Summer November 2014

Understanding the transcriptional regulation of secondary cell wall biosynthesis in the model grass Brachypodium distachyon

Pubudu Handakumbura
UMass Amherst

Understanding the transcriptional regulation of secondary cell wall biosynthesis in the model grass *Brachypodium distachyon*

A Dissertation Presented

by

PUBUDU P HANDAKUMBURA

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2014

Plant Biology Graduate Program

Understanding the transcriptional regulation of secondary cell wall biosynthesis in the model grass *Brachypodium distachyon*

A Dissertation Presented

by

PUBUDU P HANDAKUMBURA

Approved as to style and content by:

_____

Samuel P. Hazen, Chair

_____

Elsbeth L. Walker, Member

_____

Magdalena Bezanilla, Member

_____

John M. Lopes, Member

_____

Elsbeth L. Walker, Director

Plant Biology Graduate Program

DEDICATION

To my loving husband Rasitha.

ACKNOWLEDGMENTS

I would like to thank my advisor, Samuel Hazen, for his mentoring, Hazen lab members and my colleagues from the Plant Biology graduate program. I would also like to thank my family and friends for their support and encouragement throughout my graduate studies.

ABSTRACT

UNDERSTANDING THE TRANSCRIPTIONAL REGULATION OF SECONDARY CELL WALL BIOSYNTHESIS IN THE MODEL GRASS *Brachypodium distachyon*

SEPTEMBER 2014

PUBUDU P HANDAKUMBURA, B.S., UNIVERSITY OF COLOMBO, SRI LANKA

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Dr. Samuel P. Hazen

Secondary cell wall synthesis occurs in specialized cell types following completion of cell enlargement. By virtue of mechanical strength provided by a wall thickened with cellulose, hemicelluloses, and lignin, these cells can function as water-conducting vessels and provide structural support. Several transcription factor families regulate genes encoding wall synthesis enzymes. Certain NAC and MYB proteins directly bind upstream of structural genes and other transcription factors. The most detailed model of this regulatory network is established predominantly for a eudicot, *Arabidopsis thaliana*. In grasses, both the patterning and the composition of secondary cell walls are distinct from that of eudicots. These differences suggest transcriptional regulation is similarly distinct. Putative rice and maize orthologs of several eudicot cell wall regulators genetically complement mutants of *A. thaliana* or result in wall defects when constitutively over-expressed; nevertheless, aside from maize ZmMYB31, switchgrass PvMYB4, and Brachypodium BdSWN5, function has not been tested in a grass. Similar to the seminal work conducted in *A. thaliana*, gene expression profiling in

maize, rice, and other grasses implicates additional genes as regulators. Characterization of these genes in a grass species will continue to elucidate the relationship between the transcription regulatory networks of eudicots and grasses. In the context of this dissertation two cell wall genes responsible for synthesizing cellulose in the secondary cell walls were characterized. Several MYBs, a NAC and a bZIP protein was found to interact with the cellulose gene promoters. A reverse genetics approach was used to functionally characterize two of those regulators, MYB48 and GNRF. MYB48 is the first grass specific cell wall regulator found to positively regulate cell wall biosynthesis by binding to the cellulose and lignin gene promoters. It regulates above ground biomass in *B. distachyon*. GNRF, on the hand, unlike the characterized NAC proteins, was shown to repress cell wall biosynthesis. GNRF is also repressing flowering in *B. distachyon*, a novel regulatory function that has not been associated with the characterized NAC proteins to date. Further characterization of GNRF is likely to provide new insights into the pleiotropic regulatory roles of this protein.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Increasing cost of obtaining finite fossil fuels and the climate changes associated with their use has created a global need for alternative and renewable energy sources. Biofuels derived from cellulosic biomass are currently being investigated as an alternative to meet these growing energy needs. Cellulosic biomass mainly consists of plant cell walls, which are a complex mixture of polysaccharides, lignin and proteins; the former being the raw material for biofuels production. Non-food perennial grass species such as switchgrass and *Miscanthus* spp. are being developed as next generation energy crops. To rapidly improve desirable feedstock properties, we need to first understand the nature of the plant cell wall and the genes involved in synthesizing cell wall components. Additionally, we need to examine how these genes are regulated to engineer cell walls with specific functions. Surprisingly little is known about grass cell wall biosynthesis and practically nothing about its regulation. Understanding cell wall biosynthesis and its regulation will enable us to modify crops to attain desirable feedstock properties.

