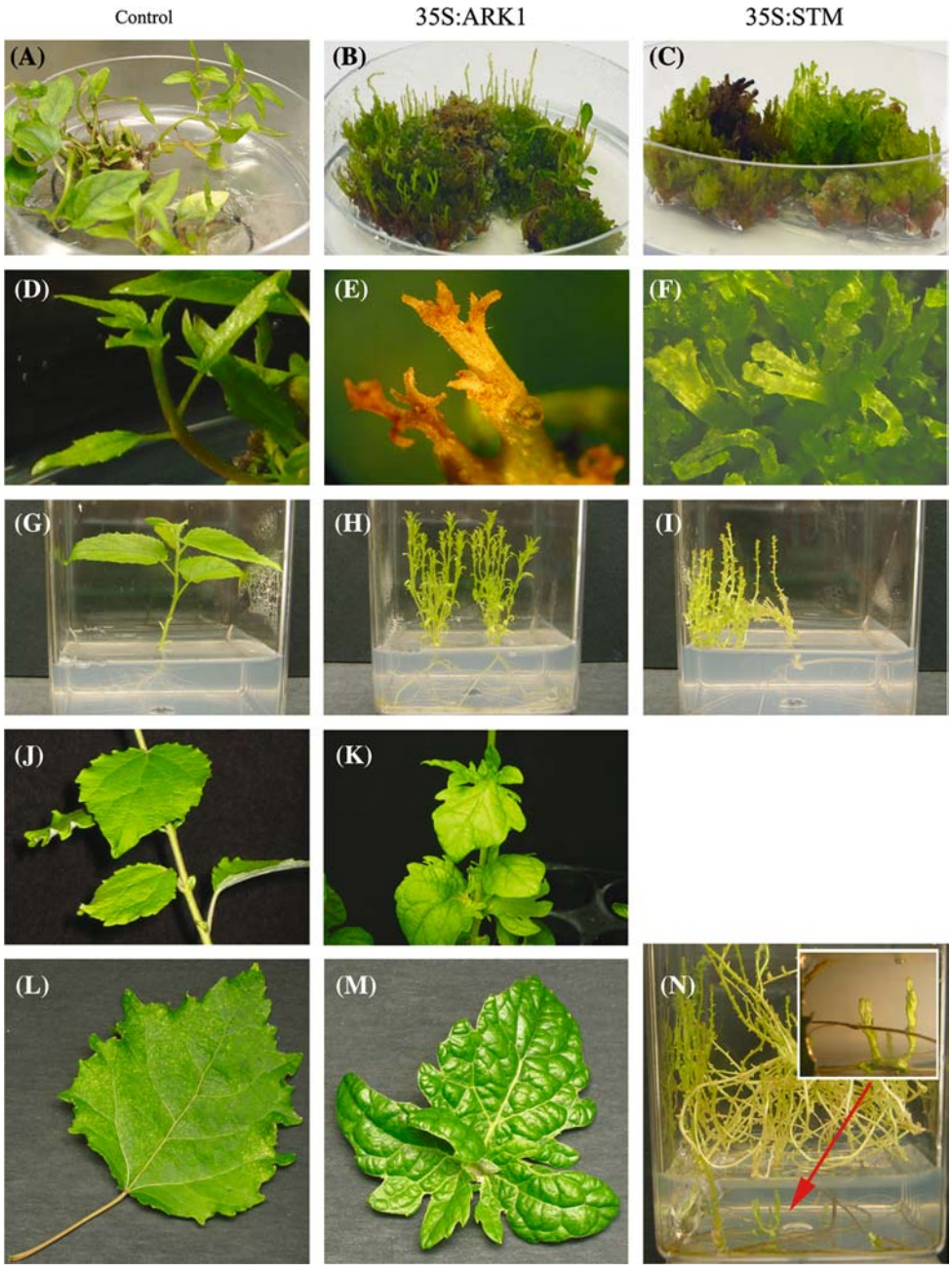

Fig. 3 ARK1 and STM overexpression phenotypes. (**A**) Gene trap control cultures with well-formed stems and leaves. (**B**) *35S:ARK1* cultures with stem-like outgrowths lacking well-formed organs. (**C**) *35S:STM* cultures with stem-like outgrowths lacking well-formed organs. (**D**) Gene trap control plantlet showing well-formed leaves and expanded internodes. (**E**) *35S:ARK1* plantlet with stem-like outgrowths lacking well-formed organs. (**F**) *35S:STM* plantlet with stem-like outgrowths lacking well-formed organs. (**G**) Gene trap control plant approximately 8 weeks after transfer to fresh medium. (**H**) *35S:ARK1* plantlet stems with highly reduced leaves. (**I**) *35S:STM* plantlet stems with highly reduced leaves. (**J**) Greenhouse grown gene trap control plant transferred to soil. (**K**) Greenhouse grown *35S:ARK1* plant with highly lobed leaves. (**L**) Wild type leaf morphology. (**M**) Morphology of *35S:ARK1* plant showing lobing and highly reduced petiole. (**N**) *35S:ARK1* culture with adventitious root shoots (arrow). Inset shows same root shoot indicated at higher magnification

35S:ARK1, *35S:STM*, and control plant stems were cross-sectioned and stained with phloroglucinol to reveal fully lignified secondary cell walls associated with differentiated treachery elements and fibers. In the wild type controls, vascular bundles associated with primary growth appear in predictable positions around the stem and are associated with leaf traces (Fig. 5A). Transition to secondary growth is relatively uniform and rapid among cultured wild type plants. Under the conditions used here a continuous vascular cambium and significant secondary xylem develops between the 5th and 8th internode (Fig. 5C and E). To promote more extensive secondary growth, cultured plants were cut back (hedged) and the base of the stems sectioned after 8 weeks additional growth. Wild type stems with more extensive secondary growth show a smooth boundary defining the interface of the cambium and cell files of the secondary xylem, reflecting similar rates of cambium initial divisions and rates of differentiation for daughter cells in adjacent cell files (Fig. 5I). A single ring of evenly-spaced groups of lignified phloem fibers appear during the transition to secondary growth in wild type plants (Fig. 5C and E). In older stems, two rings of phloem fibers are present (Fig. 5G).

