EXPLORING THE TRANSCRIPTIONAL NETWORK REGULATING SECONDARY CELL WALL AND WOOD FORMATION: NEW KEY PLAYERS AND EVOLUTIONARILY CONSERVED REGULATORY MECHANISMS

by

RYAN LEHMAN MCCARTHY

(Under the Direction of Zheng-Hua Ye)

ABSTRACT

Secondary cell wall formation is essential to plant mechanical structure and vascular system function. The evolution of the plant vascular system allowed for efficient water transport and was an essential step in the colonization of land by plants. Wood, composed of cells with roles in water transport and mechanical support, is primarily composed of secondary cells walls and is a material with a dense concentration of biopolymers that can be processed to produce liquid biofuels. By better understanding the process of secondary cell wall formation we can not only improve our understanding of land plant evolution but also develop knowledge and tools for engineering enhanced bioenergy crops for biofuel production. Secondary cell wall formation in *Arabidopsis* is transcriptionally regulated by a set of homologous secondary wall NAC (SWN) domain proteins that act as master regulators capable of activating the entire pathway. Target gene activation by SWNs is mediated through their binding to secondary wall NAC binding elements (SNBEs). MYB46 has been identified as a master regulator of secondary cell wall formation and a direct target of the SWNs. Here we identify the transcription factor MYB83 as a new master regulator of secondary cell wall formation that functions redundantly with MYB46

as a direct target of the SWNs. Double T-DNA knockout of MYB46 and MYB83 blocked secondary cell wall formation indicating that these genes are necessary for activation of secondary cell wall formation. Expanding our findings to the bioenergy crop poplar we identified PtrMYB3 and PtrMYB20 as functional orthologs of MYB46 and MYB83 with a conserved role as master regulators of secondary cell wall formation indicating an evolutionary conservation of the transcriptional network. To further investigate this evolutionary conservation we analyzed SNBE site conservation and distribution across multiple plant genomes. All vascular plants included in the analysis demonstrated a significantly enriched concentration of SNBE sites in promoters of secondary cell wall related genes. Together, these results identify new master regulatory proteins controlling secondary cell wall formation and demonstrate evidence that the transcriptional network regulating secondary cell wall formation is well conserved in the vascular plant lineage.

INDEX WORDS: Secondary Cell Wall, Wood Formation, Transcription Factors, Regulation, MYB83, PtrMYB3, PtrMYB20

EXPLORING THE TRANSCRIPTIONAL NETWORK REGULATING SECONDARY CELL WALL AND WOOD FORMATION: NEW KEY PLAYERS AND EVOLUTIONARILY CONSERVED REGULATORY MECHANISMS

by

RYAN LEHMAN MCCARTHY

B.S., Harvey Mudd College, 2007

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2012

EXPLORING THE TRANSCRIPTIONAL NETWORK REGULATING SECONDARY CELL WALL AND WOOD FORMATION: NEW KEY PLAYERS AND EVOLUTIONARILY CONSERVED REGULATORY MECHANISMS

by

RYAN LEHMAN MCCARTHY

Major Professor:	Zheng-Hua Ye
Committee:	Gregory Schmidt
	Wolfgang Lukowitz
	Xiaoyu Zhang
	Mike Deom

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2012

DEDICATION

To Nya whose love has always kept me moving towards my goals even when they seemed out of reach. And to my parents who provided never ending support to all my crazy plans.

ACKNOWLEDGEMENTS

I am indebted to the many people whose help and guidance has supported me on this journey. First and foremost I'd like to thank my major professor Zheng-Hua Ye. He has taught me innumerable important lessons and instilled in me the importance of setting the highest standards for one's own work. I am also extremely grateful to my committee members Gregory Schmidt, Wolfgang Lukowitz, Xiaoyu Zhang and Mike Deom for their guidance and insights. I would like to thank Beth Richardson for teaching me the histology and microscopy techniques that made this work possible. Many thanks go to Ruiqin Zhong whose expert advice with molecular techniques has often helped me immensely. I am grateful to Brigitte Bruns, Sue Wessler and Jim Burnette who all helped me to share my love of biology with future generations of students. Finally I wish to thank my family and friends, especially my fiancée Nya and my parents. They have provided never ending support and their love has kept me focused and sane during this endeavor.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

2 MYB83 IS A DIRECT TARGET OF SND1 AND ACTS REDUNDANTLY WITH MYB46 IN THE REGULATION OF SECONDARY CELL WALL BIOSYNTHESIS IN ARABIDOPSIS 29

3 THE POPLAR MYB TRANSCRIPTION FACTORS, PtrMYB3 AND PtrMYB20, ARE INVOLVED IN THE REGULATION OF SECONDARY WALL BIOSYNTHESIS 69

4 FROM SELAGINELLA TO POPLAR: THE SECONDARY WALL NAC DOMAIN PROTEIN BINDING SITE AS A CONSERVED CIS-ELEMENT IMPORTANT FOR SECONDARY CELL WALL FORMATION 98

5 CONCLUSIONS 141

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

The evolution of plants onto land required numerous adaptations; each addressing a problem posed by their new terrestrial environment and ultimately responsible for producing the diverse forms of extant vascular plants. The formation of vascular tissue, next to the development of seeds and a waxy cuticle, may be the most significant advancement in allowing plants to colonize nearly every land ecosystem. Vascular tissue enabled plants to transport water and nutrients throughout their bodies, permitting increased tissue specialization and reducing the reliance on water immersion. While some nonvascular plants possess water conducting tissue structurally similar to vascular tissue, the water conducting xylem vessels of vascular plants are distinguished by the formation of a lignified secondary cell wall (Ligrone et al., 2000). Secondary cell walls are also formed in other cell types including anther endothecium as well as xylem fibers which form alongside xylem vessels to compose vascular bundles in the stem and provide structural support. Competition among land plants for sunlight created selective pressure driving plants to grow taller. To support this increased height plants evolved the ability to produce wood.

Wood formation

Wood represents the world's most abundant form of biomass (Bhalerao et al., 2003). The formation of woody tissue sequesters carbon fixed through photosynthesis and decreases atmospheric CO_2. The properties of wood have made it an important resource for lumber and

paper industries and a potential source of biomass for production of liquid biofuel. Wood in dicot species is composed of xylem fiber, vessel and ray parenchyma cells which possess thick secondary cell walls that form after cessation of cell expansion. The vascular tissue in stems, leaves and roots is formed through cell division and differentiation from the meristematic procambium which is generated by the shoot apical meristem during primary growth. The procambium cells divide and differentiate to produce primary xylem and phloem tissues (Jung and Park, 2007). During primary growth, when the stem is elongating, clusters of primary xylem and phloem fiber and vessel cells derived from the fascicular procambium form to make vascular bundles which are arrayed in a circular array around the stem axis (Carlsbecker and Helariutta, 2005). After differentiation and cell expansion the fibers and vessels of the xylem and phloem tissues undergo secondary cell wall formation and programmed cell death. In the transition from primary to secondary growth, the fascicular and interfascicular cambiums merge to form a continuous ring of cells that is the vascular cambium halting stem elongation but initiating stem expansion. The vascular cambium is composed of two cell types, fusiform cambial cells (FCCs) and ray cambial cells (RCCs) from which all cells produced in secondary growth derive. During secondary growth FCCs divide adaxially to produce xylem mother cells and abaxially to produce phloem mother cells. The xylem and phloem mother cells subsequently differentiate into vessel and fibers which expand radially and axially before undergoing secondary cell wall formation followed by programmed cell death (Nieminen et al., 2004). The RCCs divide to produce phloem and xylem ray parenchyma cells which elongate radially relative to the stem axis proceeded by secondary cell wall deposition. Xylem ray parenchyma cells further differentiate into contact cells, which form connections with adjacent vessel elements and facilitate transport between vessel elements, or isolation cells, which facilitate radial transport but do not form

connections with vessel elements (Murakami et al., 1999). Interconversion of FCCs and RCCs is commonly observed and believed to occur in response to an unknown positional signal (Mellerowicz et al., 2001). In tree species the production of xylem and phloem from the vascular cambium occurs in annual cycles to generate wood and bark respectively. The formation of secondary cell walls in the xylem and phloem cells enables the fibers to provide strong support and the vessels to facilitate transport throughout the plant.

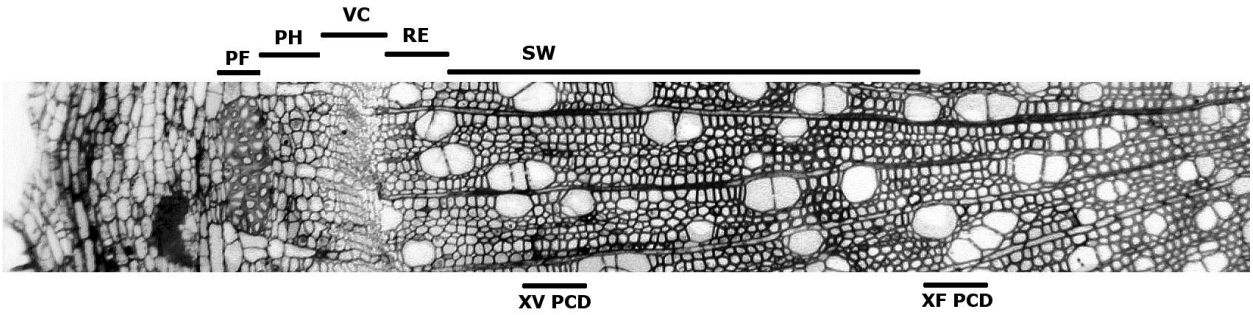

Figure 1. Transverse section of poplar stem showing cells differentiated from the vascular cambium during secondary growth. The vascular cambium (VC) divides abaxially to produce the phloem (PH) and phloem fibers (PF). The xylem cells divide adaxially from the VC, undergo a period of radial expansion (RE) followed by secondary cell wall formation (SW). Xylem vessel programmed cell death (XV PCD) occurs quickly following secondary cell wall formation while xylem fiber programmed cell death (XF PCD) occurs comparatively later.

Secondary cell wall formation

The principle components of secondary cell walls are cellulose, hemicellulose and lignin. The exact proportion of these components is tightly linked to the cell's function and can vary

between cell types. The secondary cell wall is formed between the plasma membrane and the primary cell wall by deposition of three layers called S1, S2 and S3 with S1 being the thinnest and S2 being the thickest. The formation of the secondary cell wall begins with cellulose deposition, followed closely by hemicellulose and finally by the initiation of lignification before the cell undergoes programmed cell death after which lignification may continue (Plomion et al., 2001).

Cellulose

Cellulose is a linear polymer of β-linked glucose subunits. Hydrogen bonding between parallel chains of cellulose brings them together to form microfibrils. The biosynthesis of cellulose is performed by several members of the cellulose synthase (CesA) gene family. The proteins encoded by the CesA genes form a large protein complex made up of 36 CesA proteins with the individual proteins being grouped into a rosette shaped hexamer which are further grouped into a larger rosette shaped six member ring (Somerville, 2006). In *Arabidopsis thaliana* there are ten CesA genes. Three of the CesA genes, CesA4, CesA7 and CesA8, are involved in the process of cellulose biosynthesis during secondary cell wall formation while a separate set of CesA genes perform cellulose biosynthesis for the primary cell wall (Cosgrove, 2005). Mutations in these three genes are responsible for a mutant phenotype characterized by weak xylem vessels that collapse under the negative pressure induced by transpiration and are labeled the *irregular xylem* (*irx*) mutants (Taylor, 2008). Additional mutant alleles of CesA7 and CesA8 produced the *fragile fiber* (*fra*) mutants, *fra5* and *fra6* respectively, with reduced secondary cell wall strength and cellulose content in fiber cells (Zhong et al., 2003). Genome duplication in the poplar lineage have led to an expansion of the CesA gene family to 17

members, 5 of which are believed to specifically function in secondary cell wall formation (Kumar et al., 2009).

The CesA complex is anchored spanning the plasma membrane where it secretes its cellulose microfibril outside the membrane. The deposition pattern of the microfibrils is guided by cortical microtubules with cellulose synthase interactive protein 1 facilitating the interaction between the CesA complex and the guiding microtubules (Li et al., 2012). Whether the movement of the CesA complex along microtubules is driven by motor proteins or through the force exerted by cellulose microfibril biosynthesis remains unresolved.

The cellulose microfibrils in the secondary cell wall are arranged in a parallel fashion. The S1, S2 and S3 layers differ in the angle between the microfibril and long axis of the cell. The angle at which the microfibrils are arranged is important to the mechanical rigidity and elasticity of the cell's secondary cell wall (Plomion, Leprovost and Stokes, 2001). The microfibril angle of deposition found in the layers of the secondary cell wall can vary between cell types and in response to environmental factors such as mechanical stress (Barnett and Bonham, 2004).

Hemicellulose

In stark contrast to the simple linear homoglycan structure of cellulose, hemicellulosic polysaccharides represent a wide array of complex polysaccharides that often possess sidechains. The biosynthesis of hemicellulose occurs in the Golgi by a variety of carbohydrate active enzymes (CAZy) which synthesize the unbranched polysaccharide backbone while simultaneously adding sidechains (Cosgrove, 2005). Following biosynthesis, hemicellulose is transported to the plasma membrane via Golgi vesicles and exported through exocytosis. After

being exported to the secondary cell wall, hemicellulose functions by associating with and cross linking the parallel cellulose microfibrils (Lerouxel et al., 2006). Xylan is the primary hemicellulose in dicotyledonous plant secondary cell walls while arabinoxylans are the most abundant hemicellulose in monocotyledonous plants demonstrating some divergence in cell wall formation (Pauly and Keegstra, 2008). Mannans comprise the major hemicellulose of bryophyte primary cell walls and are a significant component in the secondary cell walls of lycopods, but only exist at low concentrations in higher vascular plants (Scheller and Ulvskov, 2010). Mutation of a CAZy member of glycosyltransferase (GT) family 47 induced a specific error in the biosynthesis of xylan that produced a fragile fiber phenotype, denoted *fra8* (Zhong et al., 2005). Additional GT and CAZy proteins have been implicated in hemicellulose biosynthesis for secondary cell walls but the complicated structure and resulting complicated biosynthesis of hemicelluloses has left many details to still be discovered.

Lignin

Lignin is a phenolic heteropolymer composed of monolignols that impart increased strength and waterproofing to the secondary cell wall. Monolignols are synthesized from the amino acid phenylalanine in a process that also supplies precursors for the phenylpropanoid pathway. The biosynthetic steps involved in the conversion of phenylalanine to monolignols have been extensively studied, and enzymes in the biosynthetic pathway have been well characterized (Umezawa, 2009). The mechanism for lignin export through the plasma membrane remains unsolved, but evidence suggests that the export is facilitated by an ATP dependent membrane transporter (Kaneda et al., 2008; Miao and Lui, 2010). Lignin polymerization in the secondary cell wall occurs through a radical coupling reaction. Several

classes of proteins including peroxidases, laccases and phenol oxidases have been proposed to facilitate lignin polymerization but direct evidence has yet to be provided due to the large size of these protein families making mutational analysis difficult (Marjamaa, et al., 2009; Vanholme et al., 2010). In the secondary cell wall, lignin polymers associate with the polysaccharide matrix and provide additional mechanical strength. The hydrophobic nature of lignin makes the secondary cell wall impenetrable to water, which facilitates the transport function of vessel cells that characteristically have high lignin contents. Lignin also plays an important role in pathogen resistance since its highly variable structure and order of subunits makes enzymatic degradation difficult (Donaldson, 2001). Monolignols are classified as H, G or S depending on their chemical structure. While lignin-like molecules are found even in some unicellular algae, the polymerization of lignin in the secondary cell wall distinguishes vascular plants from the rest of the plant kingdom (Popper et al., 2011). Intriguingly, lignified secondary cell walls were found in the red algae *Calliarthron cheilosporioides* and are believed to have independently arisen through convergent evolution (Martone et al., 2009). H and G lignin occur ubiquitously in secondary cell walls across all vascular plant lineages, although the ratio of each component varies among species. However, the occurrence of S lignin is less uniform being notably absent in gymnosperms. Interestingly, the biosynthetic machinery for producing S monolignols appears to have evolved separately from cytochrome P450 dependent monooxygenases of independent origin in angiosperms, the lycopod *Selaginella moellendorffii* and the red algae *Calliarthron cheilosporioides* (Weng et al., 2008; Martone et al., 2009). A handful of ferns and gymnosperms have been reported in older literature to possess S monolignols but these claims have yet to be evaluated using modern techniques (Weng and Chapple, 2010).

Programmed cell death

To perform their functions in water transport and structural support many of the cell types that constitute wood must undergo programmed cell death (PCD). Xylem vessels, fibers and ray parenchyma cells all can undergo PCD; however the prevalence and developmental timing of this process can vary widely between these cell types. The process of PCD in xylem vessels immediately follows secondary cell wall deposition and leaves a hollow cell with perforated ends that link end to end with other xylem vessels to create a conduit for water transport. PCD in fibers occurs in a more delayed manner often lagging behind secondary cell wall deposition by up to one month (Fig. 1)(Bollhoner et al., 2012). Xylem ray parenchyma cells often remain alive, postponing program cell death for several years or indefinitely and function in both radial transport and nutrient storage (Nakaba et al., 2006).

The process of PCD occurs with some variations between xylem vessels and fibers. During their differentiation from the vascular cambium both cell types accumulate a vacuolar mixture of autolytic enzymes including cysteine proteases, meta-caspases and nucleases.

The swelling of the vacuole accompanied by a change in tonoplast permeability are the first steps of PCD in xylem vessels and prefaces vacuolar collapse (Kuriyama, 1999). The autolytic enzymes released from the vacuole rapidly degrade cytoplasmic organelles and nuclear DNA. In the final step of xylem vessel PCD enzymes hydrolyze regions of the primary cell wall not protected by lignified secondary cell wall. This primary cell wall hydrolysis is particularly essential at the ends of coupled xylem vessels where the absence of lignified secondary cell walls allows the creation of the perforation plate, an opening between the now hollow xylem vessels that allows them to function as a continuous tube (Fukuda, 1997). The perforation plate of

poplar xylem vessels are simple as they form a single unobstructed opening between cells whereas compound perforation plates are composed of multiple openings (Groover et al., 2010).

The first indication of PCD in xylem fibers is the degradation of nuclear DNA which occurs well before the cell undergoes autolysis. This degradation is accompanied by a slow hydrolysis of the cytoplasmic contents. Only once these processes have significantly progressed does the vacuole swell and burst rapidly completing the autolysis of any undegraded cellular contents. While the cellular debris of xylem vessels is rapidly removed following PCD, in xylem fibers the debris may remain for an extended period before being removed.

Cysteine proteases, metacaspases and nucleases are the primary proteins involved in the autolytic processes carried out during PCD; however many details remain poorly understood. XYLEM CYSTEINE PEPTIDASE1 (XCP1) and XCP2 are cysteine proteases that are expressed specifically in xylem vessel elements. XCP1 and XCP2 are vacuole localized and participate both in micro-autolysis of cellular components trafficked to the vacuole and cytoplasmic autolysis following vacuolar bursting. Consistent with this role, *xcp1 xcp2* double mutants demonstrated a xylem vessel specific accumulation of intact cellular components (Avci et al., 2008). The *Arabidopsis* metacaspase AtMC9 is expressed specifically in developing xylem cells. Interestingly the two poplar homologs of AtMC9 appear to have undergone sub-functionalization with one being expressed specifically in xylem vessels and the other expressed in both xylem vessels and fibers (Bollhoner et al., 2012). The absence of XCP1, XCP2 or MC9 expression in *Arabidopsis* xylem fibers accounts for the fact that xylem fibers of *Arabidopsis* generally do not undergo PCD and demonstrates an interesting developmental difference between herbaceous and woody dicots. The degradation of nuclear DNA in developing xylem is

believed to be accomplished in *Arabidopsis* by the action of BFN1, a nuclease that also functions in leaf senescence and fruit abscission (Farage-Barhom et al., 2008).

Hormonal regulation of wood formation

The plant hormones auxin, cytokinin and ethylene all have been implicated as important to the process of wood formation primarily through their regulation of vascular cambium maintenance and activity. Exogenous application of plant hormones has long been known to effect secondary growth through induction of cambial cell proliferation, but only recently have studies begun to utilize hormone signaling mutants to identify the components of this pathway.

Auxin

The hormone auxin is integrally important to numerous plant processes including vascular patterning and xylogenesis. Polar transport of auxin by the transporter protein PIN1 down from the shoot apical meristem induces formation of the procambium (Carlsbecker and Helariutta, 2005). Similarly, the flow of auxin through the leaf establishes procambium which will divide and differentiate to produce a network of veins (Benjamins and Scheres, 2008). Defects in auxin polar transport or auxin sensing are sufficient to create patterns of discontinuous veins in leaves and severe reductions in vascular cambium cell division (Nilsson et al., 2008). The distribution of the auxin indole-3-acetic acid (IAA) in a stem cross section is characterized by a concentration maximum in the vascular cambium. The auxin concentration gradient rapidly decreases in the differentiating phloem and decreases more gradually in the developing xylem tissue, slowly diminishing concomitant with xylem cell secondary cell wall formation (Moyle et al., 2002). While disruptions of this auxin gradient have been demonstrated to impact both

vascular cambium activity and xylem development the molecular details of auxin's role remains poorly understood (Mellerowicz et al, 2001).

Gibberellin

Gibberellins (GAs) are known regulators of many plant processes but have also been implicated as regulators of cambial activity. GA and IAA function cooperatively to stimulate cambial activity as simultaneous application of both hormones induces cambial cell divisions at a greater level than treatment with IAA or GA alone. Since GA treatment has also been observed to stimulate polar auxin transport it is unclear whether GA effects on cambial activity are direct or mediated through its effect on auxin (Bjorklund et al., 2007). Gibberellin concentration is highest in developing xylem tissue raising the possibility that it may be the signal responsible for activating the polar auxin transport that maintains high auxin concentration in the vascular cambium (Israelsson et al., 2005).

Cytokinin

Although auxin is an important agent for vascular tissue formation and meristem maintenance it acts in coordination with other plant hormones including cytokinin. The fact that auxin concentration in the vascular cambium remains constant during the seasonal fluctuations in cambial activity indicates the necessity of another signal for modulating cambial activity (Nilsson et al., 2008). Evidence has suggested that cytokinin concentration supplies this other signal as *Arabidopsis* mutants with TDNA insertions in cytokinin biosynthetic genes which exhibit decreased cytokinin concentration demonstrate a reduction in secondary xylem and phloem due to reduced vascular cambium activity. Furthermore, overexpressing of cytokinin

biosynthetic genes caused heightened vascular cambium activity producing elevated secondary growth correlating to the expression level of the transgene (Matsumoto-Kitano et al., 2008). Similar findings that cytokinin induces vascular cambium activity, increasing cell proliferation to produce additional secondary growth have been observed in poplar (Nieminen et al., 2008)

Ethylene

Ethylene has long been recognized as important in many plant processes including abscission and senescence. Recent findings indicate that ethylene plays an important role in wood formation by promoting secondary xylem growth through cambial cell division. Interestingly while exogenous and endogenous ethylene resulted in an increased quantity of secondary xylem cells, the cell size was decreased especially in xylem vessels (Love et al., 2009). Whether the cell size decrease is a direct effect of ethylene signaling or a secondary effect precipitated by the rapid cambial proliferation triggering a shortening of the radial expansion period is currently unknown.

Transcriptional regulation of secondary wall formation

Proper formation of secondary cell walls requires a multitude of enzymes in several biosynthesis pathways to operate in concert and in a highly cell specific manner to meet the functional requirements of each cell type. Such choreography requires precise regulation involving a network of transcription factors responding to positional cues. With the first players of this transcriptional network only being recently discovered we are just beginning to appreciate the vastness of this regulatory network.

Of chief importance to our current understanding of secondary cell wall transcriptional regulation are a number of transcription factors belonging to the NAM, ATAF and CUC (NAC) and MYELOBLASTOSIS (MYB) domain families of transcription factors.

The NAC domain family of transcription factors is unique to plants and participates in a wide range of plant developmental processes including shoot apical meristem formation and abiotic stress response. The majority of NAC proteins possess a conserved N-terminal dimerization domain that facilitates homo and heterodimerization between some members of the NAC domain family. The regulation of NAC domain proteins can be rather complex as examples of regulation by miRNA as well as ubiquitin mediated proteolysis has been illustrated (Olsen et al., 2005).

The closely related NAC domain proteins NAC SECONDARY WALL THICKENING PROMOTING FACTOR1 (NST1), NST2, NAC SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1 (SND1; also called NST3), VASCULAR-RELATED NAC-DOMAIN6 (VND6), and VND7 activate similar direct genetic targets and have been shown to be master regulators able to induce all pathways required for secondary wall formation (Kubo et al., 2005; Mitsuda et al., 2005, 2007; Zhong et al., 2006, 2007, 2008). These NAC domain proteins function as redundant pairs in distinct cell types with SND1 and NST1 acting in fibers (Mitsuda et al., 2007; Zhong et al., 2006, 2007b), NST1 and NST2 acting in anther endothecium cells (Mitsuda et al., 2005) and VND6 and VND7 acting in root and stem vessels (Kubo et al., 2005; Zhong et al., 2008). Collectively these NAC domain master regulators of secondary cell wall formation are referred to as the Secondary Wall NACs (SWNs). Whether a similar mechanism of NAC domain protein pairs also governs secondary cell wall formation in other tissues forming secondary cell walls such as the phloem vessels has yet to be explored.

While NST1 and SND1 single knockout mutants exhibit wild-type secondary cell wall formation in fiber cells, nst1/snd1 double knockout mutants lack nearly all fiber cell secondary cell wall formation and as a result display a severe reduction in stem rigidity. Similar to the effect that knocking out NST1 and SND1 has on fibers, double knockout of NST1 and NST2 abolishes secondary cell wall formation in the anther endothecium leading to an indehiscent phenotype (Mitsuda et al, 2005).

The MYB domain family of transcription factors is represented in all major eukaryotic lineages but has demonstrated significantly greater expansion in plant relative to animal genomes (Stracke et al., 2001). MYB transcription factors can be subdivided into three subgroups with the R2R3 MYB proteins, possessing two adjacent repeats in the MYB domain, being the most prevalent. The R2R3 MYB domain protein MYB46 is a direct target of the NAC domain master regulator proteins and is itself an activator protein capable of inducing all pathways required for secondary wall formation (Zhong et al., 2007a; Zhong and Ye, 2009). While constitutive overexpression of MYB46 produces ectopic secondary cell wall formation, dominant repression of MYB46 causes dramatic secondary wall thinning (Zhong and Ye, 2007; Zhong and Ye, 2009). However, MYB46 TNDA knockout mutants exhibited no phenotype indicating that, similar to all of the NAC domain master regulatory proteins, MYB46 may function redundantly with another protein. The role of MYB46 in secondary cell wall formation in all cell types provides compelling evidence that this gene operates as a universal switch controlling secondary cell wall formation.

In addition to MYB46, the SWNs directly target a number of transcription factors. These targets of the master regulators of secondary wall biosynthesis collectively activate the biosynthetic pathways, including cellulose, hemicellulose and lignin, required for secondary cell

wall formation. Individually many of these direct targets induce a subset of the genes necessary for secondary cell wall formation. MYB54 and MYB52 activate genes involved in cellulose, xylan and lignin biosynthesis but are unable to induce ectopic secondary cell wall formation indicating that genes involved in the transport and deposition are not activated. The homologous MYB transcription factors MYB58 and MYB63 are specific activators of lignin biosynthesis and deposition. Both MYB 58 and MYB63 are directly bound and activated by the SWNs and MYB46 establishing a feed-forward loop in the network topology regulating MYB58 and MYB63 (Zhou et al., 2009). MYB85 also acts specifically as an activator of the lignin pathway capable of inducing ectopic lignification (Zhong et al., 2008). Three additional SWN direct targets, MYB103, SND2 and SND3 are specific activators of cellulose biosynthesis.

The activation of the SWN direct targets is facilitated by SWN binding to the Secondary wall NAC Binding Element (SNBE). The SNBE is an imperfect palindromic 19-bp sequence present in the promoters of many genes known to be involved in secondary cell wall formation. While all SWNs demonstrated an ability to bind to the SNBE sites, AtSND1 and AtVND7 induce different, but overlapping, sets of target genes. Additionally, although some genes not induced by SWNs were found to contain SNBE sites in their promoters, the number of SNBE sites correlated with activation by SWNs (Zhong et al., 2010b).

Interestingly it appears that the process of programmed cell death is at least partially co-regulated with secondary cell wall formation as both SND1 and VND7 are capable of inducing a number of genes responsible for programmed cell death including XCP1, XCP2, MC9 and BFN1 (Zhong et al., 2010b). Differences in the degree to which SND1 and VND7 activate these cell death genes may account for the disparity in timing of programmed cell death between xylem vessels and fibers.

Our current understanding of the transcriptional network controlling secondary cell wall formation has enabled us to propose a model where a tissue specific positional signal selectively activates a subset of the cell type specific activators which in turn activate a mostly shared set of transcription factors which turn on all the downstream targets required for proper secondary cell wall formation (Fig. 2). The exact nature of the tissue specific positional signal responsible for activating the cell type specific SWNs and whether it is linked to the hormone signals regulating vascular cambium activity is unknown.

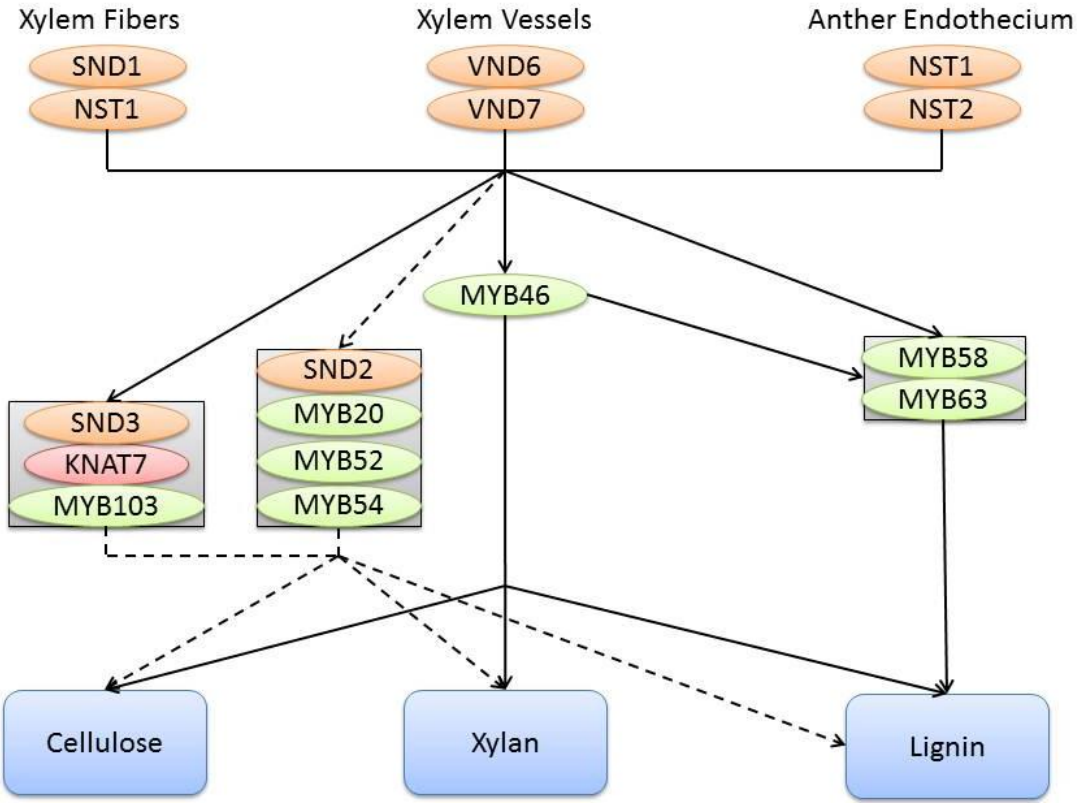

Figure 2. Diagram of the transcriptional network regulating secondary cell wall formation. Solid lines indicate experimentally verified direct activation. Dashed lines indicate indirect activation or unverified activation.

Secondary wall composition and biofuel production

The increase in atmospheric CO_2 and the resulting climate change has created a major push encouraging the development of a carbon neutral fuel source. One possible solution to meet this demand is biofuel derived from wood biomass. Biofuel has the potential to be a carbon neutral fuel source where the CO_2 released from its combustion is offset by the CO_2 fixed from the atmosphere through photosynthesis. First generation biofuels are produced by processing sugar rich foodstock, such as ears of corn, into biofuels including ethanol. This process is fairly simple due to the availability of fermentable sugars in corn kernels. However, the production of first generation biofuel is unsustainable due to relatively low biomass yields and the detrimental effect on food prices. Second generation biofuels refers to biofuels derived from processing whole plant biomass such as woody tissue (Naik et al., 2010). The principle advantage of the second generation biofuel approach is that the usable biomass yield is much higher than first generation biofuels. However, the material for producing second generation biofuels is more recalcitrant than foodstocks and requires additional processing before it can be fermented into a combustible fuel. To make second generation biofuels a sustainable and commercially viable alternative to fossil fuels it will be necessary to develop energy crops that rapidly produce a large quantity of biomass that is modified to simplify its processing into biofuel (Carroll and Somerville, 2009). One potential energy crop for second generation biofuels is the poplar tree. Poplar trees rapidly produce biomass in the form of wood and grow well in dense formations.

The function of the secondary wall does not rely on its independent constituents. Instead its strong compact structure is produced through the cross-linking of cellulose microfibrils, hemicellulose chains and lignin polymers. The cellulose bound in this compact structure is termed crystalline cellulose and is characterized by its recalcitrant nature. The release of

fermentable sugars from the cell wall is achieved by cellulases and glycosyl hydrolases which break down cellulose and hemicellulose into monosaccharides; with cellulose being the primary contributor. The inaccessibility of crystalline cellulose to enzymatic hydrolysis greatly decreases digestion efficiency and increases processing costs due to the necessity of pretreatment to render the cellulose more accessible (Joeh et al., 2007). Several studies employing genetic and chemical approaches have successfully decreased the level of cellulose crystallinity which improves the efficiency of enzymatic digestion of biomass into fermentable sugars by making cellulose microfibrils more accessible to cellulase (Harris et al., 2009; Abramson et al., 2010).

Arabidopsis and Poplar as models for investigating wood formation

Arabidopsis has been successfully utilized as a model for studying wood formation in many studies due to the high similarity between xylem fiber and vessel cells found in the herbaceous *Arabidopsis* to those that are found in the woody stems of trees. Structurally and developmentally xylem vessels and fibers show high similarity between poplar and *Arabidopsis* both in how they differentiate from plant meristems and the processes of secondary cell wall formation (Chaffery et al., 2002). Furthermore, the hormone signals involved in vascular formation and many gene members of the xylem transcriptome are highly conserved between *Arabidopsis* and woody species such as poplar (Zhang et al., 2011). The extensive number of genes involved in cellulose, xylan and lignin biosynthesis identified and characterized in *Arabidopsis* provides an ideal platform for identifying the genes responsible for regulating these downstream factors. Despite the similarities between xylem tissue in *Arabidopsis* and woody dicots there are physiological differences which may be the result of key regulatory and developmental differences. To truly understand the process of wood formation it is essential that

studies be conducted in woody species. As the first woody dicot genome to be sequenced *Populus trichocarpa* is an excellent model organism for investigating wood formation (Tuskan et al., 2006). The relatively close phylogenetic relationship between *Arabidopsis* and *Populus* and corresponding high percentage of homologous genes (Jansson and Douglas, 2007) increases the potential for information gained in *Arabidopsis* to inform our investigations of wood formation in poplar.