1.1 Plant cell walls

Plant cell walls are diverse in their polymer composition, which vary among plant species as well as cell types within a species. The primary cell wall, consisting mostly of cellulose, hemicelluloses, pectin, and proteins demarcates each plant cell. Secondary cell walls, composed mostly of cellulose, hemicelluloses, and lignin, are deposited between the plasma membrane and primary wall in specialized cell types following cessation of cell enlargement. Even though all plant cell walls contain a cell wall matrix made up of

cellulose microfibrils, hemicelluloses, and other cell wall polymers, there are significant differences in polymer types and their relative abundance between dicots and monocots (Vogel, 2008). For instance, relatively high amounts of xyloglucan, pectins, and structural proteins are typically present in the primary walls of eudicots, non-commelinoid monocots, and gymnosperms. The commelinoid monocots like grasses contain glucuronoarabinoxylans, small quantities of pectin and structural proteins and high levels of hydroxycinnamates (Vogel, 2008). Apart from wall composition, eudicots such as *Arabidopsis thaliana* and monocots such as rice (*Oryza sativa* L.) exhibit distinct morphological characteristics in addition to their namesake double and single cotyledons. Pinnate or palmate venation is characteristic of eudicots while monocots possess parallel venation. Other morphological and anatomical distinctions exist in their vasculature, tissues highly enriched in secondary cell walls. Stem radial thickening in eudicots is derived from a specialized cell layer called the cambium that differentiates into the phloem and xylem. On the other hand, monocots lack a specialized cambium layer and therefore do not undergo secondary growth (Fig. 1.1). The vascular bundles of grasses are often well defined by a single layer of bundle sheath cells surrounding the xylem and phloem and the organization of the bundles is distinct from eudicots. In *A. thaliana*, the vascular bundles are arranged as a ring along the periphery of the stem in a pattern known as eustelic, whereas monocot bundles posses an atactostele arrangement characterized by several circles around the periphery of a stem, as in rice and *Brachypodium distachyon* (Fig. 1.1), or scattered throughout the stem as in maize (*Zea mays* L.; (Kiesselbach, 1949)

With distinctions and similarities abound, it is decidedly unclear how the transcriptional regulation of secondary cell wall biosynthesis in eudicots and grasses

