EXPRESSION AND FUNCTION OF ARABINOGALACTAN-PROTEINS IN XYLEM OF LOBLOLLY PINE

C.A. Loopstra[1], E.-G. No[2], and R.R. Sederoff[3]

Abstract. Genomic and cDNA clones of two genes encoding xylem-specific proteins were previously isolated and characterized. Transcripts of these genes are extremely abundant in differentiating xylem, much less abundant in needles, and are present at very low or non-detectable levels in embryos and megagametophytes. Both genes appear to encode arabinogalactan-proteins (AGPs). AGPs are highly glycosylated proteins thought to play important roles in plant development. AGPs have been found to be abundant in differentiating xylem of loblolly pine. We are attempting to produce genetically engineered loblolly pine plantlets with reduced amounts of these AGPs in order to examine their function in xylem development. Possible roles include cell-cell signalling or cellular interactions, transport of cell wall components, autolysis, or they may act as gums or humectants.

We have isolated both promoters and found them to be functional in bombarded pine tissues and in transgenic tobacco, poplar, and white spruce. We are currently searching for the elements responsible for xylem-specificity and for the high levels of expression observed. These elements will be of value in future attempts to genetically engineer wood properties.

Key words: *Pinus taeda,* xylem development, arabinogalactan-proteins, gene expression.

INTRODUCTION

In recent years, progress has been made towards understanding the molecular basis of xylem development in coniferous species. Several enzymes involved in lignin biosynthesis have been purified and cloned (Whetten and Sederoff, 1992; O'Malley et al., 1992; Bao et al., 1993), an extensin-like cell wall protein has been purified (Bao et al., 1992) and regulatory elements of genes expressed in xylem have been isolated. There is however, much we do not yet understand about the genes involved in the formation of wood. There are genes involved in xylem differentiation with unknown functions. We also have a very poor understanding of the regulation of gene expression in xylem. In this paper, we describe work in progress to characterize two genes preferentially expressed in differentiating xylem, to determine their function, and to examine their regulation.

Isolation and characterization of putative xylem-specific arabinogalactan-proteins

The isolation and characterization of two genes preferentially expressed in newly differentiating xylem of loblolly pine (PtX3H6 and PtX14A9) have previously

[1] Assistant Professor and [2] Graduate Student, Dept. of Forest Science, Crop Biotechnology Center, Texas A & M University, College Station TX 77845-2135.
[3] Professor, Dept. of Forestry, North Carolina State University, Raleigh, NC 27695

been described (Loopstra and Sederoff, 1995) and are summarized here. A cDNA library constructed using RNA isolated from newly differentiating xylem was differentially screened to identify clones of genes with xylem-specific expression. Of the sixteen clones isolated, thirteen were found to be the same (PtX3H6) and three were the same (PtX14A9). From this, we conclude that although it is likely other genes with preferential expression in xylem exist, transcripts of PtX3H6 and PtX14A9 are the most abundant.

Northern blot analyses were used to look at transcript abundance in various tissues and at different stages of development. Transcripts of both PtX3H6 and PtX14A9 are much more abundant in xylem than in needles, embryos, or megagametophytes (Figure 1). Some hybridization to needle RNA was observed. This was not unexpected due to the presence of vascular tissues in needles. RNAs isolated from 6 week, 1 year, 2 year, and 10 year old trees as well as from earlywood and latewood were also examined. The only difference observed with tree age was a lower level of expression in the stems of 6 week old trees, possibly because pure xylem could not be isolated. No differences were detected between earlywood and latewood.

