Functional Analysis of the *PgCesA3* White Spruce Cellulose Synthase Gene Promoter in Secondary Xylem

Juliana Stival Sena[1,2], Denis Lachance[2], Isabelle Duval[2], Thi Thuy An Nguyen[2], Don Stewart[2], John Mackay[3] and Armand Séguin[2]*

[1] Department of Wood and Forest Sciences, Université Laval, Quebec City, QC, Canada, [2] Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Quebec City, QC, Canada, [3] Department of Plant Sciences, University of Oxford, Oxford, United Kingdom

Cellulose is an essential structural component of the plant cell wall. Its biosynthesis involves genes encoding cellulose synthase enzymes and a complex transcriptional regulatory network. Three cellulose synthases have been identified in conifers as being potentially involved in secondary cell wall biosynthesis because of their preferential expression in xylem tissues; however, no direct functional association has been made to date. In the present work, we characterized the white spruce [*Picea glauca* (Moench) Voss] cellulose synthase *PgCesA3* gene and 5′ regulatory elements. Phylogenetic analysis showed that *PgCesA1-3* genes grouped with secondary cell wall-associated Arabidopsis cellulose synthase genes, such as *AtCesA8*, *AtCesA4*, and *AtCesA7*. We produced transgenic spruce expressing the GUS reporter gene driven by the *PgCesA3* promoter. We observed blue staining in differentiating xylem cells from stem and roots, and in foliar guard cells indicating that *PgCesA3* is clearly involved in secondary cell wall biosynthesis. The promoter region sequence of *PgCesA3* contained several putative MYB cis-regulatory elements including AC-I like motifs and secondary wall MYB-responsive element (SMRE); however, it lacked SMRE4, 7 and 8 that correspond to the sequences of AC-I, II, and III. Based on these findings and results of previous transient trans-activation assays that identified interactions between the *PgCesA3* promoter and different MYB transcription factors, we performed electrophoretic mobility shift assays with MYB recombinant proteins and cis-regulatory elements present in the *PgCesA3* promoter. We found that PgMYB12 bound to a canonical AC-I element identified in the *Pinus taeda* PAL promoter and two AC-I like elements. We hypothesized that the PgMYB12 could regulate *PgCesA3* in roots based on previous expression results. This functional study of *PgCesA3* sequences and promoter opens the door for future studies on the interaction between PgMYBs and the *PgCesA3* regulatory elements.

Keywords: conifer, cellulose synthase phylogeny, plant genetic transformation, functional analysis, electrophoretic mobility shift assay, MYB regulatory binding sites

INTRODUCTION

Cellulose is a fundamental constituent of the plant cell wall. It represents half of the mass of wood and is a crucial resource for the production of pulp, biomaterials and biofuels (Keijsers et al., 2013). Plants and trees accumulate large amounts of crystalline cellulose during the formation of secondary cell walls, which provide mechanical strength and are the main hallmark of secondary xylem and woody tissues. A protein complex formed of cellulose synthase enzymes that are specific to the secondary cell wall is responsible for building the cellulose microfibrils that make up the extracellular matrix (Zhong and Ye, 2015). In conifers, three cellulose synthases have been identified that are putatively involved in secondary cell wall formation but no direct functional association has been made to date. The regulation of cellulose synthase genes remains only partially understood in plants and trees in spite of their central role in development.

Several of the secondary cell wall cellulose synthases have been identified in angiosperm plants and their molecular function has been the subject of several reports. In *Arabidopsis thaliana*, there are ten celluloses synthase genes, three of which make up the cellulose synthase complex for secondary wall deposition (Endler and Persson, 2011). A study using specific antibodies to the three catalytic subunits of the cellulose synthase complex (CSC) and a green fluorescent protein fusion indicated that all three subunits are required to assemble a functional complex in the developing xylem in Arabidopsis (Gardiner et al., 2003). The CSC is composed of a hexamer of catalytically active CesA trimmers, with each of CesA in equimolar amounts (Hill et al., 2014). Mutations in any of these three genes caused a reduction of cellulose content and secondary wall thickening, and a collapsed xylem phenotype (Taylor et al., 2003; Zhong and Ye, 2015). The same phenotype was observed in *CesA* mutants from other vascular plants such as rice and poplar (Tanaka et al., 2003; Joshi et al., 2011; Handakumbura et al., 2013).

In plants, other than model species there is a growing interest in understanding the role of cellulose synthases. Recent studies have deciphered the expression patterns of cellulose synthases in tobacco (Xu and Kong, 2017), flax (Galinousky et al., 2017), and soybean (Nawaz et al., 2017). In conifers, *CesA* genes have been identified in large-scale gene expression studies (Raherison et al., 2012; Jokipii-Lukkari et al., 2017). Three cellulose synthase genes in white spruce and loblolly pine (*CesA1*, *CesA2*, *CesA3*) were identified as being potentially involved in secondary cell wall formation based on their high expression in secondary xylem and sequence similarity to angiosperm *CesA* genes (Nairn and Haselkorn, 2005; Raherison et al., 2012). However, direct functional association with secondary cell wall formation is lacking and their precise expression patterns remain to be tested through *in situ* hybridization, promoter/reporter studies or other techniques.

A network of transcription factors regulates secondary cell wall biosynthesis in plants. Closely related NAC domain transcription factors activate several downstream transcription factor genes that directly trigger secondary cell wall biosynthesis genes (Zhong et al., 2010). Many of the downstream transcriptional regulators are R2R3-MYB transcription factors (Zhong et al., 2010), including AtMYB46, AtMYB52, AtMYB54, AtMYB58, AtMYB63, AtMYB83, AtMYB85, and AtMYB103 (Zhong and Ye, 2014). In conifers, molecular switches that regulate secondary cell wall were identified. For example, the *Pinus pinaster* transcription factor *PpNAC1*, a functional ortholog of the Arabidopsis *SND1* and *NST1*, activated the R2R3-MYB transcription factor *PpMYB4*, which activated *PpMYB8* (Pascual et al., 2018). In white spruce, Duval et al. (2014) identified the *PgNAC7* that could play a master regulatory role and activated *PgMYB8*, which activated *PgMYB1*. They also have reported a positive interaction of several MYB transcription factors with a white spruce *CesA3* promoter. The transcription factors that trans-activated the spruce *CesA3* promoter included *Picea glauca* PgMYB11, PgMYB12, and PgMYB22, and *Pinus taeda* PtMYB1 and PtMYB8 (Duval et al., 2014). The *Pt/PgMYB8* genes were implicated in the regulation of the secondary cell wall formation in conifers and its closest homolog in Arabidopsis is *AtMYB46* (Bomal et al., 2008). Transgenic spruces that overexpressed loblolly pine *PtMYB8* presented ectopic secondary cell wall deposition and an upregulation of CesA genes and other genes associated with cell organization and biogenesis (Bomal et al., 2008).

MYB transcription factors control the expression of target genes by interacting with specific DNA sequences in their promoter regions (Weirauch et al., 2014). Studies have shown that MYB transcription factors bind to cis-regulatory elements that are enriched in adenosine and cytosine residues, such as the AC-elements ACC[A/T]A[A/C][T/C] and ACC[A/T][A/C/T][A/C/T], and to the MYB-core sequences [C/T]NGTT[G/A] (Prouse and Campbell, 2012). Other transcription factors as AtMYB88 and AtMYB124 recognize a different consensus sequence, such as the [A/T/G][A/T/G]C[C/G][C/G] (Xie et al., 2010).

In this study, we investigated the developmental and tissue regulation of a white spruce *CesA3* gene to gain insight into its role in secondary xylem formation. Our specific objectives were to: (1) analyze the phylogeny of *CesA* gene sequences related to secondary cell wall formation in vascular development in conifers and angiosperms, (2) isolate a white spruce *CesA3* promoter and identify putative MYB regulatory binding sites, (3) analyze the tissue expression pattern of a *CesA3* promoter and GUS reporter gene fusion in transgenic white spruce plants, and (4) study promoter – transcription factor interactions by electrophoretic mobility shift assays (EMSA) of MYB recombinant proteins and cis-regulatory elements in the *PgCesA3* promoter.

MATERIALS AND METHODS

Phylogenetic Tree

We searched for conifer CesA proteins that were related to secondary cell wall biosynthesis in the non-redundant protein NCBI database (nr) using Blastp, with an e-value threshold of $1e^{-20}$, and white spruce CesA1, CesA2, and CesA3 sequences as a query. We retained conifer CesA's with at least a 75% amino acid sequence similarity and 90% sequence coverage with the white spruce CesA1, CesA2, and CesA3 proteins.

Our phylogenetic analysis included only three white spruce (*P. glauca*) cellulose synthases, PgCesA1, PgCesA2, and PgCesA3 because the other CesA protein sequences were incomplete in the databases that we searched.

Angiosperm CesA proteins linked to primary and secondary cell wall formation were identified based on phmmer (*e*-value < 0.01)[1], which provides more accurate detection of remote homologs than BLAST. We utilized the white spruce CesA1, CesA2, and CesA3 to search the UniProt database (The UniProt Consortium, 2015) for Arabidopsis (*Arabidopsis thaliana*), maize (*Zea mays*), rice (*Oryza sativa*), eucalyptus (*Eucalyptus grandis*), shinybark birch (*Betula luminifera*), poplar (*Populus trichocarpa*), and coniferales CesA proteins. We also verified in the literature that all of the published CesA proteins from these plant species were included in our sequence search results. See **Supplementary Table S1** for accession numbers.

We aligned sequences using MAFFT version 7.0 and FTT-NS-I (iterative refinement method; 1000 iterations) strategies (Katoh and Standley, 2013). We constructed the phylogenetic tree using FastTree 2 (version 2.1.6, options: -boot 1000 -wag -gamma (Price et al., 2010). The tree was visualized with FigTree v1.4.2[2]. Color modifications and group indications were made with Inkscape[3].