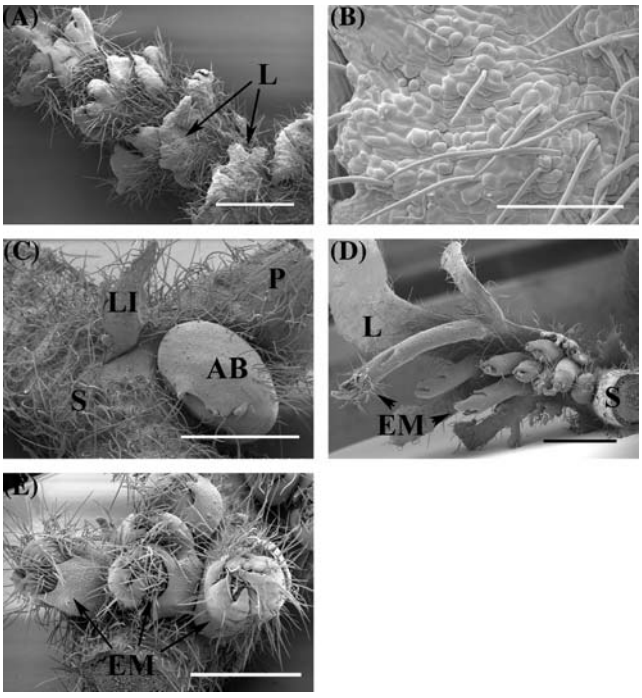

Fig. 4 Scanning electron micrographs of wild type, 35S:ARK1, and 35S:STM plants. (**A**) *35S:STM* shoot. (**B**) Abaxial surface of *35S:STM* leaf showing disorganized cells. (**C**) Wild type axil with single axillary bud. (**D**) *35S:ARK1* leaf with ectopic meristems. Note leaf lobing. (**E**) *35S:STM* stem with ectopic meristems. The associated subtending leaf is highly reduced and partially hidden from view. Abbreviations: AB, Axillary Bud (AB), Ectopic Meristem (EM), Leaf (L), Ligule (LI), Petiole (P), Stem (S). Bar in C is 200 μm. Bar in all other panels 1 mm

In contrast, the production of terminally differentiated secondary vascular cell types that stain with phloroglucinol is delayed in *35S:ARK1* and *35S:STM* plants. *35S:ARK1* plants still display primary vascular growth at internode seven (Fig. 5B), while wild type

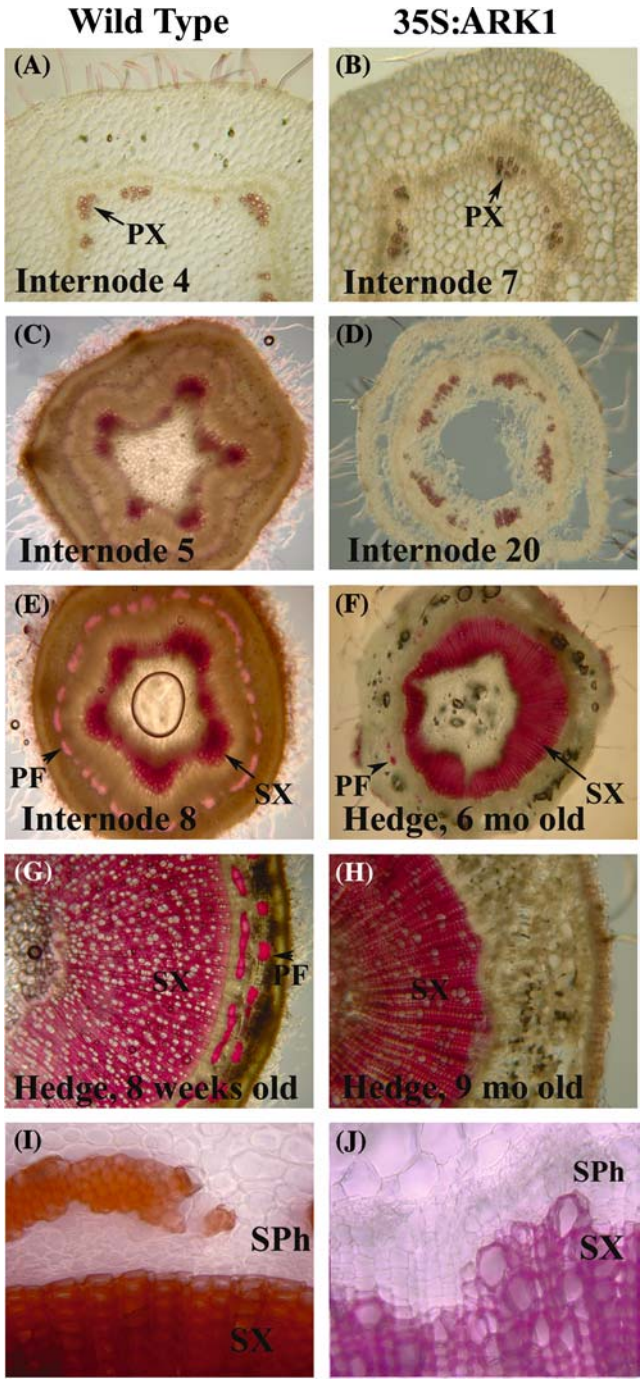

Fig. 5 Wild type and *35S:ARK1* stem sections stained with phloroglucinol to reveal lignified cells. (**A**) Eight week old wild type plant, internode four below apex with primary vascular bundles. (**B**) Eight week old *35S:ARK1* plant, internode seven below apex exhibiting primary growth. (**C**) Eight week old wild type plant, internode five below apex. (**D**) Eight week old *35S:ARK1* plant, internode 20 below apex with dispersed secondary growth. (**E**) Eight week old wild type plant, internode eight below apex with secondary growth. (**F**) Section from base of *35S:ARK1* plant hedged and grown for 6 months in culture. (**G**) Section from base of wild type plant, hedged and grown for 8 weeks in culture. (**H**) Section from base of *35S:ARK1* plant hedged and grown for 9 months in culture. (**I**) Cambium region from wildtype plant showing uniform differentiation of cambium derivatives in secondary xylem in adjacent cell files. (**J**) Cambium region of *35S:ARK1* plant showing variable differentiation of cambium derivatives in secondary xylem in adjacent cell files. Abbreviations: Phloem Fibers (PF), Primary Xylem (PX), Secondary Phloem (SPh), Secondary Xylem (SX)