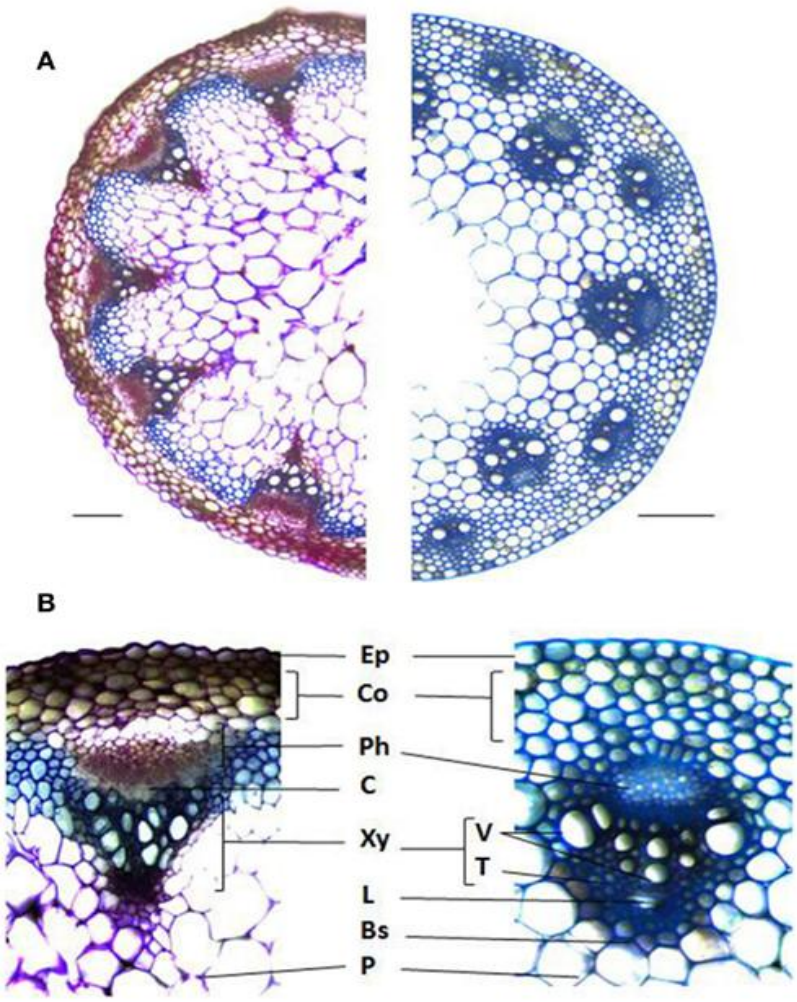

Figure 1.1. Stem cross sections illustrating the different cell types and arrangements between dicots and monocots. (A) *Arabidopsis thaliana* (left) and *Brachypodium distachyon* (right) stained with Toluidine blue. (B) Vascular bundle anatomy of *A. thaliana* (left) and *B. distachyon* (right). Ep, Epidermis; Co, Cortex; Ph, Phloem; C, Cambium; Xy, Xylem; V, Vessels; T, Tracheads; L, Lacuna; Bs, Bundle Sheath; P, Pith. Bars=0.1mm.

relate. Plant cell wall biosynthesis is regulated at different molecular and cellular levels. Current evidence supports a complex regulatory network consisting of a handful of proteins from only a small portion of over 65 different transcription factor families (Fig. 1.2). Phylogenetic analysis has identified close homologs of *A. thaliana* regulators from both vascular and non-vascular plants and some of those regulators were capable of complementing *A. thaliana* cell wall mutants (Zhong et al., 2011). These findings suggest the evolutionary conservation of the transcriptional regulators in secondary cell wall biosynthesis. On the other hand, due to the pronounced differences between eudicot and monocot secondary wall composition and anatomy, there are likely unique aspects of the regulatory network yet to be resolved.

1.2 Transcriptional regulation in *A. thaliana*

Transcriptional regulation is one of the most important processes controlling plant cell wall biosynthesis, mediated by the interaction and interplay of *cis*-regulatory DNA elements and the *trans*-acting transcription factor proteins. Recent evidence suggests the involvement of an AP2 family protein *SHINE/WAX INDUCER 1* (*SHN*) as a global level regulator of cell wall biosynthesis (Ambavaram et al., 2011). Constitutive overexpression of *A. thaliana SHN* in rice results in the activation of cellulose and other cell wall-associated genes and the repression of lignin pathway genes. This protein is also capable of activating key *NAC* and *MYB* regulators and electrophoretic mobility shift assays (EMSA) demonstrate the direct binding of SHN protein to promoters of rice cell wall-associated transcription factors (Ambavaram et al., 2011). A WRKY family transcription factor, WRKY12, acts as a global repressor of secondary wall biosynthesis (Wang et al.,

2010). Transcripts of this gene are abundant in the cortex and pith cells of *A. thaliana* that lack secondary walls. Loss-of-function *wrky12* mutants exhibit increased expression of transcription factors associated with secondary wall biosynthesis as well as ectopic depositions of lignin, cellulose, and xylan and consequently, an overall increase in plant biomass (Wang et al., 2010). In the presence of WRKY12 protein, pith stem cells are maintained parenchymatous and the deposition of secondary cell walls is repressed. MYB32 similarly acts as a repressor of secondary wall biosynthesis, but in cells where this pathway has been activated; thus, it may provide negative feedback (Preston et al., 2004). Transgenic overexpression of *MYB32* resulted in the repression of *SECONDARY WALL-ASSOCIATED NAC-DOMAIN PROTEIN 1* (*SND1*), a higher order activator of secondary wall biosynthesis (Wang et al., 2010). Interestingly, the SND1 protein directly binds the *MYB32* promoter to activate gene expression (Wang et al., 2011). A group of NAC-domain transcription factors, *NAC SECONDARY WALL THICKENING FACTOR 1* (*NST1*), *NST2, SND1* (also known as *NST3*), *VASCULAR-RELATED NAC-DOMAIN* 6 (*VND6*), and *VND7*, collectively known as the secondary wall NACs (SWNs) are implicated as positively acting master regulators in a variety of tissues (Fig. 1.2) (Demura and Fukuda, 2007; Zhong and Ye, 2007). Among the NACs, *SND1* functions as a key switch governing the regulation of all secondary wall polymers (Zhong et al., 2006; Mitsuda et al., 2007). Overexpression of *SND1* leads to ectopic deposition and activation of cellulose, hemicellulose, and lignin biosynthesis genes. Conversely, dominant repression of *SND1* results in the absence of secondary wall development in vascular and interfascicular fibers (Zhong et al., 2006; Mitsuda et al., 2007). It also directly activates itself and is repressed by MYB transcription factors under the direct influence of *MYB46*