Gene Structure

To determine gene structure (introns and exons), we utilized the *CesA3* cDNA sequence (BT116976) and the genomic sequence containing the *CesA3* gene (BAC contig KC860240) published by Sena et al. (2014). The gene sequence was mapped onto the BAC contig sequence using est2genome incorporated in the annotation software MAKER (Cantarel et al., 2008). Gene structures from Arabidopsis, maize, poplar, and eucalyptus were recovered from the databases of TAIR 10[4] and Phytozome (gene annotation of assembly v3 of *P. trichocarpa* and v2 of *E. grandis*; Tuskan et al., 2006; Goodstein et al., 2012; Myburg et al., 2014) and the Maize Genome Sequencing Project[5]. We utilized Inkscape[6] to build the figure.

Promoter Analysis

The 5′-upstream sequence of the white spruce *CesA3* gene was obtained by genome walking (Genbank, KF824520) as described by Duval et al. (2014). The transcription start site was determined by 5′RACE as described by Bedon et al. (2009). The primers used for the identification and cloning of the *CesA3* gene 5′ genomic sequence was provided by Duval et al. (2014).

Putative regulatory MYB elements in the promoter sequence were identified using the Plant Transcription Factor Database, V4.0, and FIMO, a motif search tool (Grant et al., 2011). We downloaded the MYB transcription factor binding motifs from Arabidopsis and then screened the promoters of the *P. glauca*

CesA3 gene and its orthologs *AtCesA7* from Arabidopsis and *PttiCesA7-A* from *P. trichocarpa*. The promoter and 5′ UTR regions from *P. trichocarpa* and Arabidopsis were recovered from Phytozome 12[7] (Potri.006G181900; 2000 bp upstream of the transcription start site) and the Eukaryotic Promoter Database[8] (promoter ID AT5G17420_1; 2100 bp upstream of the transcription start site), respectively.

We searched for MYB TF motifs by using two different match threshold *p*-values, 0.001. and 0.0001, and compared the outcomes. We selected the best hit of each potential MYB binding motif, first with a threshold *p*-value of 0.001 and then 0.0001, for each species. With a threshold *p*-value of 0.001, we identified 66 MYB transcription factors that could bind the promoter region of the three species. As some hits were poorly conserved among the three species, we decided to be more stringent in our analysis, using a threshold *p*-value of 0.0001. Potential MYB binding sites in the *PgCesA3* promoter region were filtered and hits with a score superior to 31.9 were retained (see **Figure 3**). We aligned the promoter region sequence of *PgCesA3* and its orthologs in Arabidopsis and poplar and compared the best match of each TF binding motif between the three species.

Transgenic Spruce Expressing the Ces3A Promoter-GUS Gene Fusion

A Pro*PgCesA3*::GUS (β-glucuronidase) fusion was created by inserting the putative promoter region into vector pMJM (Levée et al., 2009) as described in Bedon et al. (2009), digesting with Sbf1 and ligating it into the binary vector pCAMBIA2300[9]. The binary vector was subsequently transformed into *Agrobacterium tumefaciens* strain C58pMP90 (Koncz and Schell, 1986). White spruce embryonal masses of line PG653 were transformed with Agrobacterium and selected with kanamycin as described in Klimaszewska et al. (2001). Following antibiotic selection to identify transgenic lines we identified 10 lines showing good embryonal masses growth. Six lines were selected randomly for plant production for further evaluation of GUS reporter gene expression. Plant production from individual transformed lines was as described in Klimaszewska et al. (2005) with the following modification: 3-month old somatic seedlings were produced by subculturing monthly and were acclimated in a growth chamber by removing the parafilm from the germination plates for the final 2–4 weeks of culture. Small transformed seedlings were evaluated for GUS reporter gene expression and all showed vascular specific expression. Three lines were selected for plant production and transferred to the greenhouse. Occasional hand misting of the plants was required for the first week after planting of the seedlings. Control lines (empty vector) were also selected for somatic embryo maturation and seedling production.

Somatic seedlings were planted in 3:1:1 mix of peat, vermiculite and perlite and grown in a greenhouse under a 16/8 h photoperiod with day/night temperatures of 24/20°C. At the end of the first growing season, when the plants had set a terminal bud, they were subjected to an accelerated growth

---

[1] https://www.ebi.ac.uk

[2] http://tree.bio.ed.ac.uk

[3] https://inkscape.org

[4] http://arabidopsis.org

[5] http://www.maizesequence.org

[6] https://inkscape.org

[7] https://phytozome.jgi.doe.gov

[8] http://epd.vital-it.ch

[9] https://cambia.org/

schedule. Plants were hardened off by gradually reducing light and temperature conditions over a seven-week period to a final photoperiod of 8 h and day/night temperatures of 5°C. They were then subjected to a cold treatment by being put in a cold room at 2°C for six weeks after which they were returned to the greenhouse. Over an acclimatization period of two weeks, light and temperature were gradually increased from 8 h and 10°C to the original growing conditions. This accelerated growth schedule permits synchronous bud flush in all plants in the subsequent growth cycle.

Histochemical localization of GUS expression was evaluated in seedling tissues during both the first (data not shown) and second growth seasons. Seedling tissue was harvested from the three selected lines and subjected to GUS histochemical staining.

Histochemical Staining of Plant Tissues

Expression analysis of the Pro*PgCesA3*::GUS (β-glucuronidase) fusion and sample preparation were adapted from Hawkins et al. (1997) as described by Bedon et al. (2009). Samples of roots, needles, stem just below the apical meristem, and first and second year stem segments were cross cut to 5–50 mm in length depending on the organ and pre-treated 30 min in cold 90% acetone (Hemerly et al., 1993). Samples were rinsed twice with 100 mM potassium phosphate buffer (pH 8), and incubated in 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid, with potassium ferricyanide and potassium ferrocyanide in sodium phosphate buffer (Vigneault et al., 2007) at 37°C in the dark under vacuum for 6–12 h until a blue stain was clearly visible. All tissue samples, except for stem, were then supported in 6% low melting temperature agar before cross-sectioning (30–50µm thickness) using a vibratome (Vibratome 1500, Vibratome Co., St. Louis, MO, United States). Sections were counterstained with phloroglucinol-HCl to visualize lignin deposition (Sibout et al., 2005). All tested lines showed the same patterns of GUS reporter tissue-specific expression.

Expression and Purification of Recombinant MYB Proteins

Full-length white spruce MYB coding sequences (see accession numbers in **Supplementary Table S2**) were amplified by PCR with gene specific primers. The amplicons were cloned into the NdeI and XhoI restriction sites (New England Biolabs, Beverly, MA, United States) of the pET300/NT-DESTvector (ThermoFisher Scientific). The construct was confirmed by DNA sequencing and transformed into *E. coli* Rosetta[TM] 2 (DE3) competent cells (Millipore Sigma). Bacterial growth and protein induction were performed at 37°C in LB broth containing kanamycin (30 µg/mL) and chloramphenicol (100 µg/mL). For protein induction, a bacterial culture at an O.D. 600 nm of 0.6 was used. Isopropyl-beta-D-thiogalacto-pyranoside (IPTG) was added to a concentration of 0.1–1.0 mM and the culture was incubated for 2–5 h for optimal production of each spruce MYB protein (see **Supplementary Table S3**). The cells were harvested by centrifugation (9000 × *g*, 10 min, 4°C), and pellets were resuspended in cold lysis buffer [20 mM Tris–HCl, pH 7.9, 100 mM NaCl, 5 mM imidazole plus one tablet of cOmplete,

Mini, EDTA-free Protease Inhibitor Cocktail (Millipore Sigma)]. Recombinant proteins were purified by affinity chromatography using the HisBind Purification Kit (Millipore Sigma). Where protein recovery from inclusion bodies was necessary (PgMYB12 and 22), the inclusion bodies were denatured and solubilized in 5 ml denaturation buffer [8 M Urea, 20 mM Tris–HCl, pH 7.9, 500 mM NaCl, 5 mM imidazole plus one tablet of cOmplete[TM], Mini EDTA-free Protease Inhibitor cocktail (Millipore Sigma)]. The cells were lysed using a Vibracell VC-50 sonicator and probe with a 3 mm micro tip. They were lysed on ice using 6 × 10 second bursts to prevent overheating of the samples. After sonication, the cells were incubated at room temperature for 1 h, with gentle rotation and then centrifuged at 10,000 × *g*, 30 min, 4°C. The supernatant was recovered and the recombinant proteins were purified by affinity chromatography using the His•Bind® Purification Kit (Millipore Sigma). The purified proteins were renatured by dialysis in renaturation buffer (20 mM Tris, 50 mM NaCl, 20 mM imidazole, pH 8.0) followed by dialysis buffer (20 mM Tris, 50 mM NaCl, pH 8.0) using Spectra-Por 1 dialysis membrane MWCO 6-8000 (Spectrum Labs). Protein purification was verified by SDS-PAGE analysis and Coomassie blue staining. The integrity of the purified recombinant proteins was evaluated by western-blot analyses using a mouse anti-His antibody (GE-Healthcare; **Supplementary Figures S1–S5**).

Electrophoretic Mobility Shift Assay

Purified white spruce PgMYB5, PgMYB12, PgMYB13, and PgMYB22 recombinant proteins were tested for their ability to bind to oligonucleotides from the white spruce *CesA3* promoter using EMSA. We chose to test the sequence from region 32 (**Supplementary Figure S6**) of the *PgCesA3* promoter which contains two AC- like motifs in tandem. As the positive control, we used the canonical AC-I sequence from the *PAL* promoter gene (loblolly pine, accession number U39792.1; Patzlaff et al., 2003b). The 30bp fragments of the *CesA3* promoter used as probes were AC-I like, AC-I, mAC-I like and a fragment without a cis-element (**Supplementary Table S4**). The double-stranded oligonucleotides were end-labeled with γ-32P. 375 ng of recombinant protein and 100 ng of poly(dI-dC; used as a non-specific competitor) were incubated at room temperature for 30 min along with 0.2 ng labelled probe in a 20µl reaction. The incubation was done in binding buffer with a final concentration of 20 mM Tris pH 8.0, 10 mM NaCl, 2 mM EDTA, 2 mM dithiothreitol (DTT), and 10% glycerol (v/v). Competition reactions used unlabeled double-stranded probes in a 20 or 200-fold excess. Electrophoreses of the DNA–protein complexes were performed on a 6% native polyacrylamide gel at constant current of 35 mA (1 × Tris–glycine buffer, 4°C). Migration of radiolabeled probes and complexes were detected on Kodak Biomax XAR film (Bomal et al., 2014).