Research goals and brief summary of dissertation work

Recently our understanding of the process of secondary cell wall formation has been rapidly expanding. Studies in *Arabidopsis* have successfully identified a large number of transcription factors responsible for controlling the multitude of processes necessary for proper secondary cell wall formation. Many aspects of this process remain undiscovered. To fully understand the regulatory network controlling secondary cell wall formation it is necessary to both identify currently unknown transcription factors and determine which aspects of the network are conserved and which are variable among the vascular plant lineages. My research has identified additional key transcription factors in the secondary cell wall regulatory network and investigated the conservation of this network between *Arabidopsis*, *Populus*, and diverse members of the plant kingdom.

The absence of phenotype in MYB46 TDNA knockout mutants suggested that a functionally redundant protein existed which was able to compensate for the loss of MYB46. MYB83 was identified as both an uncharacterized MYB transcription factor with high xylem specific expression and a close homolog of MYB46. Similar to MYB46, MYB83 was determined to be a master regulatory of secondary cell wall formation and a direct target of the

SWNs. Additionally, MYB46 MYB83 TDNA double knockout mutants were shown to exhibit a severely stunted phenotype produced by a complete inability to form secondary cell walls in vascular tissue. Our findings presented in Chapter 2 identify MYB83 as a key master regulator of secondary cell wall formation that acts redundantly with MYB46.

The addition of MYB83 to the group of secondary cell wall master regulatory genes, including the SWNs and MYB46, defined what appears to be the core of the regulatory network regulating secondary cell wall formation in *Arabidopsis*. Extending our knowledge from *Arabidopsis* to a woody species like poplar provides insights into the conservation between herbaceous and woody plants of the transcriptional network regulating secondary cell wall formation. Additionally, knowledge gained in poplar may have direct applications for producing wood with enhanced traits for paper, lumber or biofuel production. Four close homologs of *Arabidopsis* MYB46 and MYB83 were identified in the poplar genome. PtrMYB003 and PtrMYB020 were successfully amplified from poplar cDNA and their function was characterized in *Arabidopsis*. Both PtrMYB003 and PtrMYB020 were demonstrated to function as secondary cell wall formation master regulators in *Arabidopsis* and were capable of complementing the *myb46 myb83* double TDNA knockout mutant phenotype. These findings presented in Chapter 3 provided strong evidence that the roles of the MYB domain master regulatory proteins in the secondary cell wall regulatory network is highly conserved between *Arabidopsis* and poplar.

The discovery of functional equivalents to the *Arabidopsis* MYB46 and MYB83 genes in poplar established that at least a portion of the core master regulatory proteins for secondary cell wall formation were conserved between *Arabidopsis* and poplar. Further evidence of this conservation was provided with the discovery of the poplar WNDs which act as functional equivalents to the *Arabidopsis* SWNs (Zhong et al., 2010a; Zhong et al., 2011). The

demonstrated ability of poplar WNDs and MYB secondary cell wall master regulators to function in *Arabidopsis* indicates that their mechanism of function and target binding site sequence is conserved. Intrigued by the evidence demonstrated in *Arabidopsis* and poplar we wished to determine the extent to which the secondary cell wall regulatory network was conserved among the vascular plants. By investigating the conservation of this regulatory network we hope to gain insight into the evolution of secondary cell wall formation as it relates to the emergence and radiation of vascular plants. To this aim we analyzed numerous plant genome sequences for putative SNBE site sequences and report our findings in Chapter 4. SNBE sites were found to be significantly enriched in gene promoters in vascular plant species when compared with non-promoter sequences and gene promoters from the non-vascular moss *Physcomitrella patens*. Additionally putative SNBE sites were significantly enriched in promoters of genes homologous to *Arabidopsis* genes involved in secondary cell wall formation. Further sequence comparisons of putative SNBE sites between *Arabidopsis thaliana* and the closely related *Arabidopsis lyrata* provided statistical evidence that positive selection is acting to conserve SNBE binding sites in promoters of genes activated by SWNs.

Determining the master regulators of secondary cell wall formation in *Arabidopsis* and then demonstrating the conservation of the core of this regulatory network in poplar and other vascular plants provides insight into the evolution of plant vascular systems as well as providing tools for improving plants for applications such as biofuels. The demonstrated conservation of the network regulating secondary cell wall formation provides tools for improvement of a diverse set of important plant species. With this knowledge it will be possible to better develop plants with improved secondary cell wall characteristics for a variety of applications including biofuel production.

References

Abramson, M., Shoseyov, O., Shani, Z. (2010) Plant cell wall reconstruction toward improved lignocellulosic production and processability. Plant Sci. 178:61-72

Avci, U., Petzold, H.E., Ismail, I.O., Beers, E.P., Haigler, C.H. (2008) Cysteine proteases XCP1 and XCP2 aid micro-autolysis within the intact central vacuole during xylogenesis in *Arabidopsis*. Plant J. 56:303-315

Barnett, J.R., Bonham, V.A. (2004) Cellulose microfibril angle in the cell wall of wood fibres. Biol. Rev. 79:461-472

Benjamins, R., Scheres, B. (2008). Auxin: the looping star in plant development. Annu. Rev. Plant Biol. 59:443-465

Bhalerao, R., Nilsson, O., Sandberg, G. (2003) Out of the woods: forest biotechnology enters the genomic era. Curr. Opin. Plant Biol. 14:206-213

Bjorklund, S., Antti, H., Uddestrand, I., Moritz, T., Sundberg, B. (2007) Cross-talk between gibberellin and auxin in development of *Populus* wood: gibberellin stimulates polar auxin transport and has a common transcriptome with auxin. Plant J. 52: 499-511

Bollhoner, B., Prestele, J., Tuominen, H. (2012) Xylem cell death: emerging understanding of regulation and function. J. Exp. Bot. 63(3): 1081-1094

Carlsbecker, A., Helariutta, Y. (2005) Phloem and xylem specification: pieces of the puzzle emerge. Curr. Opin. Plant Biol. 8: 512-517

Chaffey, N., Cholewa, E., Regan, S., Sundberg, B. (2002) Secondary xylem development in *Arabidopsis*: a model for wood formation. Physiologia Plantarum 114:594-600

Cosgrove, D.J. (2005). Growth of the plant cell wall. Nature Rev. 6:850-861

Dettmer, J., Elo, A., Helariutta, Y. (2009). Hormone interactions during vascular development. Plant Mol. Biol. 69:347-360

Farage-Barhom, S., Burd, S., Sonego, L., Perl-Teves, R., Lers, A. (2008) Expression analysis of the BFN1 nuclease gene promoter during senescence, abscission, and programmed cell death-related processes. J. Exp. Bot. 59(12):3247-3258

Fukuda, H. (1997) Tracheary element differentiation. Plant Cell 9:1147-1156

Groover, A.T., Nieminen, K., Helariutta, Y., Mansfield, S.D. (2010) Wood formation in *Populus*. In: Jansson, Bhalerao, Groover, eds. Genetics and genomics of *Populus*. Dordrecht, Netherlands: Springer

Harris, D., Stork, J., Debolt, S. (2009) Genetic modification in cellulose-synthase reduces crystallinity and improves biochemical conversion to fermentable sugar. GCB Bioenergy 1:51-61

Israelsson, M., Sundberg, B., Moritz, T. (2005) Tissue-specific localization of gibberellins and expression of gibberellin-biosynthetic and signaling genes in wood forming tissues in aspen. 44: 494-504

Jansson, S., Douglas, C.J. (2007) *Populus*: A model system for plant biology. Annu. Rev. Plant Biol. 58:435-458

Jung, J.H., Park, C.M. (2007). Vascular development in plants: specification of xylem and phloem tissues. J. Plant Biol. 50:301-305

Kaneda, M., Rensing, K.H., Wong, J.C.T., Banno, B., Mansfield, S.D., Samuels, A.L. (2008). Tracking monolignols during wood development in lodgepole pine. Plant Physiol. 147:1750-1760

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., Mimura, T., Fukuda, H., Demura, T. (2005) Transcription switches for protoxylem and metaxylem vessel formation. Genes and Dev. 19:1855-1860

Kumar, M., Thammannagowda, S., Bulone, V., Chiang, V., Han, K.-H., Joshi, C.P., Mansfield, S.D., Mellerowicz, E., Sundberg, B., Teeri, T., Ellis, B.E. (2009) An update on the nomenclature for the cellulose synthase genes in Populus. Trends in Plant Sci. 14(5): 248-254

Kuriyama, H. (1999) Loss of tonoplast integrity programmed in tracheary element differentiation. Plant Physiol. 121:763-774

Lerouxel, O., Cavalier, D.M., Liepman, A.L., Keegstra, K. (2006). Biosynthesis of plant cell wall polysaccharides – a complex process. Curr. Opin. Plant Biol. 9:621-630

Martone, P.T., Estevez, J.M., Lu, F., Ruel, K., Denny, M.W., Somerville, C., Ralph, J. (2009) Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. Curr. Biol. 19:169-175

Marjamaa, K., Kukkola, E.M., Fagerstedt, K.V. (2009). The role of xylem class III peroxidases in lignification. J. Exp. Bot. 60:367-376

Matsumoto-Kitano, M., Kusumoto, T., Tarkowski, P., Kinoshita-Tsujimura, K., Vaclavikova, K., Miyawaki, K., Kakimoto, T. (2008) Cytokinins are central regulators of cambial activity. Proceedings of the National Academy of Sciences 105(50):20027-20031

Mellerowicz, E.J., Baucher, M., Sundberg, B., Boerjan, W. (2001) Unravelling cell wall formation in the woody dicot stem. Plant Mol. Biol. 47:239-274

McCarthy, R.L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., Spicer, C., Ye, Z.-H. (2010) The poplar MYB transcription factors PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. Plant Cell Physiol. 51:1084-1090

McCarthy, R.L., Zhong, R., Ye, Z.-H. (2009). MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. Plant Cell Physiol. 50:1-15

Miao, Y.-C., Lui, C.-J. (2010) ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. Proceedings of the National Academy of Sciences 107(52): 22728-22733

Mitsuda, N., Seki, S., Shinozaki, K., Ohme-Takagi, M. (2005) The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. Plant Cell 17:2993-3006

Moyle, R., Schrader, J., Stenberg, A., Olsson, O., Saxena, S., Sandberg, G., Bhalerao, R.P. (2002) Environmental and auxin regulation of wood formation involves members of the Aux/IAA gene family in hybrid aspen. Plant J. 31(6):675-685

Murakami, Y., Funada, R., Sano, Y., Ohtani, J. (1999) The differentiation of contact cells and isolatin cells in the xylem ray parenchyma of Populus maximowiczii. Annuals Bot. 84:429-435

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K. (2010) Production of first and second generation biofuels: A comprehensive review. Renewable Sustainable Energy Reviews 14:578-597

Nakaba, S., Sano, Y., Kubo, T., Funada, R. (2006) The positional distribution of cell death of ray parenchyma in a conifer, *Abies sachalinensis*. Plant Cell Rep. 25:1143-1148

Nieminen, K., Immanen, J., Laxell, M., Kauppinen, L., Tarkowski, P., Dolezal, K., Tahtiharju, S., Elo, A., Decourteix, M., Ljung, K., Bhalerao, R., Keinonen, K., Albert, V.A., Helariutta, Y. (2008) Cytokinin signaling regulates cambial development in poplar. Proceedings of the National Academy of Sciences 105(50):2032-20037

Nieminen, K.M., Kauppinen, L., Helariutta, Y. (2004) A weed for wood? Arabidopsis as a genetic model for xylem development. Plant Physiol. 135: 653-659

Nilsson, J., Karlberg, A., Antti, H., Lopez-Vernaza, M., Mellerowicz, E., Perrot-Rechenmann, C., Sandberg, G., Bhalerao, R.P. (2008) Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. Plant Cell 20: 843-855

Olsen, A.N., Ernst, H.A., Leggio, L.L., Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. Trends in Plant Sci. 10:79-87

Pauly, M., Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J. 54:559-568

Plomion, C., Leprovost, G., Stokes, A. (2001). Wood formation in trees. Plant Physiol. 127:1513-1523

Popper, Z.A., Michel, G., Herve, C., Domozych, D.S., Willats, W.G.T., Tuohy, M.G., Kloareg, B., Stengel, D.B. (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. Annu. Rev. Plant Biol. 62: 567-590

Scheller, H.V., Ulvskov, P. (2010) Hemicelluloses. Annu. Rev. Plant Biol. 61, 263-289

Somerville, C. (2006). Cellulose synthesis in higher plants. Annu. Rev. Cell Dev. Biol. 22:53-78

Tuskan, G.A. et al. (2006) The genome of black cottonwood *Populus trichocarpa*. Science 313:1596-1604

Umezawa, T., (2010). The cinnamate/monolignol pathway. Phytochem Rev. 9:1-17

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., Boerjan, W. (2010) Lignin biosynthesis and structure. Plant Physiol. 153:895-905

Weng, J.-K., Li, X., Stout, J., Chapple, C. (2008) Independent origins of syringyl lignin in vacular plants. Proceedings of the National Academy of Sciences 105(22), 7887-7892

Weng, J.-K., Chapple, C. (2010) The origin and evolution of lignin biosynthesis. New Phytologist 187, 273-285

Ye, Z. (2002) Vascular tissue differentiation and pattern formation in plants. Annu. Rev. Plant Biol. 53:182-202

Zang, J., Elo, A., Helariutta, Y. (2011) *Arabidopsis* as a model for wood formation. Curr. Opin. Biotech. 22:293-299

Zhong, R., Lee, C., Ye, Z.-H. (2010a) Functional characterization of poplar wood-associated NAC domain transcription factors. Plant Physiol. 152:1044-1055

Zhong, R., Lee, C., Ye, Z.-H. (2010b) Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. Mol. Plant 3: 1087-1103

Zhong, R., Lee, C., Zhou, J., McCarthy, R.L., 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

Zhong, R., McCarthy, R.L., Lee, C., Ye, Z.-H. (2011) Dissection of the transcriptional program regulating secondary wall biosynthesis during wood formation in poplar. Plant Physiol. 157:1452-1468

Zhong, R., Morrison, W.H., Freshour, G.D., Hahn, M.G., Ye, Z.-H. (2003) Expression of a mutant form of cellulose synthase AtCesA7 causes dominant negative effect on cellulose biosynthesis. Plant Physiol. 132(2):786-795

Zhong, R., Pena, M.J., Zhou, G.-K., Naim, C.J., Wood-Jones, A., Richardson, E.A., Morrison, W.H., Darvill, A.G., York, W.S., Ye, Z.-H. (2005). *Arabidopsis fragile fiber 8*, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. Plant Cell 17:3390-3408

Zhong, R., Richardson, E.A., Ye, Z.-H. (2007a). The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. Plant Cell 19: 2776-2792

Zhong, R., Richardson, E.A., Ye, Z.-H. (2007b). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary cell wall synthesis in fibers of Arabidopsis. Planta 225:1603-1611

Zhou, J., Lee, C., Zhong, R., 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

CHAPTER 2

MYB83 IS A DIRECT TARGET OF SND1 AND ACTS REDUNDANTLY WITH MYB46 IN THE REGULATION OF SECONDARY CELL WALL BIOSYNTHESIS IN ARABIDOPSIS[1]

[1]Ryan L. McCarthy, Ruiqin Zhong and Zheng-Hua Ye. 2009. *Plant Cell Physiology*. 50(11):1950-1964.

Reprinted here with permission of the publisher

Abstract

It has been proposed that the transcriptional regulation of secondary wall biosynthesis in Arabidopsis is controlled by a transcriptional network mediated by SND1 and its close homologs. Uncovering all the transcription factors and deciphering their interrelationships in the network are essential for our understanding of the molecular mechanisms underlying the transcriptional regulation of biosynthesis of secondary walls, the major constituent of wood and fibers. Here, we present functional evidence that the MYB83 transcription factor is another molecular switch in the SND1-mediated transcriptional network regulating secondary wall biosynthesis. MYB83 is specifically expressed in fibers and vessels where secondary wall thickening occurs. Its expression is directly activated by SND1 and its close homologs, including NST1, NST2, VND6 and VND7, indicating that MYB83 is their direct target. MYB83 overexpression is able to activate a number of the biosynthetic genes of cellulose, xylan and lignin and concomitantly induce ectopic secondary wall deposition. In addition, its overexpression upregulates the expression of several transcription factors involved in regulation of secondary wall biosynthesis. Dominant repression of MYB83 functions or simultaneous RNAi inhibition of MYB83 and MYB46 results in a reduction in secondary wall thickening in fibers and vessels and a deformation of vessels. Furthermore, double T-DNA knockout mutations of MYB83 and MYB46 cause a lack of secondary walls in vessels and an arrest in plant growth. Together, these results demonstrate that MYB83 and MYB46, both of which are SND1 direct targets, function redundantly in the transcriptional regulatory cascade leading to secondary wall formation in fibers and vessels.

Introduction

Secondary cell walls account for the bulk of biomass produced by land plants. They are the major component of wood and fibers, which have been widely used by humans for energy, lumber, pulping and paper-making, and textiles. With the dwindling of the non-renewable fossil fuels, there is an urgent need to find alternative fuels to sustain our energy consumption. Plant lignocellulosic biomass has been considered to be an important renewable and environmentally friendly source of bioenergy (Ohlrogge et al., 2009). To genetically modify the quantity and quality of wood and fibers for their better utilization, it is imperative to uncover the molecular mechanisms underlying the regulation of secondary wall biosynthesis. Secondary wall formation is a developmentally regulated process occurring in some specialized cell types, such as tracheary elements and fibers. Secondary walls, which are deposited next to the primary wall after the cessation of cell elongation, are mainly composed of cellulose, hemicellulose (xylan and glucomannan) and lignin. Although little is known about the developmental signals and the subsequent signal transduction pathways leading to the formation of secondary walls, recent molecular and genetic studies of secondary wall biosynthesis in Arabidopsis have provided important insights into the mechanisms underlying the transcriptional regulation of secondary wall biosynthesis. It has been proposed that a transcriptional network encompassing a cascade of transcription factors is involved in the coordinated regulation of secondary wall biosynthesis in fibers and vessels (Zhong and Ye, 2007). A group of closely related NAC domain transcription factors, SND1 (also called NST3/ANAC012), NST1, NST2, VND6 and VND7, function at the top of this network to activate the entire secondary wall biosynthetic program (Kubo et al., 2005, Mitsuda et al., 2005; Zhong et al., 2006; Mitsuda et al., 2007; Zhong et al., 2007b; Yamaguchi et al., 2008). These secondary wall associated NACs have been shown to regulate a battery of

downstream transcription factors, which in turn activate the biosynthetic genes for secondary wall deposition (Zhong et al., 2008). Among the SND1-regulated transcription factors, MYB46, SND3, MYB103 and KNAT7 have been demonstrated to be direct targets of SND1 and its close homologs, including NST1, NST2, VND6 and VND7. These SND1 direct targets were shown to be specifically expressed in fibers and vessels and dominant repression of their function led to a reduction in secondary wall thickening, indicating that they are involved in the regulation of secondary wall biosynthesis. In particular, MYB46 functions as another level of molecular switch able to turn on the entire secondary wall biosynthetic program (Zhong et al., 2007a). Currently, it is not known which transcription factors directly activate the biosynthetic pathways of secondary wall cellulose and xylan. Available evidence indicates that members of the MYB and LIM transcription factor families are involved in the regulation of lignin biosynthetic genes (Kawaoka et al., 2000; Patzlaff et al., 2003; Goicoechea et al., 2005). Among them, MYB58 and MYB63, which are downstream components in the SND1-mediated transcriptional network, have been demonstrated to be direct activators of the lignin biosynthetic pathway in Arabidopsis (Zhou et al., 2009). Although significant progress has been made in our understanding of the transcriptional regulatory network controlling secondary wall biosynthesis, we are still far from identifying all the players and their interrelationships in the transcriptional network. For example, dominant repression of MYB46 resulted in a reduction in the secondary wall thickening of fiber cells, but T-DNA knockout of the *MYB46* gene did not cause any discernable effects on secondary wall deposition (Zhong et al., 2007a). This finding indicates that other functionally redundant transcription factors may compensate for the loss of MYB46. In this report, we demonstrate that the MYB83 transcription factor, which is phylogenetically closely related to MYB46, is another important player in the SND1-mediated transcription

network regulating secondary wall biosynthesis in fibers and vessels. The *MYB83* gene was found to be specifically expressed in cells undergoing secondary wall thickening and directly activated by SND1 and its close homologs. We show that overexpression of MYB83 induces the expression of a number of secondary wall-associated transcription factors and secondary wall biosynthetic genes, and concomitantly results in an ectopic deposition of secondary wall components, including cellulose, xylan and lignin. Dominant repression of MYB83 functions or simultaneous RNAi inhibition of *MYB83* and *MYB46* was found to cause a reduction in secondary wall thickening in fibers and vessels. We further demonstrate that double T-DNA knockout mutations of *MYB83* and *MYB46* leads to a loss of secondary walls in vessels and a concomitant arrest in plant growth. Our results establish that MYB83 functions redundantly with MYB46 in the regulation of secondary wall biosynthesis in Arabidopsis.

Results

The expression of *MYB83* is specifically associated with secondary wall thickening and is regulated by SND1

We have previously revealed a number of MYB transcription factors as members of the SND1-mediated transcriptional network regulating secondary wall biosynthesis in Arabidopsis (Zhong et al., 2008). To investigate whether additional Arabidopsis MYBs might be involved in the regulation of secondary wall biosynthesis, we searched for MYB genes that were highly expressed in the inflorescence stems where secondary wall-containing fibers and vessels are abundant. We found that *MYB83*, a previously uncharacterized MYB transcription factor, was expressed preferentially in the inflorescence stems (Fig. 1A). Further expression analyses in laser-microdissected cells (Fig. 1B) and in transgenic plants expressing the *MYB83* gene-driven

β-glucuronidase (GUS) reporter gene (Fig. 2A–C) showed that *MYB83* was specifically expressed in fiber and vessel cells that are undergoing secondary wall thickening in stems. It was interesting to find that *MYB83* was only expressed in vessels but not in xylary fibers in the developing secondary xylem of roots (Fig. 2D). The expression of *MYB83* was found to be upregulated in SND1 overexpressors (Fig. 1C) in which ectopic secondary wall deposition occurs (Zhong et al., 2006), and conversely it was downregulated in plants with simultaneous RNAi inhibition of *SND1* and *NST1* (Fig. 1D) in which secondary wall thickening is blocked in fiber cells (Zhong et al., 2007b). Protein sequence analysis revealed that MYB83 shares 41% identity and 51% similarity with MYB46 in their overall sequences and 85% identity and 93% similarity in their R2R3 MYB DNA binding domains (Fig. 1E). Consistent with its predicted function as a transcription factor, MYB83 was shown to be targeted to the nucleus in Arabidopsis protoplasts (Fig. 2E–J) and it was able to activate the expression of the His3 and β-Gal reporter genes in yeast (Fig. 2K). These results demonstrate that MYB83 is an SND1-regulated transcriptional activator associated with secondary wall biosynthesis in fibers and vessels.

Direct activation of *MYB83* expression by SND1, NST1, NST2, VND6 and VND7

The finding that the expression of *MYB83* is regulated by SND1 prompted us to investigate whether MYB83 is a direct target of SND1. We employed the steroid receptor-based inducible system to ascertain whether the induction of MYB83 expression by SND1 is direct. This approach has been successfully applied to identify direct targets of a number of plant transcription factors (Sablowski and Meyerowitz, 1998; Wagner et al., 1999; Baudry et al., 2004; Zhong et al., 2008). SND1 fused to the regulatory region of the human estrogen receptor (HER, Zuo et al. 2000) was expressed in Arabidopsis protoplasts under the control of the cauliflower

mosaic virus (CaMV) 35S promoter (Fig. 3A). Without estradiol treatment, the SND1–HER chimeric protein did not induce *MYB83* expression due to the retention of SND1–HER in the cytoplasm. As expected, estradiol treatment resulted in activation of *MYB83* expression by SND1-HER (Fig. 3B). We next tested the effects of the protein synthesis inhibitor cycloheximide on the estradiol-activated induction of *MYB83* expression. Inhibition of new protein synthesis by cycloheximide should not affect the estradiol-activated induction of SND1 direct targets but should block the induction of further downstream genes. It was found that cycloheximide treatment did not affect the SND1-induced expression of the *MYB83* gene (Fig. 3B), indicating that SND1 directly activates *MYB83* expression without new protein synthesis. Using the steroid receptor-based inducible system, we also found that the SND1 homologs, NST1, NST2, VND6 and VND7 were all able to induce *MYB83* expression upon estradiol treatment and this induction still occurred in the presence of cycloheximide (Fig. 3B). Together, these results demonstrate that MYB83 is a direct downstream target of SND1, NST1, NST2, VND6 and VND7.

Overexpression of MYB83 causes ectopic deposition of secondary walls

The results described above suggest that MYB83 is an SND1-activated transcriptional regulator potentially involved in secondary wall biosynthesis. To test this hypothesis, we examined whether MYB83 overexpression could induce the expression of secondary wall biosynthetic genes and subsequent ectopic deposition of secondary walls. The full-length *MYB83* cDNA under the control of the CaMV 35S promoter was expressed in wild-type Arabidopsis plants. MYB83 overexpressors were selected (Fig. 4A) and found to exhibit the curly leaf phenotype (Fig. 4B), which is similar to that seen in overexpressors of SND1 (Zhong et al., 2006) and MYB46 (Zhong et al., 2007a). The curly leaf phenotype was observed in 51 out of a total of 64 transgenic plants examined. Examination of the leaf epidermis revealed that although wild-type

cells had no staining for secondary wall cellulose, xylan and lignin (Fig. 4C, D, J, K), the walls of epidermal cells in MYB83 overexpressors were apparently thickened with cellulose, xylan and lignin (Fig. 4E–I). Ectopic deposition of secondary wall components was also observed in the cortical cells of the inflorescence stems of MYB83 overexpressors (Fig. 5D–F) compared with the wild-type stems in which secondary wall deposition was only seen in interfascicular fibers and xylem cells (Fig. 5A–C). Consistent with the observed ectopic deposition of secondary walls, gene expression analysis demonstrated that there was a significant elevation in the expression of secondary wall biosynthetic genes in MYB83 overexpressors (Fig. 6A). These results suggest that MYB83 is involved in regulation of the biosynthetic pathways for all three major secondary wall components, cellulose, xylan and lignin.

Induction of secondary wall-associated transcription factors by MYB83

Since MYB83 is a direct downstream target of SND1, we next examined whether MYB83 induces the expression of other SND1-regulated transcription factors. It was found that MYB83 overexpression significantly induced the expression of *MYB42, MYB43, MYB52, MYB54, MYB63, MYB85* and *KNAT7* (Fig. 6B). A slight induction of *SND3, MYB58* and *MYB103* was also seen in MYB83 overexpressors. These MYB83-induced genes were also found to be upregulated by MYB46 overexpression albeit to different levels (Fig. 6B). These results indicate that MYB83 and MYB46 regulate the same downstream transcription factors in the SND1-mediated transcription network controlling secondary wall biosynthesis.

Simultaneous downregulation of *MYB83* and *MYB46* results in a reduction in secondary wall thickening in fibers and vessels

To further substantiate our hypothesis that MYB83 is involved in the regulation of secondary wall biosynthesis, we investigated the effects of inhibition of MYB83 functions on secondary

wall deposition. RNAi inhibition or T-DNA knockout mutation of *MYB83* did not cause any discernible changes in the secondary wall thickening of fibers and vessels in the inflorescence stems, which may be due to functional redundancy of other genes. Therefore, we resorted to the dominant repression approach to investigate the functional role of MYB83 in secondary wall biosynthesis. Transgenic Arabidopsis plants expressing MYB83 fused with the dominant EAR repressor domain (Hiratsu et al., 2004) had reduced growth with shorter inflorescence stems compared with the wild type (Fig. 7A), which is reminiscent of the phenotype exhibited by secondary wall-defective mutants (Zhong et al., 2005; Pena et al., 2007). Examination of the cross-sections of stems revealed that the vessels were severely deformed in MYB83 repressors (Fig. 7K, N) compared with the wild type (Fig. 7J, M). In addition, the secondary wall thickness of interfascicular fibers, xylary fibers and vessels was significantly reduced (Fig. 7E, H, K, N) compared with the wild type (Fig. 7D, G, J, M, Table 1). These phenotypes were observed in 32 out of a total of 64 dominant repressor plants.

Table 1 Wall thickness of fibers and vessels in the stems of wild type, MYB83 repressors, and MYB46 and MYB83 double RNAi lines

Plant	**Interfascicular fiber**	**Vessel**	**Xylary Fiber**
Wild type	1.74±0.23	0.98±0.16	0.75±0.15
MYB83 repressors	0.96±0.15	0.41±0.13	0.15±0.04
MYB46/MYB83 RNAi	0.62±0.14	0.80±0.10	0.18±0.06

Wall thickness was measured from transmission electron micrographs of fibers and vessels, and the data are mean (μm) ±SE from 15 cells

The findings that MYB83 and MYB46 induce the expression of the same set of downstream transcription factors and RNAi inhibition or T-DNA knockout mutation of either *MYB83* or *MYB46* alone does not cause any reduction in secondary wall thickening suggest that they might function redundantly in the regulation of secondary wall biosynthesis. To test this hypothesis, we first generated transgenic Arabidopsis plants with simultaneous RNAi inhibition of *MYB83* and

MYB46 (Fig. 7C). It was found that similar to the dominant repression of MYB83, simultaneous RNAi inhibition of *MYB83* and *MYB46* caused reduced growth (Fig. 7B), severe deformation of vessels (Fig. 7L, O) and reduced secondary wall thickening in interfascicular fibers and xylary fibers in stems (Fig. 7F, I, L, O). These severe phenotypes were observed in 8 out of 64 RNAi inhibition plants. The knockdown study of *MYB83* and *MYB46* by the RNAi inhibition approach allowed us to reveal their redundant roles in the regulation of secondary wall biosynthesis in both fibers and vessels. We next investigated the effects of double T-DNA knockout mutations of *MYB83* and *MYB46* on secondary wall thickening. To generate the double mutants, the T-DNA knockout line of *MYB83* (*myb83)* was crossed with two independent T-DNA knockout lines of *MYB46* (*myb46-1* and *myb46-2)* (Fig. 8A, B). Although single T-DNA knockout mutation of *MYB83* or *MYB46* did not cause any reductions in secondary wall thickening and plant growth, the double mutants exhibited severely retarded growth. Both *myb46-1 myb83* and *myb46-2 myb83* showed the same phenotypes, so only the results from *myb46-1 myb83* were shown. The double mutant plants germinated but their growth was arrested after developing one to two pairs of small leaves (Fig. 8C) followed by wilting and subsequent death. Examination of the leaves showed that although the veins in the wild type exhibited prominent secondary wall thickening (Fig. 9A), those in the double knockout mutant nearly lacked secondary wall thickening (Fig. 9B). A closer examination of leaf veins revealed severely deformed vessels with no apparent wall thickening in the double knockout mutant (Fig. 9D) compared with the thick-walled vessels in the wild type (Fig. 9C). Likewise, regular-shaped vessels with thick walls were evident in the wild-type roots (Fig. 9E), whereas in the roots of the double knockout mutant no cells with thick walls were visible in the location where vessels should be present (Fig. 9F). It was noted that the sizes of veins and steles in the double knockout mutant were much smaller than those in the wild

type, which most likely is attributed to the arrested growth. Together with the data from dominant repression and RNAi inhibition studies, these results demonstrate that MYB83 and MYB46 are essential for the normal secondary wall thickening in vessels and fibers and suggest that they act redundantly in the transcriptional regulation of secondary wall biosynthesis.

Discussion

The NAC domain transcription factor SND1 has previously been demonstrated to be a master switch activating a cascade of transcription factors involved in the regulation of secondary wall biosynthesis (Zhong and Ye, 2007). Uncovering all the transcription factors and sorting out their interrelationships in the SND1-mediated transcriptional network is imperative in understanding the molecular mechanisms underlying the transcriptional regulation of secondary wall biosynthesis. We have previously shown that SND1 directly activates the expression of four secondary wall-associated transcription factors, including MYB46, SND3, MYB103 and KNAT7 (Zhong et al., 2008). In this report, we demonstrate that MYB83 is another direct target of SND1 and it functions redundantly with MYB46 in regulating the biosynthesis of secondary wall components, including cellulose, xylan and lignin. Our finding that MYB83 is another key player in the SND1-mediated transcriptional network enriches our understanding of the transcriptional control of the biosynthesis of secondary walls, the most abundant form of biomass produced by land plants.

MYB83 is a direct target of SND1 and its close homologs

We have demonstrated that *MYB83* is directly activated by SND1 and its close homologs and its expression exhibits an overlapping pattern with that of SND1 and its homologs. These findings identify MYB83 as another direct target of SND1 and its close homologs in addition to the four

known ones, including MYB46, SND3, MYB103 and KNAT7 (Fig. 10). It is interesting to note that although SND1, VND6 and VND7 have distinctive fiber or vessel-specific expression patterns, their direct targets are expressed in both fibers and vessels. Our findings that MYB83 overexpression causes ectopic deposition of secondary walls and its dominant repression results in a reduction in secondary wall thickening demonstrate that MYB83 is involved in the regulation of the entire secondary wall biosynthetic program. Among the five known SND1 direct targets, only MYB83 and MYB46 are capable of activating the entire biosynthetic program of secondary walls and thus they are considered to be another level of master switches controlling secondary wall biosynthesis. It is currently unknown why plants evolved to have two hierarchical levels of master controls for coordinating the expression of secondary wall biosynthetic genes.