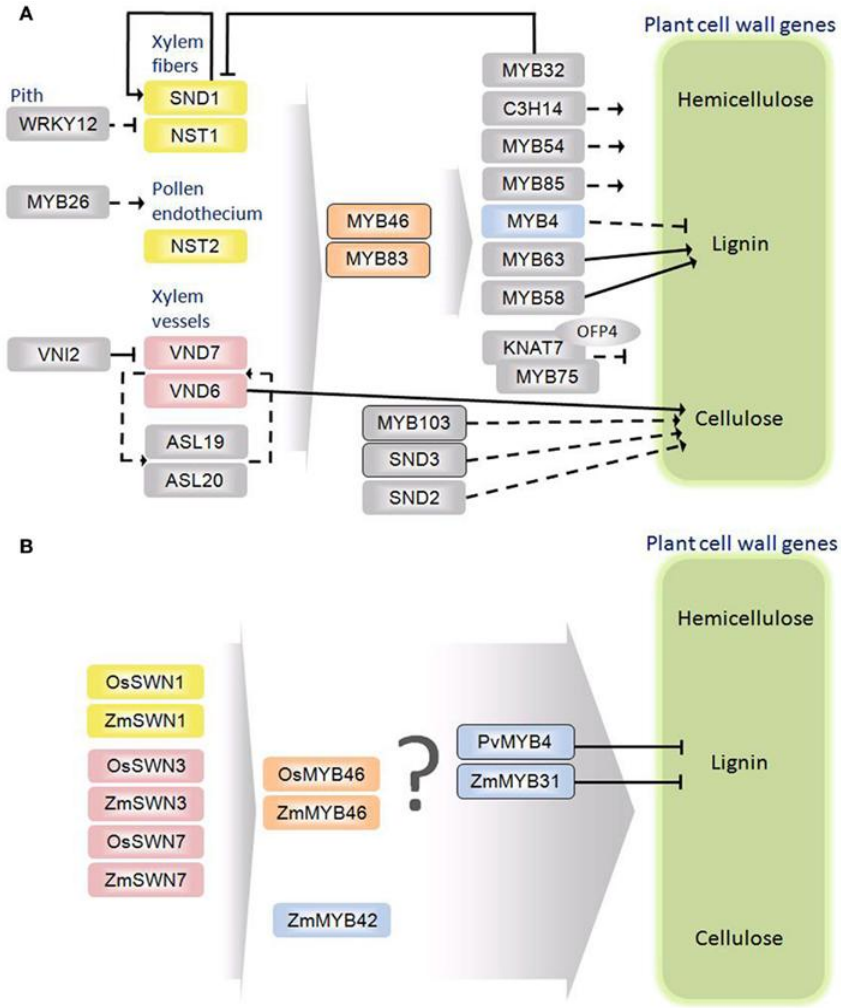

Figure 1.2. Schematic diagrams of the secondary cell wall regulatory networks in *Arabidopsis thaliana* (A) and monocots (B). Rectangles represent transcription factors. The oval indicates an interacting protein. Arrows and flat lines indicate activation and repression of cell wall genes. Solid arrows and bordered rectangles signify evidence for direct interactions. Dashed arrows indicate no evidence for direct interaction. Orthology between *A. thaliana* and grasses is denoted by color.

(Wang et al., 2011). The SND1 protein directly binds the *cis*-regulatory regions of *MYB46*, *MYB83*, and *C3H14* genes to activate their expression (Zhong et al., 2008; Ko et al., 2009; McCarthy et al., 2009). It also acts as a direct regulator of *F5H*, a gene encoding a key enzyme involved in lignin biosynthesis (Zhao et al., 2010). Another direct target of *SND1* is the KNOX type Homeodomain transcription factor, *KNAT7*, which interacts with OVATE FAMILY PROTEIN 4 to repress secondary wall biosynthesis (Li et al., 2011). Loss-of-function mutants of *knat7* display an increase in cell wall gene expression, wall thickening in interfascicular fiber cells, and an increase in lignin content (Li et al., 2011). Interestingly, a weak activator of cell wall gene expression, MYB75, was found to physically interact with KNAT7 protein to repress cell wall biosynthesis (Bhargava et al., 2010). These results imply that wall regulators can play multiple roles by interacting with different *trans*-acting factors in different cell types to provide more flexibility and complexity to the regulatory network.