RESULTS

Three white spruce cellulose synthases (*PgCesA1*, *PgCesA2*, *PgCesA3*) that are preferentially expressed in secondary xylem were identified as candidate genes during the formation

of secondary cell walls in vascular development (Raherison et al., 2012). Here, we presented sequence analyses along with functional studies of promoter and transcription factor binding to gain insight into their putative role and regulation.

Three White Spruce CesA Genes Are Homologous to Secondary Cell Wall CesA Genes of Angiosperms

Our phylogenetic analysis of cellulose synthase sequences from white spruce, other conifers and angiosperms (monocots and dicots) clustered the CesA proteins into six main clades. Monocot, dicot and gymnosperm CesA sequences clustered into different subgroups within the six clades (**Figure 1**). Previous studies found similar groupings and reported that the members of the distinct clades were associated with either primary or secondary cell wall formation (Holland et al., 2000; Djerbi et al., 2005; Huang et al., 2014). The white spruce cellulose synthases grouped in clades with angiosperm cellulose synthases previously linked to secondary cell wall formation including AtCesA8, AtCesA4, and AtCesA7 from Arabidopsis, PtiCesA8, PtiCesA4, and PtiCesA7-A from poplar, and EgCesA1, EgCesA2, and EgCesA3 from Eucalyptus (Holland et al., 2000; Tanaka et al., 2003; Xi et al., 2017).

The three white spruce CesAs clustered together with three loblolly pine CesAs that were preferentially expressed in xylem tissue (Nairn and Haselkorn, 2005), suggesting conservation of expression and supporting our hypothesis that conifer CesA proteins may be functionally orthologous to angiosperm secondary cell wall CesA proteins.

In this study we focused on the functional role and characterization of the *PgCesA3* gene because of its potential involvement in cell-wall synthesis. Additionally, in our previous work (Duval et al., 2014) we isolated the *PgCesA3* promoter region and defined a set of potential transcription factors with the ability to regulate this given promoter.

CesA3 Gene Structure and Comparative Analysis With Angiosperms

An analysis of gene structures showed a relatively conserved gene structure between the sequences from white spruce and angiosperms (**Figure 2A**). We analyzed the gene structure of *PgCesA3* by mapping the cDNA sequence (Raherison et al., 2012) on to the genomic sequence from a white spruce BAC containing the *CesA3* gene (Sena et al., 2014). We then compared the genomic sequence structure to those of the closest homologs from Arabidopsis (*AtCesA7*), maize (*ZmCesA12*), eucalyptus (*EgCesA3*) and poplar (*PtiCesA7-A*). The number of exons/introns varied slightly between species from 12 exons (11 introns) in maize and Arabidopsis to 13 exons (12 introns) in poplar, eucalyptus, and white spruce. The average exon length is similar between the four species, ranging from 240 bp in eucalyptus to 306 bp in white spruce. The total length of intronic sequences in the white spruce *PgCesA3* gene was 1.5 time (eucalyptus), two times (Arabidopsis and poplar) or three times (maize) greater than the angiosperms analyzed here.

We defined the PgCesA3 protein domains (**Figure 2B** and **Supplementary Figure S7**) as described in the literature (Richmond, 2000; Kumar and Turner, 2015; Maleki et al., 2016). As expected in CesA plant proteins, we identified the Zinc finger domain containing the "CxxC" motif, immediately after was the variable region, known as the class specific region (CSR I). There were two transmembrane domains near the N-terminal region and an other six toward the C-terminus. A second CSR was identified close to the middle of the protein. The signature motifs typical of the processive glycosyltransferases were present: motifs D (asparagine amino acid) and QxxRW (glutamine, arginine and tryptophan amino acids, respectively, as reviewed by Maleki et al. (2016).

Taken together the results of the comparative gene structure and the protein domains raise a question: is the occurrence of one more exon in the white spruce gene, compared to Arabidopsis, could be associated to a protein domain? To answer this question, we compared the exons encoding the protein domains between white spruce and Arabidopsis (as described by Richmond, 2000). We observed that the additional exon in white spruce harbors part of the CSR I region, which is longer in white spruce than in Arabidopsis. In Arabidopsis, as identified by Richmond (2000), the CSR I comprised the third and fourth exons while in white spruce it comprised the third to fifth exon. We aligned the white spruce CesA3 protein and the poplar homolog protein (PtiCesA7-A) and observed the same, poplar's additional exons harbored part of the CSR I as well. The distribution of the other domains along the exons was similar among white spruce, Arabidopsis and poplar.

CesA3 Promoter Region Contains Several Potential MYB Binding Sites

We undertook an analysis of upstream sequences to gain insight into cis-regulation of *PgCesA3* gene expression. We isolated its putative promoter region immediately upstream of its coding sequence by using genome walking (see Methods for details). A genomic fragment was obtained that included a 167 bp 5'UTR and 1950 bp 5' upstream of the putative transcription start site (+1). We also identified a Norway spruce scaffold containing a homologous *CesA3* gene and 5' upstream sequences (Nystedt et al., 2013). Pairwise alignment showed that 99% nucleotide sequence identity was present between the two spruce sequences, which had an overlap of 167 bp of 5'-UTR and 1015 bp in 5' upstream sequences; a 903 bp gap was observed in the Norway spruce sequence (**Figure 3**).