plants have clearly defined, extensive secondary growth at internode eight (Fig. 5E). Sections from internode 20 at the base of 8 week old *35S:ARK1* stems contain only discontinuous patches of lignified secondary xylem and contained no lignified phloem fibers (Fig. 5D). These discontinuous patches of xylem are associated with cambial activity, as evidenced by the cells of the xylem arranged in files. To determine if *35S:ARK1* plants have the potential to undergo more extensive secondary growth, plants were hedged (cut

back) to encourage continued growth through the release of new shoots, and thickening of the basal portion of the stem. After 6 months of growth, a continuous ring of lignified secondary xylem had formed at the stem bases (Fig. 5F). However, the boundary between the cambium and secondary xylem was wavy rather than smooth, and lignified phloem fibers were absent or found scattered in small groups. After 9 months of hedged growth (Fig. 5G), phloem fibers were still scarce, and the boundary between the cambium and secondary xylem erratic. In contrast to wild type plants, the *35S:ARK1* (Fig. 5J) cambium initials in adjacent files either divide asynchronously, or else their daughter cells are asynchronous in differentiation in adjacent files. As a consequence, the boundary between the cambium and secondary xylem in these plants is wavy. In summary, *ARK1* over-expression does not preclude secondary growth, but serves to inhibit the onset and differentiation of secondary vascular tissues at the morphological level.

ARK1 over-expression modifies expression of genes involved in secondary growth

Microarray analysis was used to identify genes mis-regulated in stems in response to *ARK1* over-expression. Two lines independently transformed with *35S:ARK1* (lines 35S32-6 and 35S32-3) with phenotypes of intermediate severity (comparable to Fig. 3H) were compared to corresponding wild type controls. Total RNA was isolated from defoliated stems of the tissue culture grown *35S:ARK1* and wild type control plants 2 months after subculturing. RNAs were labeled and hybridized to in situ synthesized oligonucleotide microarrays that include probes corresponding to 55,794 nuclear gene models from the *P. trichocarpa* genome sequence (see Materials and methods). Analysis of the resulting data sets (see Materials and methods) revealed that 120 genes were up-regulated and 173 genes were down-regulated two fold or greater in the *35S:ARK1* mutants relative to the corresponding wild type controls (see online supplementary material for annotated gene lists). Mis-expression of thirteen genes was confirmed by real time PCR (see supplementary material), indicating good performance by the oligonucleotide microarray. Notably, *ARK1* was estimated to be over-expressed six-fold in the *35S:ARK1* samples (estimated with array probes defining gw1.XI.1499.1; see supplementary material online), confirming the mutant phenotype to be the result of *ARK1* over-expression.

Of the 120 up-regulated genes, 42% are involved with extracellular matrix functions including cell wall synthesis and modification, 16% are involved in basal or secondary metabolism, 8% are involved in transcription or encode transcription factors, 7% are involved in photosynthesis and light response, and the remaining 27% of genes involved in other processes (Fig. 6). Of the down-regulated genes, 29% are involved in basal or secondary metabolism, 17% are involved in translation, 12% are involved in extracellular matrix functions including cell wall synthesis and modification, 10% are involved in hormone biosynthesis or response, and 8% are involved in photosynthesis or light response. Thus, *ARK1* mis-expression results in a change in expression of specific classes of genes. Because the plants used in this experiment express *ARK1* ectopically, some mis-expressed genes may be direct targets of *ARK1*, while other gene expression differences could reflect downstream, secondary consequences of *ARK1* mis-expression.

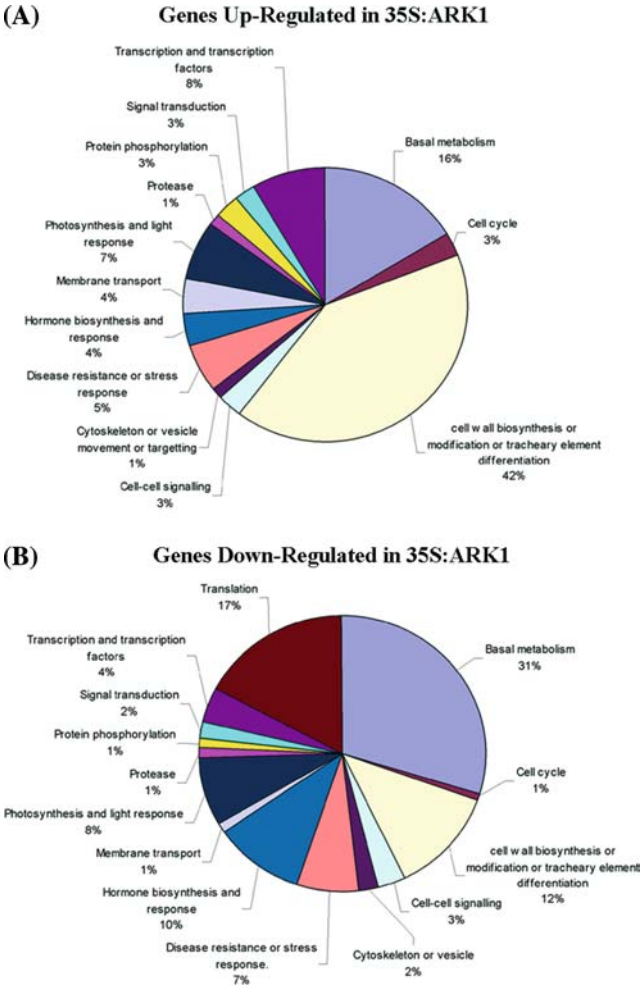

Fig. 6 Classes of genes mis-regulated in stems of *35S:ARK1* plants. Genes that could not be annotated with certainty or otherwise placed into a functional category are not included in these graphs

However, the expression of stress-related genes is actually lower in *35S:ARK1* (see supplementary material online), indicating that gene expression differences are not a general reflection of stress in response to *ARK1* over-expression.