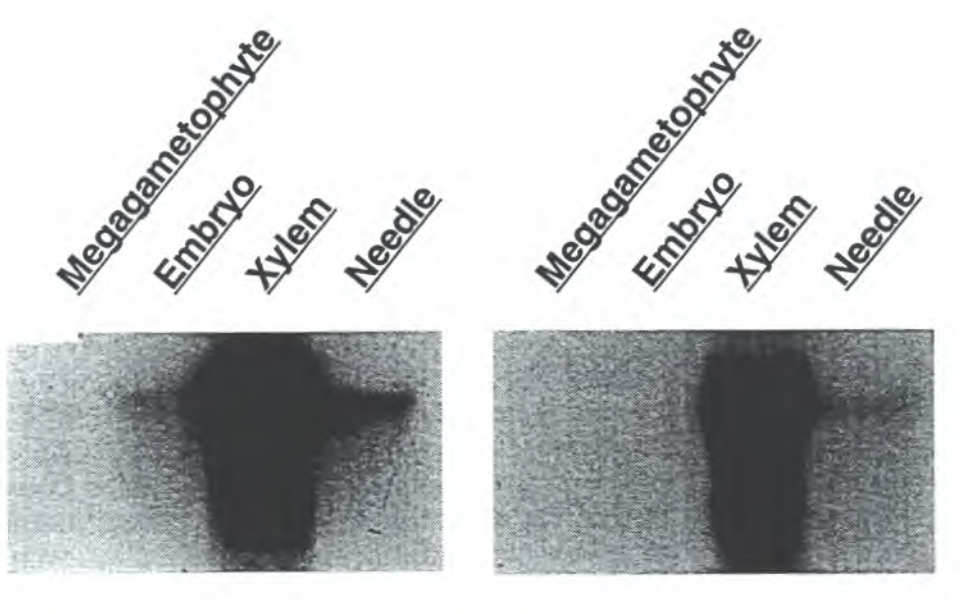

Figure 1. Northern blot hybridizations to examine abundance of the PtX3H6 and PtX14A9 transcripts in various tissues.

The relative abundances of PtX3H6, PtX14A9, phenylalanine ammonia lyase (PAL), and cinnamyl alcohol dehydrogenase (CAD) transcripts were compared by hybridizing a radioactive xylem cDNA probe to cDNA clones of the genes. PAL and CAD are both enzymes involved in lignin biosynthesis. Both enzymes have been purified from differentiating xylem of loblolly pine (Whetten and Sederoff, 1992; O'Malley et al., 1992). PtX3H6 transcripts were found to be more abundant than those of PtX14A9 and both were much more abundant than those of PAL and CAD.

The nucleotide sequences of PtX3H6 and PtX14A9 were determined and the amino acid sequences inferred. PtX3H6 is rich in prolines or hydroxyprolines and contains repeats similar to those found in proline-rich cell wall proteins (PRPs). The amino acid composition of PtX3H6 however, is much more similar to the arabinogalactan-proteins (AGPs). PtX3H6 and AGPs are rich in prolines or hydroxyprolines, alanine, threonine and serine and have few tyrosines, an amino acid usually found in cell wall proteins. The PtX14A9 sequence is not particularly similar to those previously published but does contain the sequence Ala-Pro-Ala-Pro-Ser-Pro-Ala-Ser near the amino terminus. This sequence has been found at the amino terminus of at least three AGPs. Both PtX3H6 and PtX14A9 appear to have signal peptides at the amino termini and hydrophyllic helixes at the carboxy termini. The hydrophobicity plots are very similar to that of a recently reported pear AGP (Chen et al., 1994). We are currently working under the assumption that both PtX3H6 and PtX14A9 encode AGPs.

Arabinogalactan-proteins

AGPs are a class of proteins recently receiving increased attention due to their utility in studies of angiosperm development and evolution and their potential roles in developmental processes. The properties of arabinogalactans and AGPs have been reviewed by Clarke et al. (1979) and Fincher et al. (1983). AGPs are a widely distributed class of proteoglycans and glycoproteins. Generally, only 2-10% of the weight of an AGP is made up of protein. Ninety percent or more of the molecule is carbohydrate, including galactose, arabinose, uronic acids, glucose, rhamnose, mannose, and glucosamine. The majority of the carbohydrates are attached to the protein backbone by hydroxyproline linkages and the remaining are likely to be attached to serines and threonines (van Hoist and Klis, 1981). Tissue-specific expression of AGPs has been observed and more than one AGP can be found within a tissue. AGPs are found in the extracellular matrix, associated with the plasma membrane, inserted into the cell wall, or secreted into the medium of cell cultures. They have been found in almost all tissues of higher plants and have been detected in every taxonomic group tested, including in the wood of some angiosperms and gymnosperms (Showalter and Varner, 1989). A comparison of carbohydrate-protein complexes from cell walls and cytoplasm of Siberian larch xylem revealed AGPs to be associated with the primary cell walls (Antonova and Stasova, 1990). AGPs are also abundant in xylem of loblolly pine (approximately 0.1% of fresh weight) and are extractable (R. Whetten, personal communication). AGPs have also been extracted from Douglas-fir and loblolly pine tissues collected at different stages of development (Bobalek and Johnson, 1983).