MYB83 and MYB46 function redundantly in the transcriptional regulation of secondary wall biosynthesis

We have demonstrated that the double T-DNA knockout mutations of *MYB83* and *MYB46* lead to a loss of secondary walls in vessels and concomitantly an arrest in plant growth. This finding provides direct genetic evidence that MYB83 and MYB46 perform redundant functions in the SND1-mediated transcriptional regulation of secondary wall biosynthesis. Such a gene redundancy is common for transcription factors involved in the regulation of secondary wall biosynthesis. For example, SND1 and NST1 function redundantly in the regulation of secondary wall thickening in fibers, whereas VND6 and VND7 function redundantly in the regulation of secondary wall thickening in vessels. It is possible that vascular plants evolved to have multiple genes, which could result from genome duplications or other mechanisms, to safeguard each step in the transcriptional network controlling secondary wall biosynthesis because deposition of

secondary walls in vessels is essential for the survival of vascular plants. The arrest of growth of the *myb46 myb83* double mutant is most likely attributed to a lack of functional fluid-conducting vessels due to the loss of secondary walls. This growth arrest phenotype is reminiscent of that of several other mutants defective in secondary wall biosynthesis albeit with a greater severity. For example, double knockout mutations of the Arabidopsis *FRA8* and its paralog *F8H,* which are involved in xylan biosynthesis in secondary walls of fibers and vessels, result in a severe decrease in secondary wall thickening and a collapse of vessels, and concomitantly stunted plant growth at the seedling stage (Lee et al., 2009). Likewise, overexpression of the Arabidopsis NAC domain protein XND1 causes a loss of secondary wall thickening in vessels and a dwarfed plant phenotype (Zhao et al., 2008). It should be noted that while the study of double knockout mutants of *MYB83* and *MYB46* allowed us to provide direct genetic evidence for their redundant roles in the regulation of secondary wall biosynthesis in vessels, the growth arrest of the mutant plants prevented us from investigating their redundant roles in the regulation of secondary wall biosynthesis in fibers. However, this difficulty was overcome by the knockdown study using the RNAi inhibition approach showing that both MYB83 and MYB46 are required for normal secondary wall thickening in fibers, which is consistent with their expression patterns. It should be cautioned that RNAi inhibition of *MYB83* and *MYB46* results in impaired plant growth with a shorter inflorescence due to the collapsed vessels, which could indirectly affect secondary wall thickening in fibers. However, previous studies on the *fra1* mutant showed that reduced growth with a shorter inflorescence does not affect secondary wall thickening in fibers (Zhong et al., 2002), indicating that there is no strict correlation between the inflorescence height and secondary wall thickening in fibers. Since the xylem bundles in the double RNAi plants are well developed but the secondary wall thickening is severely reduced in xylary fibers (Fig. 7L**)**, we

conclude that the observed secondary wall thickening defect in the fibers of the double RNAi plants is attributed to the downregulation of *MYB83* and *MYB46*. MYB83 and MYB46 were demonstrated to upregulate a subset of SND1-regulated, secondary wall-associated transcription factors (Figs. 6B, 10). Among them, MYB85, MYB58 and MYB63 have previously been proved to be involved in the regulation of lignin biosynthesis (Zhong et al., 2008; Zhou et al., 2009), and MYB52, MYB54 and KNAT7 have been shown to be important for regulating secondary wall thickening in fibers (Zhong et al., 2008). It is interesting to note that MYB83 and MYB46 also regulate the expression of other SND1 direct targets, including MYB103, KNAT7 and SND3 (Figs. 6B, 10). This type of hierarchical network structure is referred to as feed-forward loop, where one transcription factor regulates another and then they together regulate their common downstream targets (Yu and Gerstein, 2006). The finding that MYB83 and MYB46 regulate the expression of other SND1 downstream targets provides a basis for further understanding of the interrelationships of transcription factors in the SND1-mediated transcriptional network. In summary, we have revealed that MYB83 acts redundantly with MYB46 as another level of molecular switches regulating secondary wall biosynthesis in the SND1-mediated transcriptional network. Because MYB83 or MYB46 alone is able to activate the entire secondary wall biosynthetic program, MYB83 and MYB46 can potentially be used as an important tool for manipulating the quantity of secondary walls in fibers and wood. It has recently been proposed that burning plant lignocellullosic biomass to generate electricity for battery-driven vehicles captures more biomass energy than converting it to bioethanol and therefore, it is imperative to find means to increase plant biomass yields (Ohlrogge et al., 2009). Since secondary walls in the form of fibers and wood are the most abundant biomass produced by land plants, further understanding of the transcriptional network regulating secondary wall biosynthesis will have

important implications in our attempt to manipulate biomass yields in trees and other biofuel crop species.

Materials and Methods

Gene expression analysis

Total RNA from different Arabidopsis organs was isolated using a Qiagen RNA isolation kit and subjected to gene expression analysis. The seedlings used were 2 weeks old. Mature leaves were from 6-week-old plants. Mature roots were from 8-week-old plants. Stems from 6-week-old plants were divided into top, middle and bottom parts, which represent the rapidly elongating internodes, internodes near cessation of elongation and non-elongating internodes, respectively. Interfascicular fiber cells, xylem cells and pith cells were isolated from inflorescence stems of 6-week-old Arabidopsis plants using the PALM microlaser system (PALM Microlaser Technologies) and further used for RNA isolation and subsequent gene expression analysis (Zhong et al., 2006). Quantitative PCR analysis was carried out using the first strand cDNA as template with the QuantiTect SYBR Green PCR Kit (Clontech). The PCR primers for *MYB83* are 5'-ggagaataccaacgtcattgcttg-3' and 5'-atcgacttggaaatcaagaag-3', those for *MYB46* are 5'-atggtatctatggagtaaacg-3' and 5'-tatgctttgtttgaagttgaagta-3', and those for other genes were described previously (Zhong et al., 2006; Zhou et al., 2009). The relative mRNA levels were determined by normalizing the PCR threshold cycle number of each gene with that of the *EF1α* reference gene. The expression level of each gene in the wild type or control was set to 1 or 100 and the data were the average of three biological replicates.

GUS reporter gene analysis

The *MYB83* gene containing a 3 kb 5′ upstream sequence, the entire coding region and a 2 kb 3′ downstream sequence was used for the GUS reporter gene analysis. The GUS reporter gene was inserted in-frame right before the stop codon of the *MYB83* gene and then cloned into the binary vector pBI101 (Clontech) to create the GUS reporter construct. The construct was transformed into wild-type *Arabidopsis* plants by *Agrobacterium*-mediated transformation (Bechtold and Bouchez, 1994) and the resulting transgenic plants were examined for GUS activity as described previously (Zhong et al., 2005).

Subcellular localization

The full-length MYB83 cDNA was fused in-frame with the yellow fluorescent protein (YFP) cDNA and ligated between the CaMV 35S promoter and the nopaline synthase (NOS) terminator in pBI221 (Clontech). The construct was cotransfected with SND1–CFP (cyan fluorescent protein) into *Arabidopsis* leaf protoplasts by PEG-mediated transfection (Sheen, 2001). After 20 h incubation, the transfected protoplasts were examined for fluorescence signals using a Leica TCs SP2 spectral confocal microscope (Leica Microsystems). Images were saved and processed with Adobe Photoshop Version 7.0 (Adobe Systems).

Transcriptional activation analysis in yeast

The full-length MYB83 cDNA fused in-frame with the GAL4 DNA binding domain in the pAS2-1 vector (Clontech) was transformed into the yeast strain CG-1945 containing the *His3* and *LacZ* reporter genes. The transformed yeast cells were grown on SD plates with or without His and subjected to β-galactosidase activity assay.

Identification of direct targets using the estrogen inducible system

Estrogen-inducible expression vectors were created by fusing the full-length cDNA of SND1 or its close homologs with the regulatory region of the human estrogen receptor (Zuo et al., 2000),

which was ligated between the CaMV 35S promoter and the NOS terminator in the pBI221 vector (Clontech). After transfection of the expression constructs into Arabidopsis leaf protoplasts, the protoplasts were treated with 2 μM estradiol (Sigma) for 6 h before being harvested for RNA isolation. New protein synthesis was inhibited by treating the protoplasts for 30 min with the protein synthesis inhibitor cycloheximide (2 μM) before addition of estradiol. Under this condition, new protein synthesis was completely inhibited as tested by GUS reporter gene analysis (Zhong et al., 2008). After treatment, total RNA was isolated from the transfected protoplasts and used for subsequent quantitative RT–PCR analysis. The expression level of *MYB83* gene in the control protoplasts without estradiol treatment was set to 1, and the data were the average of three biological replicates.

Dominant repression, overexpression, RNAi inhibition and T-DNA lines

The MYB83 dominant repression construct was made by fusing its full-length cDNA in-frame with the dominant EAR repression sequence (5'-ctggatctggatctagaactccgtttgggtttcgct-3'; Hiratsu et al., 2004), which was ligated downstream of the CaMV 35S promoter in pBI121 (Clontech). The MYB83 overexpression construct was created by cloning its full-length cDNA downstream of the CaMV 35S promoter in pBI121. The *MYB83* RNAi construct was generated by ligating *MYB83* cDNA without the MYB domain (PCR primers: 5'-aacaacaatacttcatcagga-3' and 5'-atcgacttggaaatcaagaag-3') into modified pBI121 in opposite orientations on both sides of the GUS spacer, which is located between the CaMV 35S promoter and the NOS terminator. The GUS spacer contains the GUS sequence from nucleotide positions 783–1789. The *MYB46* RNAi expression cassette was created by ligating *MYB46* cDNA excluding the MYB domain (PCR primers: 5'-aagatgtccgatacctccaac-3' and 5'-tatgctttgtttgaagttgaagta-3') into pBI121 in opposite orientations on both sides of the GUS spacer. The *MYB83* and *MYB46* double RNAi construct

was made by ligating the *MYB46* RNAi expression cassette (containing the CaMV 35S promoter, *MYB46* inverted repeats with the GUS spacer and the NOS terminator) into the pBI121 vector, which already contains the *MYB83* RNAi expression cassette. The constructs were introduced into wild-type Arabidopsis plants by *Agrobacterium* -mediated transformation. Transgenic plants were selected on kanamycin and the first generation (a total of 64 plants for each construct) was used for phenotypic characterization. The phenotypes caused by overexpression, dominant repression and RNAi inhibition were examined in at least eight independent transgenic plants and the representative results were presented. T-DNA knockout mutants for MYB83 (SALK_093099C) and MYB46 (SALK-088514 and SALK_100993C) (Fig. 8A) were obtained from the Arabidopsis Biological Resource Center (ABRC; Columbus, OH, USA). The homozygous T-DNA insertions lines were crossed for generation of the double *myb83 myb46* knockout mutants. Seventeen double knockout mutant plants were identified from a total of 288 F2 plants from a cross between *myb83* and *myb46-1,* and 16 double knockout mutant plants were identified from a total of 288 F2 plants from a cross between *myb83* and *myb46-2.*

Histology

Tissues were fixed in 2% formaldehyde and embedded in low viscosity (Spurr's) resin (Electron Microscopy Sciences) as described (Burk et al., 2006). For light microscopy, 1 μm thick sections were cut with a microtome and stained with toluidine blue. For transmission electron microscopy, 85 nm thick sections were cut, post-stained with uranyl acetate and lead citrate, and observed using a Zeiss EM 902A transmission electron microscope (Carl Zeiss). Lignin in the stems was visualized by staining 50 μm thick sections with phloroglucinol-HCl, which was shown as bright red color. For lignin autofluorescence visualization, leaves were cleared in methanol and examined using a UV fluorescence microscope (Zhong et al., 2006). Secondary

wall cellulose staining was carried out by incubating 1 μm thick sections with 0.01% Calcofluor White (Hughes and McCully, 1975). Xylan was detected by using the monoclonal LM10 antibody against xylan and fluorescein isothiocyanate-conjugated goat antirat secondary antibodies according to McCartney et al. (2005). For each construct used for transformation, at least eight transgenic plants with representative phenotypes were examined for their wall phenotypes.

Statistical analysis

The experimental data of quantitative PCR and cell wall thickness measurement were subjected to statistical analysis using the Student's *t* test program (http://www.graphpad.com/quickcalcs/ttest1.cfm), and the quantitative difference between the two groups of data for comparison was found to be statistically significant ($P < 0.001$).

Funding

This work was supported by the National Science Foundation (ISO-0744170).

Acknowledgments

We thank the ABRC (Columbus, OH, USA) for the T-DNA knockout lines for *MYB83* and *MYB46,* and the editor and the two anonymous reviewers for their constructive comments and suggestions on the revision of the manuscript.

References

Baudry, A., Heim, M.A., Dubreucq, B., Caboche, M., Weisshaar, B. and Lepiniec, L. (2004) TT2, TT8, and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant J. 39: 366–380

Bechtold, N. and Bouchez, D. (1994) In planta *Agrobacterium*–mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. In Gene Transfer to Plants. Edited by Potrykus, I. and Spangenberg, G. pp. 19–23. Springer, Berlin

Burk, D.H., Zhong, R., Morrison, W.H., III and Ye, Z.-H. (2006) Disruption of cortical microtubules by overexpression of green fluorescent protein-tagged α-tubulin 6 causes a marked reduction in cell wall synthesis. J. Integr. Plant Biol. 48: 85–98

Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., et al. (2005) EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. Plant J. 43: 553–567

Hiratsu, K., Mitsuda, N., Matsui, K. and Ohme-Takagi, M. (2004) Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in Arabidopsis. Biochem. Biophys. Res. Commun. 321: 172–178

Hughes, J. and McCully, M.E. (1975) The use of an optical brightener in the study of plant structure. Stain Technol. 50: 319–329

Kawaoka, A., Kaothien, P., Yoshida, K., Endo, S., Yamada, K. and Ebinuma, H. (2000) Functional analysis of tobacco LIM protein Ntlim1 involved in lignin biosynthesis. Plant J. 22: 289–301

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005) Transcription switches for protoxylem and metaxylem vessel formation. Genes Dev. 19: 1855–1860

Lee, C., Teng, Q., Huang, W., Zhong, R. and Ye, Z.-H. (2009) The F8H glycosyltransferase is a functional paralog of FRA8 involved in glucuronoxylan biosynthesis in Arabidopsis. Plant Cell Physiol. 50: 812–827

McCartney, L., Marcus, S.E. and Knox, J.P. (2005) Monoclonal antibodies to plant cell wall xylans and arabinoxylans. J. Histochem. Cytochem. 53: 543–546

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (2007) NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. Plant Cell 19: 270–280

Mitsuda, N., Seki, M., Shinozaki, K. and Ohme-Takagi, M. (2005) The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickening and are required for anther dehiscence. Plant Cell 17: 2993–3006

Ohlrogge, J., Allen, D., Berguson, B., Dellapenna, D., Shachar-Hill, Y. and Stymne, S. (2009) Driving on biomass. Science 324: 1019–1120

Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L.J., Smith, C., et al. (2003) Characterization of a pine MYB that regulates lignification. Plant J. 36: 743–754

Pena, M.J., Zhong, R., Zhou, G.-K., Richardson, E.A., O'Neill, M.A., Darvill, A.G., et al. (2007) Arabidopsis *irregular xylem8* and *irregular xylem9:* implications for the complexity of glucuronoxylan biosynthesis. Plant Cell 19: 549–563

Raes, J., Rohde, A., Christensen, J.H., Peer, Y.V. and Boerjan, W. (2003) Genome-wide characterization of the lignification toolbox in Arabidopsis. Plant Physiol. 133: 1051–1071

Sablowski, R.W.M. and Meyerowitz, E.M. (1998) A homolog of *NO APICAL MERISTEM* is an immediate target of the floral homeotic genes *APETALA3/PISTILLATA. Cell* 92: 93–103

Sheen, J. (2001) Signal transduction in maize and Arabidopsis mesophyll protoplasts. Plant Physiol. 127: 1466–1475

Taylor, N.G., Gardiner, J.C., Whiteman, R. and Turner, S.R. (2004) Cellulose synthesis in the Arabidopsis secondary cell wall. Cellulose 11: 329–338

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673–4680

Wagner, D., Sablowski, R.W.M. and Meyerowitz, E.M. (1999) Transcriptional activation of APETALA1 by LEAFY. Science 285: 582–584

Yamaguchi, M., Kubo, M., Fukuda, H. and Demura, T. (2008) Vascular related NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and shoots. Plant J. 55: 652–64

Yu, H. and Gerstein, M. (2006) Genomic analysis of the hierarchical structure of regulatory networks. Proceedings of the National Academy of Sciences *USA* 103: 14724–14731

Zhao, C., Avci, U., Grant, E.H., Haigler, C.H. and Beers, E.P. (2008). XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulose synthesis and programmed cell death in xylem. Plant J. 53: 425–436

Zhong, R., Burk, D.H., Morrison, W.H., III and Ye, Z.-H. (2002) A Kinesin like protein is essential for oriented deposition of cellulose microfibrils and cell wall strength. Plant Cell 14: 3101–3117

Zhong, R., Demura, T. and Ye, Z.-H. (2006) SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. Plant Cell 18: 3158–3170

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

Zhong, R., Pena, M.J., Zhou, G.-K., Nairn, C.J., Wood-Jones, A., Richardson, E.A., et al. (2005) Arabidopsis *Fragile Fiber8,* which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. Plant Cell 17: 3390–3408

Zhong, R., Richardson, E.A. and Ye, Z.-H. (2007a) The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. Plant Cell 19: 2776–2792

Zhong, R., Richardson, E.A. and Ye, Z.-H. (2007b) Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. Planta 225: 1603–1611

Zhong, R. and Ye, Z.-H. (2007) Regulation of cell wall biosynthesis. Curr. Opin. Plant Biol. 10: 564–572

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

Zuo, J., Niu, Q.-W. and Chua, N.-H. (2000) An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. Plant J. 24: 265–273

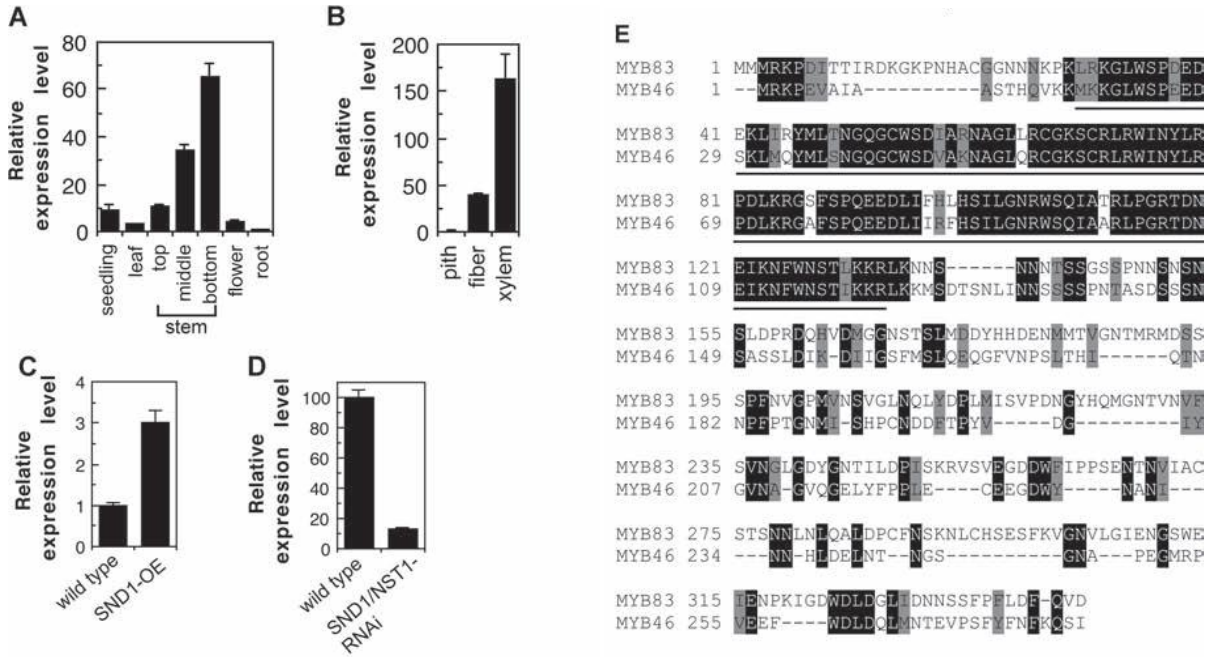

Figure 1. Expression pattern and amino acid sequence analysis of the *MYB83* gene. The expression of *MYB83* gene was examined using real-time quantitative PCR. Error bars denote SE of three biological replicates. (A) Preferential expression of *MYB83* in Arabidopsis inflorescence stems. Its expression in roots was set to 1. (B) *MYB83* was expressed predominantly in xylem and interfascicular fibers. Different types of cell were isolated by laser microdissection from Arabidopsis inflorescence stems. Its expression level in pith cells was set to 1.
(C) The expression level of *MYB83* in leaves was elevated by SND1 overexpression (SND1-OE) compared with the wild type (set to 1). (D) The expression level of *MYB83* in stems was reduced by RNAi inhibition of SND1 and NST1 (SND1/NST1-RNAi) compared with the wild type (set to 100). (E) Alignment of the deduced amino acid sequences of MYB83 and MYB46. The sequence alignment was performed using the ClustalW program (Thompson et al. 1994). The predicted R2R3 MYB DNA binding domain is underlined. Identical and similar amino acid residues are shaded with black and grey, respectively. Gaps (marked with dashes) are used to maximize the sequence alignment.

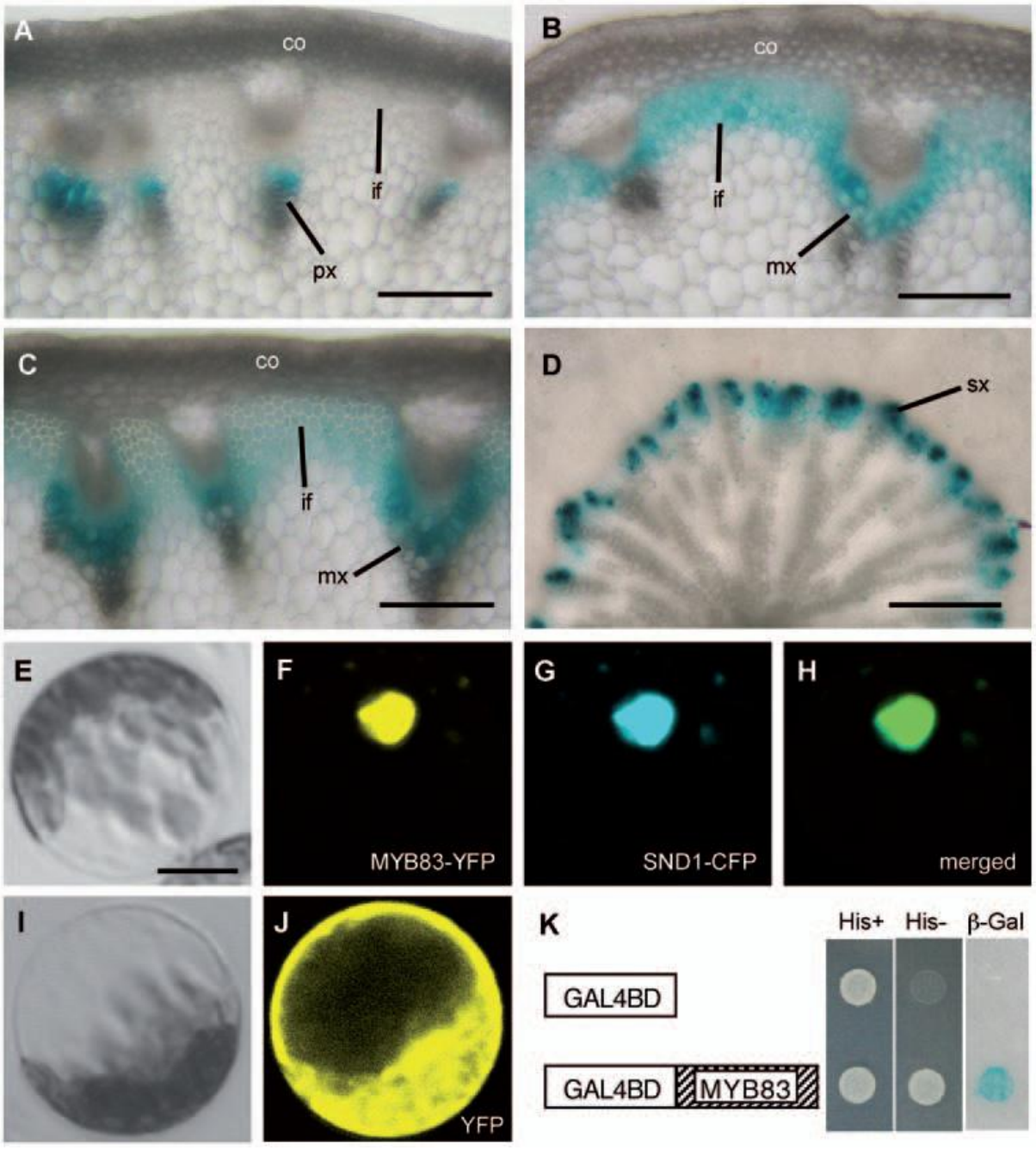
A
co
if
px
B
co
if
mx
C
co
if
mx
D
sx
E
F
MYB83-YFP
G
SND1-CFP
H
merged
I
J
YFP
K
GAL4BD
GAL4BD
MYB83
His+
His-
β-Gal

Figure 2. Developmental expression pattern, subcellular localization and transactivation analysis of MYB83. (A)–(D) The expression pattern of *MYB83* in stems and roots of transgenic Arabidopsis plants was examined using the GUS reporter gene: co, cortex; if, interfascicular fiber; mx, metaxylem; px, protoxylem; sx, secondary xylem. Bars in (A)–(D) = 120 μm. (A) Cross-section of an elongating internode showing GUS staining (blue) only in developing vessels in the protoxylem. (B) Cross-section of an internode near the cessation of elongation showing GUS staining in both developing interfascicular fibers and metaxylem. (C) Cross-section of a non-elongating internode showing GUS staining in interfascicular fiber cells undergoing secondary wall thickening and in vessels and xylary fibers in the metaxylem. (D) Cross-section of a root showing GUS staining in developing secondary xylem. (E)–(H) An Arabidopsis leaf protoplast [(E) differential interference contrast (DIC) image] co-expressing YFP-tagged MYB83 (F) and CFP-tagged SND1 (G). Note the co-localization of MYB83 and SND1 in the nucleus (H). (I), (J) An Arabidopsis protoplast (I; DIC image) expressing YFP alone (J). (K) Transcriptional activation analysis of MYB83 fused with the GAL4 DNA binding domain (GAL4BD) showing that MYB83 was able to activate the expression of His3 and β-Gal reporter genes.

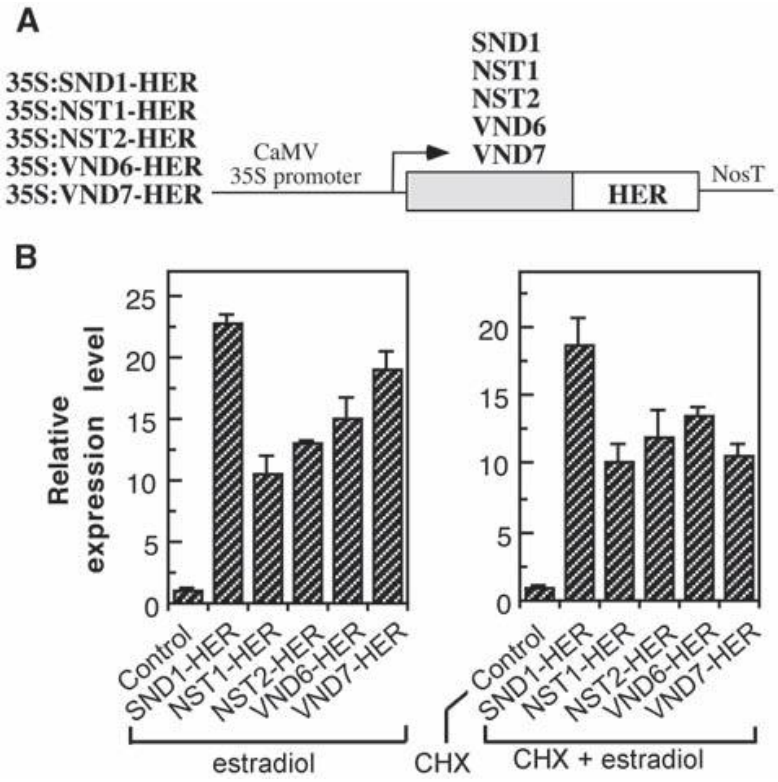

Figure 3. Induction of the expression of *MYB83* by SND1 and its close homologs. (A) Diagram of the fusion of SND1 and its close homologs with the regulatory region of the human estrogen receptor (HER) used for direct target analysis. (B) Quantitative PCR analysis showing that the estradiol-activated SND1 or its close homologs were able to activate expression of the *MYB83* gene [left panel: the expression level in the mock-treated (control) protoplasts was set to 1], and this activation remained in the presence of the protein synthesis inhibitor cyclohexmide (CHX) during estradiol treatment (right panel: the expression level in the CHX-treated protoplasts was set to 1). The SND1 or its close homolog fusion constructs were introduced into Arabidopsis protoplasts. After treatment with estradiol and CHX, the protoplasts were used for RNA isolation and subsequent expression analysis of endogenous *MYB83* gene by quantitative PCR. Error bars represent SE of three biological replicates.

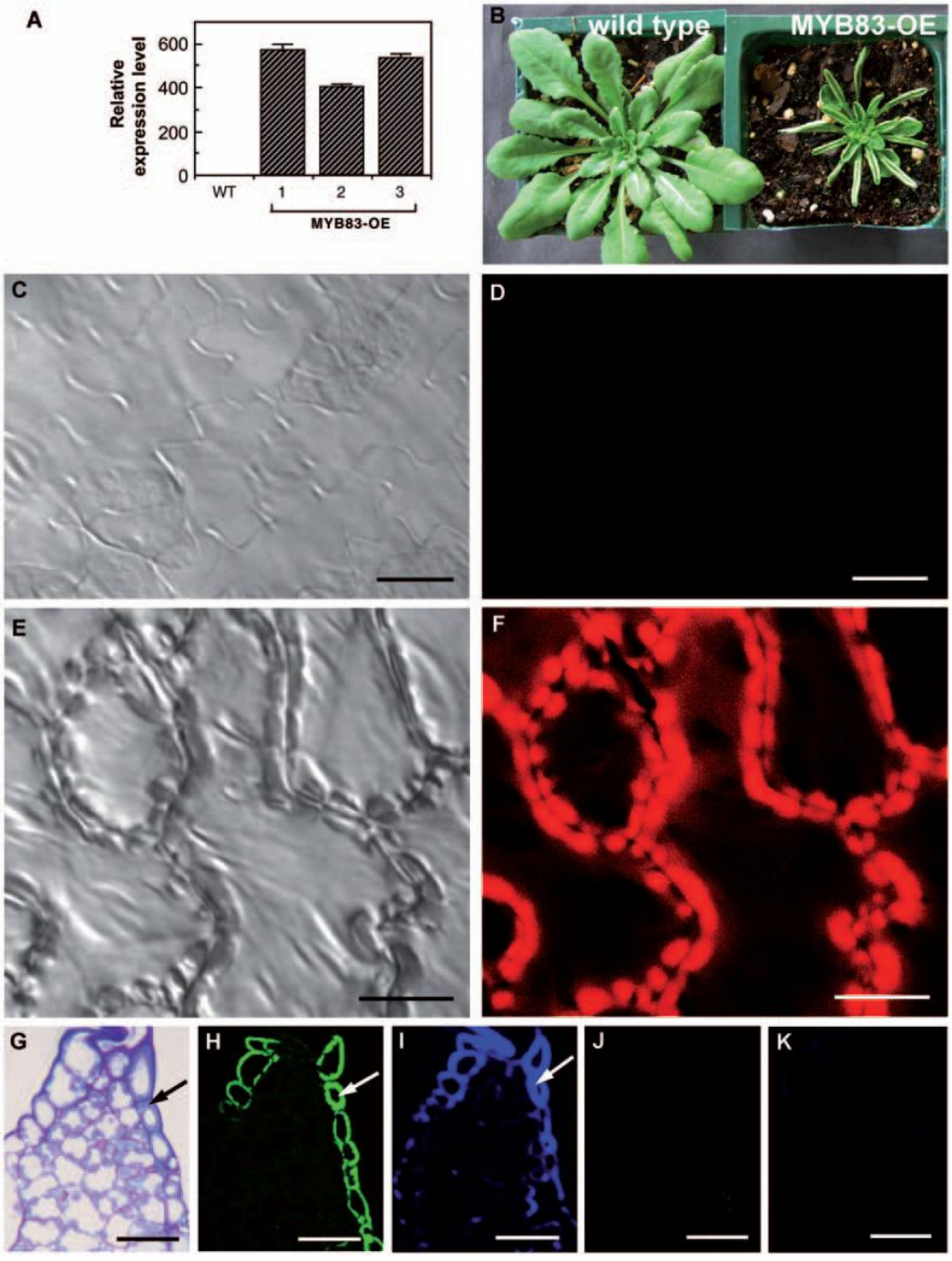
A
Relative expression level
600
400
200
0
WT
1
2
3
MYB83-OE
B
wild type
MYB83-OE
C
D
E
F
G
H
I
J
K

Figure 4. Ectopic deposition of secondary walls in the epidermis of leaves of MYB83 overexpressors. The leaves of 3-week-old MYB83 overexpressors were examined for ectopic deposition of cellulose, xylan and lignin. Bars = 22 μm in (C)–(F) and 190 μm in (G)–(J). (A) Quantitative PCR analysis showing the overexpression of *MYB83* (MYB83-OE) in the seedlings of representative transgenic lines. The expression level in the wild type was set to 1. Error bars represent SE of three technical replicates. (B) Three-week-old seedlings of the wild type (left) and a MYB83 overexpressor with curly leaves (right). (C) and (D) Differential interference contrast (DIC) (C) and lignin autofluorescence (D) images of the epidermis of a wild-type leaf. (E) and (F) DIC (E) and lignin autofluorescence (F) images of the leaf epidermis of a MYB83 overexpressor showing ectopic secondary wall thickening and lignin autofluorescence signals. (G)–(I) Sections of the leaf of a MYB83 overexpressor stained for anatomy with toluidine blue (G), xylan with the LM10 xylan antibody (H) and cellulose with Calcofluor White (I). Note the thickened epidermal cell walls with ectopic deposition of xylan and cellulose (arrows). (J), (K) Sections of the wild-type leaf stained for xylan (J) and for cellulose (K).

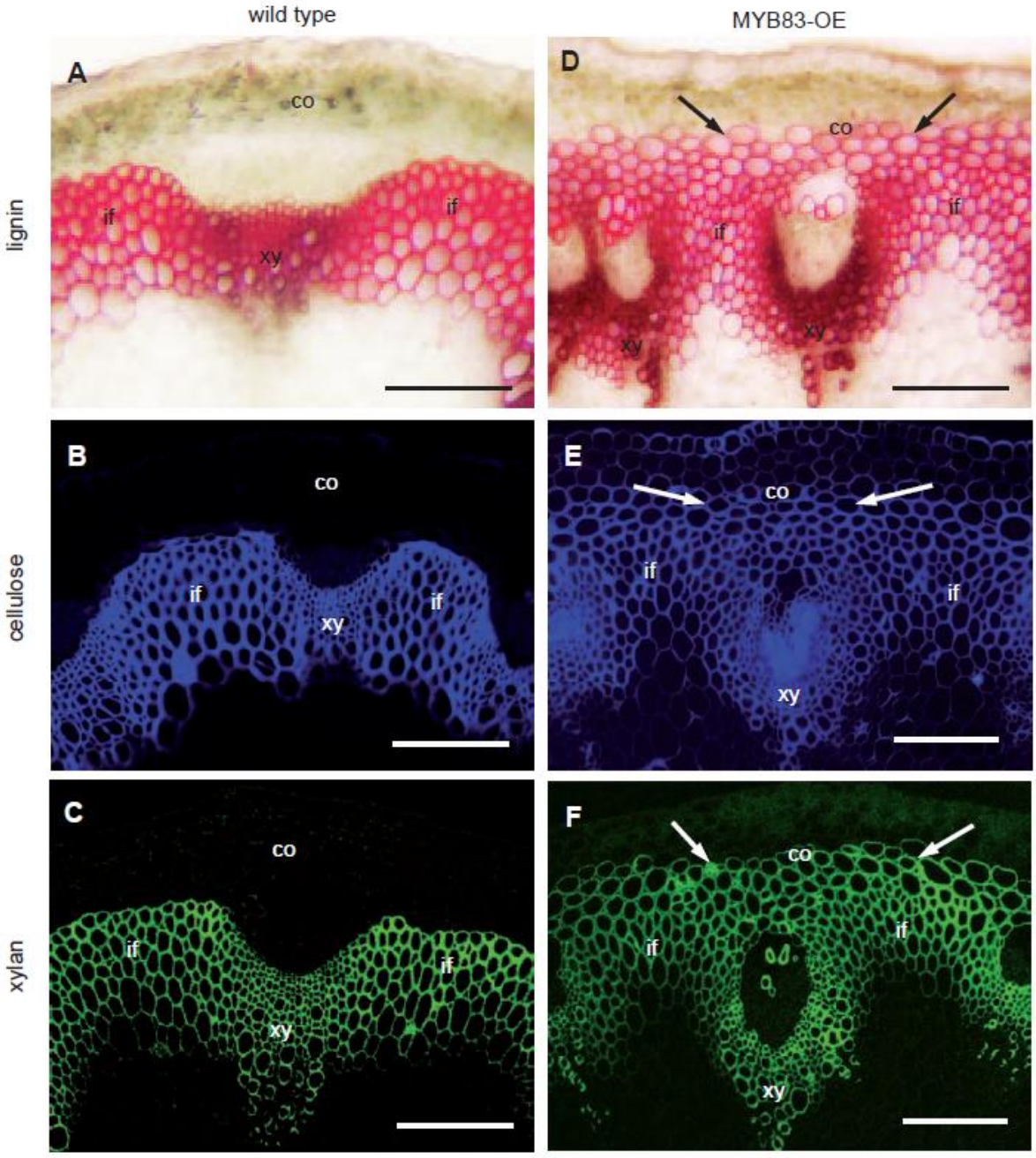
wild type
MYB83-OE
lignin
cellulose
xylan
A
B
C
D
E
F
co
if
xy

Figure 5. MYB83 overexpression induces ectopic deposition of secondary wall components in the cortex of stems. Cross-sections of stems of the wild type (A)–(C) and a MYB83 overexpressor (D)–(F) stained for lignin with phloroglucinol HCl (A, D), cellulose with Calcofluor White (B, E), and xylan with the LM10 xylan antibody (C, F). Note the ectopic deposition of lignin, cellulose and xylan in the cortical cells (arrows) of MYB83 overexpressors: co, cortex; if, interfascicular fiber; xy, xylem. Bars = 121 μm.