Table 1 shows the largest class of genes over-expressed in *35S:ARK1* stems, which are involved with extracellular matrix functions. Some of the genes identified encode proteins that have been implicated in cell identity or cell signaling functions of the extracel-

Table 1 Genes up or down-regulated in 35S:ARK1 relative to wild type controls that encode proteins involved in extracellular matrix synthesis or modification, or tracheary element differentiation

JGI gene accession[a]	Fold change[b]	Arabidopsis[c]	Definition line[d]
estExt_Genewise1_v1.C_LG_IX4802	15.8	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
eugene3.00130132	9.495	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
eugene3.00191024	8.427	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
gw1.684.7.1	7.58	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
estExt_fgenesh4_pm.C_LG_VIII0405	7.09	At1g48130.1	peroxiredoxin
eugene3.00091518	6.633	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
fgenesh4_pg.C_LG_XIX001005	5.955	At5g03170.1	fasciclin-like arabinogalactan-protein (FLA11)
gw1.1081.8.1	5.576	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
eugene3.10810001	4.933	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
fgenesh4_pg.C_LG_XIX001011	4.556	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
fgenesh4_pg.C_LG_II000045	3.413	At1g20850.1	cysteine proteinase XCP2
eugene3.00151077	3.291	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
estExt_Genewise1_v1.C_LG_X0543	3.281	At3g16920.1	glycosyl hydrolase family 19 (chitinase)
grail3.0008017001	3.137	At1g27440.1	unknown protein/similar to multiple exostoses type II
estExt_Genewise1_v1.C_LG_I1792	3.123	At5g17420.1	cellulose synthase catalytic subunit (IRX3)
gw1.XIX.2840.1	3.086	At5g60490.1	fasciclin-like arabinogalactan-protein (FLA12)
fgenesh4_pm.C_LG_XIV000298	3.064	At1g02640.1	glycosyl hydrolase family 3/similar to beta-xylosidase
eugene3.00002636	2.777	At4g18780.1	cellulose synthase—like protein
estExt_fgenesh4_pg.C_LG_XV0035	2.518	At5g54160.1	O-methyltransferase 1
estExt_fgenesh4_pm.C_LG_VI0293	2.503	At5g60020.1	laccase (diphenol oxidase)
eugene3.00041101	2.446	At5g15630.1	(COBRA-like 4) putative phytochelatin synthetase
gw1.III.1613.1	2.357	At1g62510.1	similar to 14KD proline-rich protein DC2.15 precursor
gw1.XI.3218.1	2.337	At2g21770.1	putative cellulose synthase catalytic subunit
eugene3.00021675	2.294	At5g54160.1	O-methyltransferase 1
estExt_fgenesh4_pm.C_LG_IV0566	2.217	At3g12500.1	glycosyl hydrolase family 19 (basic endochitinase)
gw1.XII.749.1	2.216	At5g61750.1	germin-like protein—like
estExt_fgenesh4_pg.C_LG_I2215	2.199	At4g24780.1	putative pectate lyase
estExt_Genewise1_v1.C_LG_V2210	2.193	At1g20850.1	cysteine proteinase XCP2
estExt_Genewise1_v1.C_LG_II3807	2.187	At3g23730.1	xyloglucan endotransglycosylase
eugene3.00130509	2.148	At5g39110.1	Germin-like protein
estExt_fgenesh4_pm.C_LG_I1023	2.143	At4g34050.1	caffeoyl-CoA 3-O-methyltransferase
estExt_fgenesh4_pg.C_280066	2.04	At3g43190.1	sucrose synthase
gw1.XIV.2207.1	2.04	At1g02460.1	polygalacturonase
gw1.20008.1.1	2.002	At4g28250.1	beta-expansin pollen allergen protein
gw1.351.9.1	0.487	At4g36220.1	ferulate-5-hydroxylase (FAH1)
fgenesh4_pm.C_LG_X000251	0.455	At1g70370.1	aromatic rich glycoprotein
eugene3.00060481	0.443	At3g57520.1	glycosyl hydrolase family 36
estExt_fgenesh4_pg.C_LG_III0902	0.438	At1g12240.1	glycosyl hydrolase family 32/identical to beta-fructosidase
fgenesh4_pg.C_LG_XIII000294	0.429	At1g26770.1	expansin
gw1.I.452.1	0.41	At3g16370.1	putative APG protein/similar to anter-specific proline-rich
gw1.XIV.2235.1	0.409	At5g19730.1	pectinesterase family
estExt_fgenesh4_pm.C_LG_X0551	0.404	At1g02640.1	glycosyl hydrolase family 3/similar to beta-xylosidase
estExt_fgenesh4_pg.C_LG_XIV0665	0.383	At4g02290.1	glycosyl hydrolase family 9/similar to endo-1,4-beta glucanase
gw1.III.2590.1	0.375	At1g72610.1	germin-like protein
gw1.IV.57.1	0.342	At1g26770.1	expansin
estExt_fgenesh4_pm.C_LG_XIII0455	0.224	At2g40610.1	expansin
gw1.7167.1.1	0.111	At5g25980.1	glycosyl hydrolase family 1/similar to myrosinase precursor