The functions of AGPs are not known but many potential roles have been postulated. Antibodies have been used to show that AGP carbohydrate epitopes display developmentally regulated patterns of cell surface expression directly reflecting cell fate in both root and floral meristems of angiosperms (Knox et al., 1991; Pennell and Roberts, 1990; Pennell et al., 1991). It has been suggested that AGPs may not only be markers of development, but may have a role in cell-cell interactions or cellular signalling during morphogenetic processes. There is also a theory that AGPs may be involved in programmed cell death. We are interested in determining the roles of AGPs in xylogenesis and the development of wood and in using AGPs to gain a greater

understanding of the molecular events leading to the differentiation of xylem. Due to recent developments in pine transformation methodologies and the isolation of pine AGP clones, we are now in an excellent position to pursue these interests. Experiments in progress to examine AGP functions in xylem are described in the results and discussion section.

Promoter analyses

In order to obtain the promoter elements thought to be involved in controlling the expression of PtX3H6 and PtX14A9, a genomic library was constructed and screened. Genomic clones of each were isolated and characterized. Primer extension analyses were used to identify the transcription start sites. Approximately 953 bp of PtX3H6 5' flanking sequence and 750 bp of PtX14A9 5' flanking sequence were determined. Both promoters contain regions of high Air content. The PtX3H6 promoter contains a 467 bp region containing 86.5% AfT. These promoter segments may contain scaffold attachment regions (SARs). The PtX3H6 promoer contains a pair of 63 bp direct repeats with four nucleotide differences and a pair of 36 bp repeats with two nucleotide differences and a one base insertion. Other smaller repeats are also found. Several sequences of 8 to 10 bp are found in both promoters. It is not known if this is due to chance or if they are conserved functional elements. The sequence CTGCATG is found in both promoters and in the promoters of two vascular-specific genes from bean (Keller and Baumgartner, 1991). In the GRP1.8 promoter, this sequence has been shown to be part of a negative regulatory element required for vascular-specific expression. It seems unlikely that this element would appear by chance in two vascular-specific bean promoters and two xylem-specific pine promoters. It is likely this element is involved in the xylem-specific expression observed but it is also likely that other elements are also involved. Experiments used to test the isolated promoters are described below.

RESULTS AND DISCUSSION

Determining the roles of AGPs in xylem development

It may be possible to gain insight into the function of AGPs in differentiating xylem by reducing their levels in genetically engineered pine plantlets and comparing the wood produced to that of normal control plants. Antisense constructs are being made by inserting the cDNA clones for PtX3H6 and PtX14A9 in an orientation opposite to normal relative to a promoter (Figure 2). We have developed two cassettes useful for the production of antisense constructs. Each contains a gene for kanamycin resistance *(npt* II) under the control of the nopaline synthase promoter (NOS) for selecting transformed cells and an antisense cassette containing the cauliflower mosaic virus 35S promoter with enhancer elements followed by the restriction sites Xho I, Kpn I, and Sacl and a NOS terminator. The order of the restriction sites relative to the promoter allows any cDNA clone isolated from a Lambda Zap library (Stratagene) to be inserted in an antisense orientation. Both constructs also contain the 35S promoter driving a reporter gene. One contains the gene for J3-glucuronidase (GUS) and the other contains a modified green fluorescent protein gene (mGFP) from jellyfish. The above pieces are located between *Agrobacterium tumefaciens* right and left borders in plasmids derived from pBin19.

We may also produce constructs containing the gene in a sense orientation. Introduction of extra copies of a gene has frequently been shown to reduce rather than enhance expression by way of cosuppression. The plasmids containing the constructs will be introduced into *Agrobacterium* strains known to be virulent in pines such as EHA105 (E. Hood) via electroporation. The *Agrobacterium* containing the plasmids will be used to inoculate loblolly pine apical shoot meristems using a technique developed by Jean Gould at Texas A & M University and transformed plants will be regenerated.