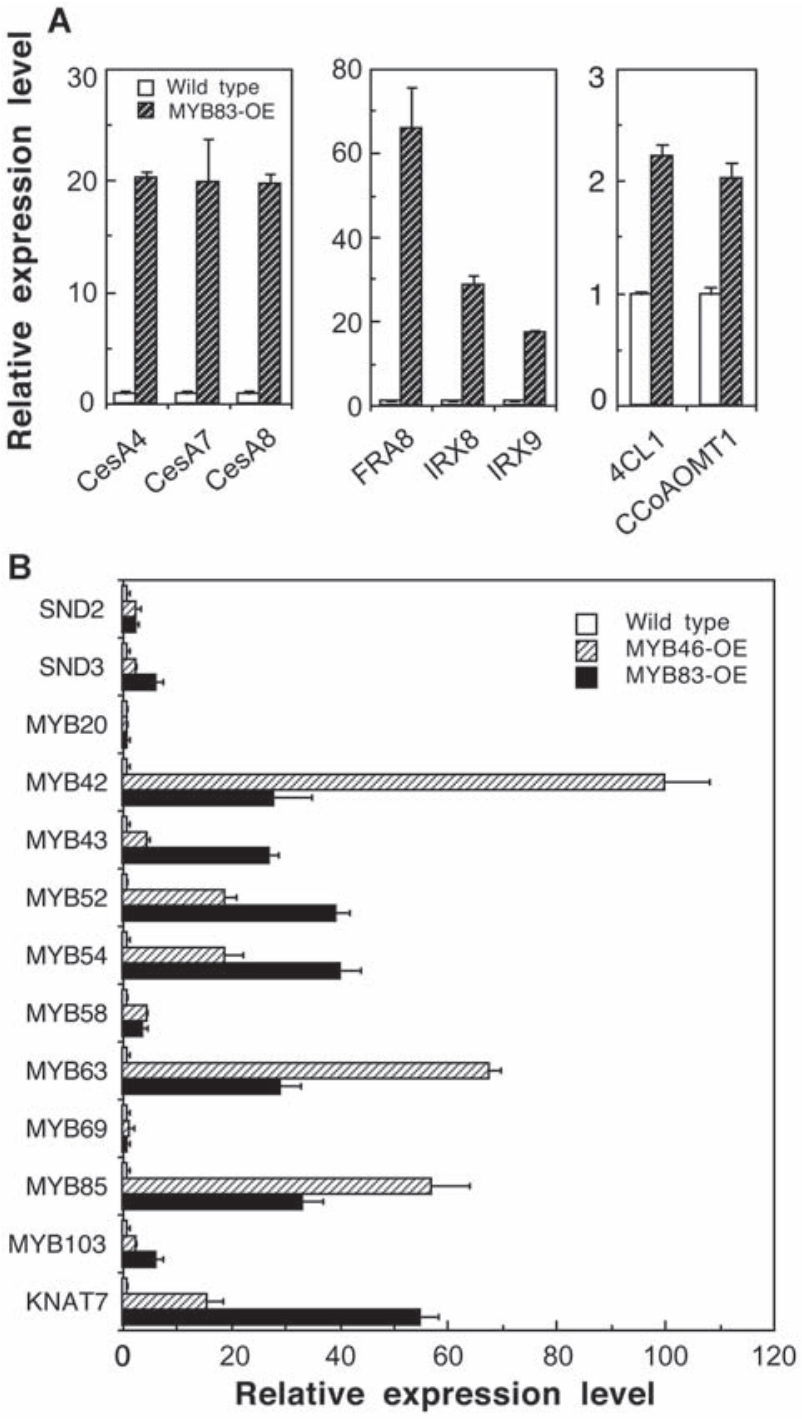
A
Relative expression level
Wild type
MYB83-OE
30
20
10
0
80
60
40
20
0
3
2
1
0
CesA4
CesA7
CesA8
FRA8
IRX8
IRX9
4CL1
CCoAOMT1
B
SND2
SND3
MYB20
MYB42
MYB43
MYB52
MYB54
MYB58
MYB63
MYB69
MYB85
MYB103
KNAT7
Wild type
MYB46-OE
MYB83-OE
0
20
40
60
80
100
120
Relative expression level

Figure 6. Induction of the expression of secondary wall biosynthetic genes and secondary wall-associated transcription factors by MYB83 overexpression. Seedlings of 3-week-old wild type and MYB83 overexpressors were examined for the expression of genes of interest using quantitative PCR. The expression level of genes of interest in the wild type was set to 1. Error bars denote SE of three biological replicates. (A) MYB83 overexpression induced the expression of secondary wall biosynthetic genes for cellulose (*CesA4, CesA7* and *CesA8;* Taylor et al. 2004), xylan (*FRA8, I RX8* and *I RX9;* Pena et al. 2007) and lignin (*4 CL1* and *CCoAOMT1;* Raes et al. 2003). (B) MYB83 overexpression induced the expression of some of the SND1-regulated, secondary wall associated transcription factors.

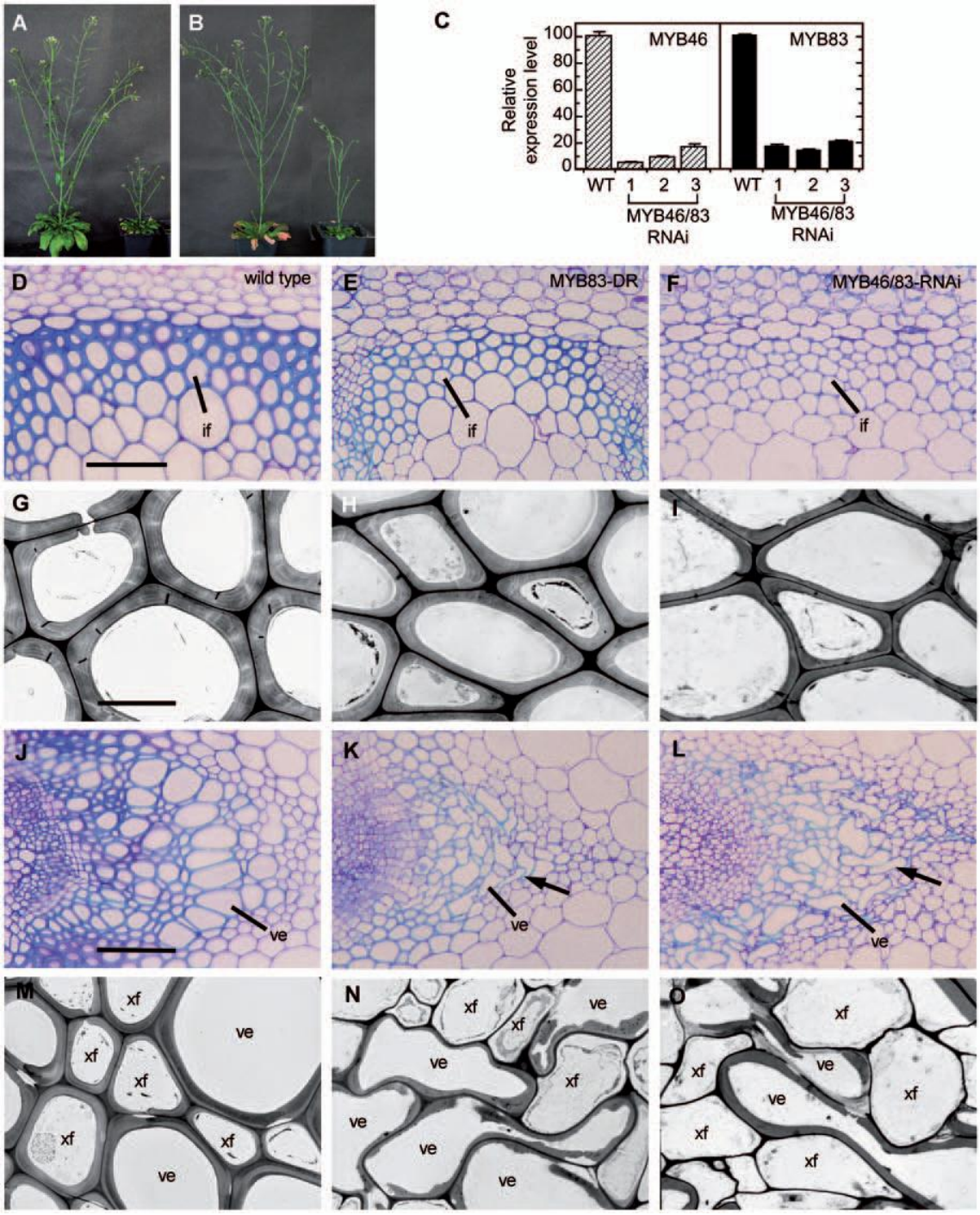
A
B
C
Relative expression level
100
80
60
40
20
0
MYB46
MYB83
WT
1
2
3
WT
1
2
3
MYB46/83 RNAi
MYB46/83 RNAi
D
wild type
if
E
MYB83-DR
if
F
MYB46/83-RNAi
if
G
H
I
J
ve
K
ve
L
ve
M
xf
ve
xf
xf
xf
xf
ve
N
xf
xf
ve
ve
xf
ve
ve
ve
O
xf
ve
xf
ve
xf
xf
xf

Figure 7. Reduction in secondary wall thickness and deformation of vessels by dominant repression of MYB83 (MYB83-DR) or simultaneous RNAi inhibition of *MYB83* and *MYB46* (MYB83/46-RNAi). (A) Wild-type plant (left) and MYB83 dominant repressor (right). (B) Wild-type plant (left) and transgenic plant with simultaneous RNAi inhibition of *MYB83* and *MYB46* (right). (C) Quantitative PCR analysis showing a reduction in the mRNA levels of *MYB46* and *MYB83* in the stems of three representative *MYB83* and *MYB46* RNAi lines (MYB83/46 RNAi) compared with the wild type. The expression level of each gene in the wild type was set to 100. Error bars represent SE of three technical replicates. (D)–(F) Cross-sections of interfascicular regions showing interfascicular fibers with reduced wall thickening in MYB83-DR (E) and MYB83/46-RNAi (F) compared with the wild type (D). (G)–(I) Transmission electron micrographs of interfascicular fiber walls of the wild type (G), MYB83-DR (H) and MYB83/46-RNAi (I). (J)–(L) Cross-sections of vascular bundles showing severely deformed vessels in MYB83-DR (K) and MYB83/46-RNAi (L) compared with the wild type (J). Transmission electron micrographs of vessels and xylary fibers with reduced wall thickness in MYB83-DR (N) and MYB83/46-RNAi (O) compared with the wild type (M). The bottom internodes of the wild type and transgenic plants with dominant repression or RNAi inhibition were examined for secondary wall thickening in fibers and vessels by light microscopy [(D)–(F) and (J)–(L)] and transmission eletron microscopy [(G)–(I) and (M)–(O)]: if, interfascicular fiber; ve, vessel; xf, xylary fiber. Bars = 43 μm in (D) for (D)–(F) and (J)–(L) and 8.5 μm in (G) for (G)–(I) and (M)–(O).

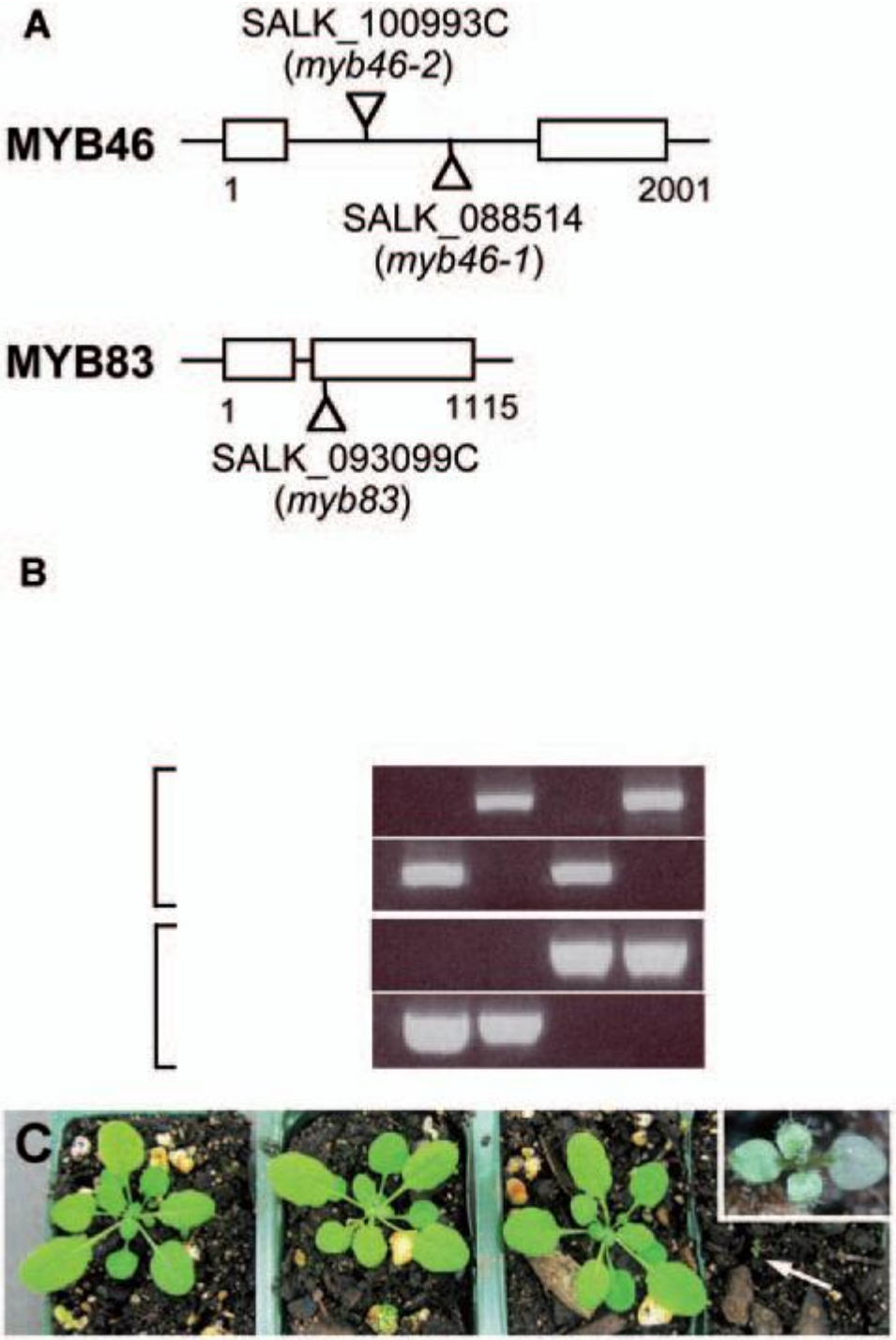
A
SALK_100993C
(myb46-2)
MYB46
1
2001
SALK_088514
(myb46-1)
MYB83
1
1115
SALK_093099C
(myb83)
B
C

Figure 8. Double T-DNA knockout mutations of *MYB83* and *MYB46* cause growth arrest. (A) Insertion sites of T-DNA in the *MYB46* and *MYB83* genes. (B) Identification of homozygous T-DNA insertions in the *myb46-1 myb83* double mutant by PCR amplification. T-DNA denotes the amplified DNA fragment with a T-DNA left border primer and a *MYB46* or *MYB83* primer flanking the T-DNA insertion site. Endogenous denotes the amplified *MYB46* or *MYB83* DNA fragment with primers spanning the T-DNA insertion site. (C) Three-week-old plants of wild type, single mutants of *myb46-1* and *myb83,* and the double mutant *myb46-1 myb83.* The *myb46-1 myb83* double mutant only developed a pair of leaves (higher magnification in inset) and was arrested in growth (arrow).

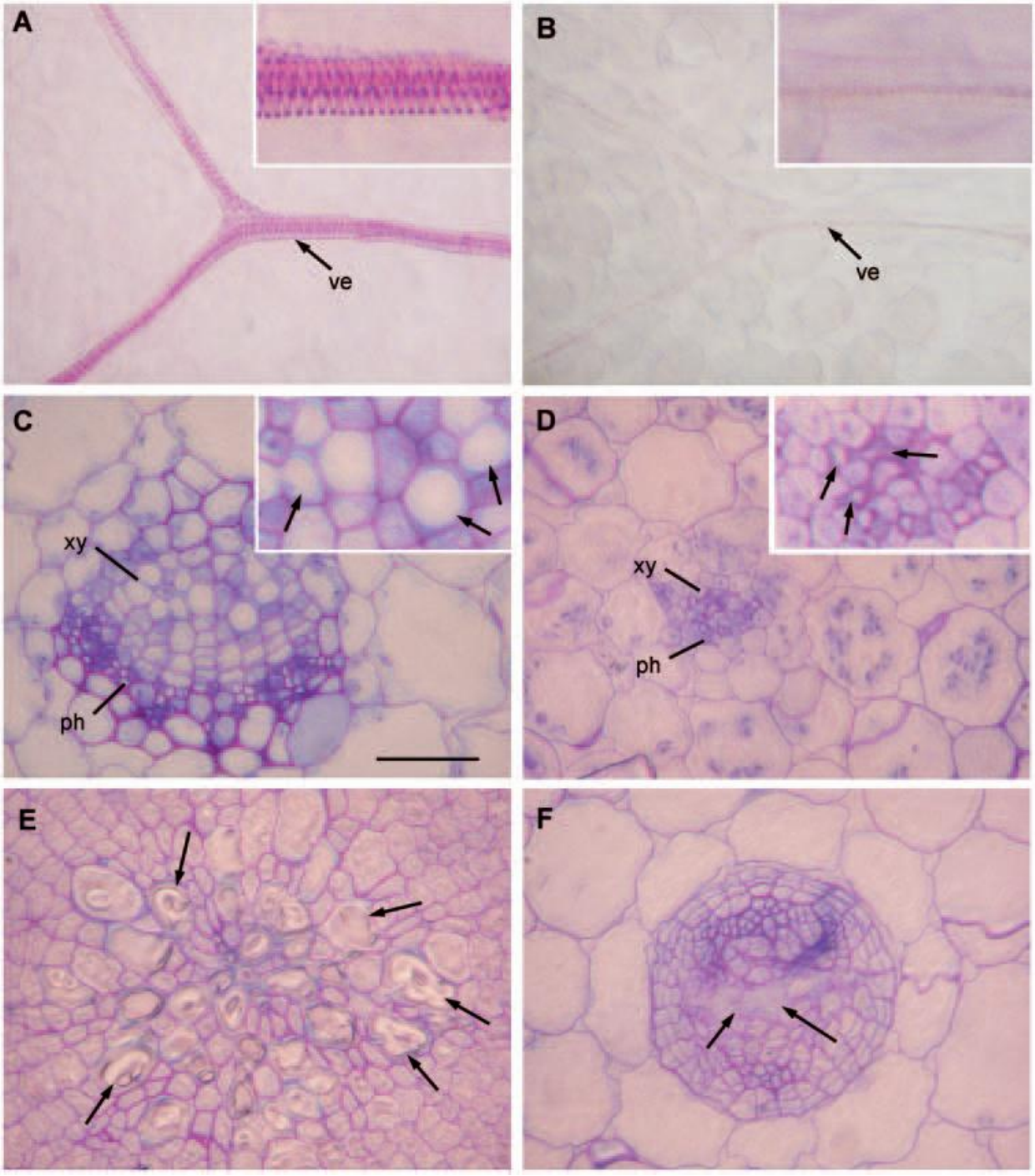
A
ve
B
ve
C
xy
ph
D
xy
ph
E
F

Figure 9. Double T-DNA knockout mutations of *MYB83* and *MYB46* lead to a loss of secondary walls in vessels. (A), (B) Lignin staining of leaves showing a nearly lack of secondary wall thickening in the veins of *myb46-1 myb83* (B) compared with the helical secondary wall thickening in those of the wild type (A). Insets show higher magnifications of veins. (C), (D) Cross-sections of leaf veins showing severely deformed vessels with no apparent secondary wall thickening in *myb46-1 myb83* (D) compared with the wild type (C). Insets show higher magnifications of vessels (arrows) in the vein. (E), (F) Cross-sections of roots showing the absence of discernable vessels in *myb46-1 myb83* (F) compared with the wild type (E). Arrows in (E) point to vessels with thick walls, and those in (F) point to the location where vessels are normally developed: ph, phloem; ve, vessel. Bar in (C) = 27 μm for (C)–(F).

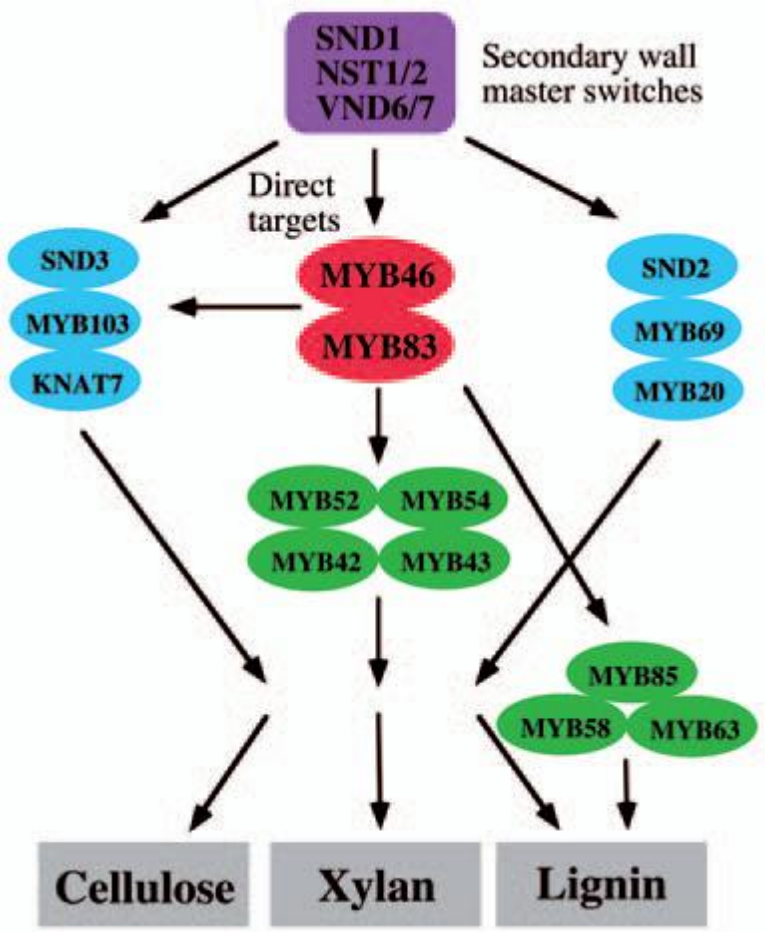

Figure 10. Diagram of the transcriptional regulatory network controlling secondary wall biosynthesis. SND1 and its close homologs are master switches at the top of the network activating a cascade of downstream transcription factors. Among them, MYB83 and MYB46 are direct targets of SND1 and its close homologs and they regulate the expression of a number of further downstream secondary wall associated transcription factors.

CHAPTER 3

THE POPLAR MYB TRANSCRIPTION FACTORS, PTRMYB3 AND PTRMYB20, ARE INVOLVED IN THE REGULATION OF SECONDARY WALL BIOSYNTHESIS[1]

[1]Ryan L. McCarthy, Ruiqin Zhong, Scott Fowler, David Lyskowski, Hasanthi Piyasena, Kara Carleton, Charles Spicer and Zheng-Hua Ye. 2010. *Plant Cell Physiology*. 51(6):1084-1090
Reprinted here with permission of the publisher

Abstract

Wood represents the world's most abundant biomass and a valuable renewable resource. Dicot wood is mainly composed of cellulose, xylan and lignin formed together into the secondary walls of xylem cells. Secondary cell wall formation requires the coordinated regulation of cellulose, xylan and lignin biosynthesis. While many transcription factors regulating this process have been identified in *Arabidopsis* we know very little about how it is regulated in poplar. In this report, we demonstrate that the poplar wood-associated MYB transcriptional activators, PtrMYB3 and PtrMYB20, activate the biosynthetic pathways of cellulose, xylan and lignin when overexpressed in Arabidopsis and they are also able to activate the promoter activities of poplar wood biosynthetic genes. We also show that PtrMYB3 and PtrMYB20 are functional orthologs of the *Arabidopsis* secondary cell wall master regulators MYB46 and MYB83. Additionally the expression of PtrMYB3 and PtrMYB20 is directly activated by poplar PtrWND2, a functional ortholog of *Arabidopsis* SND1, suggesting their involvement in the regulation of wood formation in poplar and providing evidence of evolutionary conservation between herbaceous and woody dicots of the transcriptional network controlling secondary cell wall formation.

Introduction

Vascular plants are the primary sink of fixed carbon dioxide on earth, and a large proportion of this fixed carbon is stored as lignocellulosic polymers in plant secondary cell walls. Lignocellulosic polymers are the principal components of wood, which represents the most abundant biomass produced by land plants. Wood is primarily comprised of xylem vessel and fiber cells which possess thick secondary cell walls composed of mainly cellulose, xylan and lignin. Because of the importance of wood in many industrial applications, such as pulping and paper-making, tremendous efforts have been put into understanding how lignocellulosic polymers in wood are synthesized in the hope of genetically engineering trees with altered wood composition. Recently, due to the potential of lignocellulosic polymers as a renewable source of biofuel production, study of wood formation has gained a renewed interest. To utilize lignocellulosic biomass efficiently for biofuel production, it is imperative to have a complete picture of how individual polymers in wood are biosynthesized and how their biosynthesis is transcriptionally regulated. Uncovering the transcriptional regulators controlling wood biosynthesis will probably provide novel tools to alter the biosynthetic pathways of wood components based on our needs. Transcriptome profiling of wood formation has identified a number of transcription factors that are preferentially expressed in developing wood. Some of these transcription factors show close homology to Arabidopsis meristem identity genes and were proposed to regulate the cambial activity (Schrader et al., 2004), whereas others were suggested to mediate the regulation of the biosynthesis of lignocellulosic polymers in wood (Bedon et al., 2007; Wilkins et al. 2009). Due to the difficulty of genetic studies of gene functions in tree species, most of these wood-associated transcription factors have not yet been subjected to functional characterization. Among the few well-characterized transcription factors

are several MYBs from pine (PtMYB1, PtMYB4 and PtMYB8) and eucalyptus (EgMYB2). These pine and eucalyptus MYB genes have been shown to cause ectopic deposition of lignin or altered phenylpropanoid metabolism when overexpressed in tobacco or spruce (Patzlaff et al., 2003; Goicoechea et al., 2005; Bomal et al., 2008). Because of their abilities to bind to the AC elements that are commonly present in the promoters of lignin biosynthetic genes, these pine and eucalyptus MYBs have been proposed to regulate lignin biosynthesis. However, it is not clear whether they also regulate the biosynthesis of other secondary wall components during wood formation. PtMYB4 and EgMYB2 are phylogenetically closely related to Arabidopsis MYB46 and MYB83 (Fig. 1A), which are direct targets of a group of secondary wall NAC (SWN) domain master regulators, including SND1, NST1, NST2, VND6 and VND7, controlling secondary wall biosynthesis in Arabidopsis (Kubo et al., 2005; Mitsuda et al., 2005; Zhong et al., 2006; Mitsuda et al., 2007; Zhong et al., 2007b; Zhou et al. 2009). MYB46 and MYB83 have been shown to be another level of key switches capable of activating the entire secondary wall biosynthetic program and they function redundantly in regulating the secondary wall biosynthesis in fibers and vessels (Zhong et al., 2007a; McCarthy et al., 2009). Although recent progress has greatly improved our understanding of the transcriptional network regulation secondary cell wall formation in *Arabidopsis* we know very little about secondary cell wall regulation in woody dicots such as poplar. In order to effectively utilize poplar as a bioenergy crop we need to improve our understanding by identifying regulators of secondary wall formation. In this report, we demonstrate that the poplar wood-associated MYB transcription factors, PtrMYB3 and PtrMYB20, are functional orthologs of Arabidopsis MYB46 and MYB83, and capable of activating the biosynthetic pathways of cellulose, xylan and lignin, suggesting

that they are involved in the regulation of the biosynthesis of all three major wood components in poplar.

Results

PtrMYB3 and PtrMYB20 are transcriptional activators expressed in secondary cell wall forming tissue

Although *Arabidopsis* and poplar physiology is considerably different the process of secondary cell wall formation as it occurs on a cellular level is believed to be very similar (Nieminen et al., 2004). This similarity presents the possibility that elements of the transcriptional network regulating secondary cell wall formation may be evolutionarily conserved. A search of the poplar (*Populus trichocarpa)* genome for close homologs of Arabidopsis MYB46 and MYB83 revealed four MYB transcription factors, PtrMYB2, PtrMYB21, PtrMYB3 and PtrMYB20 (Fig. 1A; Wilkins et al. 2009). We attempted to PCR-amplify their full-length cDNAs from RNAs isolated from woody tissues of poplar stems and were successful with only two of them, *PtrMYB3* and *PtrMYB20*. This result is consistent with poplar microarray data from Wilkins et al., (2008) which demonstrated PtrMYB3 and PtrMYB20 to exhibit high expression specifically in poplar xylem. Therefore, these two MYBs were subjected to functional characterization. In situ hybridization revealed that *PtrMYB3* and *PtrMYB20* were expressed preferentially in vessels and fibers in developing wood (Fig. 2A–C), indicating their close association with wood formation. In addition, their expression was evident in phloem fibers and ray parenchyma cells (Fig. 2A, B). PtrMYB3 and PtrMYB20 expression in xylem fibers, vessels and ray parenchyma was confirmed in transgenic poplar lines expressing the GUS reporter gene driven by the PtrMYB3 and PtrMYB20 promoters (Fig. 3). This expression was coincident with the expression

of wood biosynthetic genes PtrCesA8, PtrGT43B and PtrCCoAOMT which are involved in cellulose, xylan and lignin biosynthesis respectively (Fig. 2D–G). PtrMYB3 and PtrMYB20 were targeted to the nucleus when expressed in Arabidopsis protoplasts (Fig. 1B–G) and had transcriptional activation activities in yeast (Fig. 1H), indicating that they are transcriptional activators.

Overexpression of PtrMYB3 and PtrMYB20 induces ectopic deposition of secondary walls in Arabidopsis

To investigate whether PtrMYB3 and PtrMYB20 are involved in the regulation of secondary wall biosynthesis, we generated transgenic Arabidopsis plants overexpressing PtrMYB3 or PtrMYB20. For each construct, 64 transgenic plants were inspected and at least 48 of them exhibited a curly leaf phenotype (Fig. 4N,O), which is typical of the MYB46 and MYB83 overexpressors (Zhong et al., 2007a; McCarthy et al., 2009). Examination of leaves revealed ectopic deposition of lignin in the walls of epidermal cells in these overexpression lines compared with the wild type, in which lignin was only present in veins (Fig. 4A–F). It was apparent that the lignified walls of these epidermal cells were significantly thicker than non-lignified walls. Further examination of other wall components demonstrated that cellulose and xylan were also ectopically deposited in the walls of these cells (Fig. 4G–L). Similarly, ectopic deposition of cellulose, xylan and lignin was also seen in the walls of cortical cells or pith cells in the stems of PtrMYB3 and PtrMYB20 overexpressors (Fig. 5A–I). These results provide direct evidence demonstrating that PtrMYB3 and PtrMYB20 are capable of activating the biosynthetic pathways of all three principal components of secondary walls.

Activation of Arabidopsis secondary wall-associated transcription factors by PtrMYB3 and PtrMYB20

Consistent with this finding, quantitative PCR analysis revealed that overexpression of PtrMYB3 or PtrMYB20 resulted in an elevation in the expression of secondary wall biosynthetic genes for cellulose (*CesA4, CesA7* and *CesA8),* xylan (*FRA8, IRX8* and *IRX9)* and lignin (*CCoAOMT1* and *4CL1)* (Fig. 7A). It has been shown that Arabidopsis MYB46 and MYB83 up-regulate the expression of a number of secondary wall-associated transcription factors (Zhong et al., 2007a; McCarthy et al., 2009). We next investigated whether PtrMYB3 and PtrMYB20 were also able to induce the expression of these transcription factors. It was found that the expression levels of these transcription factors were elevated in both PtrMYB3 and PtrMYB20 overexpressors (Fig. 7B). We further demonstrated that PtrMYB3 and PtrMYB20 were able to complement the growth arrest and the vessel wall-thickening defect in the *myb46 myb83* double mutant (Fig. 6A, B), suggesting that PtrMYB3 and PtrMYB20 are functional orthologs of MYB46 and MYB83.

PtrMYB3 and PtrMYB20 are directly activated by PtrWND2 and induce expression of poplar secondary wall biosynthetic genes

To substantiate further the roles of PtrMYB3 and PtrMYB20 in the regulation of secondary wall biosynthesis, we examined whether they were capable of activating the promoters of poplar wood biosynthetic genes. Transactivation analysis in Arabidopsis protoplasts showed that PtrMYB3 and PtrMYB20 were able to activate the promoters of representative biosynthetic genes for cellulose (*PtrCesA8),* xylan (*PtrGT43B)* and lignin (*PtrCCoAOMT1)* (Fig. 8A), indicating that PtrMYB3 and PtrMYB20 are involved in regulation of the expression of wood biosynthetic genes in poplar. In addition, we found that the promoters of *PtrMYB3* and *PtrMYB20* could be directly activated by PtrWND2 (Fig. 8B), a poplar close homolog of

Arabidopsis SND1 (Zhong et al., 2010), suggesting that PtrMYB3 and PtrMYB20 are direct targets of PtrWND2.

Overexpression of PtrMYB3 and PtrMYB20 induces ectopic secondary cell wall formation in Poplar

To determine whether PtrMYB3 and PtrMYB20 are capable of functioning as master regulators of secondary cell wall formation in poplar as well as *Arabidopsis* we generated transgenic overexpression poplar lines PtrMYB3-OE and PtrMYB20-OE. More than 50 independent transgenic poplar lines were generated for each gene and lines were screened for ectopic secondary cell wall formation. Ectopic secondary cell wall formation was assayed by lignin staining using Phloroglucinol-HCL. Lignin staining of poplar transformed with empty vector demonstrated lignin formation restricted to secondary xylem and phloem fibers (Fig. 9A). PtrMYB3-OE and PtrMYB20-OE lines exhibited ectopic lignin deposition in cortex and phloem parenchyma cells (Fig. 9B; unpublished data). Further examination of cells with ectopic lignin deposition revealed thick secondary walls indicating that these cells likely also exhibit ectopic cellulose and xylan deposition as lignin alone cannot produce significant secondary wall thickening.