[a] JGI gene accession refers to the accession number assigned to the assayed gene model by the Joint Genome Institute (http://www.genome.jgi-psf.org)

[b] Fold Change is expressed as the ratio of gene expression in 35S:ARK1 to wild type control

[c] Arabidopsis refers to the accession number of the best Arabidopsis BLAST return using the JGI gene model as query

[d] Definition line is from the Arabidopsis accession

lular matrix, including arabinogalactan proteins (e.g. Acosta-Garcia and Vielle-Calzada 2004) and glycosyl hydrolases (e.g. Zhong et al. 2002). Such proteins could potentially play roles in determining cell fate or identity, or signaling within the meristem. Surprisingly, also within this class are up-regulated genes encoding proteins involved in secondary cell wall synthesis, which is an irreversible, terminal processes. For example, some of the genes up-regulated in *35S:ARK1* have Arabidopsis orthologs for which loss of function mutant phenotypes indicate a positive role in promoting secondary cell wall synthesis, *COBRA*-like 4 (At5g15630), the glycosyl transferases *IRREGULAR XYLEM 1* (*IRX1;* At4g18780), *9* (At2g21770) and *10* (*IRX10*; At1g27440), and xyloglucan transferase (At3g23730).

35S:ARK1 stems also showed up-regulation of genes involved in lignification of cell walls, another terminal differentiation process. Lignin biosynthesis genes up-regulated in *35S:ARK1* include O-methyltransferase (estExt_fgenesh4_pg.C_LG_XV0035), caffeoyl-CoA 3-O-methyltransferase (CCOMT; estExt_fgenesh4_pm.C_LG_I1023), O-methyltransferase 1 (COMT; eugene3.00021675), and a laccase (estExt_fgenesh4_pm.C_LG_VI0293). Additional genes involved in lignin biosynthesis were also up-regulated, but at levels just below the two-fold cutoff used for this analysis (see below). In contrast, a putative *FAH1* ortholog (gw1.351.9.1) encoding ferulate 5-hydroxylase (F5H) is down-regulated in *ARK1* over-expressing stems. F5H is responsible for shunting lignin precursors destined to produce either *p*-hydroxyphenyl (H), guaiacyl (G) or syringyl (S) monolignols to the S-specific pathway (Franke et al. 2000).

35S:ARK1 stems have altered cell wall chemistries

To determine if the mis-expression of genes encoding extracellular matrix proteins resulted in a biochemical phenotype, cell wall chemistries were compared for *ARK1*-over-expressing and wild type plants (see Materials and methods). To provide an effective comparison of the cell wall chemistries, plantlets from the same genotypes used for microarray analysis were grown until secondary xylem was evident in both the wild type and over-expression plants. A comparison of the structural carbohydrates (cellulose and hemicellulose) indicated that, on a dry weight basis, four of the five neutral sugars (arabinose, galactose, glucose and mannose) were slightly reduced in the *35S:ARK1* plants, while the xylan composition showed a >30% increase (Table 2). Although the expression data indicate the up-regulation of putative genes involved in cellulose biosynthesis (cellulose synthase-like eugene3.00002636 was 2.8-fold upregulated; putative cellulose synthase catalytic subunit gw1.XI.3218.1 was 2.3-fold upregulated) it is possible that the up-regulation of genes such as the xyloglucan endotransglycosylase encoded by estExt_Genewise1_v1.C_LG_II3807 may be responsible for the observed alteration in structural xylan carbohydrate composition.

The available soluble sugars (glucose, fructose and sucrose) were also quantified from developing cell walls and found to be very different between wild type and *35S:ARK1* plants (Table 2). The soluble carbohydrate composition of wild type plants was comprised primarily of sucrose and a minor amount of glucose. In contrast, the *35S:ARK1* plants had a reduced concentration of sucrose (~34%), and comparable levels of both glucose and fructose. The glucose content was also three-fold greater than that available in the wild type plants. It is plausible that increased soluble glucose content, the appearance of fructose and the concurrent drop in sucrose concentrations in the stems is associated with the elevated expression of a sucrose synthase-like gene (SuSy-like: aspen_CL8Contig3) which showed a two-fold increase in expression (not shown: see NCBI GEO Accession GSE2929). The liberation of higher levels of fructose could plausibly be the result of slightly elevated expression levels of genes encoding fructose metabolizing enzymes, including fructokinase (gw1.XVI.3447.1—1.638×) and fructose bisphosphatase (estExt_Genewise1_v1.C_939-0008—1.515×).

The total lignin content (Table 3) was quantified and found to be slightly higher in the *35S:ARK1* lines (as much as a 10% increase by weight). The increased total lignin content is primarily associated with an increase in the acid soluble fraction, not the more highly conjugated, high molecular weight acid-insoluble lignin, which likely accounts for lower phloroglucinol

Table 2 Structural and soluble carbohydrate composition of wild type and *35S:ARK1* poplar

	Structural carbohydrates (μg/mg tissue)					Soluble carbohydrates (μg/mg tissue)		
	Arabinose	Galactose	Glucose	Xylose	Mannose	Sucrose	Glucose	Fructose
Wild type	18.4 (±2.4)	13.7 (±2.0)	258.7 (±1.1)	23.6 (±8.3)	6.9 (±0.7)	28.0(±2.5)	2.9 (±0.2)	ND
35S:ARK1	14.2 (±0.7)	12.1 (±0.7)	242.1 (±18.0)	30.3 (±5.0)	5.8 (±2.4)	18.3 (±1.0)	10.8 (±0.1)	9.3 (±0.1)