We identified 68 regions that harbor potential MYB binding sites in the *PgCesA3* promoter sequence by using a permissive threshold (p-value of 0.001) to facilitate the identification of motifs that are less conserved. We labeled putative binding sites for MYB transcription factors according to the unique identifiers in Arabidopsis binding site database (Plant Transcription Factor Database, V4.0; **Supplementary Figure S6** and **Supplementary Table S5**). The regions labeled as 7, 8, 10, 11, 28, 29, 41, 42, 43, 44, 59, and 61 in the *PgCesA3* promoter (**Supplementary Figure S6**) contained five of the eight
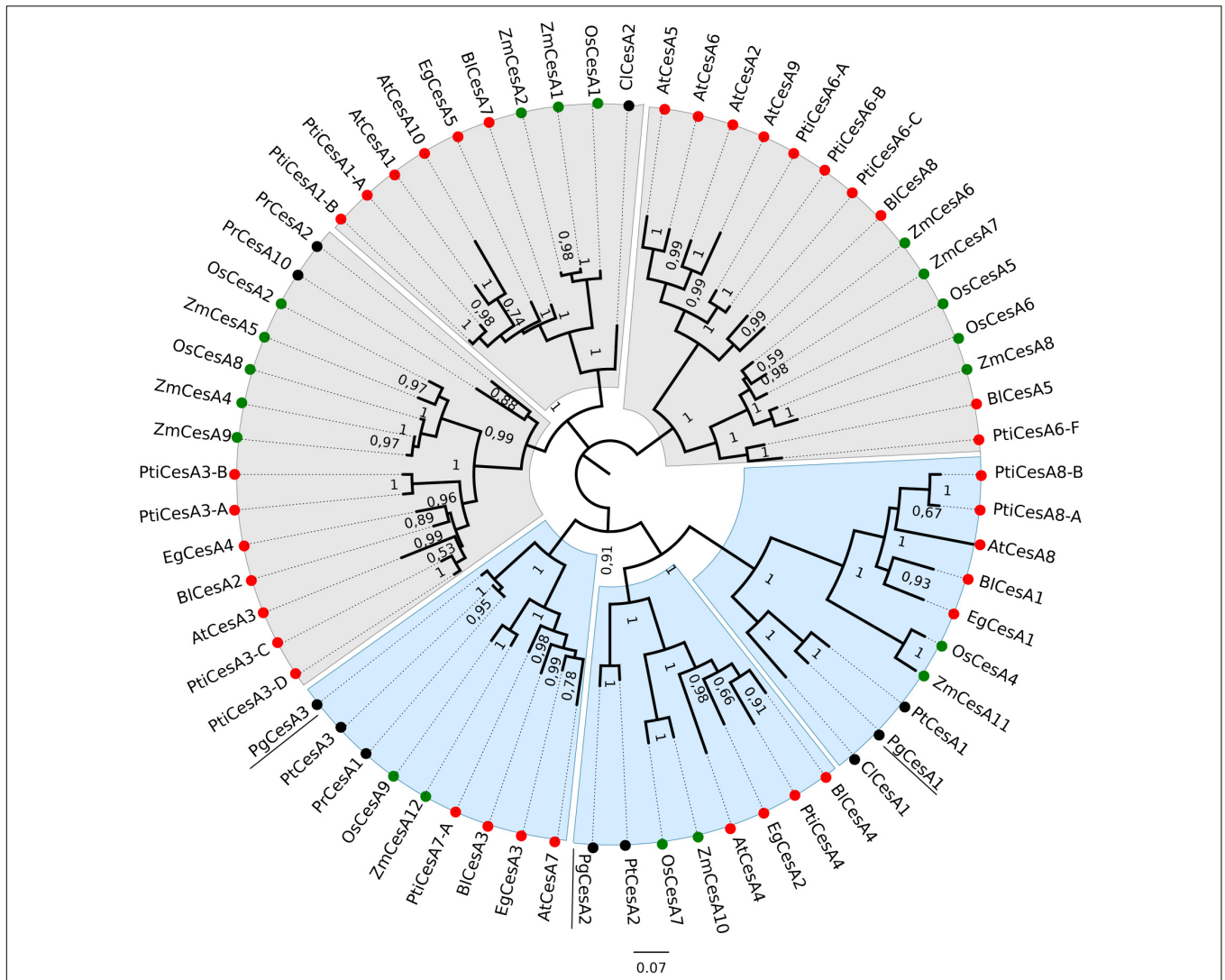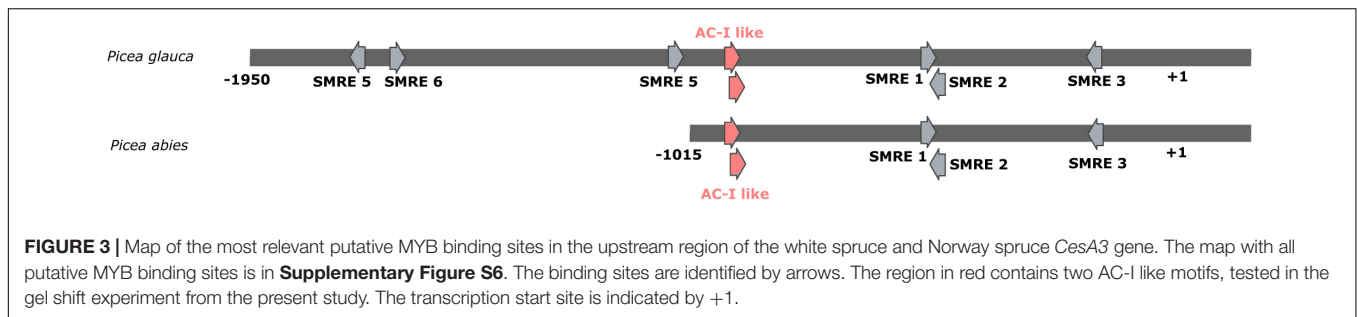
FIGURE 1 | Phylogenetic tree of angiosperm and gymnosperm cellulose synthases. The unrooted tree was created with an alignment of 66 CesA protein sequences and FastTree. Clades containing CesA proteins mainly associated with primary cell wall synthesis are shaded in grey, and clades related to secondary cell wall biosynthesis are shaded in blue. Monocots are indicated with green circles, dicots with red circles and gymnosperms with black circles. Bootstrap values > 0.50 are indicated at branch points. The white spruce CesA genes are underlined. Species included in this analysis: Os: *Oryza sativa subsp. japonica*; Zm: *Zea mays*; At: *Arabidopsis thaliana*, Eg: *E. grandis*; Pti: *Populus trichocarpa*; Bl: *Betula luminifera*; Pr: *Pinus radiata*; Pt: *Pinus taeda*; Cl: *Cunninghamia lanceolate*; Pg: *Picea glauca*.
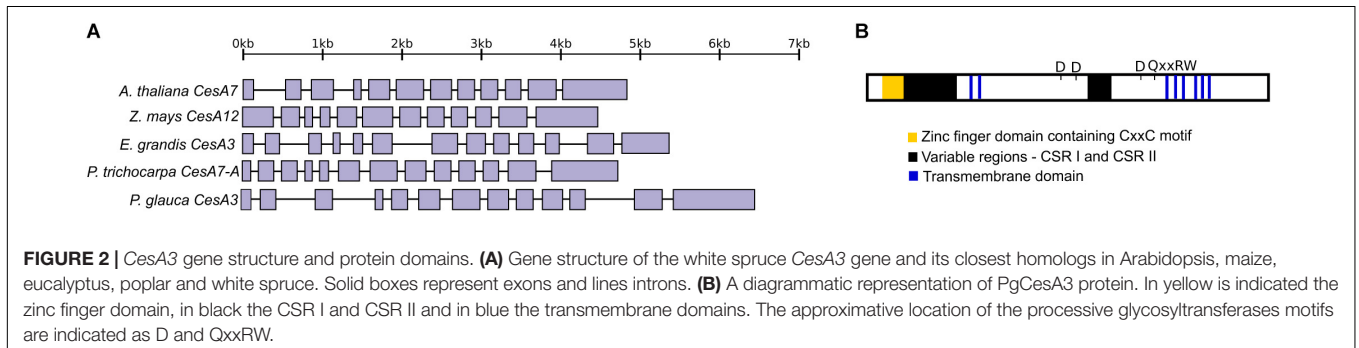
FIGURE 2 | *CesA3* gene structure and protein domains. **(A)** Gene structure of the white spruce *CesA3* gene and its closest homologs in Arabidopsis, maize, eucalyptus, poplar and white spruce. Solid boxes represent exons and lines introns. **(B)** A diagrammatic representation of PgCesA3 protein. In yellow is indicated the zinc finger domain, in black the CSR I and CSR II and in blue the transmembrane domains. The approximative location of the processive glycosyltransferases motifs are indicated as D and QxxRW.

secondary wall MYB-responsive elements (SMRE), i.e., SMRE 1, 2, 3, 5, and 6, ACC(A/T)A(A/C)(T/C), described by Zhong and Ye (2012) (**Figure 3**). The *PgCesA3* promoter lacked SMRE4,

SMRE7, or SMRE8, that correspond to the sequence canonical elements AC-II, AC-III and AC-I, respectively (**Figure 3**). The other potential MYB binding sites were variations of

FIGURE 3 | Map of the most relevant putative MYB binding sites in the upstream region of the white spruce and Norway spruce *CesA3* gene. The map with all putative MYB binding sites is in **Supplementary Figure S6**. The binding sites are identified by arrows. The region in red contains two AC-I like motifs, tested in the gel shift experiment from the present study. The transcription start site is indicated by +1.

the SMRE consensus sequence (**Supplementary Table S6** and **Supplementary Figure S6**). The pairwise alignment with the available sequence of Norway spruce *CesA3* promoter showed the conservation of sequence and position of the SMRE elements (**Figure 3**). In comparison, we found the SMRE3 and SMRE5 elements in the *AtCesA7* Arabidopsis promoter region and, SMRE2, SMRE4, and SMRE5 in the poplar *PtiCesA7-A* promoter.

The potential binding sites included sites for MYB transcription factors previously found to trans-activate the white spruce *PgCesA3* promoter (Duval et al., 2014). These include: PtMYB8, a close homolog of AtMYB46/83 (regions in green and yellow; **Supplementary Figure S6**); PtMYB1, a close homolog of AtMYB20/43 (regions in black; **Supplementary Figure S6**); and PgMYB12, a close homolog of AtMYB15 (regions in blue; **Supplementary Figure S6**). We also identified putative binding sites for AtMYB103, which is involved in secondary cell wall biogenesis (Zhong et al., 2008).

Our analysis of potential MYB binding motifs identified 26 MYB transcription factors that can potentially bind motifs (threshold *p*-value of 0.0001) in the promoter region of both white spruce and Arabidopsis. Eight regions in the white spruce promoter and nine in the Arabidopsis promoter possess motifs for binding by these transcription factors (**Supplementary Table S7**). A comparison between poplar and white spruce revealed ten common MYB transcription factors that can potentially bind six and four regions, respectively (**Supplementary Table S8**). Among Arabidopsis, poplar and white spruce we identified eight MYBs in common (**Supplementary Table S9**), of which two were related to secondary cell wall formation, AtMYB83, and AtMYB43.

A higher level of conservation was found in the common potential MYB binding motifs between white spruce and Arabidopsis such as in SMRE 3 (ACCAAAC), than between white spruce and poplar where the best hit for AtMYB43 was an SMRE 5 (ACCTAAT) and in white spruce was an SMRE 2 (ACCAACT), both in the reverse strand. The common potential MYB binding motifs between poplar and white spruce did not present the exact same sequence. For example, the best hit in poplar for AtMYB43 was an SMRE 5 (ACCTAAT) and in white spruce it was an SMRE 2 (ACCAACT), both in the reverse strand. However, the majority of shared binding sites (best hits) between Arabidopsis and white spruce were SMRE 3 (ACCAAAC).
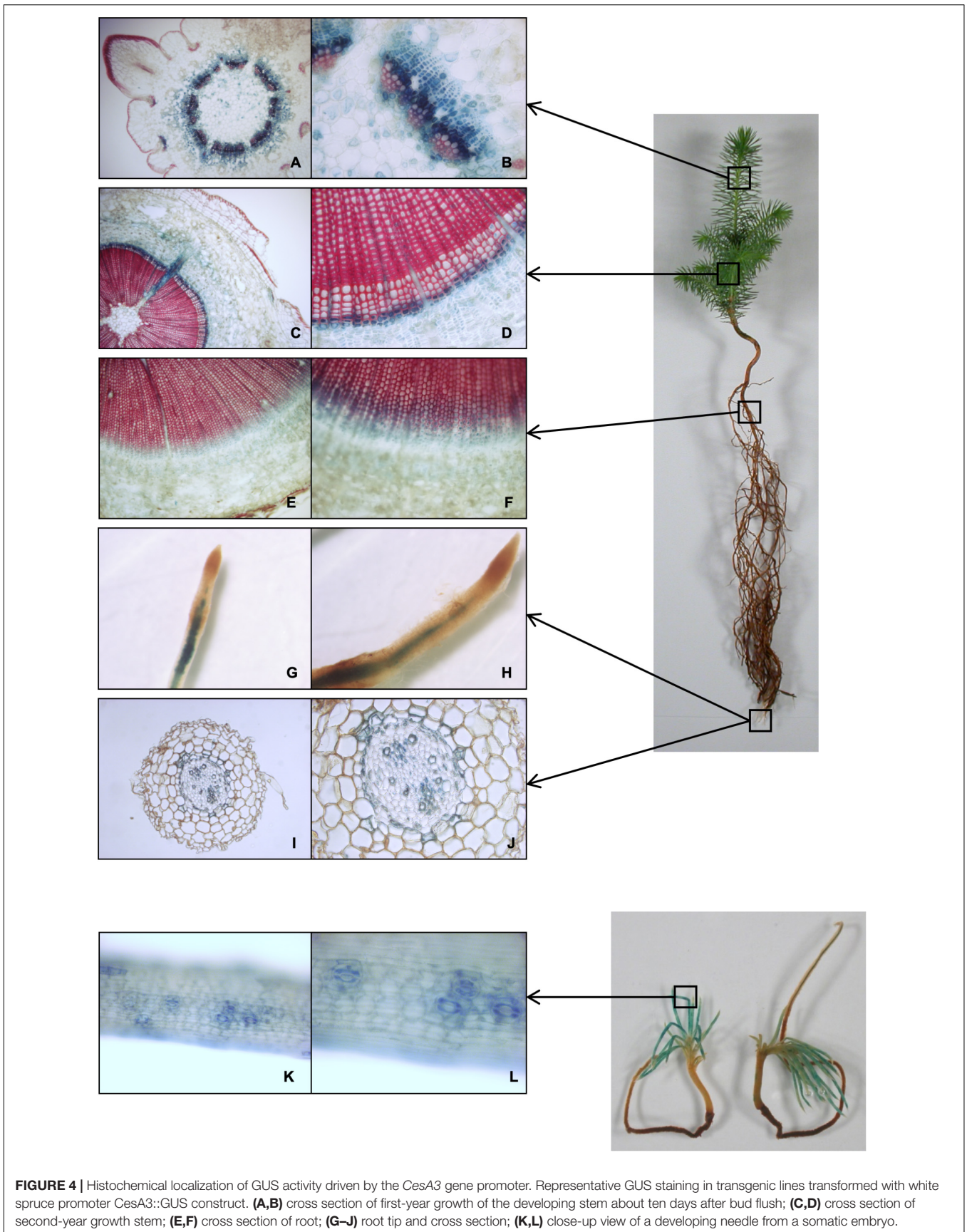
CesA3 Expression in Developing Xylem Tissues