The *SND1* homologs *NST1* and *NST2* play a crucial role in the *A. thaliana* anther endothecium (Mitsuda et al., 2005). In this tissue, these proteins are activated by *MYB26* which is essential for anther dehiscence and proper pollen release (Yang et al., 2007). Other SWNs, in particular *VND6* and *VND7*, regulate metaxylem and protoxylem development and are repressed by the *VND-INTERACTING2 NAC* protein ((Kubo et al., 2005; Yamaguchi et al., 2010). The VND proteins VND6 and VND7 driven by *SND1* promoter were capable of complementing the *snd1/nst1* mutant phenotype implying their conserved functionality (Zhong et al., 2007). *VND6* and *VND7* are positively regulated by *ASYMMETRIC LEAVES2-LIKE19* (*ASL19*) and *ASL20* (Soyano et al., 2008). Transgenic overexpression of *ASL19* and *ASL20* induces trans-differentiation of non-vascular tissues

into treachery elements and an increased cell wall thickening in mutant lines. They are also able to partially recover the dominant negative effect of the VND6 and VND7 repressor lines (Soyano et al., 2008).The SWNs, *SND1*, *NST1*, *NST2*, *VND6*, and *VND7* activate a cascade of downstream transcription factors such as *SND2*, *SND3*, *MYB103*, *MYB85*, *MYB54*, *MYB46*, *MYB69*, *MYB63*, *MYB83*, and *KNAT7* (Zhong et al., 2008; Ko et al., 2009; McCarthy et al., 2009). Some of the downstream regulators, *SND2*, *SND3*, and *MYB103* exclusively activate cellulose biosynthesis where as others such as *MYB63*, and *MYB58* regulate lignin biosynthesis (Zhong et al., 2008). Even though direct protein–DNA interactions have been shown for some of the cellulose and lignin specific regulators, further characterization of many downstream regulators is needed.

Discovery of the *trans*-acting transcription factors of cell wall biosynthesis facilitated the opportunity to identify common *cis*-elements shared among the master regulators. The tracheary-element-regulating *cis*-element (TERE) is one such 11-bp motif, CTT/(C)NAAA/(C)GCNA(T), involved in tissue specific cell wall biosynthesis and programmed cell death. First identified in the *Zinnia cysteine protease 4* promoter and it is present in numerous cell death and xylem differentiation genes such as *Cysteine protease 1* (*XCP1*), *XCP2, Serine protease 1*, and several other genes associated with wall function that include xylanases and acetyltransferases (Pyo et al., 2007). More recent studies demonstrated the physical interaction between VND6 protein and the TERE (Ohashi-Ito et al., 2010). Other SWNs bind a 19-bp imperfect palindromic sequence (T/A)NN(C/T)(T/C/G)TNNNNNNNA(A/C)GN(A/C/T)(A/T) referred to as the secondary wall NAC binding element (SNBE; (Zhong et al., 2010)). A synthetic promoter harboring six copies of the SNBE fused to a GUS reporter revealed specific

expression in xylem and interfascicular fibers, phenocopying native *MYB46* promoter behavior (Zhong et al., 2007). Other direct targets of SND1, including *MYB83*, *MYB103*, *SND3*, and *KNAT7* also posses the SNBE element (Zhong et al., 2010). A similar *cis*-element, TACNTTNNNNATGA, was identified recently in the *SND1* promoter and is the target of binding that serves as a target of positive feedback from SND1 itself (Wang et al., 2011).

A series of MYB transcription factor family proteins are also implicated in the cell wall regulatory network, a majority of which act downstream of the SWNs. One such protein, MYB46, acts subsequent the SWNs (Fig.1.2A). Lignin biosynthesis is specifically regulated by MYB63 and MYB58 interacting with the AC/Pal-box promoter sequences. This motif was first identified in the promoter of parsley *PHENYLALANINE AMMONIA-LYASE 1* and subsequently identified as three AC rich elements AC-I (ACCTACC), AC-II (ACCAACC), AC-III (ACCTAAC) involved in lignin gene regulation (Lois et al., 1989; Hatton et al., 1995; Raes et al., 2003). Binding of MYB proteins to the AC elements *trans*-activates the respective promoters thus, activating the genes in a xylem specific manner repressing the expression of the same genes in phloem or the cortical cells (Hatton et al., 1995). The consensus sequence of the AC element was recently expanded to include four more forms interchanging a T with the C at the last position; thus, ACC(T/A)A(A/C)(C/T) (Zhong and Ye, 2012). This 7-bp sequence, the secondary wall MYB responsive element (SMRE), is bound by both *MYB46* and *MYB83* proteins and is sufficient for the activation of a suite of transcription factor and cell wall biosynthetic genes (Zhong and Ye, 2012). The MYB46-responsive *cis*-element (M46RE) is an 8-bp sequence (A/G)(G/T)T(A/T)GGT(A/G) found in the *C3H14* promoter, which