Table 3 Total lignin composition of wild type and *35S:ARK1* poplar

	Total lignin (mg/100 mg)	Acid insoluble lignin (mg/100 mg)	Acid soluble lignin (mg/100 mg)	H:G:S monomer ratio	Total monomer yield (μmol/g lignin)
Wild type	24.85	21.97 (±0.16)	2.88 (±0.50)	4.8:31.6:63.6	333.8 (±4.53)
35S:ARK1	27.63	22.31 (±0.14)	5.32 (±0.65)	12.0:47.9:40.1	439.8 (±10.75)

staining in the *35S:ARK1* lines. The increased lignin content is consistent with the up-regulation of several key lignin-specific biosynthetic genes recently shown to up-regulated during developmental lignification (Ehlting et al. 2005), including *Populus* orthologs of: *PAL1* (At2g37040: 1.695×), *C4H* (At2g30490: 1.753×), *4CL*s (At1g51680: 1.773× and At5g63300: 1.581×), *CCOMT* (At4g34050: 2.143×), *CCR* (At1g15950: 1.549×), *COMT* (At5g54160: 2.454× and At4g35160: 1.730×) *CAD*s (At3g19450: 1.825× and At4g39330: 1.802×), and a laccase (At5g60020: 2.44×: Table 1, and NCBI GEO Accession GSE2929).

A comprehensive analysis of the cell wall lignin monomer fraction by thioacidolysis (Table 3) revealed that the *35S:ARK1* plants have altered lignin monomer composition. The wild type plants exhibited the expected ~2:1 ratio of S:G (63.6:31.6:4.8—S:G:H) for an angiosperm, while the *35S:ARK1* plants have a dramatically different ratio, with a higher composition of G- and H-derived lignin monomers, and a substantially reduced concentration of S-lignin (43.3:48.8:7.9, S:G:H). These findings are consistent with the down-regulation of ferulate 5-hydroxylase (F5H, encoded by putative *FAH* ortholog gw1.351.9.1) by more than two fold (Table 1), and the current up-regulation of several associated lignin biosynthetic genes in the *35S:ARK1* plants, as previously mentioned. F5H is responsible for the hydroxylation of coniferaldehyde and coniferyl alcohol to 5-hydroxy coniferaldehyde or coniferyl alcohol moieties during the production of S-lignin, and has previously been implicated in regulating lignin monomer composition in woody plants (Franke et al. 2000).

Discussion

The work presented here supports the notion that a major genetic regulatory mechanism is shared by the SAM and the vascular cambium. We cloned *ARBORKNOX1* (*ARK1*), a *Populus* ortholog of *SHOOT MERISTEMLESS* (*STM*), and found that it is expressed in both the SAM and cambium. Over-expression analysis was used to study the function of *ARK1* in *Populus* because the resulting dominant phenotypes could be scored in primary transformants, obviating the need for lengthy rounds of sexual reproduction. Importantly, like *stm*, loss-of-function *ark1* plants are expected to be seedling lethal, precluding the ability to study secondary growth which occurs late in development. Secondary growth could be studied in *ARK1* over-expressing plants because seedlings were viable.

Both *ARK1* and *STM* over-expression in *Populus* resulted in ectopic meristems forming on the adaxial side of leaves, inhibition of leaf development, shortening of internode lengths, and delay of the terminal differentiation of daughter cells derived from the cambium. The simplest explanation of these morphological phenotypes is that *ARK1* and *STM* are functional orthologs, and that these genes promote meristematic cell fate and delay terminal differentiation in both the SAM and the cambium. This interpretation would be consistent with previous interpretations of *STM* function in the SAM (discussed in Long and Barton 1998).

Microarray and biochemical analysis of *ARK1* over-expressing plants suggests *ARK1*/*STM* act in part to modify the extracellular matrix (ECM). While once considered to play only a structural role, the current concept of the ECM of plants is a dynamic structure that performs critical roles in cell–cell signaling, cell identity, and cell adhesion (Baluska et al. 2003). These same processes are fundamental to the function of the SAM and the cambium, although the genetic regulation of these processes remains obscure. Our results suggest that *STM*/*ARK1* may directly or indirectly regulate genes involved in ECM functions, and point to candidate ECM genes that could play roles in signaling or cell identity in meristematic cells and their daughters. A distinctive group of strongly up-regulated genes in *ARK1* over-expressing plants encode fasciclin class arabinogalactan proteins. Although the function of these proteins has not been fully described in plants, the fasciclin domain in animals plays an important function in cell adhesion (Johnson et al. 2003). Fasciclin class arabinogalactan genes orthologous to the ones reported here (*FLA11* and *FLA12*) have been previously described as being highly expressed in differentiating secondary xylem (Lafarguette et al. 2004), indicating these genes are normally expressed during secondary growth. Similarly, the family of glycosyl hydrolase-encoding genes (including ''chitinases'') that are up-regulated in *ARK1* over-expressing plants have

been implicated in regulating various developmental processes and presumably act through modification of the ECM (e.g. Wiweger et al. 2003). Expansins are also up-regulated in *ARK1* over-expressing stems, and have been implicated in regulating cell expansion and organ initiation in the SAM (Pien et al. 2001).

Other genes up-regulated in *ARK1* over-expressing stems have previously been shown to be involved in tracheary element differentiation and lignification, processes associated with terminal differentiation. At first glance, this result seems in conflict with the phenotype of *ARK1* over-expressing stems, whose outward anatomy and morphology is characterized by delayed appearance or differentiation of fully lignified cell types. However, the expression pattern of *STM* and *ARK1* is broad in both the SAM and cambium, and extends beyond the presumptive meristematic initials to overlap both cell layers and developmental compartments defined by expression patterns of other regulatory genes. It seems reasonable to speculate that daughter cells in the peripheral regions of *STM*/*ARK1* expression in both the SAM and cambium have already initiated programs of differentiation and are adopting identities distinct from their meristematic origins. Indeed, Long and Barton (1998) proposed that *STM* acts differently in the embryonic peripheral region and embryonic central region of developing embryos, based in part on observed overlapping expression patterns of STM and genes associated with different functions in the SAM, including genes involved in stem cell maintenance (e.g. *CLV1*), and genes involved in differentiation (e.g. *ANT*). Similarly, *ARK1* (Fig. 2) and related Class I KNOX genes (Schrader et al. 2004) are broadly expressed in the cambium region, and overlap the expression patterns of other SAM meristem regulatory genes (Schrader et al. 2004). The function of ARK1 could differ in the cambial initials versus cambial daughters based on interactions with other proteins or modifying processes. Our analysis of gene expression in *ARK1* over-expressing plants is consistent with the notion that *STM*/*ARK1* may play a dual role in both supporting gene expression required to maintain the meristem initials, as well as initiating gene expression necessary for differentiation and/or conferring new cell identity through modification of the extracellular matrix as daughter cells are displaced away from the initials.