The *CesA3* promoter sequence was fused to the GUS reporter gene and used to stably transform white spruce (see Methods, for details). Histochemical staining of GUS activity driven by the putative *PgCesA3* promoter revealed intense blue staining in developing xylem tissues, close to the cambium zone (**Figure 4**). In the stem of 1st and 2nd-year growth, we observed blue staining in the cytoplasm of differentiating xylem with no evidence of GUS staining in mature tracheids. The results were the same in roots but the blue staining was much lighter, resulting from a lower number of cell types harboring secondary cell wall development. Ray parenchyma cells also stained blue (**Figure 4**). Comparing the samples of the 2nd year stem growth ten days after bud flush to the same tissue sampled just after budset we observed more cell rows of differentiating xylem in the latter (**Supplementary Figure S8**). Blue staining was also observed in both the xylem and the endodermis cells in root tips, and in the guard cells of developing needles from somatic embryos (**Figure 4**). These results are consistent with preferential expression in the developing xylem tissue in agreement with the pattern of the *PgCesA3* endogenous gene (Raherison et al., 2012).

PgMYB12 Binds an AC-Like Sequence in the *PgCesA3* Promoter

Transient trans-activation assays revealed interactions between the *PgCesA3* promoter and several MYB transcription factors (Duval et al., 2014) based on simultaneous introduction of MYB-containing vectors and promoter-reporter gene constructs into spruce. Positive interactions were found with PgMYB12 but not with PgMYB5 and PgMYB13, both of which possess an LxLxL motif in their N-terminal region that can mediate transcriptional repression in plants (Bedon et al., 2010).

The analysis of the *PgCesA3* promoter identified several SMRE sequences, 1, 2, 3, 5 and 6, and variations of the SMRE consensus sequence but no AC elements (SMRE 4, 7, and 8). However, we identified AC- I like motifs (region 32 in **Supplementary Figure S6** and **Figure 3**) and decided to investigate their interaction with PgMYBs by EMSA. Region 32 of the *PgCesA3* promoter contains two tandem AC-I like motifs, which have a single mismatch each compared to the canonical sequence of the AC-I element (Patzlaff et al., 2003a,b). In Norway spruce we observed exactly the same AC-I like motifs.

FIGURE 4 | Histochemical localization of GUS activity driven by the *CesA3* gene promoter. Representative GUS staining in transgenic lines transformed with white spruce promoter CesA3::GUS construct. **(A,B)** cross section of first-year growth of the developing stem about ten days after bud flush; **(C,D)** cross section of second-year growth stem; **(E,F)** cross section of root; **(G–J)** root tip and cross section; **(K,L)** close-up view of a developing needle from a somatic embryo.
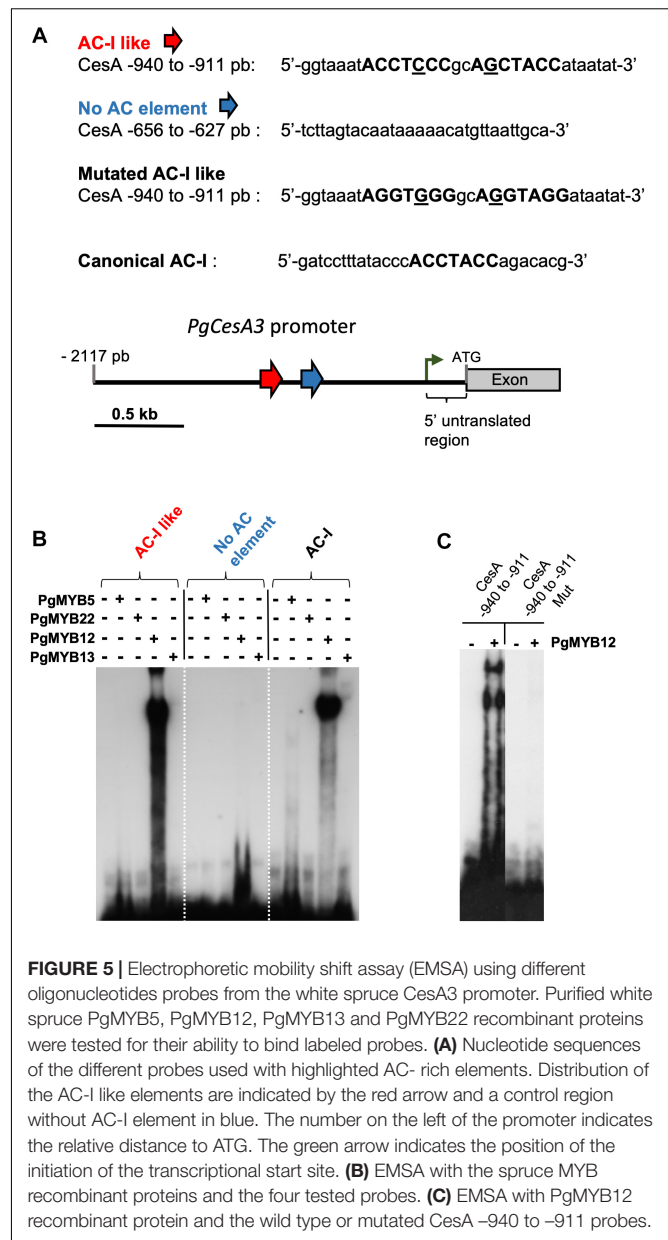
According to the *in silico* analysis of the PgCesA3 presented in the previous section (3.3) demonstrating the absence of typical AC elements (I, II, and III) we put the emphasis of our study on the identified AC-I like element for further functional analyses. We used the available full-length spruce and pine sequences to produce recombinant MYB proteins (see **Supplementary Table S2**). We obtained EMSA compatible products for PgMYB5, PgMYB8 (Bomal et al., 2014), PgMYB12, PgMYB13 and PgMYB22 (see **Supplementary Figures S1–S5** and **Supplementary Tables S2–S3**). The EMSA showed that PgMYB12 was able to bind to the AC-I like motif (region 32, **Supplementary Figure S6**) in the *PgCesA3* promoter, causing a mobility shift (**Figure 5**). PgMYB12 protein also bound to the AC-I canonical sequence, showing that it is not specific to just one binding sequence. The mutated probe of AC-I like did not induce band shifts with PgMYB12 (**Figure 5**). None of the five other MYBs were able to bind to either the AC-I element or the AC-like sequence (**Figure 5**).

DISCUSSION

White Spruce *CesA1*, *CesA2*, and *CesA3* Are Orthologous to Secondary Cell Wall CesA Proteins

We analyzed three white spruce cellulose synthases, *PgCesA1, PgCesA2, and PgCesA3,* that clustered in clades 4, 5, and 6 (**Figure 1**) with proteins known to be involved in secondary cell wall formation in angiosperms. The assembly of a functional secondary cell wall cellulose synthase complex requires three protein subunits in Arabidopsis (*AtCesA8/4/7*; Taylor et al., 2003; Hill et al., 2014) and rice (*OsCesA4/7/9*; Tanaka et al., 2003) and these different protein subunits also cluster in the clades 4, 5, and 6. The white spruce CesA proteins were orthologous to the three CesA subunit triplets in Arabidopsis, rice, eucalyptus, shinybark birch and the conifer loblolly pine. This observation suggests that there is a one-to-one orthology of the triple subunits among angiosperms and gymnosperms, which are likely to have derived from gene duplications before the split of angiosperms and gymnosperms (330 myr; Savard et al., 1994). In contrast, triplets of secondary cell wall CesA subunits were not found in the conifers Monterey pine (*Pinus radiata*) and China fir (*Cunninghamia lanceolata*), which could be due to gene loss in those species or deeper gene sampling in white spruce (Rigault et al., 2011) and loblolly pine (Nairn and Haselkorn, 2005). Poplar is the only species which presented an additional copy (CesA8 in clade 6), probably as the result of the recent whole genome duplication in the Salicaceae (Tuskan et al., 2006; Takata and Taniguchi, 2015).