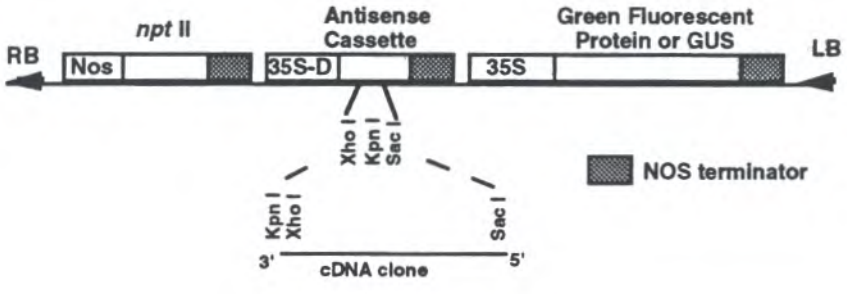

Figure 2. Diagram of the constructs being used to produce loblolly pine plantlets with reduced levels of xylem AG Ps. The segment shown is within pBin19 derivatives.

We also plan to characterize the AGPs found in differentiating loblolly pine xylem. AGPs have been shown to be abundant in this tissue and are extractable (R. Whetten, personal communication). AGPs will be purified from differentiating xylem, deglycosylated, and the numbers, sizes, and tissue-specificities determined. AGPs that appear to be the sizes expected if encoded by PtX3H6 or PtX14A9 and are xylem-specific will be further purified and the amino acid compositions will be determined. Those that remain candidate proteins will be partially sequenced to determine if they are indeed encoded by PtX3H6 or PtX14A9.

Promoter analyses using microprojectile bombardments

In order to examine the regulation of PtX3H6 and PtX14A9, their promoters were fused to the gene for p-glucuronidase in the vector pRT99-gus (Topfer, 1988) and in Bluescript KS (Stratagene). PRT99-gus contains the *gusA* gene and a gene for kanamycin resistance *(nptII),* both under control of the CaMV35S promoter. The CaMV35S promoter preceding the *gusA* gene is flanked by restriction sites making it possible to remove the promoter and insert one of interest. An approximately 3.2 kb promoter fragment from genomic subclone G3H(5.7) and a 750 bp fragment from subclone G14A9(5.1) were cloned into PRT99-gus. A promoterless *gusA* construct was made in pRT99-gus by removing the CaMV35S promoter.

A method for examining transient gene expression in differentiating xylem of pine using microprojectile bombardment has been previously reported (Loopstra et al. 1992). In this study, loblolly pine xylem (stem sections), embryos, and megagametophytes were bombarded with constructs containing putative xylem-specific promoters to look for evidence of tissue-specific expression. Bombardments using the promoterless-gusA constructs failed to produce any staining cells after incubation with 5-bromo-4-chloro-3-indolyl-p-D-glucuronic acid (X-gluc). Previous

bombardments using pUC 18 plasmid DNA or a *gusA* fusion containing an inverted Em promoter from wheat (Marcotte et al. 1988) also resulted in no observable stained cells (Loopstra et al. 1992). Bombardments of xylem samples with the 35S, PtX3H6, and PtX14A9 constructs resulted in cells exhibiting expression of the *gusA* gene, demonstrating that the isolated promoters are functional. The number of staining foci resulting from bombardments with the CaMV35S and PtXH6 constructs were similar with averages of 66.5 and 55.1 staining foci per stem section respectively. Bombardments done with PtX14A9 constructs resulted in far fewer stained foci with an average of 1.0 per stem section.

Previous bombardments and hand sections of bombarded material had demonstrated expression in at least three cell types; tracheids, ray parenchyma, and axial parenchyma associated with resin canals. In xylem bombarded with the 35S and PtX3H6 constructs, 20.4% and 18.9% of the staining cells were tracheids, similar to the 20% previously reported (Loopstra et al., 1992). Only 36 stained foci have been observed following bombardment with the PtX14A9 construct and all have been tracheids. It is not known if this is due to the sequences found in the promoters or if expression from the PtX14A9 promoter is just so low it is only noticable in the larger tracheids.