An alternative explanation is illustrated by the observation that, while other lignin-related genes are up-regulated, a putative *FAH1* ortholog (gw1.351.9.1) encoding lignin ferulate 5-hydroxylase (F5H) is down-regulated in *ARK1* over-expressing stems. F5H shunts lignin precursors destined to produce either *p*-hydroxyphenyl (H), guaiacyl (G), or syringyl (S) monolignols to the S-specific pathway, and the down-regulation of this late step in the S-monolignol pathway might result in compensation by up-regulation of other upstream components of the lignin biosynthesis pathway. In this scheme, *FAH* may be a direct target of *ARK1*, while other lignin biosynthetic genes may be affected secondarily through compensatory mechanisms. Similarly, mis-expression of ECM genes regulated directly by STM could result in modifications of cell identity that trigger expression of downstream ECM genes. Clearly, more experimentation is required to distinguish between these interpretations.

Similar experiments with another Class I KNOX gene, *BREVIPEDICELLUS* (*BP*), also found mis-regulation of genes encoding cell wall biosynthetic enzymes in response to either over-expression or loss-of-function alleles (Mele et al. 2003). BP was found to directly bind to the promoters of genes encoding the lignin biosynthetic enzymes COMT and CCoAOMT, and the experiments were interpreted to indicate a role for BP in negatively regulating genes involved in terminal differentiation and lignification. An alternative interpretation was favored by studies of *bp* loss-of-function mutants, which were interpreted to indicate that BP may act to promote vascular development and the appearance of lignified cells (Smith and Hake 2003). A possible explanation for these different interpretations is that competency to lignify could differ between the different light regimes and grown conditions used in these two studies (Rogers and Campbell 2004), or differences in sampling or assays used. Regardless, the Class I KNOX genes *BP* and *ARK1*/*STM* act in part to modify the extracellular matrix, although further experimentation is required to detail the precise mechanisms.

That *ARK1* is expressed in both the SAM and cambium (as is *STM*, J. Long and K. Barton, pers. com.) is consistent with the direct co-option (Ganfornina and Sanchez 1999) of *STM* function from the SAM during the evolution of the vascular cambium (Groover 2005). Extensive microarray-based comparison of gene expression in the SAM and cambium of *Populus* by Schrader et al. (2004) identified additional instances of genes expressed during secondary growth that have been well-characterized for their role in regulating the SAM, for example *HD-ZIP* and *KANADI* genes that pattern both the SAM and primary vasculature (Emery et al. 2003). Other regulatory genes were found to be uniquely expressed in the SAM, but had one or more paralogs expressed in the

cambium (e.g. *WUSCHEL*). Co-option of a duplicated element (Ganfornina and Sanchez 1999), for example gene duplication followed by acquisition of unique functions by paralogs, might explain the expression of paralogs of SAM genes in secondary vascular tissues. Functional characterization of these genes during secondary growth would help to discriminate between overlapping and unique aspects of cambium and SAM regulation.

Furthering our understanding of the vascular cambium and secondary growth has important practical implications. Woody plants and forest trees are the basis of forest industry, provide fuel wood in developing countries, underpin vital ecosystems, and play fundamental roles in carbon sequestration and buffering climate change. The results presented here provide an example that, by using the extensive knowledge of the genes and mechanisms regulating the SAM as a roadmap, rapid progress could be made in unraveling the mechanisms regulating the cambium and wood formation, and lead to new applications.

Acknowledgments We thank Kathy Barton for seminal discussion about *STM* and secondary growth, and sharing of unpublished data. We thank Annie Delfino-Mix for greenhouse cultivation of *Populus* trees, Bethany Heynen for assistance in tissue culture and greenhouse propagation of *Populus* trees, and UC LEADS intern Diana Cheng for outstanding technical assistance. We thank the Bowman and Sinha labs for tolerating repeated pilfering of plasmids and reagents. We thank the DOE Joint Genome Institute for sequencing the *Populus* genome and making annotation publicly available prior to publication.

References

Acosta-Garcia G, Vielle-Calzada JP (2004) A classical Arabinogalactan protein is essential for the initiation of female gametogenesis in Arabidopsis. Plant Cell 16:2614–2628

Baluska F, Samaj J, Wojtazek P, Volmann D, Menzel D (2003) Cytoskeleton-plasma membrane-cell wall continuum in plants. Emerging links revisited. Plant Phys 133:482–491

Busov V, Meilan R, Pearce D, Ma C, Rood S, Strauss S (2003) Activation tagging of a dominant gibberellin catabolism gene (GA 2-oxidase) from poplar that regulates tree stature. Plant Phys 123:1283–1291

Byrne M, Groover A, Fontana J, Martienssen R (2003) Phyllotactic pattern and stem cell fate are determined by the *Arabidopsis* homeobox gene *BELLRINGER*. Development 130:3941–3950

Chen H, Banerjee A, Hannapel D (2004) The tandem complex of BEL and KNOX partners is required for transcriptional repression of ga20ox1. Plant J 38:276–284

Chuck G, Lincoln C, Hake S (1996) KNAT1 induces lobed leaves with ectopic meristems when overexpressed in Arabidopsis. Plant Cell 8:1277–1289