The xylem preferential expression of the white spruce CesA genes (Raherison et al., 2012) supports a role in secondary cell wall formation and suggests they may be functional orthologs to CesA genes from other plant species. For example, Arabidopsis *CesA4/7/8*, eucalyptus *CesA1/2/3*, maize *CesA10/11/12*, poplar *CesA4/7-A/8-A/8-B*, shinybark birch *CesA1/3/4* and loblolly pine *CesA1/2/3*

FIGURE 5 | Electrophoretic mobility shift assay (EMSA) using different oligonucleotides probes from the white spruce CesA3 promoter. Purified white spruce PgMYB5, PgMYB12, PgMYB13 and PgMYB22 recombinant proteins were tested for their ability to bind labeled probes. **(A)** Nucleotide sequences of the different probes used with highlighted AC- rich elements. Distribution of the AC-I like elements are indicated by the red arrow and a control region without AC-I element in blue. The number on the left of the promoter indicates the relative distance to ATG. The green arrow indicates the position of the initiation of the transcriptional start site. **(B)** EMSA with the spruce MYB recombinant proteins and the four tested probes. **(C)** EMSA with PgMYB12 recombinant protein and the wild type or mutated CesA –940 to –911 probes.

all presented high expression in stem or stalk, tissues that undergo secondary cell wall synthesis in xylem cells (Appenzeller et al., 2004; Hamann et al., 2004; Nairn and Haselkorn, 2005; Ranik and Myburg, 2006; Suzuki et al., 2006; Huang et al., 2014).

The additional exon in white spruce and poplar harbored part of the CSR I region which is longer in these two species than in Arabidopsis. The CSR is variable among species and is hypothesized to be involved in the formation of the rosette-like cellulose synthase complexes, helping to stabilize the CesA complex assembly (Sethaphong et al., 2013). It is unknown if the additional exon in perennial tree species has an impact on expression or protein function; however, phylogenetic inferences together with the expression profiles

suggest conservation of function among proteins related to secondary cell wall biosynthesis.

Spruce *CesA3* Promoter Contains MYB Binding Sequences

We analyzed the promoter region of the spruce *PgCesA3* gene and identified several putative MYB cis-regulatory elements including the SMRE 1, 2, 3, 5, 6 (**Figure 3**) and variations of the SMRE consensus sequences (**Supplementary Figure S6**). The variety of MYB binding sites could have a direct impact on gene regulation. It was shown that the eight SMRE variants have different responses to activation by AtMYB46/83 (Zhong and Ye, 2012). The same was observed in poplar, eucalyptus and pine orthologs where poplar PtrMYB2/3/20/21, eucalyptus EgMYB2 and pine PtMYB4, which were shown to differentially bind to and activate the eight variants of SMRE consensus sequences (Zhong et al., 2013). We noted that many different MYBs have the potential to bind to the same regions (**Supplementary Figure S6** and **Supplementary Table S5**). The motifs that contain an SMRE can potentially bind several MYB proteins involved in the regulation of secondary wall biosynthesis genes (Zhong and Ye, 2014) as was shown for AtMYB46, AtMYB83, AtMYB43, AtMYB103, as well as MYBs AtMYB15, AtMYB62, AtMYB99, AtMYB116. Zhong and Ye (2012) showed that AtMYB46 and its paralog AtMYB83 bound to all eight variants of the SMRE consensus sequence ACC(A/T)A(A/C)(T/C). ATMYB58 and ATMYB63 are lignin-specific activators that may bind to three variants of SMRE that are identical to the AC-elements (Zhou et al., 2009), which also bound by AtMYB15 involved in wound response (Chen et al., 2006). There is considerable overlap among MYB DNA-binding sequences, probably because the R2R3 MYB DNA-binding domains are highly conserved (Stracke et al., 2001). Soler et al. (2015) showed that R2R3–MYBs controlling lignin and secondary cell wall biosynthesis were conserved across Arabidopsis, eucalyptus, poplar, grape, and rice. In eukaryotes, the overall nucleotide diversity in the promoter region is much higher than in the DNA coding sequence; however, cis-regulatory elements are conserved among orthologous genes (Lockton and Gaut, 2005). We compared the CesA promoter regions of Arabidopsis, poplar, and spruce and identified shared MYB binding sites for AtMYB43, AtMYB46, AtMYB83, and AtMYB103, which are involved in the regulation of secondary cell wall. This conservation of putative MYB cis-regulatory elements suggests that transcriptional regulation of cellulose biosynthesis appeared early in the evolution in vascular plants. We found a higher level of conservation in MYB DNA binding sites between white spruce and Arabidopsis than between white spruce and poplar. The time of divergence between gymnosperms and angiosperms is approximately 330 myr (Savard et al., 1994) and the time of divergence between poplar (Fabid) and Arabidopsis (Malvid) is approximately 82 myr (Clarke et al., 2011). This pattern may result from the slower rates of evolution in conifers than angiosperms (Buschiazzo et al., 2012) and a higher rate due to whole genome duplications in poplar and Arabidopsis, amongst other reasons. It has been shown that duplicated CesA genes in poplar diverged in a few regulatory elements, possibly causing differential expression (Takata and Taniguchi, 2015).

CesA3 Promoter in White Spruce Is Linked to Secondary Cell Wall Biosynthesis

The *PgCesA3* promoter was strongly activated in differentiating stem xylem cells according to GUS expression of the white spruce *CesA3* promoter fused to the β-glucuronidase reporter gene. Our results suggest that *PgCesA3* is involved in the deposition of the secondary cell wall during xylogenesis. One of the phases of xylem differentiation is the deposition of a thick multilayered secondary cell wall, which is coordinated and driven by numerous proteins, including cellulose synthases (Plomion et al., 2001). Some angiosperm CesA genes have been reported for their roles in the biosynthesis of the secondary cell wall. For example, the eucalyptus promoter *EgCesA3* exhibited GUS staining in Arabidopsis tissues undergoing secondary cell wall deposition (Creux et al., 2008). The poplar *PtiCesA7-A* gene was implicated in the multilaminar structure of the secondary cell wall (Xi et al., 2017) and the Arabidopsis *AtCesA7* was reported as essential for cellulose biosynthesis in secondary cell walls (Taylor et al., 1999, 2000).

In the present study, we observed that the promoter of *PgCesA3* drove xylem preferential GUS expression most strongly in differentiating xylem cells of developing stem and produced lighter blue staining in needles and roots. Our observations are consistent with microarray data showing that *PgCesA3* transcripts were most abundant in xylem tissue and detected at lower levels in foliage and roots (Raherison et al., 2012). We found that the *PgCesA3* promoter was also activated in the stomata guard cells. During stomatal movements, the cellulose microfibrils in the stomata undergo dynamic reorganization to enable aperture and closure driven by changes in turgor pressure. An Arabidopsis secondary cell wall *AtCesA7* mutant had smaller stomatal opening and smaller guard cells (Liang et al., 2010); interestingly, the guard cells were not deficient in cellulose, but displayed a loss of cellulose reorganization during stomatal movement. It is unknown if these mutant characteristics were due to collapsed xylem and compromised water transport or to the loss of cellulose reorganization (Rui and Anderson, 2016). Together, these observations suggest that *PgCesA3* could be necessary for stomatal function but additional studies will be needed to elucidate its specific role.

PgMYB12 Interacts With AC-I Like Motifs in the *PgCesA3* Promoter

The spruce *PgCesA3* promoter and its putative ortholog in Arabidopsis both lack AC-I elements corresponding to SMRE4, SMRE7, and SMRE8. The poplar ortholog also lacks SMRE7 and SMRE8 but has SMRE4 elements. Due to the absence of AC-I elements, we hypothesized that an AC-I like element may function as a MYB binding site in the *PgCesA3* promoter.

Here, we found that PgMYB12 bound to both a canonical AC-I element (from PAL promoter, Patzlaff et al., 2003b) and also a region in the *PgCesA3* promoter containing two

AC-I like elements. Previous work showed that PgMYB12 transactivated the expression of cellulose synthase, *PgCesA3* using an embryogeneic cell culture (Duval et al., 2014). PgMYB12 is a close homolog to AtMYB15 in Arabidopsis, which binds to AC-I elements in shikimate pathway genes (Chen et al., 2006). PgMYB12 could be involved in the regulation of cellulose biosynthesis in spruce and could regulate *PgCesA3* in roots based on it is preferential expression in root xylem of young and mature spruce trees (Bedon et al., 2007), which is in agreement with our GUS expression results (**Figure 4**).

It was also shown that neither PgMYB5 nor PgMYB13 were able to activate *PgCesA3* expression (Duval et al., 2014). Both MYBs are part of subgroup 4 and present an EAR core motif (Bedon et al., 2010), suggesting they could mediate transcriptional repression of specific genes. We found that they did not bind to the AC-I canonical sequence and the AC-I Iike motif. They could potentially bind to other elements in the *CesA3* promoter, competing with cognate transcription factors or sequestrate components of the transcriptional machinery away from the cis-regulatory DNA elements (Gill and Ptashne, 1988), since they present the basic helix-loop-helix (bHLH) interaction site (Zimmermann et al., 2004). Duval et al. (2014) showed that PgMYB22 was able to trans-activate the *PgCesA3* promoter as well as the promoters of genes related to the phenylpropanoid (*CAD*, *PAL*) and shikimate pathways (*DAHP*), and also those of transcription factors (*PgMYB1*, *PgMYB8*, and *PgLIM1*). In the present study we have not been able to identify specific DNA-protein interactions using EMSA for PgMYB22 nor for PgMYB8 or PgMYB15 (**Figure 5**).

In Arabidopsis, AtMYB46 binds directly to the promoters of *AtCesA4*, *AtCesA7*, and *AtCesA8* involved in secondary cell wall formation in *in vivo* and *in vitro* studies (Kim et al., 2012) and may directly regulate their expression (Ko et al., 2009). The closest ortholog to AtMYB46 in white spruce and in loblolly pine is PgMYB8 and PtMYB8, respectively (Bedon et al., 2007). As PtMYB8 transactivated white spruce *PgCesA3*, it is highly likely that Pt/PgMYB8 are functionally related to AtMYB46 and that the regulation of CesA genes involved in the secondary cell wall synthesis is conserved among angiosperms and conifers (Duval et al., 2014). Testing this hypothesis will be the next step in the understanding of the regulation of spruce cellulose synthase genes.