Discussion

We have demonstrated that PtrMYB3 and PtrMYB20 activate the expression of secondary wall biosynthetic genes and induce ectopic deposition of cellulose, xylan and lignin when overexpressed in *Arabidopsis* and poplar. Together with the fact that they are also able to activate the promoters of poplar wood biosynthetic genes, these results suggest that PtrMYB3 and PtrMYB20 are involved in the regulation of the entire secondary wall biosynthetic program

during wood formation in poplar. Additionally since they are capable of functioning in both *Arabidopsis* and poplar, PtrMYB3 and PtrMYB20 must bind to the same set of target genes and induce similar expression profiles as MYB46 and MYB83. This indicates the presence of an evolutionarily conserved cis-element bound by secondary cell wall MYB domain master regulatory proteins. The high level of observed functional conservation may indicate that the transcriptional network regulating secondary cell wall formation may be broadly conserved among vascular plants. Conversely, differences in vascular tissue and wood physiology imply the existence of some elements of the transcriptional network regulating secondary cell wall formation that are distinct between species. Compared to *Arabidopsis*, poplar wood has the added complexity of xylem ray parenchyma cells which facilitate radial and lateral transport and often undergo severely delayed secondary cell wall formation (Murakami et al., 1999). While PtrMYB3 and PtrMYB20 were observed to be expressed not only in xylem fibers and vessels but also xylem rays it is possible that novel transcription factors not present in *Arabidopsis* may be necessary to accommodate secondary cell wall formation in this distinct cell type.

Since PtMYB4 and EgMYB2 are phylogenetically closely grouped with PtrMYB3 and PtrMYB20 (Fig. 1A), it is likely that they also regulate the biosynthesis of all three major wood components. Further studies of PtMYB4 and EgMYB2 function in *Arabidopsis* and their native species are necessary to more fully understand the evolutionary conservation of secondary cell wall regulators. Our study provides molecular evidence suggesting that although Arabidopsis and poplar shared their last common ancestor >100 million years ago (Tuskan et al., 2006), the mechanism underlying the transcriptional regulation of secondary wall biosynthesis is well conserved.

A genome duplication event in the poplar lineage approximately 65 million years ago resulted in the significant expansion of many gene families (Tuskan et al., 2006) including MYB and NAC domain transcription factors (Wilkins et al., 2008; Hu et al., 2010). In *Arabidopsis* the five SWNs operate as pairs in a tissue specific manner with SND1 and NST1 functioning in fibers, NST1 and NST2 in anther endothecium and VND6 and VND7 in xylem vessels (Mitsuda et al., 2007; Zhong et al., 2006, 2007b). Presumably the subfunctionalization of the SWNs into cell type specific pathways occurred after gene duplication of an ancestral SWN. Gene duplication events are well known to contribute to gene subfunctionalization and neofunctionalization through relaxing selective pressure on redundantly functioning duplicate genes (Adams and Wendel, 2005; Flagel and Wendel, 2009). The vascular tissue of basal vascular plants such as the lycophyte *Selaginella* is comprised of only tracheids which perform dual roles in water transport and structural support in contrast to angiosperms where these roles are separately performed by xylem vessels and fibers (Schulz et al., 2010). Although we currently have no knowledge of secondary cell wall regulation in *Selaginella*, presumably fewer cell type specific pathways for activating secondary cell wall biosynthesis are present in the basal vascular plants as there are fewer cell types forming secondary walls. The expansion of the MYB domain transcription factor family in poplar creates the possibility of subfunctionalization among PtrMYB2, PtrMYB3, PtrMYB20 and PtrMYB21 similar to what has been observed with the SWNs in *Arabidopsis*.

Here we demonstrated the expression of PtrMYB3 and PtrMYB20 in xylem fibers and vessels during secondary cell wall formation (Fig. 4A, B). According to microarray data collected by Wilkins et al. (2008), PtrMYB2 is expressed primarily in poplar male catkins while PtrMYB21 exhibits a broad expression profile with peak expression in xylem tissue but

moderate expression in young leaves, roots and catkins. It is possible that PtrMYB2 may act to promote secondary cell wall formation specifically in dehiscent anthers and this function may have arisen in a manner similar to the pair of SWNs NST1 and NST2 in *Arabidopsis* (Mitsuda et al., 2005). While we were unable to amplify PtrMYB21 from poplar stem cDNA its observed expression pattern in young leaves, roots and stems (Wilkins et al., 2008) opens the possibility that it functions during primary xylem but not secondary xylem secondary cell wall formation. Since all four of these genes contain full open reading frames capable of expressing detectable transcripts with tissue specific expression it is likely that each performs some vital role in poplar. Dissecting the nature of these roles will provide greater insight into the process of wood formation in poplar and may shed light onto the process of subfunctionalization after gene duplication.

The identification of PtrMYB3 and PtrMYB20 as transcriptional regulators activating the biosynthesis of wood components not only sheds light on our understanding of the molecular control of wood formation but also potentially provides a novel means to improve wood biomass production in tree species. Improving the molecular characteristics or production quantity of poplar biomass could produce substantial gains in biofuel production efficiency which would improve its potential as a sustainable carbon neutral alternative to conventional fossil fuels.

Materials and Methods

Subcellular localization

Total RNA was isolated with a Qiagen RNA isolation kit (Qiagen, Valencia, CA, USA) from various tissues of poplar (*P. trichocarpa)* and *Arabidopsis* plants and used for reverse transcription–PCR or real-time quantitative PCR analysis (Zhong et al., 2006).

In Situ Hybridization

Poplar stems were fixed, embedded in paraffin and sectioned for in situ mRNA localization. The 200bp 3'-untranslated region of the cDNAs of *PtrMYB3, PtrMYB20, PtrCesA8, PtrGT43B* and *PtrCCoAOMT1* were used for synthesis of digoxigenin-labeled antisense and sense RNA probes with the DIG RNA labeling mix (Roche).

Subcellular Localization

The subcellular localization of PtrMYB3, PtrMYB20 and SND1 was dome by expression of YFP tagged PtrMYB3 and PtrMYB20 and CFP tagged SND1 in carrot (*Daucus carota*) protoplasts. Full length gene cDNA was fused in-frame with the YFP and CFP cDNA and ligated between the CaMV 35S promoter and the nopaline synthase terminator in pBI221 (Clontech). The YFP and CFP fusion protein constructs were transfected into carrot protoplasts according to Liu et al. (1994). The transfected protoplasts were incubated in darkness for 20 h before examination and imaging using a Leica TCs SP2 spectral confocal microscope (Leica Microsystems). Images were saved and processed with Adobe Photoshop version 7.0 (Adobe Systems).

Overexpression and Complementation

The overexpression constructs consisting of the full-length cDNAs of *PtrMYB3* and *PtrMYB20* driven by the cauliflower mosaic virus (CaMV) 35S promoter in pBI121 were introduced into wild-type Arabidopsis plants by *Agrobacterium*–mediated transformation. Transgenic plants exhibiting severe phenotypes were selected for phenotypic characterization, and representative results were presented. Leaf and stem tissues were fixed and sectioned for cellulose, xylan and lignin staining as described previously (Zhong et al., 2006). For complementation analysis, the full-length cDNAs of *PtrMYB3* and *PtrMYB20* were ligated under the 3 kb promoter of *MYB46* in pBI121 and transformed into the *myb46 myb83* (+/-) double mutant. Transgenic plants

homozygous for the T-DNA insertions in both *MYB46* and *MYB83* were selected for phenotypic analysis (McCarthy et al., 2009).

Transactivation Analysis

To test the ability of PtrMYB3 and PtrMYB20 to activate the poplar gene promoters, the reporter construct and the effector construct were co-transfected into Arabidopsis leaf protoplasts (Zhong et al., 2006). Another construct containing the firefly luciferase gene driven by the CaMV 35S promoter was also included in the transfection for determination of the transfection efficiency. The β-glucuronidase (GUS) activity was normalized against the luciferase activity in each transfection, and the data are the average of three biological replicates. To test the direct activation of the promoters of *PtrMYB3* and *PtrMYB20* by PtrWND2, the PtrWND2-HER expression construct, which was created by fusing the full-length cDNA of *PtrWND2* with the regulatory region of human estrogen receptor (HER) at the C-terminus (Zuo et al., 2000), was co-transfected with the *PtrMYB3/20* promoter-driven GUS reporter construct into Arabidopsis leaf protoplasts. The transfected protoplasts were treated with estradiol and cycloheximide, and analyzed for gene expression with quantitative PCR as described previously (Zhong et al., 2008).

GUS reporter gene analysis

The 3kb 5' upstream sequence of PtrMYB3 and PtrMYB20 were ligated in front of the GUS reporter gene start codon in the pBI101 vector to create the GUS reporter constructs. *Abrobacterium tumefaciens*-mediated transformation was used to introduce these constructs into poplar (*Populus alba* × *Populus tremula*) as described by Leple et al., (1992). GUS activity was assayed as described previously (Zhong et al., 2005). At least 30 independent transgenic lines were analyzed for GUS expression for each construct.

Overexpression of PtrMYB3 and PtrMYB20 in poplar

PtrMYB3 and PtrMYB20 overexpression constructs (PtrMYB3-OE and PtrMYB20-OE) were created by ligating the full-length PtrMYB3 or PtrMYB20 cDNA downstream of the CaMV 35S promoter in pBI121. *Abrobacterium tumefaciens*-mediated transformation was used to introduce these constructs into poplar (*Populus alba* × *Populus tremula*) as described by Leple et al., (1992). The transgenic poplar seedlings were selected on Murashige and Skoog medium containing 50mgL^{-1} kanamycin and 500 mgL^{-1} carbenicillin. Following root formation transgenic seedlings were transferred to soil and grown under greenhouse conditions. Poplar transformed with empty vector was utilized as a control for comparisons. At least 50 independent transgenic lines were used for analysis for each construct.

Funding

This work was supported by the National Science Foundation [ISO-0744170]; the US Department of Agriculture National Institute of Food and Agriculture [AFRI Plant Biology progam (#2010-65116-20468)].

Acknowledgments

We thank Dr. G. A. Tuskan for providing the stem cutting of *Populus trichocarpa.*

References

Adams, K.L., Wendel, J.F. (2005) Polyploidy and genome evolution in plants. Curr. Opin. Plant Biol. 8: 135-141

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

Bomal, C., Bedon, F., Caron, S., Mansfield, S.D., Levasseur, C., Cooke, J.E., 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

Felsenstein, J. (1989). PHYLIP—Phylogeny Inference Package (Version 3.2). Cladistics 5: 164–166

Flagel, L.E., Wendel, J.F. (2009) Gene duplication and evolutionary novelty in plants. New Phytologist 183: 557-564

Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A, et al. (2005) EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. Plant J. 43: 553–567

Hu, R., Qi, G., Kong, Y., Kong, D., Gao, Q., Zhou, G. (2010) Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. BMC Plant Biol. 10:145

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005) Transcription switches for protoxylem and metaxylem vessel formation. Genes Dev. 19: 1855–1860

Leple, J.C., Brasileiro, A.C.M., Michel, M.F., Delmotte, F., Jouanin, L. (1992) Transgenic poplars: expression of chimeric genes using four different constructs. Plant Cell Rep. 11: 137-141

Lui, Z.B., Ulmasov, T., Shi, X., Hagen, G., Guilfoyle, T.J. (1994) Soybean GH3 promoter contains multiple auxin-inducible elements. Plant Cell 6:645-657

McCarthy, R.L., Zhong, R. and Ye, Z.-H. (2009) MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in *Arabidopsis.* Plant Cell Physiol. 50: 1950–1964

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (2007) NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. Plant Cell 19: 270–280

Mitsuda, N., Seki, M., Shinozaki, K. and Ohme-Takagi, M. (2005) The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickening and are required for anther dehiscence. Plant Cell 17: 2993–3006

Murakami, Y., Funada, R., Sano, Y., Ohtani, J. (1999) The differentiation of contact cells and isolation cells in the xylem ray parenchyma of *Populus maximowiczii*. Annals Bot. 84: 429-435

Nieminen, K.M., Kauppinen, L., Helariutta, Y. (2004) A weed for wood? *Arabidopsis* as a genetic model for xylem development. Plant Physiol. 135: 653-659

Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L.J., Smith, C., et al. (2003) Characterization of a pine MYB that regulates lignification. Plant J. 36: 743–754

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

Schulz, C., Little, D.P., Stevenson, D.W., Bauer, D., Moloney, C., Stutzel, T. (2010) An overview of the morphology, anatomy, and life cycle of a new model species: the lycophyte *Selaginella apoda* (L.) spring. Int. J. Plant Sci. 171(7): 693-712

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

Wilkins, O., Nahal, H., Foong, J., Provart, N.J. and Campbell, M.M. (2009) Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. Plant Physiol. 149: 981–993

Zhong, R., Demura, T. and Ye, Z.-H. (2006) SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis.* Plant Cell 18: 3158–3170

Zhong, R., Lee, C. and Ye, Z.-H. (2010) Functional characterization of poplar wood-associated NAC domain transcription factors. Plant Physiol. 152: 1044–1055

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

Zhong, R., Pena, M.J., Zhou, G.-K., Wood-Jones, A., Richardson, E.A., Morrison III, W.H., Darvill, A.G., York, W.S., Ye, Z.-H. (2005) *Arabidopsis Fragile Fiber8*, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. Plant Cell 17: 3390-3408

Zhong, R., Richardson, E.A. and Ye, Z.-H. (2007a) The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis*. Plant Cell 19: 2776–2792

Zhong, R., Richardson, E.A. and Ye, Z.-H. (2007b) Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis.* Planta 225: 1603–1611

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

Zuo, J., Niu, Q.-W. and Chua, N.-H. (2000) An estrogen receptor based transactivator XVE mediates highly inducible gene expression in transgenic plants. Plant J. 24: 265–273

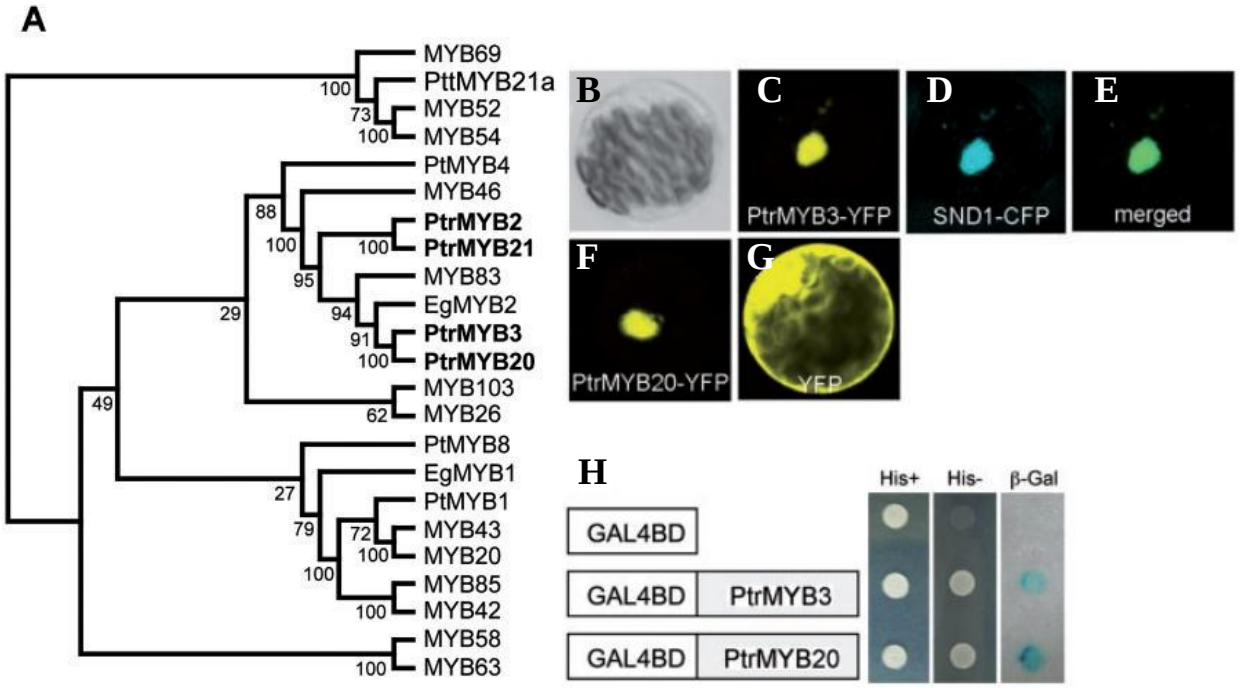

Figure 1. Phylogentic relationship, subcellular localization and transcriptional activation analysis of PtrMYB3 and PtrMYB20. (A) Phylogenetic relationship of PtrMYB3 and PtrMYB20 with other secondary wall- or lignin-associated MYBs. Sequence alignment was performed with the Clustal W program and the phylogenetic tree was constructed using the PHYLIP software (Felsenstein 1989). The bootstrap values are shown in percentages at nodes. (B–H) Subcellular localization of PtrMYB3 and PtrMYB20. An Arabidopsis protoplast (B) co-expressing PtrMYB3–yellow fluorescent protein (YFP) (C) and SND1–cyan fluorescent protein (CFP) (D) shows their co-localization in the nucleus (E; superimposed image). An Arabidopsis protoplast expressing PtrMYB20–YFP (F) or YFP alone (G) shows its localization in the nucleus or in the cytoplasm, respectively. (H) Transcriptional activation analysis showing that PtrMYB3 and PtrMYB20 fused with the GAL4 DNA-binding domain (GAL4BD) could activate the expression of the His3 and β-Gal reporter genes in yeast.

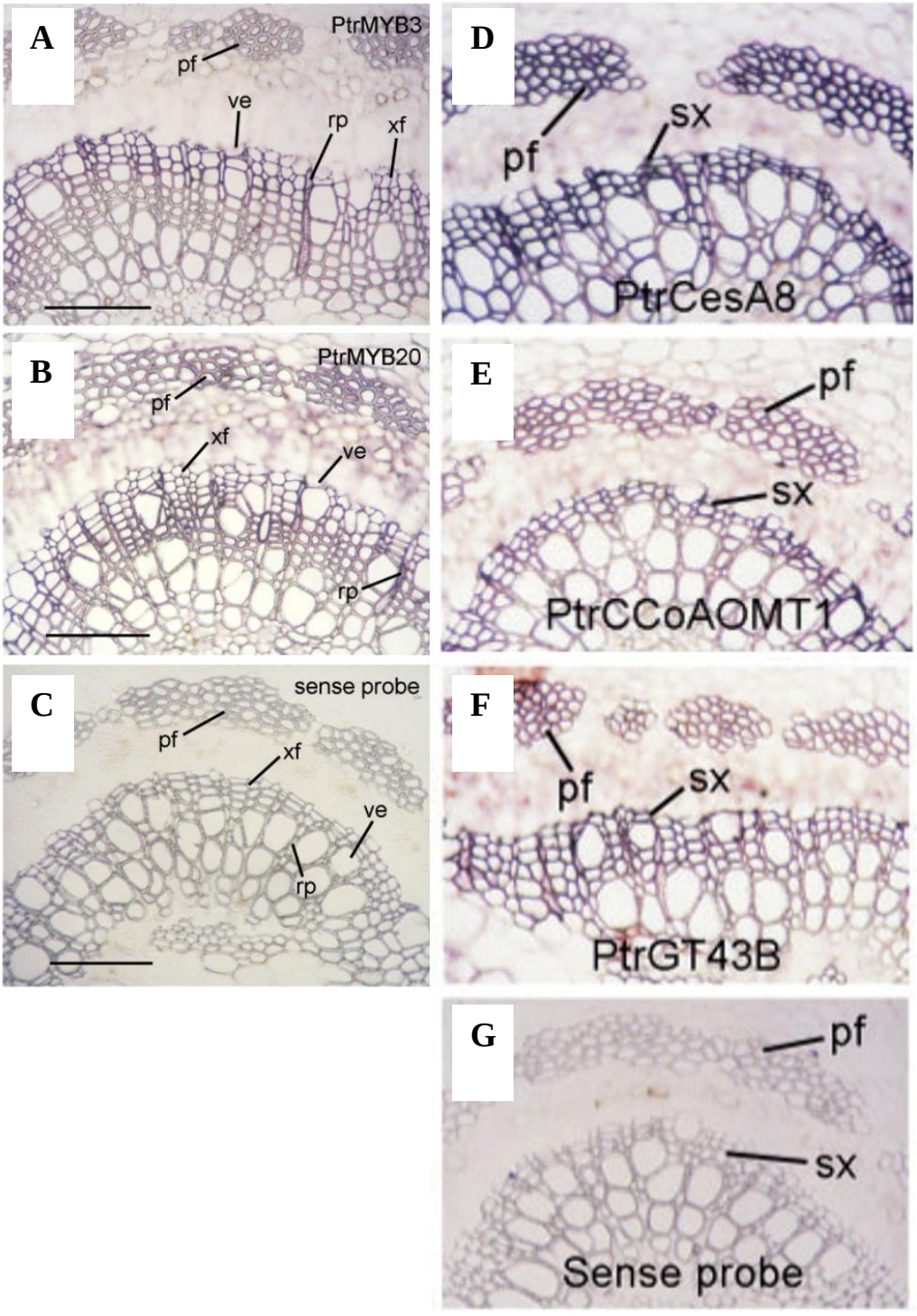
A
PtrMYB3
pf
ve
rp
xf
B
PtrMYB20
pf
xf
ve
rp
C
sense probe
pf
xf
ve
rp
D
pf
sx
PtrCesA8
E
pf
sx
PtrCCoAOMT1
F
pf
sx
PtrGT43B
G
pf
sx
Sense probe

Figure 2. Cell type expression patterns of PtrMYB3, PtrMYB20 and secondary cell wall biosynthetic genes in the developing wood of poplar stems. (A–C) In situ hybridization showing the expression patterns of *PtrMYB3* (A) and *PtrMYB20* (B) in developing wood of poplar stems. The hybridization signal is shown as purple. The control stem section was hybridized with the sense probe of PtrMYB3 (C). pf, phloem fiber; rp, ray parenchyma; ve, vessel; xf, xylary fiber. Bars in B–D = 70 μm. (D–G) In situ mRNA localization of *PtrCesA8* (D), *PtrGT43B* (E) and *PtrCCoAOMT1* (F) showing their expression in developing secondary xylem and phloem fibers in poplar stems. The sense probe of *PtrCesA8* was used as a control (G). pf, phloem fiber; sx, secondary xylem.

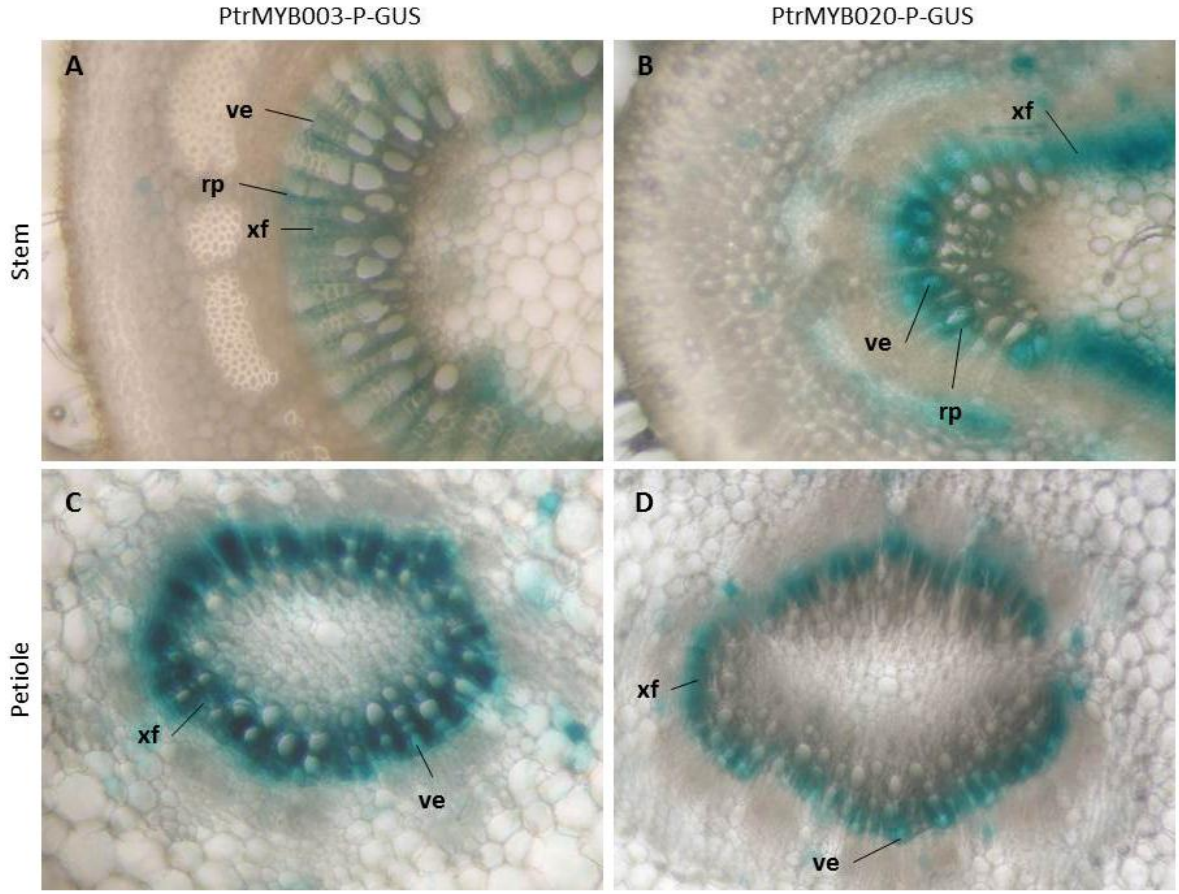

Figure 3. Expression patterns of PtrMYB003 and PtrMYB020 genes in poplar stems and petioles. Transgenic poplar plants expressing the GUS reporter gene under control of the PtrMYB003 promoter in (A) stems and (B) petioles, and the PtrMYB020 promoter in (C) stems and (D) petioles demonstrating GUS expression in xylem vessels and fibers. Stem and petiole sections from over 30 plants were sectioned and stained for GUS expression; representative data are shown. Expression of both genes was observed in developing xylem fibers, vessels and rays but not in mature xylem cells. rp, ray parenchyma; ve, vessel; xf, xylary fiber.

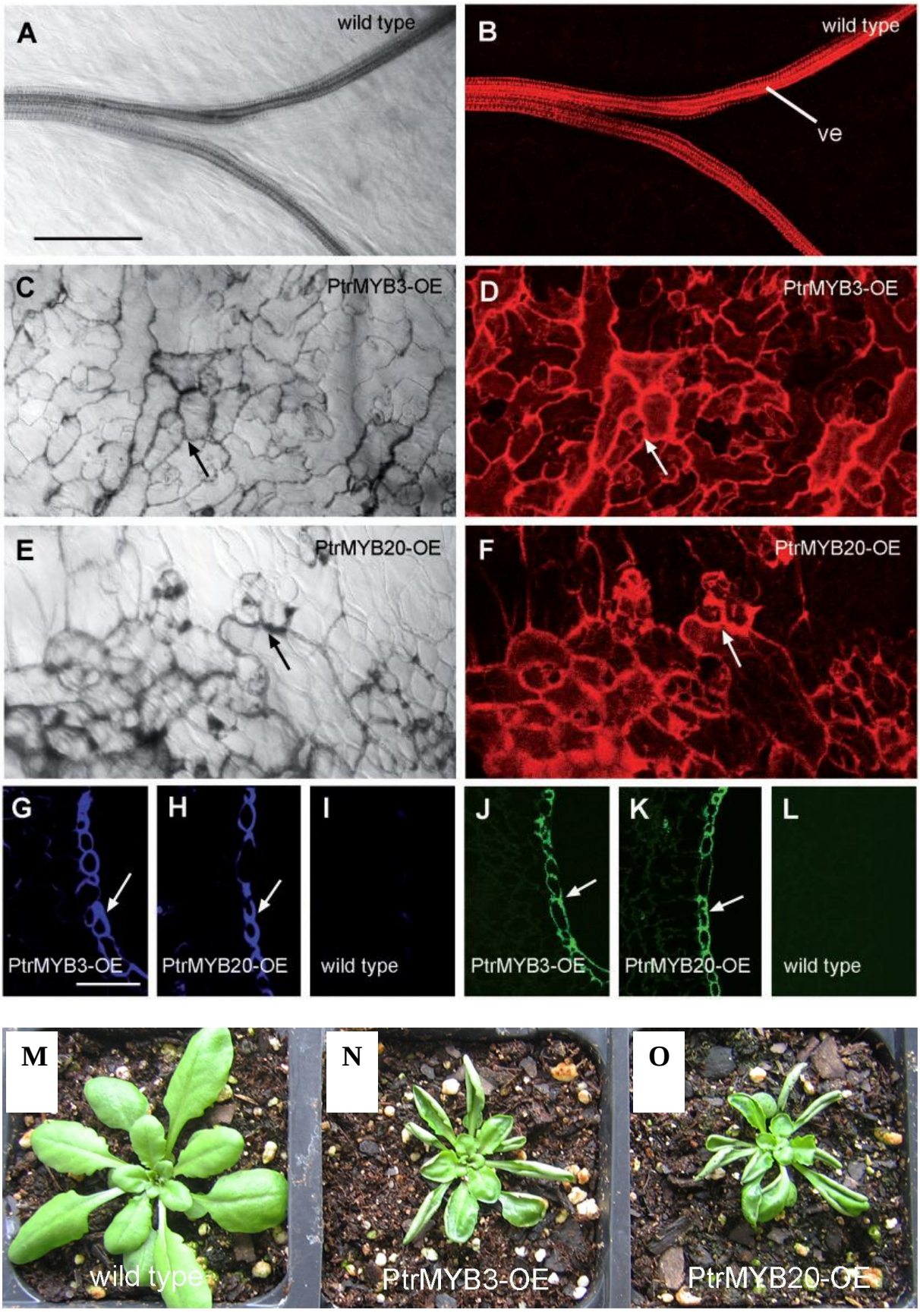
A
wild type
B
wild type
ve
C
PtrMYB3-OE
D
PtrMYB3-OE
E
PtrMYB20-OE
F
PtrMYB20-OE
G
PtrMYB3-OE
H
PtrMYB20-OE
I
wild type
J
PtrMYB3-OE
K
PtrMYB20-OE
L
wild type
M
wild type
N
PtrMYB3-OE
O
PtrMYB20-OE

Figure 4. Overexpression of PtrMYB3 (PtrMYB3-OE) or PtrMYB20 (PtrMYB20-OE) in Arabidopsis results in ectopic deposition of lignin (A–F), cellulose (G–I) and xylan (J–L) in the walls of leaf epidermal cells (arrows) of 4-week-old transgenic seedlings compared with the wild type. Lignin, cellulose and xylan were examined using UV autofluorescence, Calcofluor White staining and immunostaining with the xylan LM10 antibody, respectively. ve, vein. Bar in A = 85 μm for A–F, and bar in G = 58 μm for G–L. Overexpression of PtrMYB3 and PtrMYB20 results in a curly leaf phenotype (N,O) which is not present in wild type (M).

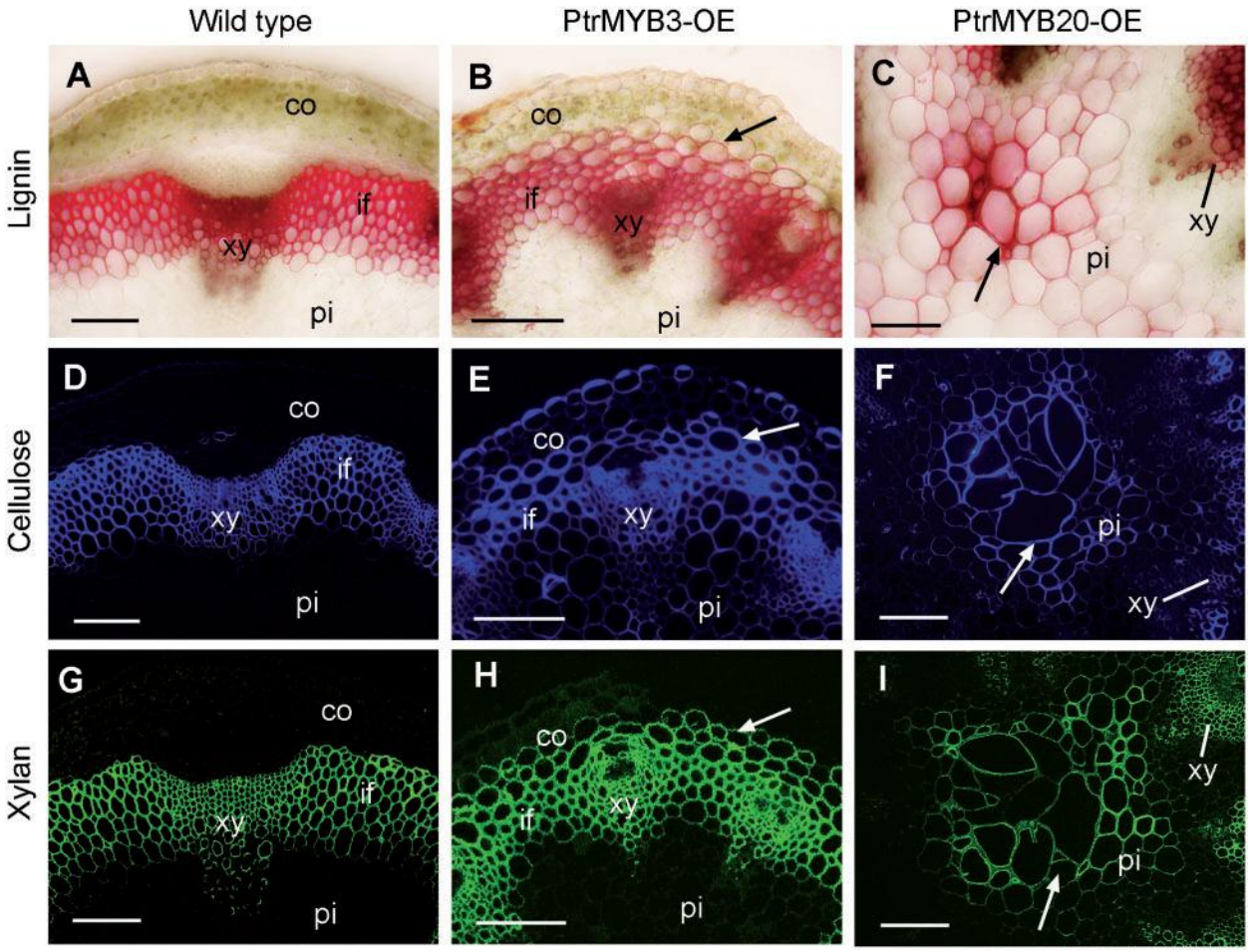

Figure 5. Overexpression of PtrMYB3 (PtrMYB3-OE) or PtrMYB20 (PtrMYB20-OE) in Arabidopsis results in ectopic deposition in *Arabidopsis* stem tissues. (A–I) Ectopic deposition of lignin, cellulose and xylan in the walls of cortical cells or pith cells (arrows) in the stems of PtrMYB3 and PtrMYB20 overexpressors compared with the wild type. co, cortex; if, interfascicular fiber; pi, pith; xy, xylem. Bars = 145 μm.