Ehlting J, Mattheus N, Aeschliman DS, Li E, Hamberger B, Cullis(IF, Zhuang J, Kaneda M, Mansfield SD, Samuels AL, Ritland K, Ellis BE, Bohlmann J Douglas CJ (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. Plant J 42:618–640

Emery J, Floyd S, Alvarez J, Eshed Y, Hawker N, Izhaki A, Baum S, Bowman J (2003) Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. Curr Biol 14:1768–1774

Franke R, McMichael CM, Meyer K, Shirley AM, Cusumano JC, Chapple C (2000) Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase. Plant J 22:223–234

Gallois J, Woodward C, Reddy G, Sablowski R (2002) Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in Arabidopsis. Development 129:3207–3217

Ganfornina M, Sanchez D (1999) Generation of evolutionary novelty by functional shift. BioEssays 21:432–439

Gleave AP (1992) A versatile binary vector system with a T-DNA organizational structure conducive to efficient integration of cloned DNA into the Plant Genome. Plant Mol Biol 6:1203–1207

Groover A (2005) What genes make a tree a tree? Trends Plant Sci 10:210–214

Groover A, Fontana J, Dupper G, Ma C, Martienssen R, Strauss S, Meilan R (2004) Gene and enhancer trap tagging of vascular-expressed genes in poplar trees. Plant Phys 134:1742–1751

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

Han K, Meilan R, Ma C, Strauss S (2000) An Agrobacterium transformation protocol effective in a variety of cottonwood hybrids (genus *Populus*). Plant Cell Rep 19:315–320

Hay A, Kaur H, Phillips A, Hedden P, Hake S, Tsiantis M (2002) The gibberellin pathway mediates KNOTTED1-Type homeobox functions in plants with diverse body plans. Curr Biol 12:1557–1565

Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield SD (2003) Improved pulping efficiency in C4H-F5H transformed poplar. J Agr Food Chem 51:6178–6183

Johnson K, Jones B, Bacic A, Schultz C (2003) The fasciclin-like arabinogalactan proteins of Arabidopsis. A multigene family of putative cell adhesion molecules. Plant Phys 133:1911–1925

Ko J, Han K (2004) Arabidopsis whole-transcriptome profiling defines the features of coordinated regulations that occur during secondary growth. Plant Mol Biol 55:433–453

Lafarguette F, Leple J, Dejardin A, Laurans F, Consta G, Lesage-Descauses M, Pilate G (2004) Poplar genes encoding fasciclin-like arabinogalactan proteins are highly expressed in tension wood. New Phyt 164:107–121

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

Larson P (1994) The vascular cambium. Springer-Verlag, NY

Laux T, Mayer K, Berger J, Jurgens G (1996) The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. Development 122:87–96

Lenhard M, Jurgens G, Laux T (2002) The WUSHEL and SHOOT MERISTEMLESS genes fufil complementary roles in Arabidopsis shoot meristem regulation. Development 129:3195–3206

Long JA, Moan E, Medford J, Barton MK (1996) A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. Nature 379:66–69

Long JA, Barton MK (1998) The development of apical embryonic pattern in *Arabidopsis*. Development 125:3027–3035

Mayer K, Schoof H, Haecker A, Lenhard M, Jurgens G, Laux T (1998) Role of *WUSHCEL* in regulating stem cell fate in the Arabidopsis shoot meristem. Cell 95:805–815

Mele G, Ori N, Sato Y, Hake S (2003) The *Knotted1*-like homeobox gene *BREVIPEDICELLUS* regulates cell differentiation by modulation metabolic pathways. Genes Dev 17:2088–2093

Rogers L, Campbell M (2004) The genetic control of lignin deposition during plant growth and development. New Phyt 164:17–30

Pien S, Wyrzykowska J, McQueen-Mason S, Smart C, Fleming A (2001) Local expression of expansin induces the entire process of leaf development and modified leaf shape. Proc Natl Acad Sci USA 98:11812–11817

Ruzin SE (1999) Plant microtechnique and microscopy. Oxford University Press, NY

Sakamoto T, Kamiya N, Ueguchi-Tanaka M, Iwahori S, Matsuoka M (2001) KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. Genes Dev 15:581–590

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

Sinha N, Williams R, Hake S (1993) Overexpression of the maize homeobox gene, *Kn1*, causes a switch from determinant to indeterminate cell fates. Genes Dev 7:787–795

Smith H, Hake S (2003) The interaction of two homeobox genes, BREVIPEDICELLUS and PENNYWISE, regulates internode patterning in the Arabidopsis inflorescence. Plant Cell 15:1717–1727

Sterky F, Bhalerao RR, Unneberg P, Segerman B, Nilsson P, Brunner AM, Charbonnel-Campaa L, Lindvall JJ, Tandre K, Strauss SH, Sundberg B, Gustafsson P, Uhlen M, Bhalerao RP, Nilsson O, Sandberg G, Karlsson J, Lundeberg J, Jansson S (2004) A *Populus* EST resource for plant functional genomics. Proc Natl Acad Sci USA 101:13951–13956

Vollbrecht E, Veit B, Sinha N, Hake S (1991) The developmental gene Knotted-1 is a member of a maize homeobox gene family. Nature 350:241–243

Williams R (1998) Plant homeobox genes: many functions stem from a common motif. BioEssays 20:280–282

Wiweger M, Farbos I, Ingouff M, Lagercrantz U, von Arnold S (2003) Expression of Chia4-Pa chitinase genes during somatic and zygotic embryo development in Norway spruce (Picea abies): similarities and differences between gymnosperm and angiosperm class IV chitinases. J Exp Bot 54: 2691–2699

Zhong R, Kays S, Schroeder B, Ye Z (2002) Mutation of a chitinase-like gene causes extopic deposition of lignin, aberrant cell shapes, and overproduction of ethylene. Plant Cell 14:165–179