The affinity of conifer R2R3-MYBs for AC elements has been reported in pine, PtMYB1 (Patzlaff et al., 2003b) and PtMYB4 (Gómez-Maldonado et al., 2004; Patzlaff et al., 2003a); in maritime pine, PpMYB8 (Craven-Bartle et al., 2013) and in spruce, PgMYB8, PgMYB14 and PgMYB15 (Bomal et al., 2014). However, the interaction of these MYBs with potential DNA MYB binding sites in the promoter of a conifer cellulose synthase had not been tested before the present study. The potential MYB binding sites identified in this work open avenues for examining the interaction between other PgMYBs and the *PgCesA3* promoter motifs. (Duval et al., 2014).

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00626/full#supplementary-material

REFERENCES

Appenzeller, L., Doblin, M., Barreiro, R., Wang, H., Niu, X., Kollipara, K., et al. (2004). Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family. *Cellulose* 11, 287–299. doi: 10.1023/B:CELL.0000046417.84715.27

Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S. D., et al. (2010). Subgroup 4 R2R3-MYBs in conifer trees: gene family expansion and contribution to the isoprenoid- and flavonoid-oriented responses. *J. Exp. Bot.* 61, 3847–3864. doi: 10.1093/jxb/erq196

Bedon, F., Grima-Pettenati, J., and Mackay, J. (2007). Conifer R2R3-MYB transcription factors: sequence analyses and gene expression in wood-forming tissues of white spruce (*Picea glauca*). *BMC Plant Biol.* 7:17. doi: 10.1186/1471-2229-7-17

Bedon, F., Levasseur, C., Grima-Pettenati, J., Séguin, A., and MacKay, J. (2009). Sequence analysis and functional characterization of the promoter of the *Picea glauca* cinnamyl alcohol dehydrogenase gene in transgenic white spruce plants. *Plant Cell Rep.* 28, 787–800. doi: 10.1007/s00299-009-0688-0

Bomal, C., Bedon, F., Caron, S., Mansfield, S. D., Levasseur, C., Cooke, J. E. K., et al. (2008). Involvement of *Pinus taeda* MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis. *J. Exp. Bot.* 59, 3925–3939. doi: 10.1093/jxb/ern234

Bomal, C., Duval, I., Giguère, I., Fortin, É., Caron, S., Stewart, D., et al. (2014). Opposite action of R2R3-MYBs from different subgroups on key genes of the shikimate and monolignol pathways in spruce. *J. Exp. Bot.* 65, 495–508. doi: 10.1093/jxb/ert398

Buschiazzo, E., Ritland, C., Bohlmann, J., and Ritland, K. (2012). Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS

in conifers compared to angiosperms. *BMC Evol. Biol.* 12:8. doi: 10.1186/1471-2148-12-8

Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., et al. (2008). MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18, 188–196. doi: 10.1101/gr.6743907

Chen, Y., Zhang, X., Wu, W., Chen, Z., Gu, H., and Qu, L.-J. (2006). Overexpression of the wounding-responsive gene AtMYB15 activates the shikimate pathway in *Arabidopsis*. *J. Integ. Plant Biol.* 48, 1084–1095. doi: 10.1111/j.1744-7909.2006.00311.x

Clarke, J. T., Warnock, R. C. M., and Donoghue, P. C. J. (2011). Establishing a time-scale for plant evolution. *New Phytol.* 192, 266–301. doi: 10.1111/j.1469-8137.2011.03794.x

Craven-Bartle, B., Pascual, M. B., Cánovas, F. M., and Ávila, C. (2013). A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine. *Plant J.* 74, 755–766. doi: 10.1111/tpj.12158

Creux, N. M., Ranik, M., Berger, D. K., and Myburg, A. A. (2008). Comparative analysis of orthologous cellulose synthase promoters from *Arabidopsis*, *Populus* and *Eucalyptus*: evidence of conserved regulatory elements in angiosperms. *New Phytol.* 179, 722–737. doi: 10.1111/j.1469-8137.2008.02517.x

Djerbi, S., Lindskog, M., Arvestad, L., Sterky, F., and Teeri, T. T. (2005). The genome sequence of black cottonwood (*Populus trichocarpa*) reveals 18 conserved cellulose synthase (CesA) genes. *Planta* 221, 739–746.

Duval, I., Lachance, D., Giguère, I., Bomal, C., Morency, M.-J., Pelletier, G., et al. (2014). Large-scale screening of transcription factor–promoter interactions in spruce reveals a transcriptional network involved in vascular development. *J. Exp. Bot.* 65, 2319–2333. doi: 10.1093/jxb/eru116

Endler, A., and Persson, S. (2011). Cellulose synthases and synthesis in *Arabidopsis*. *Mol. Plant* 4, 199–211. doi: 10.1093/mp/ssq079

Galinousky, D., Padvitski, T. Bayer, G., Pirko, Y., Pydiura, N., Anisimova, N., et al. (2017). Expression analysis of cellulose synthase and main cytoskeletal protein genes in flax (*Linum usitatissimum* L.). *Cell Biol. Int.* doi: 10.1002/cbin.10837 [Epub ahead of print].

Gardiner, J. C., Taylor, N. G., and Turner, S. R. (2003). Control of cellulose synthase complex localization in developing xylem. *Plant Cell* 15, 1740–1748.

Gill, G., and Ptashne, M. (1988). Negative effect of the transcriptional activator GAL4. *Nature* 334, 721–724. doi: 10.1038/334721a0

Gómez-Maldonado, J., Avila, C., Torre, F., Cañas, R., Cánovas, F. M., and Campbell, M. M. (2004). Functional interactions between a glutamine synthetase promoter and MYB proteins. *Plant J.* 39, 513–526. doi: 10.1111/j.1365-313X.2004.02153.x

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944

Grant, C. E., Bailey, T. L., and Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018. doi: 10.1093/bioinformatics/btr064

Hamann, T., Osborne, E., Youngs, H. L., Misson, J., Nussaume, L., and Somerville, C. (2004). Global expression analysis of CESA and CSL genes in *Arabidopsis*. *Cellulose* 11, 279–286. doi: 10.1023/B:CELL.0000046340.99925.57

Handakumbura, P. P., Matos, D. A., Osmont, K. S., Harrington, M. J., Heo, K., Kafle, K., et al. (2013). Perturbation of *Brachypodium distachyon* CELLULOSE SYNTHASE A4 or 7 results in abnormal cell walls. *BMC Plant Biol.* 13:131. doi: 10.1186/1471-2229-13-131

Hawkins, S., Samaj, J., Lauvergeat, V., Boudet, A., and Grima-Pettenati, J. (1997). Cinnamyl alcohol dehydrogenase: identification of new sites of promoter activity in transgenic poplar. *Plant Physiol.* 113, 321–325. doi: 10.1104/pp.113.2.321

Hemerly, A. S., Ferreira, P., Engler, J. de A., Montagu, M. V., Engler, G., and Inzé, D. (1993). cdc2a expression in *Arabidopsis* is linked with competence for cell division. *Plant Cell* 5, 1711–1723. doi: 10.1105/tpc.5.12.1711

Hill, J. L., Hammudi, M. B., and Tien, M. (2014). The *Arabidopsis* cellulose synthase complex: a proposed hexamer of CESA trimers in an equimolar stoichiometry. *Plant Cell* 26, 4834–4842. doi: 10.1105/tpc.114.131193

Holland, N., Holland, D., Helentjaris, T., Dhugga, K. S., Xoconostle-Cazares, B., and Delmer, D. P. (2000). A comparative analysis of the plant cellulose synthase (CESA) gene family. *Plant Physiol.* 123, 1313–1323.

Huang, H., Jiang, C., Tong, Z., Cheng, L., Zhu, M., and Lin, E. (2014). Eight distinct cellulose synthase catalytic subunit genes from *Betula luminifera* are associated with primary and secondary cell wall biosynthesis. *Cellulose* 21, 2183–2198

Jokipii-Lukkari, S., Sundell, D., Nilsson, O., Hvidsten, T. R., Street, N. R., and Tuominen, H. (2017). NorWood: a gene expression resource for evo-devo studies of conifer wood development. *New Phytol.* 216, 482–494. doi: 10.1111/nph.14458

Joshi, C. P., Thammannagowda, S., Fujino, T., Gou, J.-Q., Avci, U., Haigler, C. H., et al. (2011). Perturbation of wood cellulose synthesis causes pleiotropic effects in transgenic aspen. *Mol. Plant* 4, 331–345. doi: 10.1093/mp/ssq081

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Keijsers, E. R. P., Yılmaz, G., and van Dam, J. E. G. (2013). The cellulose resource matrix. *Carbohydr. Polym.* 93, 9–21. doi: 10.1016/j.carbpol.2012.08.110

Kim, W.-C., Ko, J.-H., and Han, K.-H. (2012). Identification of a cis-acting regulatory motif recognized by MYB46, a master transcriptional regulator of secondary wall biosynthesis. *Plant Mol. Biol.* 78, 489–501. doi: 10.1007/s11103-012-9880-7

Klimaszewska, K., Lachance, D., Pelletier, G., Lelu, M.-A., and Séguin, A. (2001). Regeneration of transgenic *Picea glauca*, *P. Mariana* and *P. abies* after cocultivation of embryogenic tissue with *Agrobacterium tumefaciens*. *Vitro Cell Dev. Biol. Plant* 37, 748–755. doi: 10.1007/s11627-001-0124-9.

Klimaszewska, K., Rutledge, R. G., and Séguin, A. (2005). "Genetic transformation of conifers utilizing somatic embryogenesis," in *Transgenic Plants: Methods and Protocols Methods in Molecular Biologytm.* ed. L. Pena (Totowa, NJ: Humana Press Inc.), 151–163.