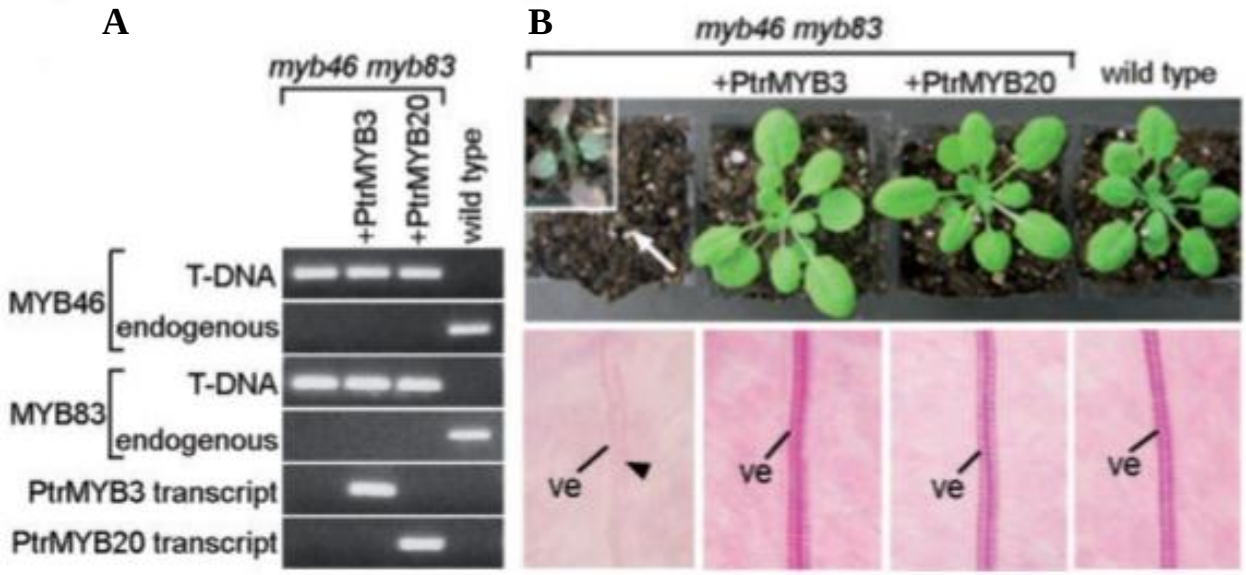

Figure 6. Complementation of the *Arabidopsis myb46 myb83* double mutant by expressiobn of PtrMYB3 and PtrMYB20. (A) Detection of the *PtrMYB3* and *PtrMYB20* transcripts in the transgenic *myb46 myb83* double mutant. (B) The growth arrest (arrow; higher magnification of *myb46 myb83* in inset) and the vessel wall-thickening defect (arrowhead) of *myb46 myb83* were complemented by the expression of PtrMYB3 or PtrMYB20. The veins (ve) of leaves stained with phloroglucinol-HCl were displayed below the corresponding plants.

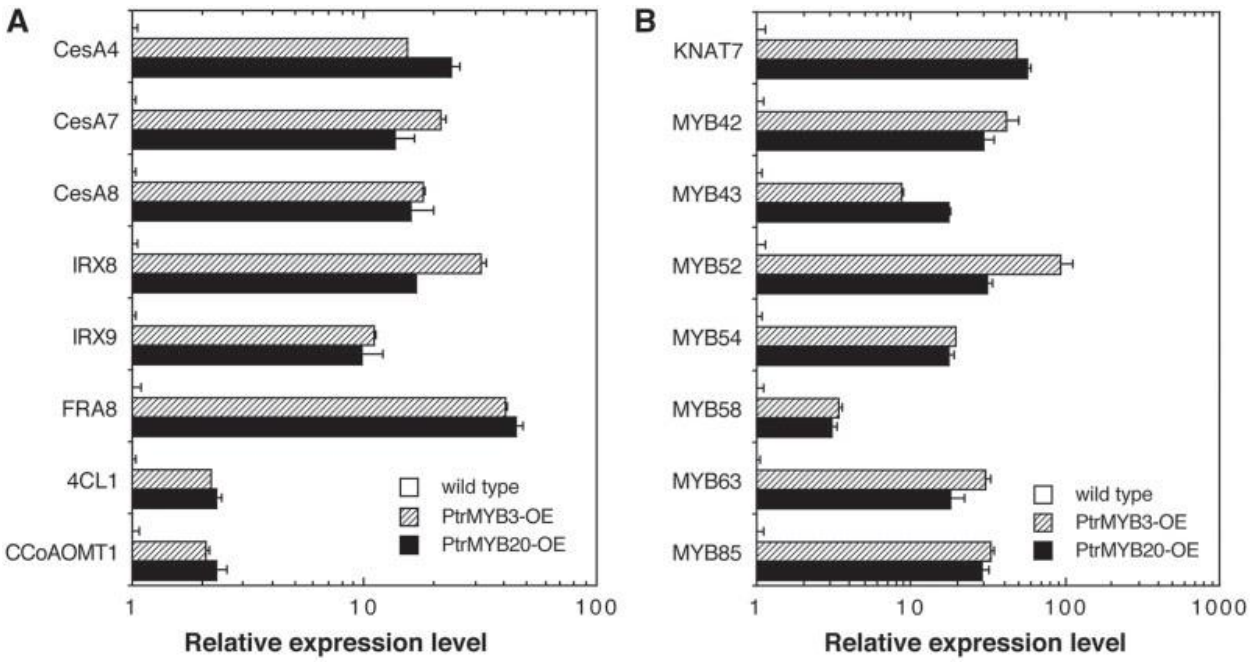

Figure 7. Induction of the expression of secondary wall biosynthetic genes (A) and secondary wall-associated transcription factors (B) by overexpression of PtrMYB3 and PtrMYB20. The expression level of each gene in the wild type was set to 1. Error bars represent the SE of three biological replicates.

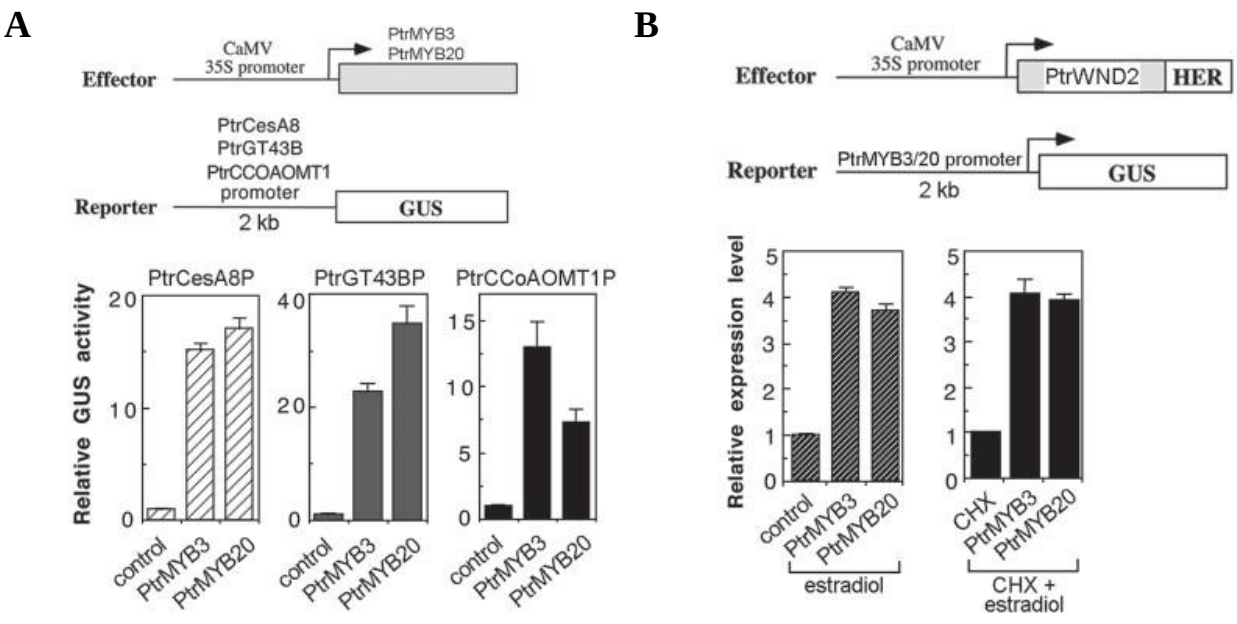

Figure 8. Activation of the promoters of poplar secondary wall biosynthetic genes by PtrMYB3 and PtrMYB20 (A). The upper panel depicts diagrams of the effector and reporter constructs. The lower panel shows the PtrMYB3- or PtrMYB20-activated expression of the GUS reporter gene driven by the promoters of poplar *PtrCesA8*, *PtrGT43B* and *PtrCCoAOMT1*. The GUS activity in protoplasts transfected with the reporter construct alone was used as a control and was set to 1. (B) Direct activation of the promoters of *PtrMYB3* and *PtrMYB20* by the poplar NAC domain transcription factor PtrWND2. The upper panel shows diagrams of the effector and reporter constructs. The lower panel shows that estradiol activation of PtrWND2 induces the expression of the *PtrMYB3/20* promoter-driven GUS reporter gene (left) and this activation remains in the presence of cycloheximide (CHX) (right). The expression level of the GUS reporter gene in the control protoplasts without addition of estradiol was set to 1. Error bars represent the SE of three biological replicates.

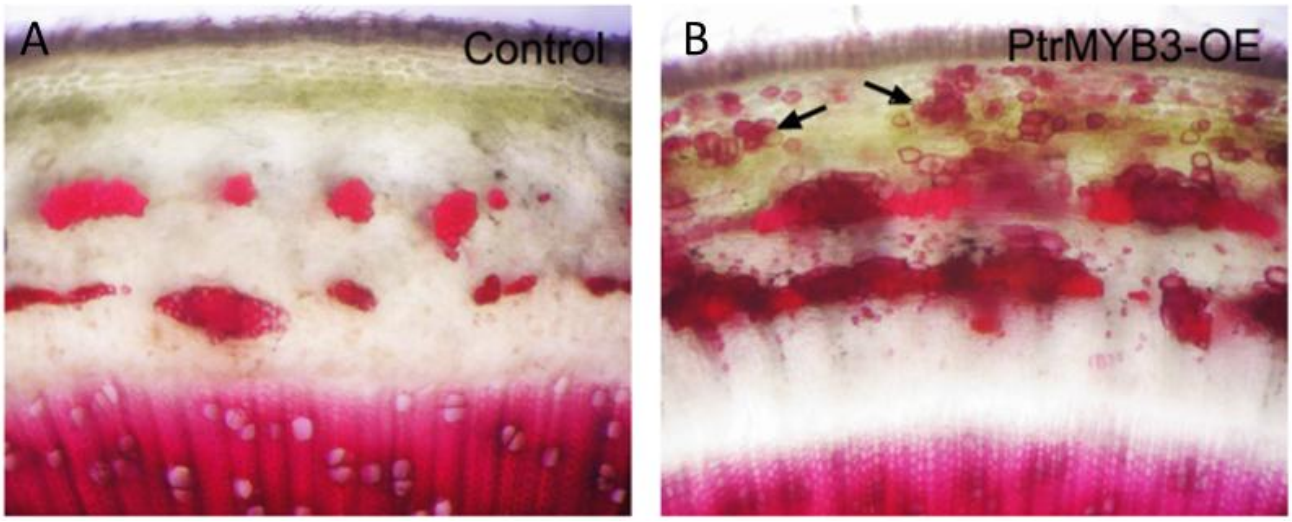

Figure 9. Overexpression of PtrMYB3 induces ectopic secondary cell wall formation in stem cortex cells assayed by lignin staining. Poplar stem sections stained for lignin with phloroglucinol-HCl for wild-type (A) and PtrMYB3 overexpressing (B) poplar. Examples of cells with ectopic lignin deposition are indicated by arrows. Cortex cells demonstrating ectopic lignin deposition also exhibit secondary wall thickening. Similar results were observed for PtrMYB20-OE transgenic lines (data not shown).

CHAPTER 4

FROM SELAGINELLA TO POPLAR: THE SECONDARY WALL NAC DOMAIN PROTEIN BINDING SITE AS A CONSERVED CIS-ELEMENT IMPORTANT FOR SECONDARY CELL WALL FORMATION[1]

[1]Ryan L. McCarthy and Zheng-Hua Ye. To be submitted to Plant Signaling Behavior

Abstract

Changes in regulatory elements guiding gene expression represent a key mechanism of evolutionary change. The evolution of vascular tissue in land plants marks a major advance allowing colonization of terrestrial environments. Formation of a secondary cell wall in vascular tissue is essential to its function in transport and support. Recent analysis has determined the target binding sequence in *Arabidopsis thaliana* for the secondary wall NAC (SWN) transcription factors that act as master regulators of secondary cell wall formation. To explore the conservation and evolution of the secondary wall NAC binding element (SNBE) we examined its distribution in the genomes of multiple plant lineages. Through whole genome sequence analysis SNBE sites were found to be enriched in the promoters of genes implicated in secondary cell wall formation for all vascular plants analyzed. SNBE sites were also identified in the promoters of many genes known to not be expressed in secondary cell wall forming tissue, indicating that additional factors beyond SNBE site sequence are important for SWN binding. Comparison of orthologous SNBE sites between *Arabidopsis thaliana* and *Arabidopsis lyrata* revealed that selective pressure is acting to preserve the binding site sequence in agreement with the previously identified SNBE consensus sequence. Our findings provide evidence that the binding of SNBE sites by SWNs is a conserved mechanism evolved early in the lineage of vascular plants to regulate the formation of the secondary cell wall.

Introduction

The evolution of plants onto land required numerous adaptations; each addressing a problem posed by their new terrestrial environment and ultimately responsible for producing the diverse forms of extant vascular plants. The formation of vascular tissue, next to the development of seeds and a waxy cuticle, may be the most significant advancement in allowing plants to colonize nearly every land ecosystem. Vascular tissue enabled plants to transport water and nutrients throughout their bodies, permitting increased tissue specialization and reducing the reliance on full body water contact required by most nonvascular land plants such as the bryophytes which must acquire their water through simple diffusion. While some bryophytes possess water conducting tissue structurally similar to vascular tissue, the water conducting xylem vessels of vascular plants are distinguished by the formation of a lignified secondary cell wall (Ligrone et al., 2000). Secondary cell walls are also formed in other cell types including anther endothecium as well as xylem fibers which form alongside xylem vessels to compose vascular bundles in the stem and provide structural support. The secondary cell wall is formed between the plasma membrane and the primary cell wall where it increases the strength and water impermeability of the cell. The primary components of the secondary cell wall include cellulose, hemicellulose and lignin with lower quantities of other polysaccharides as well as structural and enzymatic proteins. The secondary cell wall, and the underlying genes responsible for its formation, has gradually arisen and diverged as separate plant lineages recruited new genes into the process resulting in a diversity of secondary cell wall structure across the plant kingdom (Sarkar et al., 2009).

Cellulose, composed of an unbranched linear chain of D-glucose residues, is the primary component of primary and secondary cell walls in all plants and most cell wall forming algae.

However, a separate set of cellulose synthase (CesA) genes is responsible for producing secondary cell wall cellulose in vascular plants. Based on phylogenetic relationships it is believed that the divergence of primary and secondary CesA genes occurred after the divergence of vascular plants and bryophytes (Roberts and Bushoven, 2006).

Hemicellulose describes a broad category of often branched cell wall associated polysaccharides with some forms appearing early in the plant lineage but diversifying to an extent that each plant lineage has a distinct hemicellulose composition. Xylan is the primary hemicellulose of dicot secondary cell walls while grass monocots primarily contain glucuronoarabinoxylan. Mannans comprise the major hemicellulose of bryophyte primary cell walls and are a significant component in the secondary cell walls of lycopods, but only exist at low concentrations in higher vascular plants (Scheller and Ulvskov, 2010).

Lignin is a phenolic heteropolymer composed of monolignols that impart increased strength and waterproofing to the secondary cell wall. Monolignols are classified as H, G or S depending on their chemical structure. While lignin-like molecules are found even in some unicellular algae, the polymerization of lignin in the secondary cell wall distinguishes vascular plants from the rest of the plant kingdom (Popper et al., 2011). Intriguingly, lignified secondary cell walls were found in the red algae *Calliarthron cheilosporioides* and are believed to have arisen through convergent evolution (Martone et al., 2009). H and G lignin occur ubiquitously in secondary cell walls across all vascular plant lineages, although the ratio of each component varies among species. However, the occurrence of S lignin is less uniform as the biosynthetic machinery for producing S monolignols appears to have evolved separately from cytochrome P450 dependent monooxygenases of independent origin in angiosperms, the lycopod *Selaginella moellendorffii* and the red algae *Calliarthron cheilosporioides* (Weng et al., 2008; Martone et al.,

2009). A handful of ferns and gymnosperms have been reported in older literature to possess S monolignols but these claims have yet to be evaluated using modern techniques (Weng and Chapple, 2010).

The production of secondary cell walls is a highly coordinated process requiring precise regulation, biosynthesis, transport and assembly of many molecular components. While the biosynthesis of secondary cell wall components is the most well understood part of secondary cell wall formation, recent discoveries have also greatly improved our knowledge of the regulatory network responsible for the coordination of this process. In *Arabidopsis thaliana* a group of NAC and MYB domain transcription factors have been identified as master regulators capable of switching on all the biosynthetic pathways, including cellulose, xylan (a major component of hemicellulose) and lignin, necessary for secondary cell wall formation (Zhong et al., 2007a, 2007b). Investigation of these genes has revealed the architecture of the regulatory network controlling secondary cell wall formation. The closely related NAC domain proteins NAC Secondary Wall Thickening Promoting Factor1 (AtNST1), AtNST2, NAC Secondary Wall-associated NAC Domain Protein1 (AtSND1; also called AtNST3), Vascular-related NAC-Domain6 (AtVND6), and AtVND7 activate similar direct genetic targets and have been shown to be activator switches able to induce all pathways required for secondary wall formation (Kubo et al., 2005; Mitsuda et al., 2005, 2007; Zhong et al., 2006, 2007b, 2008). These NAC domain proteins function as redundant pairs in distinct cell types with AtSND1 and AtNST1 acting in fibers (Mitsuda et al., 2007; Zhong et al., 2006, 2007b), AtNST1 and AtNST2 acting in anther endothecium cells (Mitsuda et al., 2005) and AtVND6 and AtVND7 acting in root and stem xylem vessels (Kubo et al., 2005; Zhong et al., 2008).

Recent analysis has revealed the sequence of the secondary cell wall NAC binding element (SNBE) to be an imperfect palindromic 19-bp sequence present in the promoters of many genes known to be involved in secondary cell wall formation. While all SWNs demonstrated an ability to bind to the SNBE sites, AtSND1 and AtVND7 were shown to induce different, but overlapping, sets of target genes. Additionally, although some genes not induced by SWNs were found to contain SNBE sites in their promoters, the number of SNBE sites correlated with activation by SWNs (Zhong et al., 2010b).

Although originally identified in *Arabidopsis thaliana*, studies in poplar have identified functional orthologs of AtSND1 capable of acting as master regulators of secondary cell wall formation in poplar. These poplar SWNs were able to rescue the phenotype of the secondary cell wall deficient *Arabidopsis thaliana snd1 nst1* double mutant indicating that the poplar and *Arabidopsis thaliana* SWN orthologs are capable of binding to and activating the same set of target genes (Zhong et al., 2010a). These results present the possibility that the binding of SNBE sites by SWNs is well conserved between *Arabidopsis thaliana* and poplar.

Mutations in cis-regulatory sequences, including transcription factor binding sites, have been linked to significant phenotypic differences in a number of studies (Ludwig, 2002). Mutations are typically permitted to accumulate at a faster rate in non-coding regions than coding regions of the genome due to the fact that mutations in non-coding regions are less likely to be deleterious. While mutations disrupting the function of cis-regulatory elements can be harmful, mutations in the flanking sequence of such elements are considered to be neutral mutations. This high rate at which mutations accumulate in noncoding regions causes cis-regulatory sequences to arise by chance in gene promoters where they have the potential to alter a gene's expression pattern (Doniger and Fay, 2007). Mutational analyses of cis-regulatory

sequences and comparative genomics have demonstrated a large percentage of the observed interspecies and intraspecies phenotypic variations are attributable to changes in cis-regulatory sequences (Prud'homme et al., 2007). While the role of SWNs binding to SNBE sites to regulate secondary cell wall formation has begun to be revealed in *Arabidopsis thaliana*, it remains to be seen if this same regulatory mechanism is responsible for secondary cell wall formation in other vascular plants. Additionally, if this regulatory mechanism is conserved it will be important to determine to what extent differences in secondary cell wall structure correspond to differences in SNBE site position and SWN binding.

The SNBE consensus sequence and the availability of numerous sequenced plant genomes enable us to examine the conservation of SNBE sites across many lineages of the plant kingdom. In this study we investigate the prevalence and position of SNBE sites in *Arabidopsis thaliana* and a number of other plant species. We report the presence of recognizable SNBE sites in the promoter regions of genes from a representative sample of vascular plants. We analyze the positional evolution rates of nucleotides in the SNBE sites located in promoters of genes induced four fold or greater by AtSND1 and/or AtVND7 by comparing sequences between the closely related *Arabidopsis thaliana* and *lyrata*. Our analysis demonstrates that the system of interaction between SWNs and SNBE sites appeared early in vascular plant development and is pervasive throughout all vascular plant lineages investigated here. Furthermore, our evidence suggests that selection is acting to functionally conserve the SNBE motif as a binding site for SWNs in promoters of genes crucial to secondary cell wall formation.

Results

Previous studies have identified the specific target sequence bound by SWNs in *Arabidopsis thaliana* allowing us to investigate the properties of the regulatory network in *Arabidopsis thaliana* and explore the role of SWNs as master regulators of secondary cell wall formation in a wider range of vascular plants. We will further substantiate the role of SWNs and SNBEs by examining their conserved presence as regulatory elements of secondary cell wall genes in distantly related plant species as well as their sequence conservation among species of *Arabidopsis*.

Analysis of SNBE location in the *Arabidopsis thaliana* genome

The identification of the SNBE sequence now allows us to identify putative SNBE sites in genomic sequence and determine how SNBE sites are distributed in the *Arabidopsis thaliana* genome (Fig. 1). The distribution of putative SNBE sites appears to weakly correlate with gene rich regions in chromosomes one and five where a dip in putative SNBE site density matches the position of the centromere. However, in the gene rich regions of these chromosomes and in the other three chromosomes the density of putative SNBE sites is approximately 20 per 10kb; approximately the same density expected to arise by chance. SNBE sites not being highly enriched in the genome is consistent with their role as binding sites for SWNs to regulate genes for secondary cell wall formation which comprise a small percentage of the genome. Interestingly, there are peaks of high putative SNBE site density, reaching as high as 90 sites in 10kb, indicating that SNBE sites may be enriched around specific genes or clusters of genes involved in secondary cell wall formation. These genomic regions of high putative SNBE site density contain many carbohydrate active enzymes including members of the glycosyl transferases and glycosyl hydrolases gene families often involved in biosynthesis and

modification of secondary cell wall polysaccharides. Another region of high putative SNBE site density is located in middle of a cluster of nine genes annotated as cytochrome P450s. Although the function of these specific cytochrome P450 genes is unknown, a number of cytochrome P450 genes catalyze crucial steps in the pathway of lignin biosynthesis (Weng, 2010).

To further investigate the possibility of selective SNBE site enrichment we compared genes known to be induced by SWNs to all annotated genes from *Arabidopsis thaliana*. Although the direct target genes of SND1 and VND7 have previously been identified, it remains to be determined how these SWN regulated genes differ from other genes in SNBE site number and position. To see if genes demonstrated to be induced fourfold or higher by SWNs were markedly distinct either in number of SNBE sites or in the positions of SNBE sites we analyzed the sequences from all *Arabidopsis thaliana* gene models for the presence of sequences matching the SNBE consensus sequence (Fig. 2). SNBE sites were found to vary by gene region and whether the genes were SWN regulated. All *Arabidopsis thaliana* genes and the subset of genes induced fourfold or greater by SWNs possessed their highest average density of SNBE sites in the 1.5kb promoter regions. For all *Arabidopsis thaliana* genes the 1.5kb promoter region had significantly more SNBE sites per kb ($p<2.2e-16$) than all other regions examined except the 3' UTRs which were not significantly different. The 1.5kb promoter region of SWN regulated genes had a significantly greater concentration of SNBE sites then 5'UTRs, exons, introns and 3'UTRs ($p=0.001$, $p<2e-12$, $p=0.003$ and $p<5e-9$ respectively). The 1.5kb promoter regions of SWN regulated genes possessed a significantly higher concentration of SNBE sites ($p<1e-8$) than the 1.5kb promoters from all genes. Conversely the 3'UTRs of SWN regulated genes contain a significantly lower average of SNBE sites ($p<0.0005$) than the 3'UTRs from all genes. While the number or position of SBNE sites cannot be reliably used to predict a gene's ability to be

activated by SWNs, this data indicates that the SWN regulated genes do seem to preferentially accumulate putative SNBE sites in their promoters.

Comparison of SNBE site density in plant genomes

To extend our findings on the role of SWNs and SNBE sites to additional plant species we analyzed plant genomes for an enrichment of putative SNBE sites similar to what we observed in *Arabidopsis thaliana* promoters (Fig. 2). If the density of putative SNBE sites was higher than expected to accumulate by pure chance then we might find support of a functional role for these sites in other plant species. Nucleotide frequency was empirically determined from each plant genome and was used to determine the expected density of putative SNBE sites. While the expected density of putative SNBE sites varied widely with nucleotide frequencies the relation of expected and observed results was consistent among all vascular plants analyzed. For *Arabidopsis thaliana*, *Populus trichocarpa* and the basal vascular plant *Selaginella moellendorffii* the genomic density of putative SNBE sites was lower than expected while the density in the 1.5kb promoter region of SWN induced genes was greater than expected (Table 1). This difference could be the product of selective pressure towards retention of SNBE sites in the promoters of genes involved in secondary cell wall formation and negative selection acting to rid other genes of SWN control. The greatest promoter enrichment of SNBE sites was observed in *Arabidopsis thaliana* and *Oryza sative* where the observed density was 50% and 88%, repectively, greater than the expected density. This enrichment of putative SNBE sites in promoters across divergent plant species could indicate a remarkable level of conservation of the transcriptional network regulating secondary cell wall formation. The less dramatic results in *Selaginella moellendorffii* may be partially attributable to a divergence in the permitted SNBE site sequence causing some functional SNBE sites to avoid detection. Alternatively, it could

indicate the formation of new regulatory relationships between SWNs and target genes in the angiosperm lineage that had not been established before their divergence from the progenitors of *Selaginella*. The non-vascular *Physcomitrella patens* exhibited a genomic density of putative SNBE sites close to the density expected in the presence of no selective pressure. This finding is consistent with there being no functional role for SNBE sites in non-vascular plants. Curiously the density of SNBE sites in promoters of genes orthologous to *Arabidopsis thaliana* SWN induced genes was far lower than expected (Table 1).

Table 1. Average number of putative SNBE sites observed and expected per kb in plant genomes. Number of expected SNBE sites calculated based upon species specific nucleotide frequency determined from whole genome sequence.

Organism	Putative SNBE sites expected per kb	Putative SNBE sites per kb in whole genome	Putative SNBE sites per kb in 1.5kb promoter of SWN regulated genes
Arabidopsis thaliana	2.06	1.96	3.08
Populus trichocarpa	2.23	1.77	3.21
Oryza sativa ssp. Japonica	1.52	1.29	1.91
Selaginella moellendorffii	1.40	1.16	1.63
Physcomitrella patens	2.23	2.19	1.30

Comparison of SNBE sites in the promoters of secondary wall genes

The findings that putative SNBE sites may be conserved among vascular plants drove us to investigate if SNBE sites were conserved between gene orthologs. Genes demonstrated to be highly induced by SND1 or VND7 in addition to well characterized targets of SWNs from *Arabidopsis thaliana* were selected and their closest orthologs were identified in *Populus trichocarpa, Vinis vinifera, Oryza sativa ssp. japonica, Zea mays, Sorghum bicolor* and

Selaginella moellendorffii (Table 2). Analysis of the 1.5kb promoter region of this set of orthologs revealed the presence of putative SNBE sites in all but 11 out of 248 promoters. However, some genes have SNBE sites in their introns such as Os03g16980 which possesses five putative SNBE sites in its second intron and seven across all introns. Additionally, five of the promoters without an SNBE site had a close paralog with at least one putative SNBE site. Since gene duplications can reduce selective pressure on a duplicate gene it is not surprising that in some cases a redundant gene may evolve to no longer be regulated by SWNs. Surprisingly, even in the evolutionarily distant basal vascular plant *Selaginella moellendorffii* the presence of putative SNBE sites appears to be conserved.

Although the presence of putative SNBE sites appears to be well conserved in promoters we wanted to determine the extent to which the position of putative SNBE sites was conserved in promoters across the plant kingdom. We selected several direct targets of SWNs including AtMYB46, AtMYB83, AtSND3, AtKNAT7 and AtMYB103. These genes were selected since they are well characterized as direct targets activated by SWNs and are key regulators of secondary cell wall formation. Orthologs of these genes were identified in a number of plant species; their promoter sequences were analyzed for putative SNBE sites and were related to each other through the phylogenetic relationship interpreted from their gene's protein sequences. Promoter sequence of genes from closely related species demonstrated conservation of putative SNBE position. Comparisons of promoter sequences from *Arabidopsis thaliana* genes AtMYB46, AtMYB83, SND3 and MYB103 with their *Arabidopsis lyrata* orthologs demonstrates a high level of putative SNBE site conservation (Fig. 4, Fig. 5, Fig. 6, Fig. 7). Additionally, several paralogs including the pairs PtrMYB002 and PtrMYB021 as well as PtrMYB003 and PtrMYB020 exhibit highly conserved putative SNBE site position (Fig. 4).

This similarity in SNBE site position in PtrMYB020 and PtrMYB003 promoters is consistent with the finding that the AtSND1 close homolog PtrWND2 induces expression of both genes at similar levels (McCarthy et al., 2010). In the monocot lineages we see conservation of putative SNBE sites between the *Oryza sativa*, *Sorghum bicolor*, *Setaria italica* and *Zea mays* orthologs of AtSND3. These four promoters share a set of two putative SNBE sites on the DNA top strands which are well conserved in sequence but seem to shift in position due to indel mutations (Fig. 4). Between more distantly related species there are no clearly identifiable conserved SNBE sites. These observations are consistent with high rates of transcription factor binding site turnover observed between closely related species of *Saccharomyces* and *Drosophila* (Moses et al., 2006; Doniger and Fay, 2007).

As the SNBE consensus sequence was determined with electrophoretic mobility shift assay (EMSA) using *Arabidopsis* protein (Zhong et al., 2010b) there is the possibility that binding affinities of SWNs may differ from other species or *in vivo Arabidopsis*. The sequence of SNBE sites in promoters of *Arabidopsis* genes known to be direct targets activated by SWNs were collected and used to generate a sequence logo (Fig. 8A). At position 15 there is a marked preference for adenine over cytosine in a ratio of nearly five to one. This result is consistent with previous EMSA competition analysis which show that SNBE sequences with cytosine at position 15 are only weakly competitive for AtSND1 binding compared to SNBE sequences with adenine at position 15 (Zhong et al., 2010b). To determine if nucleotide preference in SNBE sites varied among plant lineages we generated additional sequence logos using putative SNBE sites from monocots (Fig. 8B) or dicots (Fig. 8C). The nucleotide preference is similar between monocots and dicots, except at position 15 where putative SNBE sites from monocots exhibit a much higher ratio of cytosine to adenine occurrence than observed in dicots, 0.4 and .15 respectively,

potentially indicating that a slight divergence in binding site preference may exist between plant lineages allowing monocots to more easily tolerate cysteine at position 15.

Determining if selective pressure is maintaining SNBE sites

The apparent high turnover of SNBE sites in putative SWN regulated genes brings into question the essential nature of regulation by SWNs for proper secondary wall formation. Determining if selective pressure is acting to conserve SNBE site sequence could provide evidence whether these genes are being maintained under SWN regulation. The availability of high quality genome sequence for *Arabidopsis thaliana* and *Arabidopsis lyrata* allowed us to use the experimentally determined SBNE consensus sequence to computationally identify putative SNBE sites in gene promoters. The recent divergence and corresponding high sequence homology between the closely related *Arabidopsis thaliana* and *Arabidopsis lyrata* provides an ideal situation for quantifying the evolution rate of each nucleotide position in the SNBE site and ascertaining whether selective pressure is maintaining SNBE sites. The promoters of *Arabidopsis thaliana* genes induced more than fourfold by SND1 and/or VND7 and the promoters of their *Arabidopsis lyrata* homologs were analyzed for the presence of putative SNBEs matching the SNBE consensus sequence. By aligning the promoter sequences, the high sequence homology allowed identification of 156 putative SNBE site pairs that diverged from a single putative SNBE site in the common ancestor. Only pairs where one or both of the putative SNBE sites fit the consensus sequence were included in the analysis as pairs where neither site matched the consensus sequence were indistinguishable from the background. To establish if selective pressure was maintaining the SNBE site sequences we simulated the divergence of 156 pairs of SNBE sites from common ancestral SNBE sites under no selective pressure. Simulation

parameters including SNBE site nucleotide frequency and base pair substitution rate were empirically determined from the *Arabidopsis thaliana* and *Arabidopsis lyrata* sequences. The frequency of substitutions that abolished agreement with the SNBE consensus sequence, termed degenerative substitutions, were compared between the observed SNBE site pairs and the distribution of degenerative substitution rates of 10,000 simulation replicates (Fig. 9). Individual positions 1, 14 and 19 of observed SNBE sites exhibited degenerative substitution rates significantly lower ($p<0.05$) than the simulated SNBE sites under neutral selection. Interestingly the degenerative substitution rate for position 16 in observed SNBE site pairs is higher than the rate for position 16 in 88% of the simulated SNBE site datasets. Total observed SNBE site sequences from *Arabidopsis thaliana* and *Arabidopsis lyrata* exhibited significantly lower rates of degenerative substitutions than the simulated SNBE sites under neutral selection ($p<6e-11$). Comparisons of non-degenerative and total substitution rates between observed and simulated SNBE site pairs revealed a trend towards higher substitution rates in the observed SNBE site pairs when compared to the simulated SNBE site pairs (unpublished data). This indicates that the base pair substitution rate used for the simulation may be an underestimate and may signify that the SNBE site sequence is actually more highly conserved then our data suggests.

Discussion

Previous research has successfully identified a number of SWNs that operate as master regulatory switches of secondary cell wall formation in several species of vascular plants. Recent study of the DNA binding affinity of SWNs identified the target sequence required for SWN binding in *Arabidopsis thaliana*. This analysis of putative SNBE sites across many lineages of the plant kingdom provides supporting evidence that the mechanism of SWNs

binding SNBE sites to activate expression of secondary cell wall genes is conserved across the lineage of vascular plants.