To determine if the promoter-gusA fusions are expressed in a xylem-specific manner following bombardment, megagametophytes and embryos were also bombarded. The 35S, PtX3H6, and PtX14A9 constructs all resulted in staining foci when bombarded into megagametophytes and embryos. To determine if the expression observed was transient due to wounding, bombarded megagametophytes and embryos were maintained on non-selective media before staining. Blue staining was observed in tissues bombarded with all three constructs for at least 3 weeks after bombardment, arguing against a transient wound response.

Expression of pine promoters in transgenic tobacco

Transgenic tobacco plants were produced to determine if the PtX3H6 and PtX14A9 promoters would direct expression of the *gusA* gene in a non-woody angiosperm and to determine if any expression observed is xylem-specific. Constructs containing the pine promoters, a CaMV35S promoter, or no promoter along with the reporter gene *gusA* were made in pBin19. The plasmids were moved into *Agrobacterium tumefaciens* strain LBA4404 using triparental matings. Transgenic *Nicotiana tabacum* plants were produced using leaf disc transformation and plant regeneration methods. Stem and leaf samples were analyzed by histochemical staining with X-gluc and by fluorescence produced by homogenization and incubation with 4-methylumbelliferyl-3-D-glucuronide (MUG) in microtiter plate assays. Wells containing buffer and MUG only were included as well as samples from nontransformed plants.

The PtX3H6 promoter was found to function in at least 93% of the plants surviving selection on kanamycin. Expression was not however, restricted to xylem tissue. MUG assays using leaf tissue from plants transformed with the PtX3H6 construct resulted in lower levels of fluorescence in microtiter plates than assays of plants containing the 35S construct. However, the percent of plants giving a positive

MUG result were similar. The relative levels of fluorescence in stems vs leaves were similar in PtX3H6 and 35S plants. A positive MUG result was obtained in only three of the 54 plants assayed that were transformed with the PtX14A9 construct. In two of the cases, the fluorescence was low. In the third, it was quite high, possibly due to the position where the DNA was integrated. No fluorescence was seen in samples taken from non-transformed plants or when buffer alone was incubated with MUG. The results of the MUG assays on leaf tissue are given in Table 1.

Table 1. Analyses of loblolly pine xylem-specific promoters in transgenic tobacco.

Promoter	Plants harvested [a]	Plants assayed [b]	Positive plants
CaMV35S	206	38	35 (92%)
Promoterless	127	**26**	0 (0%)
14A9	229	54	3 (5.6%)
3H6	180	57	53 (93%)

[a] Number of plants harvested at end of experiment.
[b] Number of plants assayed for GUS activity in leaf tissue using a MUG assay.

Expression of pine promoters in transgenic poplar and spruce

The 3.2 kb PtX3H6 promoter, the 750 bp PtX14A9 promoter, an inverted PtX14A9 promoter, and the CaMV35S promoter were cloned in front of the *gusA, lux,* and *luc* genes in the *Agrobacterium* binary vectors ppCV812, ppCV813, and ppCV814. The plasmid was moved into *Agrobacterium* strain 8145 using electroporation. Poplar stem sections were co-cultivated with the *Agrobacterium* and plantlets regenerated in the laboratory of Dr. 011e Olsson at Swedish Agricultural University in Umea Sweden. The patterns of GUS expression driven by the promoters were determined by staining with X-gluc. In some regenerated plants, expression of the *gusA* gene was greatest in differentiating xylem. However, in various plants, many patterns of expression were observed.

Three lines of transgenic spruce were regenerated by Dr. David Ellis (U. of Wisc.) using the PtX3H6-gusA construct in pRT99 and microprojectile bombardment of white spruce somatic embryos as described by Ellis et al., (1993). All three lines appear to have very different patterns of blue staining.

Since we have not observed xylem-specific expression of the gusA gene under the control of the PtX3H6 or PtX14A9 promoters using four different systems, it is likely that our constructs are lacking an element required for xylem-specificity. It is possible that important elements are located further upstream or that they are located within the transcribed or 3' portions of the genes. Experiments are now underway to make and test constructs containing a reporter gene and various putative regulatory regions. Initial transformations are likely to be in tobacco or poplar due to the relative ease of transformation. Once isolated, the elements responsible for xylem-specificity and those responsible for the high levels of expression observed will be valuable for genetic engineering of wood properties.