is a direct target of MYB46 (Kim et al., 2012). *Trans*-activation assays coupled with EMSA reveled M46RE is required and sufficient for the activation of *C3H14*. The 8-bp core sequence was present in nearly 43% of the genes in the *A. thaliana* genome but was enriched in the downstream genes activated by MYB46 along with secondary cell wall related structural genes (Kim et al., 2012).

Apart from the key SWNs and MYBs, a handfull of downstream MYBs, NACs, and transcription factors from other families are involved in this complex cell wall regulatory network. A closely related homolog of *MYB32*, *MYB4*, functions as a repressor of *CINNAMATE-4-HYDROXYLASE* (C4H; (Jin et al., 2000). Loss-of-function mutants of *MYB4* exhibited elevated levels of sinapoyl malate, a component in the lignin pathway, and an increase in *C4H* expression. Collectively, these findings suggest a complex and hierarchical transcription regulatory network for eudicot cell wall biosynthesis (Fig. 1.2A). While this dissertation primarily discusses discoveries in *A. thaliana*, it should be noted that a number of regulators have been characterized in other eudicot species such as *Populus trichocarpa, Eucalyptus gunnii, Nicotiana tabacum, Antirrhinum majus, Pinus taeda, Vitis vinifera*, and *Medicago truncatula*.

1.3 Grass cell wall regulators

From the time of eudicot and monocot divergence 140–150 million year ago, the transcription factor families have disproportionately expanded and the regulatory networks have likely diverged; thus, the existing eudicot network model for transcription regulation is not wholly generalizable to monocots (Chaw et al., 2004; Shiu et al., 2005).Conversely, recent functional characterization of grass transcription factors implies

great commonality in how a similar network could regulate grass cell wall biosynthesis. While many cell wall genes have been characterized in grasses, almost nothing is known about the regulation of walls in monocots. The *A. thaliana* model, which is by far the best developed, is admittedly nascent. This stands in stark contrast to the model for grasses that consists of only a few genes (Fig. 1.2B). Maize *CAFFEIC ACID-O-METHYL TRANSFERASE* (*COMT*) is a key lignin pathway gene with an AC-III element recognized by R2R3-MYB transcription factors (Vignols et al., 1995; Fornalé et al., 2010). Group four R2R3-MYB transcription factors are described as repressors and based on sequence homology to the known *A. thaliana* MYB repressors, five maize transcription factors, ZmMYB31, ZmMYB42, ZmMYB2, ZmMYB8, and ZmMYB39 were identified as candidates for direct repression of *ZmCOMT* (Fornalé et al., 2006).When overexpressed in *A. thaliana*, *ZmMYB31* and *ZmMYB42* resulted in down regulation of lignin associated genes and subsequently reduced lignin content (Fornalé et al., 2010). Overexpression of *ZmMYB42* caused a reduction in leaf size, an adaxial curvature indicative of less tertiary vein formation, reduction of the syringyl lignin monomers, and dwarfism in *A. thaliana* (Sonbol et al., 2009).The absence of results in a monocot is likely due to the relative recalcitrance of crop species to genetic study; thus, *A. thaliana* serves as an imperfect heterologous system to study grass gene function. The maize gene, *ZmMYB31* is the first wall specific regulator characterized in a grass ((Fornalé et al., 2010). Chromatin immunoprecipitation demonstrated the direct interaction between ZmMYB31 and *ZmCOMT* promoter and an AC element similar to AC-II was identified as the binding motif. The switchgrass (*Panicum virgatum* L.) protein PvMYB4, an ortholog to *A. thaliana* MYB4, is another recently characterized