Distribution of SNBE sites in *Arabidopsis thaliana* suggests additional complexities in regulation by SWNs

SNBE sites were found to be enriched in promoters of genes induced by SWNs, but the number of SNBE sites found in promoters of other genes and distributed throughout the genome is consistent with the number expected to arise by chance (Fig. 1). This finding indicates that SNBE sites established in promoters of genes not induced by SWNs are tolerated and not selected against. However, if all detected SNBE sites were functional the expression of a large number of misexpressed genes in secondary cell wall forming cells would likely be disadvantageous. Therefore it is reasonable to assume that many of the genes with putative SNBE sites in their promoters are not activated by SWNs. This is consistent with previous findings that the SWNs AtSND1 and AtVND7 induce only a small subset of genes possessing SNBE sites (Zhong et al., 2010b). The number of SNBE sites in a gene's promoter loosely correlates with activation by SWNs. In contrast some genes with a single SNBE site were activated while other genes with many SNBE sites were not. Whole genome sequence analysis of *Arabidopsis thaliana* identified gene promoters possessing high numbers of SNBE sites (Sup. Table 2). Many of these genes had not been detected as induced by SWNs in a SWN direct activation assay detected by genome-wide microarray (Zhong et al., 2010b). This finding indicates that additional factors beyond SNBE site sequence control which SNBE sites are bound by SWNs to activate gene expression. The binding of SWNs could be prevented by epigenetic factors such as establishment of heterochromatin making the SNBE site inaccessible. Alternatively, activation of some genes may require an unknown binding cofactor. There is also

the possibility that some target genes may be expressed at low levels easily missed by microarray analysis. To determine whether genes located in regions with high SNBE site density are regulated by SWNs and expressed in secondary cell wall forming tissue or sheltered from SWN activation by a closed chromatin structure or currently unknown mechanism will require further experimental analysis.

The existence of SNBE sites in plant lineages has implications for the evolution of vascular tissue and secondary cell wall formation

We have demonstrated that promoters of genes induced by SWNs in *Arabidopsis thaliana* contain significantly more SNBE sites than promoters of genes not regulated by SWNs, consistent with previous findings that SNBE number correlates with activation by SWNs (Zhong et al., 2010b). Close orthologs of the *Arabidopsis thaliana* SWN induced genes were also shown to possess a significantly higher density of putative SNBE sites in their promoter regions when compared to the promoter regions of all genes for the vascular plants analyzed (Fig. 3). Based on this result it seems likely that the regulatory mechanism of SWNs binding SNBE sites to induce secondary cell wall formation is at least partially conserved across all major lineages of vascular plants included in this study. Additionally, variation within the SNBE site consensus sequence was minimal between separate plant lineages with the greatest variation observed between monocots and dicots for nucleotide preference at position 15 (Fig. 8). Analysis of individual putative orthologs from herbaceous and woody dicots, monocots and lycophytes of genes induced by SWNs in *Arabidopsis thaliana* revealed that nearly all possessed SNBE sites in their promoters indicating a conserved mechanism of regulation (Table 2). Relative to the other genomes analyzed, *Selaginella moellendorffii* possessed the most modest increase in putative SNBE site density. As *Selaginella moellendorffii* is a member of the lycophytes, which diverged

from the angiosperms approximately 400 million years ago, some differences in the process of secondary cell wall formation is expected. One explanation is that the SNBE site sequence in *Selaginella moellendorffii* is sufficiently divergent to allow functional SNBE sites that do not match our SNBE consensus sequence and thus are not identified in the analysis. It could also be argued that the vascular structure of lycophytes does not require the same number of SWN activated secondary cell wall genes. While these explanations likely contribute, a more interesting possibility is that many of the genes induced by SWNs in *Selaginella moellendorffii* may not be readily identifiable as orthologs of SWN regulated genes from *Arabidopsis thaliana* because they evolved from independent gene origins. There is precedent for convergent evolution in the formation of secondary cell walls with both angiosperms and *Selaginella moellendorffii* separately evolving biosynthetic machinery to convert G monolignol intermediates to S monolignols. Although mostly functionally interchangeable, the genes responsible for the first committed step of this conversion in *Arabidopsis thaliana* and *Selaginella moellendorffii*, AtF5H and SmF5H respectively, are believed to have evolved from separate genes (Weng et al., 2008). Although not known to be induced by SWNs, both AtF5H and SmF5H possess several putative SNBE sites in their promoters (unpublished data). It is possible that the regulatory mechanism of SWNs binding SNBE sites to regulate secondary cell wall formation may be conserved between angiosperms and lycophytes, but many of the downstream targets which evolved following the divergence of angiosperms and lycophytes may differ. Further analysis of the genes involved in secondary cell wall formation in *Selaginella moellendorffii* will be required to fully determine the differences that have arisen over 400 million years, but the occurrence of SNBE sites in the promoters of secondary cell wall genes appears to have remained.

In stark comparison to the vascular plants, in the non-vascular land plant *Physcomitrella patens* no difference was observed in SNBE site density between promoters of the closest homologs to *Arabidopsis thaliana* SWN induced genes and promoters of all *Physcomitrella patens* genes (Fig. 3). As *Physcomitrella patens* does not produce secondary cell walls this is an unsurprising result, but it has implications for the evolution of the regulation of secondary cell wall formation. The absence of SNBE sites from promoters of genes in non-vascular plants suggests that following the divergence of vascular plants from bryophytes there was a proliferation of SNBE sites into the promoters of genes that would assume roles in secondary cell wall formation before lycophytes diverged from the vascular plant lineages.

New genes and gene networks evolve by repurposing existing genetic resources. Members of the NAC domain family transcription factors are found in *Physcomitrella patens* indicating a possible origin of the SWNs. The existence of a handful of SNBE site containing genes in *Physcomitrella patens* that are close homologs to genes induced by SWNs in *Arabidopsis thaliana* enables us to propose a model describing the origin of the SWN regulatory network found in vascular plants. Of the *Physcomitrella patens* genes with abnormally high numbers of SNBE sites that are similar to *Arabidopsis thaliana* genes induced by SWNs, many are involved in either microtubule formation and trafficking, or polysaccharide modification. All these genes could be involved in vesicle trafficking of polysaccharides along microtubules by motor proteins, which is a key process in both primary and secondary cell wall formation. Conceivably the non-vascular plant precursors to SWNs were involved in the regulation of the biosynthesis and transport of polysaccharides to the primary cell wall. This small gene network could be repurposed, possibly following a gene duplication event, by evolving tissue specific expression for the SWN precursors and gradually recruiting more genes. Tissue specific

expression in the non-secondary cell wall forming water conducting cells found in some bryophytes could be an important step in the evolution of secondary cell walls and true vascular cells. Given the thick primary cell walls formed in a cell type specific manner in the water conducting tissue of bryophytes (Ligrone et al., 2000) it is clear that primary cell wall biosynthesis genes are being differentially regulated in this tissue. It would be of great interest to determine if a NAC domain transcription factor is responsible for this increased thickening.

The establishment of an SNBE site in the promoter of a gene may alter the expression pattern to include vascular tissue, thus changing the selective pressure being exerted upon the gene. If this gene proves beneficial to secondary cell wall formation, it will undergo selective pressure to refine its role. If the expression of the gene in vascular tissue is deleterious it could be quickly selected against and removed from the population. The high mutation rate of non-coding relative to coding sequences may therefore act as a mechanism for gene expression pattern to rapidly evolve to fit tissue specific roles.

Tissue specific transcription factors, such as SWNs, may have the potential to act as recruiters of genes that can then be shaped by natural selection to form a complex developmental pathway. It is possible that the establishment of the interaction between SWNs and their target SNBE sites in the earliest of vascular plants was the first step to incorporate additional genes into both the regulatory and biosynthetic networks involved in secondary cell wall formation. Additional characterization of the regulatory networks and biosynthetic genes involved in secondary cell wall formation in the basal vascular plants as well as a better understanding of the roles of NAC domain transcription factors in non-vascular plants will be required to determine if our proposed model for the evolution of secondary cell wall formation is accurate. If a similar regulatory network controlled by NAC master regulatory transcription factors also controls the

independently evolved version of secondary cell wall formation in the red algae *Calliarthron cheilosporioides* (Martone et al., 2009) it might be concluded that any genes of common origin involved in both lineages were uniquely suited to be adapted to the process of secondary cell wall formation.

SNBE sites are evolutionarily conserved between closely related species but exhibit a high rate of turnover across greater evolutionary distance

Another key finding of our study is that nucleotides essential for SWN binding in SNBE sites accumulate mutations incongruous with the SNBE consensus sequence at a significantly lower rate than what would be expected under neutral evolution, indicating that selective pressure is acting to conserve SNBE function (Fig. 9). To our knowledge our study is the first to demonstrate evolutionary conservation of transcription factor binding sites in plants. The rates of evolution for nearly all nucleotide positions are consistent with the expected functional constraint based on the SNBE consensus sequence. The unexpected high rate of evolution observed at position 16 of the SNBE site sequence is intriguing and could be caused by several factors. As the SNBE consensus sequence was determined by evaluating which single nucleotide mutations eliminated SWN binding it is possible that mutations of multiple nucleotides, while not allowed individually, can have compensatory effects allowing SWNs to bind with sequences outside the identified SNBE consensus. Alternatively, as the SNBE consensus sequence was determined using AtSND1 as a binding partner, other SWNs may tolerate other nucleotides at this position. The high rate of binding site turnover is consistent with previous findings in yeast and vertebrates that even functional transcription factor binding sites can exhibit high rates of turnover which may either be compensatory or contribute to interspecies variation (Doniger and Fay, 2007; Schmidt et al., 2010). The relatively high chance

of random mutations producing new transcription factor binding sites reducing selective pressure maintaining other sites may allow previously critical sites to deteriorate. This high rate of turnover between all but the most closely related promoter sequences was apparent in our comparison of promoters from AtMYB46 and its homologs where only gene promoters from closely related genes shared common putative SNBE sites (Fig. 4). Experiments in yeast have demonstrated that transcription factor binding sites were more highly conserved between species in promoters of target transcription factors than in promoters of other target genes (Borneman et al., 2007). Presumably changes to expression patterns of target transcription factors produces a domino effect changing the expression of numerous downstream genes whereas mutations in promoters of non-regulatory genes have a much lower capacity to effect gene expression. AtMYB46 and AtMYB83 are direct targets of SWNs and are master regulators of secondary cell wall formation influencing expression of many downstream target genes (McCarthy et al., 2009). Consistent with findings from yeast, orthologs of AtMYB46 and AtMYB83 exhibited a high degree of SNBE site conservation between species likely due to their key importance in secondary cell wall formation (Fig. 4).

Interspecies variations of SNBE site positions may be partially responsible for the observed diversity of secondary cell walls

Secondary cell wall architecture and content is extremely diverse owing to the involvement of an overlapping but different gene set in separate species. Which genes are turned on in secondary cell wall forming tissue depends largely upon cis-regulatory elements such as transcription factor binding sites. As the furthest upstream regulators of secondary cell wall formation, SWNs exert tremendous control over the entire secondary cell wall pathway through their regulation of not only biosynthetic and transport related proteins but also other transcription factors. Differences

in transcription factor binding site and positions can effect levels of gene expression and influence phenotypic variation (Wray, 2007). It is likely that much of the phenotypic variation observed in the composition of plant secondary cell walls is attributable to variations in gene regulation. However, in some cases regulatory sequence can change while conserving gene expression (Weirauch and Hughes, 2009) making it impossible to accurately predict the changes to expression patterns caused by SNBE site variation without additional experimental evidence. We expect that the interspecies variation in SNBE site number and position partially contributes to the variation in secondary cell wall composition among vascular plant species.

Methods

Identification of putative SNBE binding sites

Whole genome sequence as well as sequence for the 1.5kb region upstream of the translational start sites, 5' UTRs, exons, introns and 3' UTRs for all annotated *Arabidopsis thaliana* gene models were retrieved from The Arabidopsis Information Resource (www.arabidopsis.org). Current genome assemblies and gene model coordinates for *Populus trichocarpa*, *Vinis vinifera, Oryza sativa ssp. japonica*, *Zea mays*, *Sorghum bicolor*, *Selaginella moellendorffii* and *Physcomitrella patens* were downloaded from Phytozome (www.phytozome.org). Gene model coordinates were relied upon to retrieve sequence of gene regions for analysis. Putative SNBE sites were identified based upon sequence agreement with the SBNE consensus sequence from Zhong et al., 2010b. Putative SNBE sites with overlapping sequence were assumed to represent two distinct transcription factor binding opportunities and counted as two sites except in the case of palindromic sequences which were counted as a single site. Genomic nucleotide frequency for calculating the expected frequency of SNBE sites per 1kb was determined empirically from

whole genome sequence. Statistical comparisons of putative SNBE site density in gene regions were performed by unpaired two tailed t-test in R version 2.11.0 (www.r-project.org).

Comparison of binding site position to phylogenetic relationship

Putative orthologs of *Arabidopsis thaliana* genes induced by SWNs were determined by performing reciprocal BLAST using protein sequences retrieved from the Phytozome database (www.phytozome.org). Promoter sequences included in the analysis represent the 1.5kb sequence upstream of the annotated start codon. Phylogenetic relationships between the protein sequences of putative orthologs were predicted using MEGA 4.0.2 (Tamura et al., 2007) to construct maximum parsimony phylogenetic trees. 500 bootstrap replicates were performed for each tree.

Positional analysis of the evolution rate of SNBE sites

Protein sequences from the 83 *Arabidopsis thaliana* genes identified by Zhong et al., 2010b as being induced at least fourfold by SND1 and/or VND7 were used as queries to perform protein BLAST against the *Arabidopsis lyrata* genome. Gene pairs from *Arabidopsis thaliana* and *Arabidopsis lyrata* were assumed to be orthologous if they were each other's highest reciprocal BLAST hits. The 1.5kb promoter sequences of the paired orthologs were aligned using ClustalW2 (www.ebi.ac.uk/Tools/msa/clustalw2/). Locations of putative SNBE sites in both promoter sequences were used along with the alignment to identify pairs of putative SNBE sites with identical or near identical flanking sequence. SNBE sites were paired only if they appeared to have diverged from a common sequence as indicated by high homology of their flanking sequence. A total of 156 putative SNBE pairs with either one or both SNBE sites fitting within the constraints of the consensus sequence were selected as the dataset for the analysis. The

evolution rate for a nucleotide position in the SNBE motif was calculated as the frequency of single nucleotide polymorphisms (SNPs) between putative SNBE pairs at that position.

The exclusion from the dataset, due to detection difficulties, of degenerate SNBE site pairs where both sites no longer match the SNBE consensus sequence creates a likelihood of underestimating mutation rates for positions constrained by the SNBE site sequence and was anticipated to bias the data towards false positive results. To avoid false positives, the bias introduced by the sampling procedure was compensated for by comparing the position specific rates of evolution between the observed putative SNBE site pairs and simulated SNBE site pairs which diverged under neutral mutation and were sampled using the same method. Background evolution rate was determined from the SNBE sites' 50bp flanking sequence and downstream sequence. Nucleotides aligning with gaps due to indel mutations were not included in the calculation of evolution rate. The rate of mutation was calculated separately for each possible base pair substitution due to variation in nucleotide conversion. For estimating the background evolution rate it was assumed that the nucleotide frequencies observed in *Arabidopsis thaliana* and *Arabidopsis lyrata* do not deviate substantially from their common ancestor and that a maximum number of one mutation occurred at each position. Given the relatively recent divergence of *Arabidopsis thaliana* and *Arabidopsis lyrata* as well as a low occurrence of SNPs in the aligned sequence, approximately 8.7%, this was believed to be a reasonable assumption. Given nucleotides $x, y, z, w \in \{A,T,G,C\}$, the probability $P_{x\to(x,y)}$ that nucleotide x at position a in the common ancestor will mutate to generate a SNP between *Arabidopsis thaliana* and *Arabidopsis lyrata* where position a is x in one sequence and y in the other is given by

$$P_{x\to(x,y)} = \frac{N_{xy}\left(\frac{n_x}{n_x+n_y}\right)}{N_{xy}+ N_{xy}\left(\frac{n_x}{n_x+n_y}\right)+N_{xz}\left(\frac{n_x}{n_x+n_z}\right)+N_{xw}\left(\frac{n_x}{n_x+n_w}\right)} \quad (1)$$

where N_{ij}, for $i, j \in \{x,y,z,w\}$, is the number of occurrences of nucleotide i aligned with nucleotide j and n_i is the total number of nucleotide i in the aligned promoter sequences. The numerator of equation 1 estimates the number of SNPs between x and y nucleotides that originated from an x in the ancestral sequence. The denominator of equation 1 estimates the total number of nucleotide x present in the ancestral sequence. The probability $P_{x\to(x,y)}$ is related to the probability of a single nucleotide x mutating to y, $P_{x\to y}$, by the equation

$$P_{x\to(x,y)} = 1 - (P_{x\to y}^2 + (1 - P_{x\to y})^2) \quad (2)$$

$P_{x\to y}$ was solved for numerically to within 1e-9 of its true value. The putative SNBE sites matching the SNBE site consensus sequence from the 156 aligned SNBE site pairs were used to calculate the frequency of each nucleotide at each position in the 19-bp sequence. These frequencies were used to generate a set of artificial SNBE sites to act as ancestral SNBE sites in the simulation. The estimated probabilities, $P_{x\to y}$, for each nucleotide were used to simulate the divergence of the artificial SNBE sites over the same evolutionary time separating *Arabidopsis thaliana* and *Arabidopsis lyrata*. 156 pairs of simulated SNBE sites that had diverged from a common sequence were selected excluding pairs where neither site matched the SNBE site consensus sequence. The evolution rate for a nucleotide position in the SNBE motif was calculated as the frequency of SNPs between simulated SNBE pairs at that position. 1,000 iterations of the simulation were performed. Statistical comparisons between individual positions of the simulated and observed SNBE sites were performed by one sample two tailed t-tests in R version 2.11.0.

References

Borneman, A.R., Gianoulis, T.A., Zhang, Z.D., Yu, H., Rozowsky, J., Seringhaus, M.R., Wang, L.Y., Gerstein, M., Snyder, M. (2007) Divergence of transcription factor binding sites across related yeast species. Science 317: 815-819

Doniger, S.W., Fay, J.C. (2007) Frequent gain and loss of functional transcription factor binding sites. PLoS Comput. Biol. 3(5): e99

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., Mimura, T., Fukuda, H., Demura, T. (2005) Transcription switches for protoxylem and metaxylem vessel formation. Genes and Dev. 19:1855-1860

Ligrone, R., Duckett, J.G., Renzaglia, K.S. (2000) Conducting tissues and phyletic relationships of bryophytes. Phil. Trans. R. Soc. Lond. B Biol. Sci. 355: 795-813

Ludwig, M.Z. (2002) Functional evolution of noncoding DNA. Curr. Opin. Genet. Dev. 12: 634-639

Martone, P.T., Estevez, J.M., Lu, F., Ruel, K., Denny, M.W., Somerville, C., Ralph, J. (2009) Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. Curr. Biol. 19: 169-175

McCarthy, R.L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., Spicer, C., Ye, Z.-H. (2010) The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. Plant Cell Physiol. 51(6): 1084-1090

McCarthy, R.L., Zhong, R., Ye, Z.-H. (2009). MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. Plant Cell Physiol. 50:1-15

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., Ohme-Takagi, M. (2007) NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. Plant Cell. 19:270-280

Mitsuda, N., Seki, M., Shinozaki, K., Ohme-Takagi, M. (2005) The NAC transcription factors NST1 and NST2 of Arabidopsis regulated secondary wall thickenings and are required for anther dehiscence. Plant Cell 17:2993-3006

Moses, A.M., Chiang, D.Y., Kellis, M., Lander, E.S., Eisen, M.B. (2003) Position specific variation in the rate of evolution in transcription factor binding sites. BMC Evolutionary Biol. 3(19): 1-13

Moses, A.M., Pollard, D.A., Nix, D.A., Iyer, V.N., Li, X.-Y., Biggin, M.D., Eisen, M.B. (2006) Large-scale turnover of functional transcription factor binding sites in Drosophila. PLoS Comp. Biol. 2(10): e130

Popper, Z.A., Michel, G., Herve, C., Domozych, D.S., Willats, W.G.T., Tuohy, M.G., Kloareg, B., Stengel, D.B. (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. Annu. Rev. Plant Biol. 62: 567-590

Prud'homme, B., Gompel, N., Carroll, S.B. (2007) Emerging principles of regulatory evolution. Proceedings of the National Academy of Sciences 104: 8605-8612

Roberts, A.W., Bushoven, J.T. (2007) The cellulose synthase (CESA) gene superfamily of the moss Physcomitrella patens. Plant Mol. Biol. 63: 207-219

Sarkar, P., Bosneaga, E., Auer, M. (2009) Plant cell walls throughout evolution: towards a molecular understanding of their design principles. J. Exp. Bot. 60(13): 3615-3635

Scheller, H.V., Ulvskov, P. (2010) Hemicelluloses. Annu. Rev. Plant Biol. 61, 263-289

Schmidt, D., Wilson, M.D., Ballester, B., Schwalie, P.C., Brown, G.D., Marshall, A., Kutter, C., Watt, S., Martinez-Jimenez, C.P., Mackay, S., Talianidis, I., Flicek, P., Odom, D.T. (2010) Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. Science 328: 1036-1040

Tamura, K., Dudley, J., Nei, M., Kumar, S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24(8):1596-1599

Weirauch, M.T., Hughes, T.R. (2009) Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. Trends in Genet. 26(2): 66-74

Weng, J.-K., Li, X., Stout, J., Chapple, C. (2008) Independent origins of syringyl lignin in vacular plants. Proceedings of the National Academy of Sciences 105(22): 7887-7892

Weng, J.-K., Chapple, C. (2010) The orignin and evolution of lignin biosynthesis. New Phytologist 187: 273-285

Wray, G.A. (2007) The evolutionary significance of cis-regulatory mutations. Nature Rev. Genet. 8: 206-216

Zhong, R., Lee, C., Ye, Z.-H. (2010a) Functional characterization of poplar wood-associated NAC domain transcription factors. Plant Physiol. 152: 1044-1055

Zhong, R., Lee, C., Ye, Z.-H. (2010b) Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. Mol. Plant 3: 1087-1103

Zhong, R., Richardson, E.A., Ye, Z.-H. (2007a). The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. Plant Cell 19: 2776-2792

Zhong, R., Richardson, E.A., Ye, Z.-H. (2007b). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary cell wall synthesis in fibers of Arabidopsis. Planta 225:1603-1611

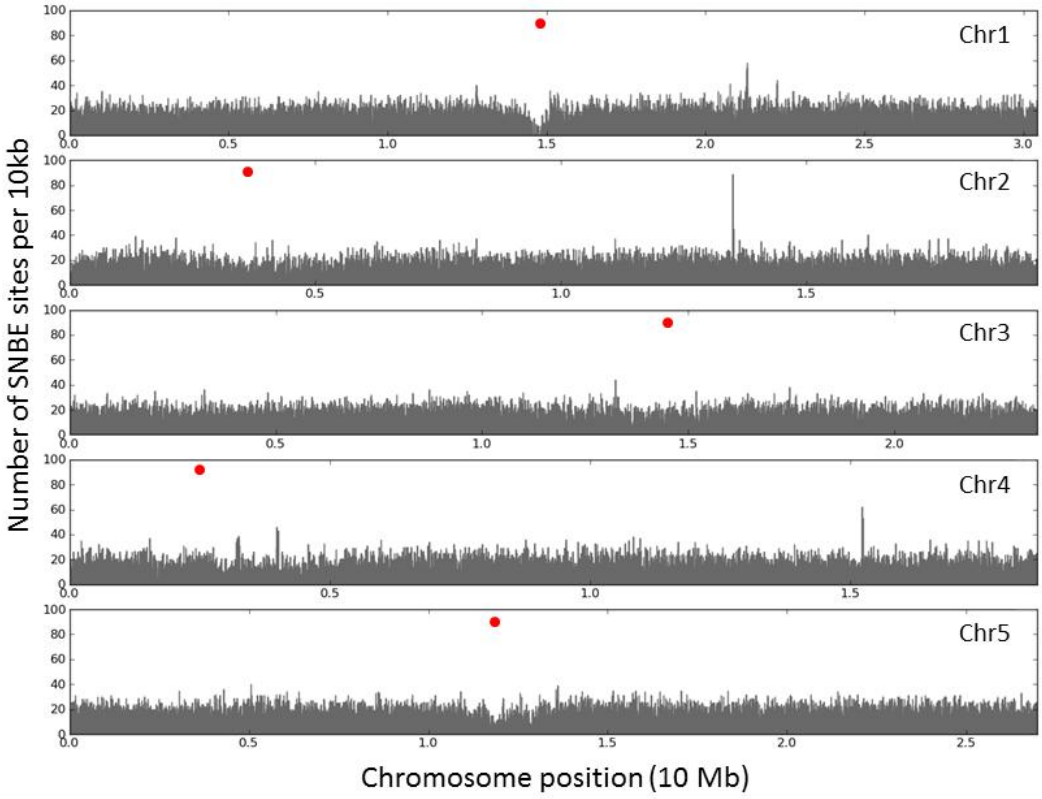

Figure 1. Genomic distribution of putative SNBE sites across *Arabidopsis thaliana* chromosomes. Red circles indicate approximate centromere position.

The *Arabidopsis thaliana* genomic sequence was analyzed for regions matching the SNBE consensus sequence. Locations of putative SNBE sites were recorded and the number of sites per 10kb was plotted for each chromosome.

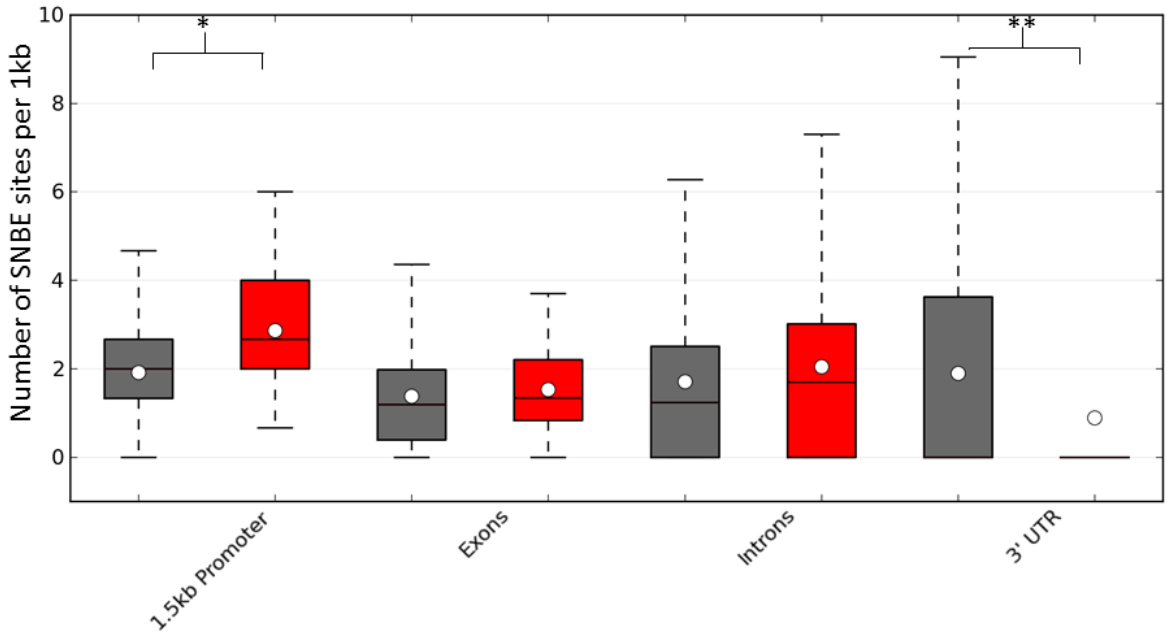

Figure 2. Number of putative SNBE sites per kb in the 1.5kb promoter, 5' UTR, exon, intron and 3' UTR regions of all *Arabidopsis thaliana* gene models (gray) and all gene models induced fourfold or more by SND1 or VND7 (red). The sequence of all annotated *Arabidopsis thaliana* gene models and their 1.5kb promoters were analyzed for sequences matching the SNBE consensus sequence. Boxplot distributions showing the upper and lower quartiles (box), median (solid black line), mean (white circle) and the 87.5%-12.5% distribution (whiskers).

*, $p<1e-8$; **, $p<0.0005$; unpaired t-test.

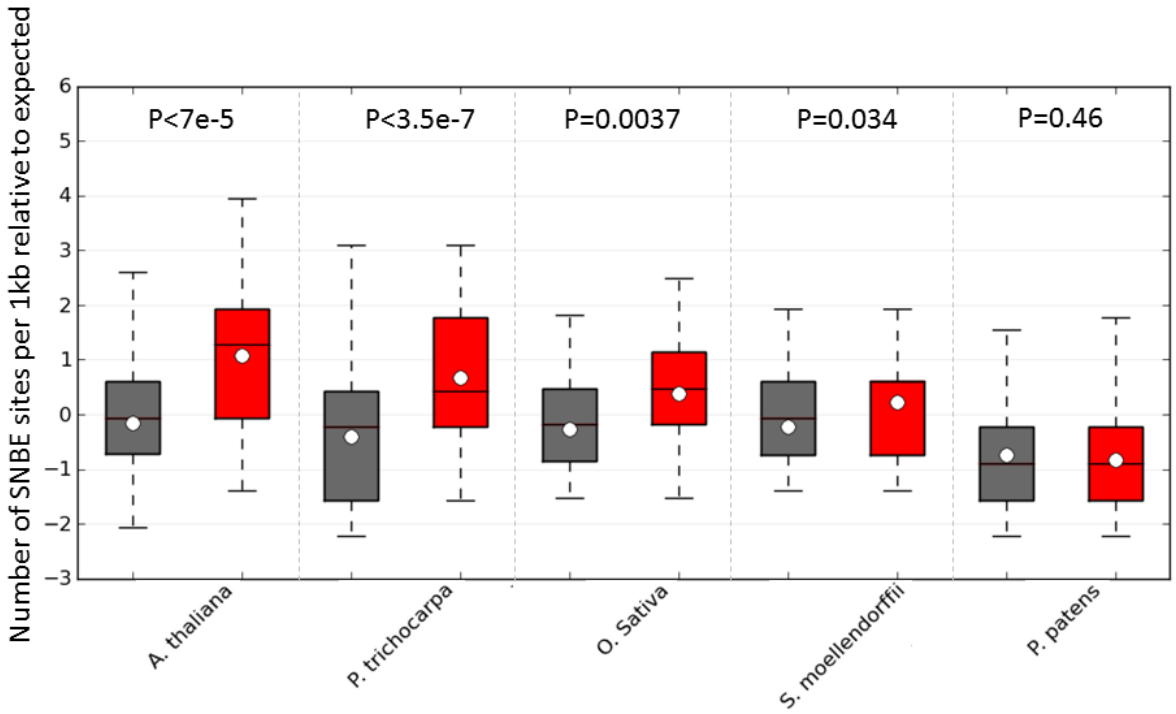

Figure 3. Comparison of the number of putative SNBE sites per 1kb in the 1.5kb promoter located in all gene promoters (gray) and promoters of the subset of genes orthologous to genes highly induced by SWNs (red) relative to the probabilistically expected number of SNBE site sequences for each species. The sequence of all annotated gene models from five plant species were analyzed for sequences matching the SNBE consensus sequence. Boxplot distributions showing the upper and lower quartiles (box), median (solid black line), mean (white circle) and the 87.5%-12.5% distribution (whiskers). P values reported are from unpaired two tailed t-tests between distributions.

Table 2. Gene models retrieved from Phytozome (www.phytozhome.org) identified as orthologous to genes from *Arabidopsis thaliana* highly induced by AtSND1 and/or AtVND7. The number of putative SNBE sites present in the 1.5kb sequence upstream of the transcriptional start site the gene models is reported in parenthesis. Multiple genes are listed in order of protein sequence similarity to their Arabidopsis ortholog.