LITERATURE CITED

Antonova G.F. and V.V. Stasova. 1990. Location of arabinogalactan proteins in xylem of Siberian larch. Soviet Plant Physiology 37 (4,1): 547-554.

Bao W., D.M. O'Malley, and R.R. Sederoff. 1992. Wood contains a cell-wall structural protein. Proc. Natl. Acad. Sci. USA 89: 6604-6608.

Bao W., D.M. O'Malley, R. Whetten and R.R. Sederoff. 1993. A laccase associated with lignification in loblolly pine xylem. Science 260: 672-674.

Bobalek, J.F. and M.A. Johnson. 1983. Arabinogalactan-proteins from Douglas fir and loblolly pine. Phytochemistry 22 (6): 1500-1503.

Chen, C.-G., Z.-Y. Pu, R.L. Moritz, R.J. Simpson, A. Bacic, A.E. Clarke, and S.-L. Mau. 1994. Molecular cloning of a gene encoding an arabinogalactan-protein from pear (Pyrus cornunis) cell suspension culture. PNAS 91 (22): 10305-10309.

Clarke A.E., R.L. Anderson, and B.A. Stone. 1979. Form and function of arabinogalactans and arabinogalactan-proteins. Phytochemistry 18: 521-540.

Ellis D.D., D.E. McDabe, S. McInnis, R. Ramachandran, D.R. Russell, K.M. Wallace, B.J. Martinell, D.R. Roberts, K.F. Raffa and B.H. McCown. 1993. Stable transformation of *Picea glauca* by particle acceleration. Bio/Technology 11: 84-89.

Fincher G.B., B.A. Stone, and A.E. Clarke. 1983. Arabinogalactan-proteins: structure, biosynthesis and function. Ann. Rev. Plant Physiol. 34: 47-70.

Keller, B. and Baumgartner, C. (1991) Vascular-specific expression of the bean GRP 1.8 gene is negatively regulated. Plant Cell 3: 1051-1061.

Knox J.P., P.J. Linstead, J. Peart, C. Cooper and K. Roberts. 1991. Developmentally regulated epitopes of cell surface arabinogalactan proteins and their relation to root tissue pattern formation. Plant J 1(3): 317-326.

Loopstra, C.A., Weissinger, A.K., and Sederoff, R.R. (1992) Transient gene expression in differentiating pine wood using microprojectile bombardment. Can. J. For. Res. 22: 993-996.

Loopstra C. and R.R. Sederoff. 1995. Xylem-specific gene expression in loblolly pine. Plant Mol Biol:27(2):277-291.

Marcotte, W.R. Jr., Bayley, C.C., and R.S. Quatrano. (1988) Regulation of a wheat promoter by abscisic acid in rice protoplasts. Nature 335: 454-457.

O'Malley, D.M., S. Porter, and R.R. Sederoff. 1992. Purificiation, characterization and cloning of cinnamyl alcohol dehydrogenase in loblolly pine (Pinus taeda L.). Plant Physiol. 98:1364-1371.

Pennell R.I. and K. Roberts. 1990. Sexual development in the pea is presaged by altered expression of arabinogalactan protein. Nature 344: 547-549.

Pennell, R.I., L. Janniche, P. Kjellbom, G.N. Scofield, J.M. Peart, and K. Roberts. 1991. Developmental regulation of a plasma membrane arabinogalactan proteins epitope in oilseed rape flowers. Plant Cell 3: 1317-1326.

Showalter A.M. and J.E. Varner. 1989. Plant hydroxyproline-rich glycoproteins. In: Marcus A. (ed) The Biochemistry of Plants 15: 485-519. Academic Press, NY.

TOpfer, R., Schell, J. and Steinbiss, H.-H. (1988) Versatile cloning vectors for transient gene expression and direct gene transfer in plant cells. Nucleic Acids Res. 16: 8725.

van Ho1st G.-J. and FM Klis. 1981. Hydroxyproline glycosides in secretory arabinogalactan-protein of *Phaseolus vulgaris* L. Plant Physiol. 68: 979-980.

Whetten R. and R.R. Sederoff. 1992. Phenylalanine ammonia-lyase in loblolly pine: purification of the enzyme and isolation of complementary DNA clones. Plant Physiol. 98: 80-386.