repressor of lignin (Shen et al., 2012). Ectopic expression of *PvMYB4* in switchgrass resulted in a reduction in total lignin and altered lignin monomer ratio (Shen et al., 2012). The AC element is also implicated as the binding site of PvMYB4, which results in repression of lignin pathway genes (Shen et al., 2012). In addition, rice and maize orthologs of *A. thaliana* SWNs and MYB46 were shown to activate secondary wall biosynthesis when overexpressed in *A. thaliana* (Zhong and Ye, 2012). Moreover, OsSWNs and ZmSWNs were able to complement and partially rescue the pendant stem phenotype of the *A. thaliana snd1*/*nst1* double mutant. Similarly, *OsMYB46* and *ZmMYB46* under the control of *AtMYB46* promoter were able to complement the loss of helical secondary wall thickening in vessels of *A. thaliana myb46/myb83* double mutant. In addition, SNBEs were identified in *OsMYB46* and *ZmMYB46* promoters and were bound and activated by the rice and maize SWNs in *A. thaliana* transient protoplast assays (Zhong and Ye, 2012). While studies in a heterologous system have proven informative, further functional characterization in a grass species is necessary. Contrary to expectations of distinct aspects of monocot cell wall regulation, the existing model is populated exclusively by homologs of known *A. thaliana* genes.

Gene expression profiling was critical in identifying many of the candidates that formed the foundation of the existing *A. thaliana* cell wall regulatory network (Oh et al., 2003; Ehlting et al., 2005; Kubo et al., 2005; Zhao et al., 2005). Similar tools can be applied to grasses to identify candidates for functional characterization, especially those specific to monocots. Accordingly, a comparison of the transcriptome of maize elongating and non-elongating internodes revealed several transcription factors that are feasibly involved in cell wall regulation (Bosch et al., 2011). Likewise, the expression of

three barley (*Hordeum vulgare* L.) NACs, *HvNAC033*, *HvNAC034*, and *HvNAC039*, were significantly greater in stem tissue where extensive secondary cell wall biosynthesis occurs. One of these proteins, HvNAC033, is the closest barley homolog to *A. thaliana* NST1, further supporting the possible role of this protein in grass cell wall regulation (Christiansen et al., 2011). Comparison of expression networks across species similarly reinforces the notion of shared features across eudicots and monocots and revealed potential distinctions. Genes with expression patterns similar to secondary wall *CELLULOSE SYNTHASE A* genes in rice included those most similar to *A. thaliana MYB63*, *MYB103*, *NST1*, *SND2*, and *KNAT7* (Ruprecht et al., 2011). Conversely, several rice transcription factors were co-expressed with structural genes that do not share sequence homology to co-expressed or characterized *A. thaliana* genes. These include MYB and NAC as well as bZip and AP2 family genes (Ruprecht et al., 2011). As with *ZmMYB42* and *PvMYB4*, candidates identified using sequence similarity and expression profiling require further characterization in the native systems in order to solidify both overlap and divergence between eudicot and grass cell wall regulatory networks.

1.4 *Brachypodium distachyon:* a new model system for grasses

The grass family *Poaceae* is the fourth largest in the plant kingdom with more than 10,000 species distributed around the world including many of the most important agricultural commodities: maize, rice, wheat and sugarcane (Kellogg, 2001). Species currently under investigation as bioenergy crops including grasses like *Miscanthus spp*. and switchgrass, are also part of the *Poaceae* (Rooney et al., 2007; Schmer et al., 2008). *Arabidopsis thaliana* has served as the most advanced plant model system for over two

decades (Somerville and Koornneef, 2002). Although, it is an ideal species for genetic analysis of dicots, it is not the best model for monocots due to their fundamental differences with dicots (Brkljacic et al., 2011). For instance, dicots have two cotyledons and form a taproot whereas monocots only contain a single cotyledon and form lateral roots. Other anatomical distinctions exist in their vascular bundles, which are cell types particularly relevant to biofuel feedstock accumulation. In general, dicots contain a cambium layer within the vascular bundle, in between the phloem and xylem and the vascular bundles are arranged in a circle along the periphery of the stem. On the other hand, monocots do not have a cambium. They also have xylem and phloem arranged in vascular bundles and these bundles can be arranged in several circles around the periphery of a stem, as in rice and *B. distachyon*, or scattered throughout the stem as in maize. Due to the significant differences between monocots and dicots, a grass species would be the best model system to study traits that will be important for monocots.