Arabidopsis thaliana	Populus trichocarpa	Vinis vinifera	Oryza sativa ssp. japonica	Zea mays	Sorghum bicolor	Selaginella moellendorffii
At1g65570 (8)	POPTR_0010s18500 (3) POPTR_0010s18510 (4)	GSVIVG01004006001 (2)	Os11g14410 (4)			
At1g70500 (6)	POPTR_0001s46730 (7)	GSVIVG01013772001 (5)	Os07g10700 (6)	GRMZM2G111609 (4)	Sb04g002530 (4)	85947 (3) 89085 (1)
At2g04780 (1) FLA7	POPTR_0014s16100 (3) POPTR_0002s22020 (4)	GSVIVG01025660001 (4)	Os09g30486 (2) Os08g39270 (2) Os04g39600 (1)	AC234156.1_FG005 (2) GRMZM2G081017 (1)	Sb07g028140 (3) Sb02g028120 (2)	
At3g16920 (6) CTL2	POPTR_0010s15150 (1) POPTR_0014s14370 (8)	GSVIVG01035029001 (5) GSVIVG01031685001 (3)	Os09g32080 (2)	GRMZM2G168364 (2)	Sb02g028770 (2) Sb07g027310 (4)	227948 (3)
At3g47400 (4)	POPTR_0012s14620 (4) POPTR_0015s14740 (5)	GSVIVG01018620001 (5)				
At4g08160 (3)	POPTR_0002s11380 (7) POPTR_0002s11370 (6)	GSVIVG01009272001 (4)	Os01g04300 (4)	GRMZM2G002260 (3)	Sb02g010990 (4)	111304 (1)

At5g15050 (3)	POPTR_0017s11200 (4)	GSVIVG01015525001 (5)	Os12g44240 (4)	GRMZM2G423476 (3)	Sb08g023170 (2)	74856 (2)
At5g54670 (3) ATK3	POPTR_0011s13530 (5) POPTR_0001s44200 (2)	GSVIVG01014585001 (2)	Os04g53760 (2)	GRMZM2G436981 (2)	Sb02g000560 (5)	84710 (0)
At5g67230 (6) IRX14	POPTR_0005s18500 (4)	GSVIVG01022064001 (4)	Os06g47340 (2)	GRMZM2G113655 (2)	Sb10g027970 (0)	12082 (1)
At1g11190 (3) BFN1	POPTR_0011s04430 (2)	GSVIVG01005841001 (1)	Os04g54390 (0) Os01g03730 (3)	GRMZM2G168744 (2)	Sb06g030020 (0)	98354 (3)
At5g60490 (3) FLA12	POPTR_0006s13120 (3) POPTR_0016s09010 (6)	GSVIVG01025665001 (8)	Os01g06580 (4)	GRMZM2G177242 (2)	Sb09g028480 (3)	
At1g27440 (7) IRX10	POPTR_0001s12940 (1) POPTR_0012s11150 (4)	GSVIVG01008661001 (4) GSVIVG01001097001 (7)	Os01g70200 (4) Os02g32110 (8)	GRMZM2G000581 (1) GRMZM2G100143 (3) GRMZM2G056702 (4) GRMZM2G059845 (3)	Sb03g044530 (1) Sb03g044520 (4)	442111 (3)
At1g62990 (3) KNAT7	POPTR_0001s08550 (1)	GSVIVG01019880001 (3)	Os03g03164 (1)	GRMZM2G159431 (2)	Sb04g005620 (5)	90744 (0)
At2g40470 (5) LBD15	POPTR_0013s08040 (1) POPTR_0019s14790 (3)	GSVIVG01032714001 (2) GSVIVG01024592001 (4)	Os01g60960 (2)	GRMZM2G092517 (2)	Sb03g038410 (3)	37169 (2)
At4g00220 (5) LBD30	POPTR_0002s15000 (4) POPTR_0014s06600 (6)	GSVIVG01027621001 (3)	Os03g14270 (3)	GRMZM2G150594 (2)	Sb01g041130 (4)	59642 (4)

At5g04200 (5) MC9	POPTR_0016s02510 (2) POPTR_0006s02730 (2)	GSVIVG01025829001 (8)	Os11g04010 (2)	GRMZM2G022799 (1) GRMZM2G083016 (4)	Sb04g023560 (1) Sb04g007090 (3)	152567 (2)
At3g13100 (3) MRP7	POPTR_0001s37340 (4)	GSVIVG01037789001 (2)	Os01g07870 (8)	GRMZM2G355491 (2) GRMZM2G032218 (3)	Sb06g026830 (3)	402326 (5)
At5g12870 (9) MYB46	PtrMYB003 (7) PtrMYB020 (5) PtrMYB002 (8) PtrMYB021 (7)	GSVIVG01032088001 (2)	Os12g33070 (3)	GRMZM2G052606 (1)	Sb08g016620 (2)	
At3g08500 (9) MYB83		GSVIVG01025269001 (4)				437219 (2)
At1g09540 (5) MYB61	PtrMYB055 (6) PtrMYB216 (3) PtrMYB121 (4) PtrMYB170 (3)	GSVIVG01031341001 (4) GSVIVG01028235001 (6) GSVIVG01028235001 (3)	Os01g18240 (1) Os05g04820 (4)	GRMZM2G171781 (4) GRMZM2G017520 (2)	Sb03g011640 (3)	
At1g63910 (4) MYB103	PtrMYB128 (4) PtrMYB010 (4)	GSVIVG01019410001 (2)	Os08g05520 (1)	GRMZM2G325907 (0)	Sb07g003320 (1)	
At1g19300 (6) PARVUS	POPTR_0002s13400 (7) POPTR_0014s03980 (6)	GSVIVG01009639001 (3) GSVIVG01019033001 (4)	Os03g47530 (3) Os03g24510 (5)	GRMZM2G149024 (1) GRMZM2G300692 (4)	Sb01g012060 (3)	234804 (4)
At5g66390 (1) Peroxidase72	POPTR_0007s13420 (2) POPTR_0005s12070 (4)	GSVIVG01018865001 (2)	Os01g36240 (0) Os01g15830 (2)	GRMZM2G041308 (5)	Sb03g024460 (0) Sb03g010250 (3)	230146 (0) 430498 (3)

At1g26820 (5) RNS3	POPTR_0008s08650 (4) POPTR_0010s17570 (3)	GSVIVG01019974001 (6)	Os08g33710 (4)	GRMZM2G161274 (2)	Sb07g021330 (2)	270532 (1)
At1g28470 (8) SND3	POPTR_0004s04900 (5)	GSVIVG01020609001 (4)	Os01g48130 (5)	GRMZM2G031200 (3)	Sb09g028430 (3)	67638 (5)
At1g50010 (2) TUA2	POPTR_0001s29670 (3)	GSVIVG01033415001 (3)	Os11g14220 (1)	GRMZM2G153292 (2)	Sb01g009570 (2)	178975 (3)
At5g59290 (7) UXS3	POPTR_0010s21420 (3) POPTR_0008s05320 (3) POPTR_0001s24380 (6)					
At2g28760 (7) UXS6		GSVIVG01025003001 (5)	Os03g16980 (0)	GRMZM2G044027 (1) GRMZM2G165357 (2)	Sb01g039350 (1)	438351 (3)
At4g35350 (4) XCP1	POPTR_0004s21740 (8)	GSVIVG01023863001 (11) GSVIVG01013420001 (5)	Os02g48450 (2)	GRMZM2G066326 (3)	Sb09g000960 (4) Sb03g047290 (3)	230713 (3)
At1g20850 (3) XCP2	POPTR_0005s27730 (7)		Os05g01810 (3) Os01g73980 (3)	GRMZM2G367701 (1)		78855(5)
At5g64530 (2) XND1	POPTR_0003s01720 (5) POPTR_0001s21440 (8)	GSVIVG01029709001 (4) GSVIVG01026468001 (5)	Os02g34970 (3) Os04g35660 (2)	GRMZM2G316840 (3) GRMZM2G094067 (1)	Sb06g017190 (3) Sb04g022672 (2)	

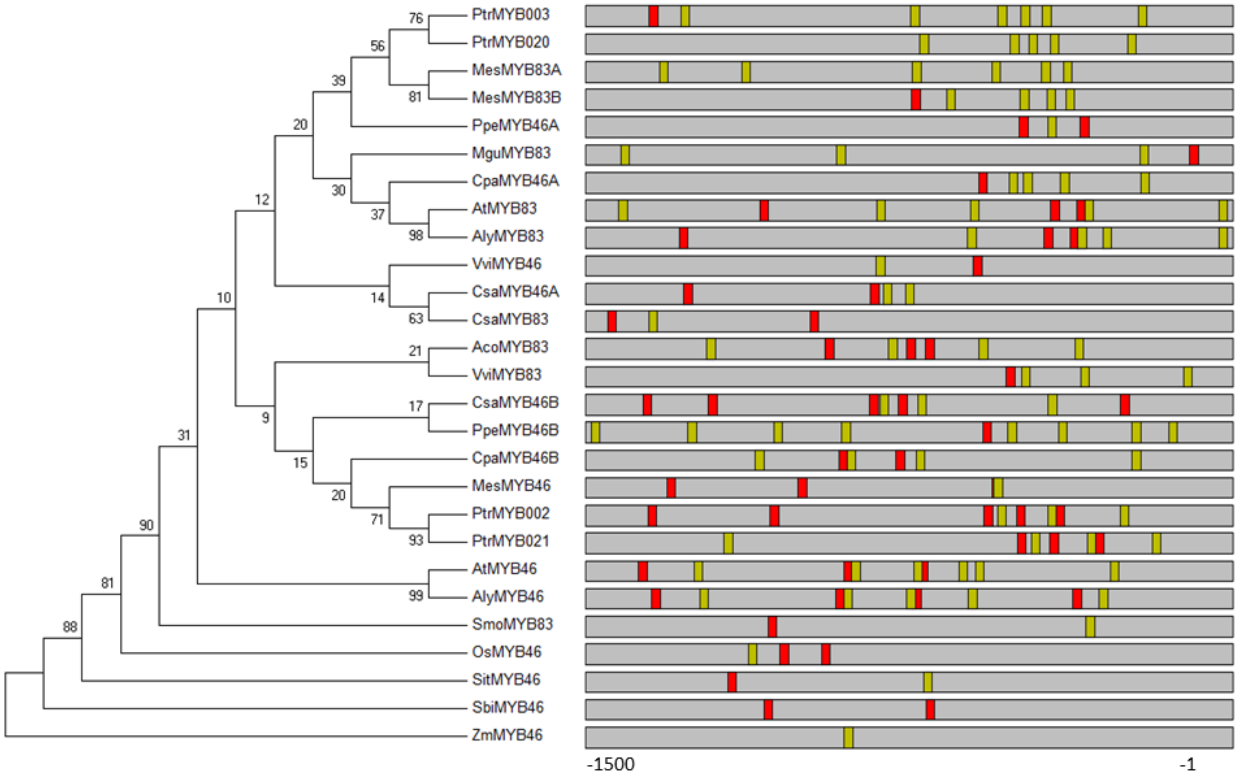

Figure 4. Positions of putative SNBE sites located on the top (yellow) and bottom strands (red) of the 1.5kb promoter region of genes orthologous to AtMYB46 and AtMYB83 grouped by the maximum parsimony phylogenetic relationship based upon amino acid sequence.

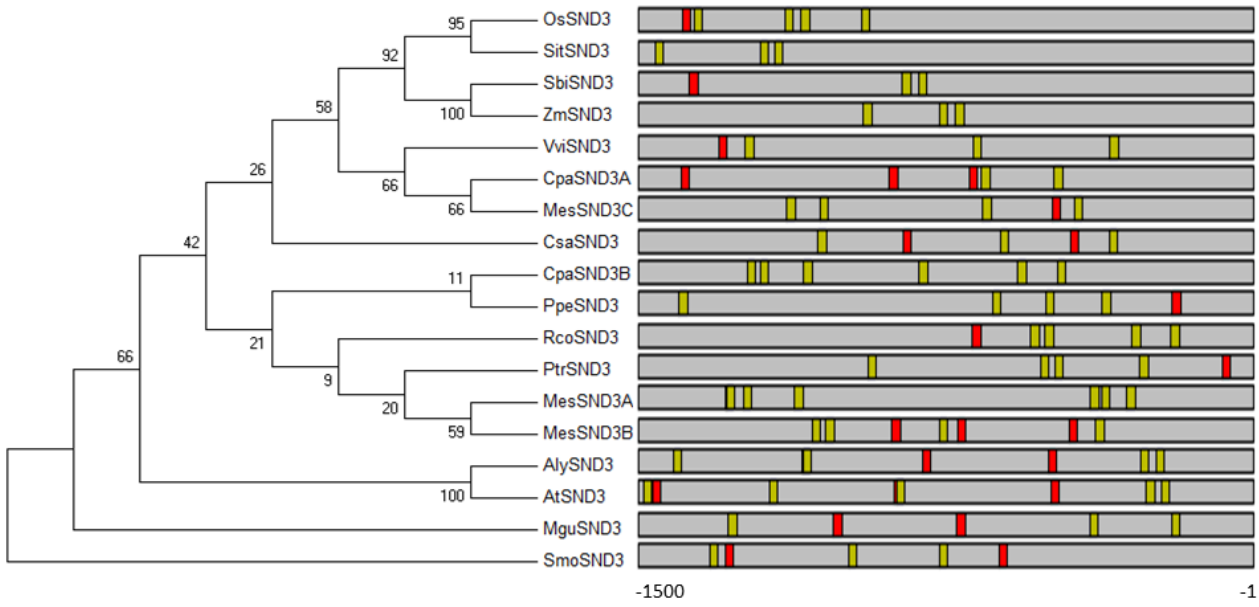

Figure 5. Positions of putative SNBE sites located on the top (yellow) and bottom strands (red) of the 1.5kb promoter region of genes orthologous to AtSND3 grouped by the maximum parsimony phylogenetic relationship based upon amino acid sequence.

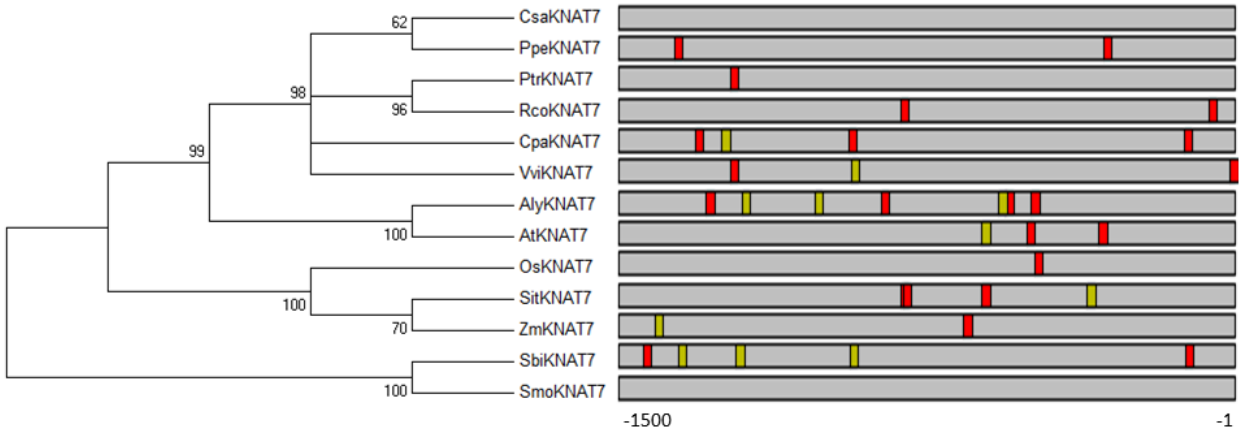

Figure 6. Positions of putative SNBE sites located on the top (yellow) and bottom strands (red) of the 1.5kb promoter region of genes orthologous to AtKNAT7 grouped by the maximum parsimony phylogenetic relationship based upon amino acid sequence.

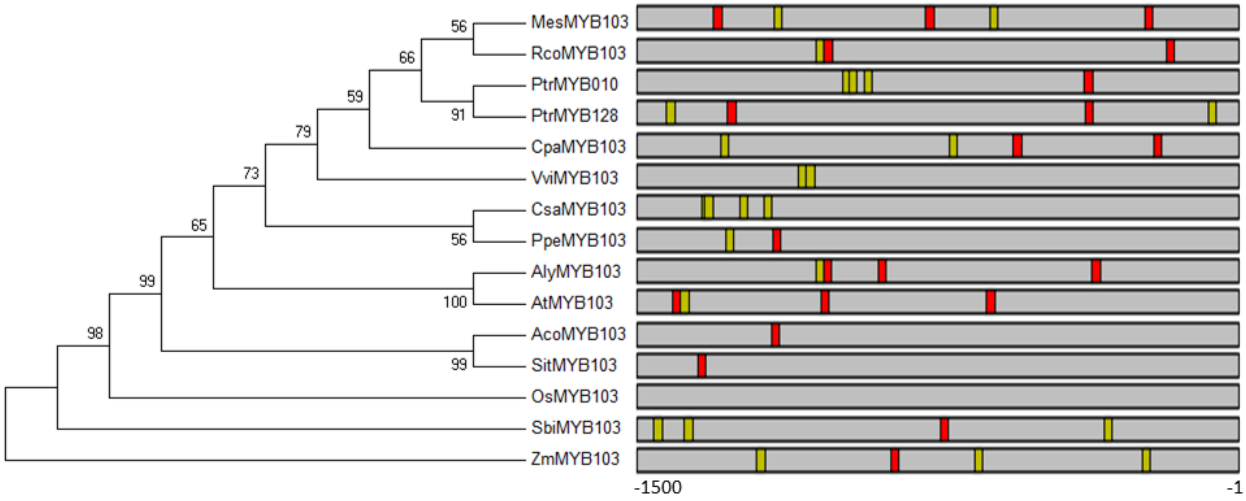

Figure 7. Positions of putative SNBE sites located on the top (yellow) and bottom strands (red) of the 1.5kb promoter region of genes orthologous to AtMYB103 grouped by the maximum parsimony phylogenetic relationship based upon amino acid sequence.

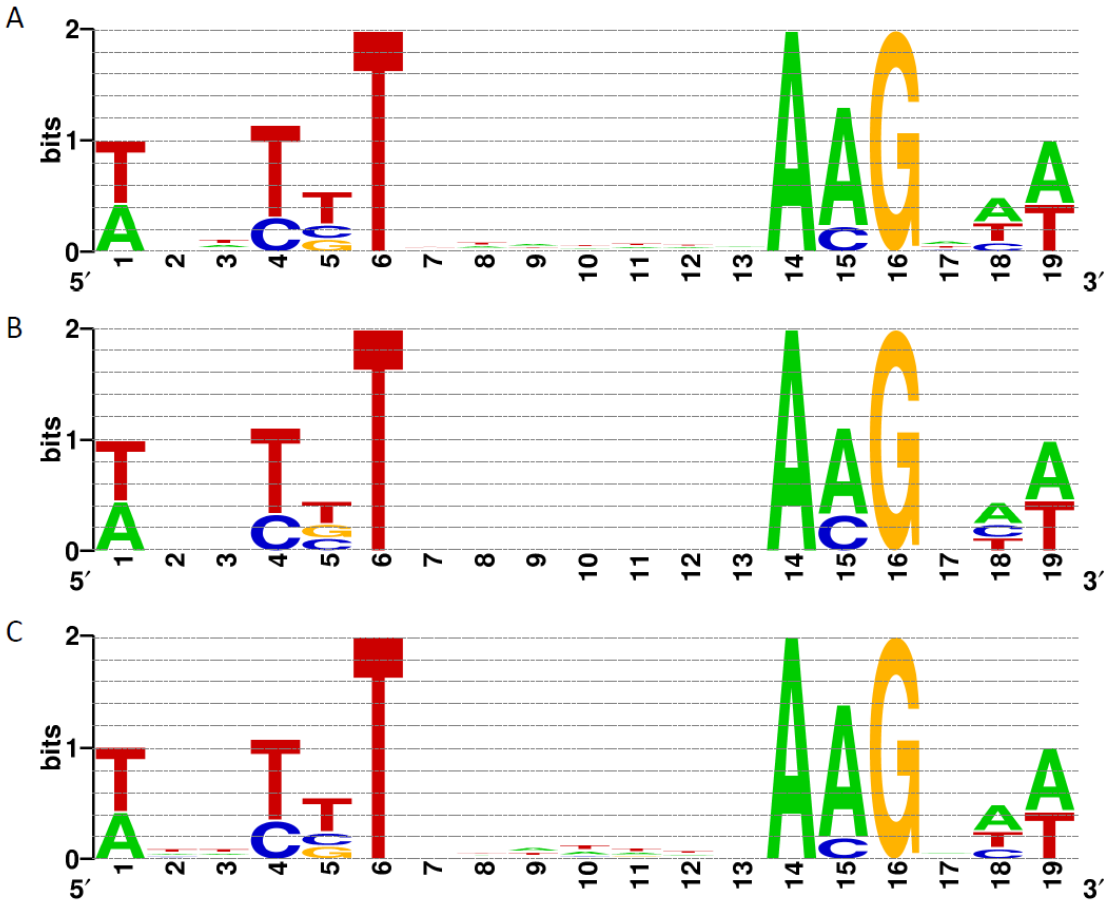

Figure 8. Consensus sequence logo generated with putative SNBE sites located in promoters of AtMYB46, AtMYB83, AtSND3, AtKNAT7 and AtMYB103 orthologs. The height of the stack is proportional to the sequence conservation at that position with higher stacks indicating greater sequence conservation. The height of each letter is proportional to the relative frequency of the nucleotide at that position. (A) SNBE sites present in promoters of *Arabidopsis* genes determined to be directly activated by SWNs by Zhong et al., 2010b. (B) Putative SNBE sites obtained from the promoters of genes listed in Table 2 from monocots. (C) Putative SNBE sites obtained from the promoters of genes listed in Table 2 from dicots.

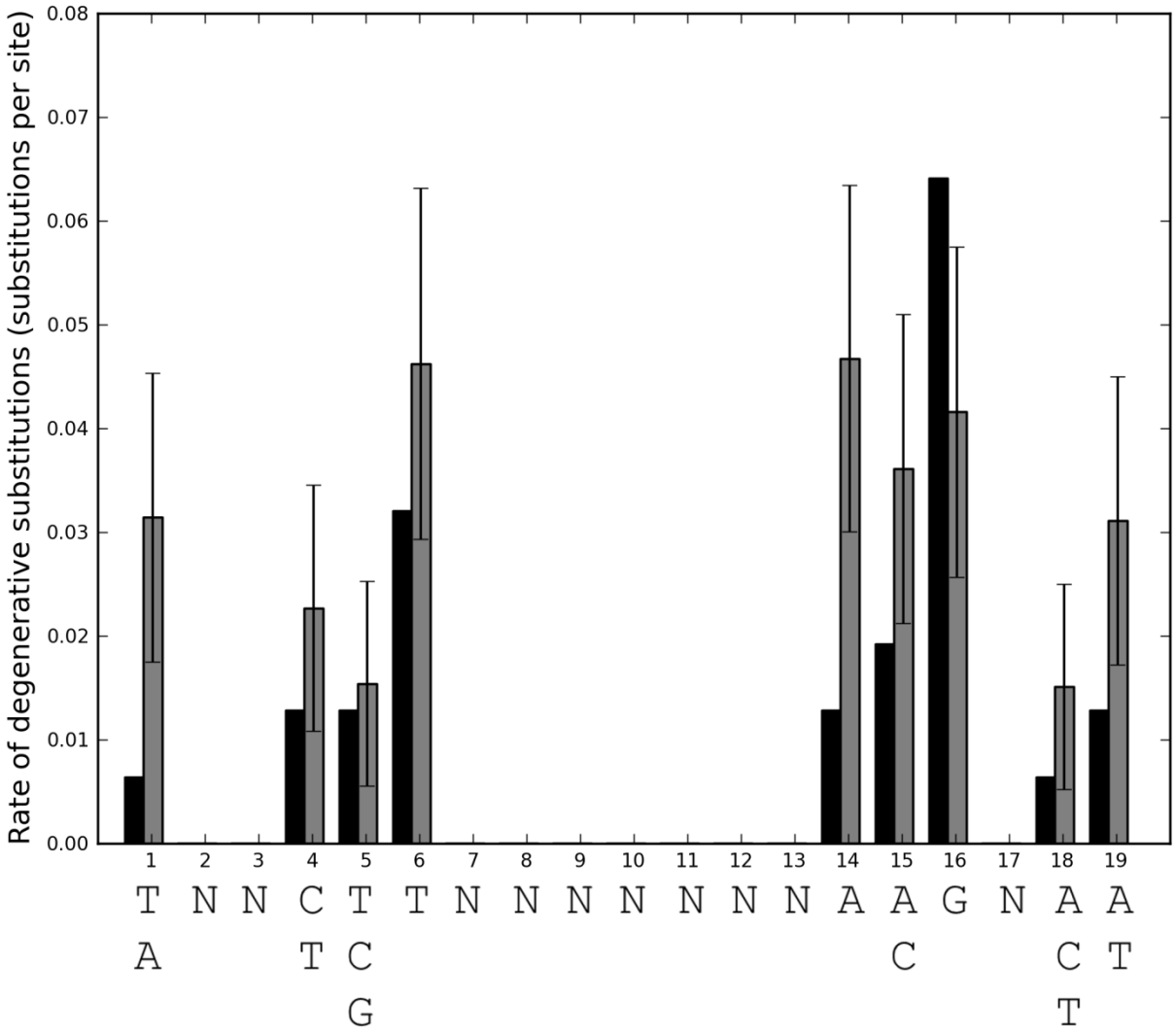

Figure 9. Positional rate of degenerative substitutions for nucleotides in pairs of SNBEs from *Arabidopsis thaliana* and *Arabidopsis lyrata* for observed (black bars) and simulated (grey bars) promoter sequences. Error bars indicate standard deviation of degenerative substitution rate for positions in simulated SNBE sites over 10,000 replicates. The SNBE consensus sequence for each position is listed along the bottom.

CHAPTER 5

CONCLUSIONS

MYB83 acts redundantly with MYB46 as key master regulators of secondary cell wall formation

MYB46 was previously shown to be master regulatory protein acting downstream of the secondary wall NAC (SWN) domain master regulatory proteins. Our findings presented in Chapter 2 reveal that MYB83, the closest homolog of MYB46, acts redundantly with MYB46 in the transcriptional network regulating secondary cell wall formation. The function of either MYB46 or MYB83 appears to be essential to secondary cell wall formation in all cell types as *myb46/myb83* double knockout mutants demonstrated a complete loss of secondary cell wall formation in all cell types analyzed. However, since plant growth does not progress past the early seedling stage it was impossible to evaluate the effect of this double mutation in mature *Arabidopsis*. The vital function of these genes in mature *Arabidopsis* was confirmed by double RNAi knockdown of MYB46 and MYB83 which severely reduced secondary wall thickening in both xylem vessels and fibers. MYB46 and MYB83 act as second level master regulators that are directly activated by the SWNs and are essential for the SWNs to initiate secondary cell wall formation. SWNs directly regulate a number of genes with functions in secondary cell wall biosynthesis and deposition but without the action of MYB46 and MYB83 it appears that this is insufficient for proper secondary cell wall formation. Identification of the set of genes directly activated by SWNs through the secondary wall NAC binding element (SNBE; Zhong et al., 2010b) as well as the genes directly activated by MYB46 and MYB83 through the secondary wall MYB-resposive element (SMRE; Zhong and Ye, 2012) shed light on how these two level of

master regulators operate in secondary cell wall formation. While the SWNs, MYB46 and MYB83 share a number of common target genes the majority of target genes are uniquely activated by either SWNs or by MYB46 and MYB83. With loss of MYB46 and MYB83 the SWNs no longer have the capacity to activate these essential downstream genes. Intriguingly, since neither MYB46 nor MYB83 activate the SWNs their ability to induce ectopic secondary cell wall formation indicates two possibilities. The first possibility is that the genes directly activated by SWNs but not by MYB46 or MYB83 are nonessential for secondary cell wall formation but may be necessary in certain cell types or under specific conditions. The second possibility is that the genes directly activated by SWNs are also indirectly activated by MYB46 and MYB83 through the activity of their target transcription factors. Dissecting how the two levels of secondary cell wall master regulators cooperatively function will require a better understanding of not only the targets of other transcription factors in the regulatory network but also the functions of genes directly involved in secondary wall biosynthesis and deposition. Exploring the evolutionary conservation of the transcriptional network regulating secondary cell wall formation may provide vital insights by revealing how the network structure compares between plant species with key physiological differences.

PtrMYB3 and PtrMYB20 are poplar orthologs of MYB46 and MYB83 with evolutionarily conserved function.

Herbaceous and woody angiosperms share many key similarities and differences. The process of secondary cell wall formation is believe to be largely conserved between the herbaceous *Arabidopsis* and the woody poplar based upon the shared similar cell morphology of xylem fibers and vessels (Zhang et al., 2012). The cellular similarities are reinforced by our findings presented in Chapter 3 that demonstrate that the poplar genes PtrMYB3 and PtrMYB20

are functionally conserved with their close *Arabidopsis* homologs MYB46 and MYB83. PtrMYB3 and PtrMYB20 demonstrated the capability to function as master regulators of secondary cell wall formation in both *Arabidopsis* and poplar. Additionally, PtrMYB3 and PtrMYB20 were capable of fully rescuing the *myb46/myb83* double mutant phenotype in *Arabidopsis*. These results demonstrate that not only are the functions of these genes conserved but their relationships in the transcriptional network regulating secondary cell wall formation, mediated through cis-element binding by transcription factors, are also conserved. Several poplar functional equivalents of the *Arabidopsis* SWNs have been identified and demonstrated to be master regulators of secondary cell wall formation in poplar (Zhong et al., 2010a; Zhong et al., 2011). We have shown that PtrWND2, one of the poplar SWNs (PtrSWN), is capable of binding the promoter and directly activating PtrMYB3 and PtrMYB20. The collective evidence built in poplar demonstrates that a two level set of secondary cell wall master regulators operate as the core of the regulatory network. The ability of polar transcription factors to functionally substitute for their *Arabidopsis* homologs indicates that both the transcription factor DNA binding domain and the target cis-elements are sufficiently conserved to permit interaction. The SNBE and SMRE sequences identified in *Arabidopsis* likely must be highly conserved in poplar in order to permit the observed interactions. As non-coding regions typically exhibit rapid rates of turnover the continued existence of SNBE and SMREs indicates the action of significant selective pressure.

SNBE sites represent an evolutionarily conserved mechanism for gene regulation by secondary wall NACs

The evidence for selective pressure maintaining transcription factor interactions between *Arabidopsis* and poplar brought about the possibility that this conservation might extend to a

wider proportion of the vascular plants. Utilizing the SNBE consensus sequence determined in *Arabidopsis* we analyzed the genomic sequence of many vascular plants for evidence supporting the widespread existence of SNBE sites. Analysis of the *Arabidopsis* genome unsurprisingly revealed that SNBE sites were significantly enriched in gene promoters, particularly in the promoters of genes known to be activated by SWNs. The enrichment of SNBE sites in promoters of SWN activated gene homologs proved to be a commonality shared by all vascular plant genomes analyzed including the basal vascular plant *Selaginella moellendorffii*. Analysis of the genomic sequence from the nonvascular bryophyte *Physcomitrella patens* revealed that the enrichment of SNBEs in gene promoters was not observed outside the vascular plant lineage. This apparent sharp distinction between vascular and nonvascular plants likely indicates that the core of the regulatory network controlling secondary cell wall formation coevolved with the plant vascular system and has been largely maintained in the vascular plant lineage.

Discussion and Prospects for future analysis

Our findings demonstrated in Chapters 2 help to define the core of the regulatory network controlling secondary cell wall biosynthesis in *Arabidopsis*. The discoveries elucidated in Chapters 3 and 4 demonstrate the evolutionary conservation of this network across diverse vascular plants. Collectively this work provides insight into the structure and evolution of the transcriptional network controlling secondary cell wall formation and establishes tools for creating plants with enhanced secondary cell wall properties for biofuel applications.

The identification of MYB83 as a functionally redundant homolog of MYB46 greatly increased our understanding of the network structure and function. The universal activity of MYB46 and MYB83 in both xylem vessels and fibers is in contrast to the tissue specific functions of SWN pairs and may provide insight into how the transcriptional network regulating

secondary cell wall formation changed as the plant vascular system evolved during the divergence of gymnosperms and angiosperms. The gymnosperm tracheid with dual roles in both structural support and water transport is believed to be similar to the ancestral cell form that led to the specialized fiber and vessel cells in angiosperms (Patzlaff et al., 2003). This assertion is backed up by the presence of tracheid like vascular tissue in ferns and fern allies (Friedman and Cook, 2000). The division of support and transport roles to xylem fibers and vessels respectively also required many specific changes to cell shape and secondary wall characteristics. It is possible that the function of the MYB master regulators is to establish expression of the genes required for complete secondary cell wall formation but the unique targets of SWNs provide cell type specific alterations. If this is true we would expect the cell type specific expression of SWN pairs to have evolved concurrently with the emergence of distinct xylem fibers and vessels. To evaluate this possibility we will need to expand our understanding of secondary cell wall regulation in gymnosperms. Currently very little is known about this process in gymnosperms with only the transcription factors PtMYB1, PtMYB4 and PtMYB8 from *Pinus taeda* and their corresponding white spruce homologs being implicated in the processes of phenylpropanoid biosynthesis and secondary cell wall formation (Patzlaff et al., 2003; Bomal et al., 2008). However, while these genes are homologous to MYB46 and MYB83 their roles as master regulators of secondary cell wall biosynthesis remain unconfirmed. Efforts to identify potential genes involved in gymnosperm secondary cell wall regulation would be greatly advanced with the completion of a fully sequenced gymnosperm genome. Although sequencing projects in the conifers *Pinus taeda, Pseudotsuga menziesii* and *Pinus lambertiana* are currently underway the exceptionally large genomes of these species (18.7-33.5 Mb) makes the process a long term goal. Meanwhile overexpression and complementation studies in *Arabidopsis* similar to those that

helped to elucidate the functions of PtrMYB3 and PtrMYB20 should be undertaken to further our understanding of these pine MYB genes.

PtrMYB3 and PtrMYB20 demonstrate a conserved mechanism between woody and herbaceous angiosperms for secondary cell wall regulation but physiological differences likely indicate that key dissimilarities exist. Both NAC and MYB domain families of transcription factors have greatly expanded in poplar since its divergence from *Arabidopsis* approximately 100 million years ago (Wilkins et al., 2009; Hu et al., 2010). This expansion has been fueled by gene and genome duplications including the duplication of many poplar homologs of *Arabidopsis* NAC and MYB domain transcription factors demonstrated to have roles in secondary cell wall formation. As genome duplication often reduces selective pressure it is often followed by subfunctionalization of the duplicated genes (Adams and Wendel, 2005). While several of the poplar wood-associated NAC domain transcription factors (PtrWNDs) have been demonstrated to be master regulators of secondary cell wall biosynthesis akin to the SWNs in *Arabidopsis* current results suggest all are expressed and function in both xylem fibers and vessels (Zhong et al., 2010a). This discrepancy with the cell type specific expression and function of SWNs in *Arabidopsis* may indicate a key difference between herbaceous and woody species. However, the many differences between xylem vessels and fibers must necessarily have an underlying cause. Unraveling the mechanism for cell type specific patterns of secondary cell wall formation and whether it reveals subfunctionalization of the PtrWNDs will be crucial to improve our understanding of this process. Additionally, characterizing the WND regulated transcription factors will likely reveal both conserved and novel secondary cell wall regulators in poplar (Zhong et al., 2011).

The evolution of vascular tissue was a gradual process with many extant plants providing examples of prototypical intermediate forms. While they lack true vascular tissue with lignified secondary cell walls, many bryophytes possess water conducting cells that form a secondary layer of primary cell wall material and undergo programmed cytoplasmic lysis which can be viewed as a precursor to vascular tissue (Ligrone et al., 2000). Our analysis of the evolutionary conservation of SNBE site distribution throughout the vascular plant lineage demonstrated a stark contrast between bryophytes and vascular plants. While all vascular plants analyzed demonstrated a significant enrichment of SNBE sites in gene promoters the bryophyte *Physcomitrella patens* exhibited no enrichment. This distinction may underlie key evolutionary changes required for the formation of vascular tissue with lignified secondary cell walls. Exploring the origin of the transcriptional network regulating secondary cell wall formation may reveal whether a similar network is responsible for the vascular cell like processes observed in the water conducting cells of bryophytes. By better understanding how the regulatory network controlling secondary cell wall formation originated we will develop a better understanding of the origin of vascular tissue and its role in the plant kingdom's expansion to land.

Gaining a better understanding of secondary cell wall formation in diverse plant species, especially bioenergy crops such as poplar, gives us the capability to utilize this knowledge to improve biofuel production. The capacity of the NAC and MYB secondary cell wall master regulatory proteins to ectopically activate secondary cell wall formation creates the possibility for creating poplar trees with greatly enhanced quantities of secondary cell wall biomass. Since constitutive overexpression and subsequent ectopic secondary cell wall formation leads to aberrant phenotypes and stunted growth (Zhong et al., 2011) the key will be determining a way to increase secondary cell wall biomass in a manner that does not impede normal plant growth.

Future efforts to accomplish this goal have the capability of greatly decreasing the costs of biofuel production and making it a more viable alternative to conventional fossil fuels. Migrating to renewable sources of energy such as biofuels has the potential to drastically decrease atmospheric CO_2 and mitigate the effects of climate change.

References

Adams, K.L., Wendel, J.F. (2005) Polyploidy and genome evolution in plants. Curr. Opin. Plant Biol. 8: 135-141

Bomal, C., Bedon, F., Caron, S., Mansfield, S.D., Levasseur, C., Cooke, J.E.K., Blais, S., Tremblay, L., Morency, M.-J., Pavy, N., Grima-Pettenati, J., Sequin, A., MacKay, J. (2008) Involvment of Pinus taeda MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis. J. Exp. Bot. 59(14): 3925-3939

Friedman, W.E., Cook, M.E. (2000) The origin and early evolution of tracheids in vascular plants: integration of palaeobotanical and neobotanical data. Phil. Trans. R. Soc. Lond. B. 355: 857-868

Hu, R., Qi, G., Kong, Y., Kong, D., Gao, Q., Zhou, G. (2010) Comprehensive analysis of NAC domain transcription factor gene family in Populus trichocarpa. BMC Plant Biol. 10:145

Ligrone, R., Duckett, J.G., Renzaglia, K.S. (2000) Conducting tissues and phyletic relationships of bryophytes. Phil. Trans. R. Soc. Lond. B. 355:795-813

Patzlaff, A., Mclnnis, S., Courtney, A., Surman, C., Newman, L.J., Smith, C., Bevan, M.W., Mansfield, S., Whetten, R.W., Sederoff, R.R., Campbell, M.M. (2003) Characterization of a pine MYB that regulates lignification. Plant J. 36: 743-754

Wilkins, O., Nahal, H., Foong, J., Provart, N.J., Campbell, M.M. (2009) Expansion and diversification of the Populus R2R3-MYB family of transcription factors. Plant Physiol. 149: 981-993

Zang, J., Elo, A., Helariutta, Y. (2011) *Arabidopsis* as a model for wood formation. Curr. Opin. Biotech. 22:293-299

Zhong, R., Lee, C., Ye, Z.-H. (2010a) Functional characterization of poplar wood-associated NAC domain transcription factors. Plant Physiol. 152:1044-1055

Zhong, R., Lee, C., Ye, Z.-H. (2010b) Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. Mol. Plant 3: 1087-1103

Zhong, R., McCarthy, R.L., Lee, C., Ye, Z.-H. (2011) Dissection of the transcriptional program regulating secondary wall biosynthesis during wood formation in poplar. Plant Physiol. 157:1452-1468

Zhong, R., Ye, Z.-H. (2012) MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary cell wall biosynthetic genes. Plant Cell Physiol. 53(2): 368-380