Ko, J.-H., Kim, W.-C., and Han, K.-H. (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in *Arabidopsis*. *Plant J.* 60, 649–665. doi: 10.1111/j.1365-313X.2009.03989.x

Koncz, C., and Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204, 383–396. doi: 10.1007/BF00331014

Kumar, M., and Turner, S. (2015). Plant cellulose synthesis: CESA proteins crossing kingdoms. *Phytochemistry* 112, 91–99. doi: 10.1016/j.phytochem.2014.07.009

Levée, V., Major, I., Levasseur, C., Tremblay, L., MacKay, J., and Séguin, A. (2009). Expression profiling and functional analysis of *Populus* WRKY23 reveals a regulatory role in defense. *New Phytol.* 184, 48–70. doi: 10.1111/j.1469-8137.2009.02955.x

Liang, Y.-K., Xie, X., Lindsay, S. E., Wang, Y. B., Masle, J., Williamson, L., et al. (2010). Cell wall composition contributes to the control of transpiration efficiency in *Arabidopsis thaliana*. *Plant J.* 64, 679–686. doi: 10.1111/j.1365-313X.2010.04362.x

Lockton, S., and Gaut, B. S. (2005). Plant conserved non-coding sequences and paralogue evolution. *Trends Genet.* 21, 60–65. doi: 10.1016/j.tig.2004.11.013

Maleki, S. S., Mohammadi, K., and Ji, K. (2016). Characterization of cellulose synthesis in plant cells. *Sci. World J.* 2016:8641373 doi: 10.1155/2016/8641373

Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of *Eucalyptus grandis*. *Nature* 510, 356–362. doi: 10.1038/nature13308

Nairn, C. J., and Haselkorn, T. (2005). Three loblolly pine CesA genes expressed in developing xylem are orthologous to secondary cell wall CesA genes of angiosperms. *New Phytol.* 166, 907–915. doi: 10.1111/j.1469-8137.2005.01372.x

Nawaz, M. A., Rehman, H. M., Baloch, F. S., Ijaz, B., Ali, M. A., Khan, I. A., et al. (2017). Genome and transcriptome-wide analyses of cellulose synthase gene superfamily in soybean. *J. Plant Physiol.* 215, 163–175. doi: 10.1016/j.jplph.2017.04.009

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211

Pascual, M. B., Llebrés, M.-T., Craven-Bartle, B., Cañas, R. A., Cánovas, F. M., and Ávila, C. (2018). PpNAC1, a main regulator of phenylalanine biosynthesis and utilization in maritime pine. *Plant Biotechnol. J.* 16, 1094–1104. doi: 10.1111/pbi.12854

Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L. J., Smith, C., et al. (2003a). Characterisation of a pine MYB that regulates lignification. *Plant J.* 36, 743–754. doi: 10.1046/j.1365-313X.2003.01916.x

Patzlaff, A., Newman, L. J., Dubos, C., Whetten, R. W., Smith, C., McInnis, S., et al. (2003b). Characterisation of MYB1, an R2R3-MYB from pine xylem. *Plant Mol. Biol.* 53, 597–608. doi: 10.1023/B:PLAN.0000019066.07933.d6

Plomion, C., Leprovost, G., and Stokes, A. (2001). Wood formation in trees. *Plant Physiol.* 127, 1513–1523. doi: 10.1104/pp.010816

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490

Prouse, M. B., and Campbell, M. M. (2012). The interaction between MYB proteins and their target DNA binding sites. *Biochim. Biophys. Acta Gene Regul. Mech.* 1819, 67–77. doi: 10.1016/j.bbagrm.2011.10.010.

Raherison, E., Rigault, P., Caron, S., Poulin, P.-L., Boyle, B., Verta, J.-P., et al. (2012). Transcriptome profiling in conifers and the piceagenexpress database show patterns of diversification within gene families and interspecific conservation in vascular gene expression. *BMC Genomics* 13:434. doi: 10.1186/1471-2164-13-434

Ranik, M., and Myburg, A. A. (2006). Six new cellulose synthase genes from *Eucalyptus* are associated with primary and secondary cell wall biosynthesis. *Tree Physiol.* 26, 545–556.

Richmond, T. (2000). Higher plant cellulose synthases. *Genome Biol.* 1, reviews3001.1–reviews3001.6.

Rigault, P., Boyle, B., Lepage, P., Cooke, J. E. K., Bousquet, J., and MacKay, J. J. (2011). A white spruce gene catalog for conifer genome analyses. *Plant Physiol.* 157, 14–28. doi: 10.1104/pp.111.179663

Rui, Y., and Anderson, C. T. (2016). Functional analysis of cellulose and xyloglucan in the walls of stomatal guard cells of *Arabidopsis*. *Plant Physiol.* 170, 1398–1419. doi: 10.1104/pp.15.01066

Savard, L., Li, P., Strauss, S. H., Chase, M. W., Michaud, M., and Bousquet, J. (1994). Chloroplast and nuclear gene sequences indicate late pennsylvanian time for the last common ancestor of extant seed plants. *Proc. Natl. Acad. Sci. U.S.A.* 91, 5163–5167.

Sena, J. S., Giguère, I., Boyle, B., Rigault, P., Birol, I., Zuccolo, A., et al. (2014). Evolution of gene structure in the conifer *Picea glauca*: a comparative analysis of the impact of intron size. *BMC Plant Biol.* 14:95. doi: 10.1186/1471-2229-14-95

Sethaphong, L., Haigler, C. H., Kubicki, J. D., Zimmer, J., Bonetta, D., DeBolt, S., et al. (2013). Tertiary model of a plant cellulose synthase. *PNAS* 110, 7512–7517. doi: 10.1073/pnas.1301027110

Sibout, R., Eudes, A., Mouille, G., Pollet, B., Lapierre, C., Jouanin, L., et al. (2005). Cinnamyl alcohol dehydrogenasE-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of *Arabidopsis*. *Plant Cell* 17, 2059–2076. doi: 10.1105/tpc.105.030767

Soler, M., Camargo, E. L. O., Carocha, V., Cassan-Wang, H., San Clemente, H., Savelli, B., et al. (2015). The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. *New Phytol.* 206, 1364–1377. doi: 10.1111/nph.13039

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* 4, 447–456.

Suzuki, S., Li, L., Sun, Y.-H., and Chiang, V. L. (2006). The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*. *Plant Physiol.* 142, 1233–1245. doi: 10.1104/pp.106.086678.

Takata, N., and Taniguchi, T. (2015). Expression divergence of cellulose synthase (CesA) genes after a recent whole genome duplication event in *Populus*. *Planta* 241, 29–42. doi: 10.1007/s00425-014-2217-9.

Tanaka, K., Murata, K., Yamazaki, M., Onosato, K., Miyao, A., and Hirochika, H. (2003). Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. *Plant Physiol.* 133, 73–83.

Taylor, N. G., Howells, R. M., Huttly, A. K., Vickers, K., and Turner, S. R. (2003). Interactions among three distinct CesA proteins essential for cellulose synthesis. *PNAS* 100, 1450–1455. doi: 10.1073/pnas.0337628100

Taylor, N. G., Laurie, S., and Turner, S. R. (2000). Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in *Arabidopsis*. *Plant Cell* 12, 2529–2539. doi: 10.1105/tpc.12.12.2529

Taylor, N. G., Scheible, W.-R., Cutler, S., Somerville, C. R., and Turner, S. R. (1999). The irregular xylem3 locus of *Arabidopsis* encodes a cellulose synthase required for secondary cell wall synthesis. *Plant Cell* 11, 769–779. doi: 10.1105/tpc.11.5.769

The UniProt Consortium (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* 43, D204–D212. doi: 10.1093/nar/gku989

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691

Vigneault, F., Lachance, D., Cloutier, M., Pelletier, G., Levasseur, C., and Séguin, A. (2007). Members of the plant NIMA-related kinases are involved in organ development and vascularization in poplar, *Arabidopsis* and rice. *Plant J.* 51, 575–588. doi: 10.1111/j.1365-313X.2007.03161.x.

Weirauch, M. T., Yang, A., Albu, M., Cote, A., Montenegro-Montero, A., Drewe, P., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* 158, 1431–1443. doi: 10.1016/j.cell.2014.08.009

Xi, W., Song, D., Sun, J., Shen, J., and Li, L. (2017). Formation of wood secondary cell wall may involve two type cellulose synthase complexes in *Populus*. *Plant Mol. Biol.* 93, 419–429. doi: 10.1007/s11103-016-0570-8

Xie, Z., Lee, E., Lucas, J. R., Morohashi, K., Li, D., Murray, J. A. H., et al. (2010). Regulation of cell proliferation in the stomatal lineage by the *Arabidopsis* MYB four lips via direct targeting of core cell cycle genes. *Plant Cell* 22, 2306–2321. doi: 10.1105/tpc.110.074609

Xu, Z.-C., and Kong, Y. (2017). Genome-wide identification, subcellular localization and gene expression analysis of the members of CESA gene family in common tobacco (*Nicotiana tabacum* L.). *Yi Chuan* 39, 512–524. doi: 10.16288/j.yczz.16-428

Zhong, R., Lee, C., and Ye, Z.-H. (2010). Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends Plant Sci.* 15, 625–632. doi: 10.1016/j.tplants.2010.08.007.

Zhong, R., Lee, C., Zhou, J., McCarthy, R. L., and Ye, Z.-H. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell* 20, 2763–2782. doi: 10.1105/tpc.108.061325

Zhong, R., McCarthy, R. L., Haghighat, M., and Ye, Z.-H. (2013). The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PLoS One* 8:e69219. doi: 10.1371/journal.pone.0069219

Zhong, R., and Ye, Z.-H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.* 53, 368–380. doi: 10.1093/pcp/pcr185

Zhong, R., and Ye, Z.-H. (2014). Complexity of the transcriptional network controlling secondary wall biosynthesis. *Plant Sci.* 229, 193–207. doi: 10.1016/j.plantsci.2014.09.009

Zhong, R., and Ye, Z.-H. (2015). Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant Cell Physiol.* 56, 195–214. doi: 10.1093/pcp/pcu140

Zhou, J., Lee, C., Zhong, R., and Ye, Z.-H. (2009). MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell* 21, 248–266. doi: 10.1105/tpc.108.063321

Zimmermann, I. M., Heim, M. A., Weisshaar, B., and Uhrig, J. F. (2004). Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J.* 40, 22–34. doi: 10.1111/j.1365-313X.2004.02183.x