Rice serves as a model system for monocots. However, unlike many potential energy crop species it is not a temperate grass and it lacks small stature and a rapid life cycle like *A. thaliana*. In addition, as a semi-aquatic tropical grass, with specially evolved mechanisms for nutrient uptake, its growth conditions can be unnecessarily demanding for laboratory research. A smaller, non-food crop like *B. distachyon*, which is more closely related to cereals and temperate grasses (Kellogg, 2001) would be a better model to represent temperate grasses.

Brachypodium distachyon is a small annual grass native to the Middle East and Mediterranean (Garvin et al., 2008). It has a fully sequenced genome (272 Mbp) slightly larger than *A. thaliana* with similar growth requirements and generation time (8- 12

weeks) (Initiative, 2010). This grass belongs to the tribe Brachypoieae which only contains the genus *Brachypodium* (Bevan et al., 2010). Currently various laboratories are developing *B. distachyon* resources for forward and reverse genetics approaches. Over 10,000 T-DNA lines are currently available (http://www.brachytag.org/known-genes.htm, http://brachypodium.pw.usda.gov/). An efficient *Agrobacterium* mediated transformation protocol (Vogel et al., 2006; Vogel and Hill, 2008), efficient crossing techniques (http://brachypodium.pw.usda.gov/), microarrays and web resources (http://www.brachytag.org/, http://www.phytozome.net/, http://www.modelcrop.org/) are now available for this new model system. With the aid of these resources, *B. distachyon* is now being studied to understand a variety of grass specific questions. Cell wall composition analysis between select dicots and monocots demonstrates a close resemblance between *B. distachyon* non-cellulosic monosaccharides and that of other grasses like wheat, barley and *Miscanthus spp*. (Gomez et al., 2008; Rancour et al., 2012). This similarity is important for cell wall research as it provides evidence for the suitability of *B. distachyon* as a model for grass cell wall research. Current studies on *B. distachyon* along with the genetic attributes of this small grass emphasize its suitability as a model for biomass crop research.

My objective in this dissertation is to uncover the transcriptional regulators of grass secondary cell wall biosynthesis. I hypothesize grass secondary cell wall biosynthesis is regulated by a complex network of transcription factors. This regulatory network will include unique grass proteins, which will define the features specific to monocot cell walls. I hypothesize transcription factors highly expressed in specific tissues rich in secondary cell walls will have important regulatory roles. I have used a reverse

genetics approach to characterize two cell wall genes and two transcription factors associated with grass secondary cell wall biosynthesis. Together my findings reveal a grass specific transcriptional activator capable of regulating cell walls and feedstock attributes and a repressor with pleiotropic regulatory roles linking cell wall biosynthesis and flowering.

CHAPTER 2

PERTURBATION OF *Brachypodium distachyon CELLULOSE SYNTHASE A4* OR *7* RESULTS IN ABNORMAL CELL WALLS

2.1 Introduction

With continued consumption of fossil fuels, humankind faces a growing challenge of finding renewable sources of energy. Although plant derived biomass contains appreciable energy, to serve as a fuel for transportation it must be chemically or biologically liquefied, a conversion that is the subject of strenuous efforts to make economical. Success will presumably require not only advances in chemistry such as better catalysts, but also improved plants to serve as feedstock. As an ideal input, grasses are receiving considerable attention because certain species grow to great density, are perennial, and can require little if any fertilizer or irrigation (Heaton et al., 2008). But their very size, longevity, and genome complexity makes these species difficult subjects to study and breed.

For studying grasses, whether as sources for biofuel or for any other grass-specific question, an emerging model is *Brachypodium distachyon.* Of particular note for biofuels research, *B. distachyon* has cell walls that are similar compositionally to that of other grasses, like wheat (*Tritium aestavum*), barley (*Hordeum vulgare*), and *Miscanthus* (Gomez et al., 2008; Christensen et al., 2010; Guillon et al., 2011). Insofar as the cell wall constitutes almost the entirety of the input for converting biomass to biofuel, this similarity, along with the genetic attributes of this small grass, emphasize its suitability as a model for grass-related, biomass crop research.

Within the cell wall, as a target for optimization, cellulose is pre-eminent. Cellulose is the most abundant of any single wall component and is made exclusively of glucose, a tractable and energy rich molecule. Cellulose comprises long polymers of (1-4) β-linked glucose that are synthesized at the plasma membrane and associate laterally into a microfibril. Because of the configuration of the glucose residues, hydrogen bonds form at great density both within and between chains, a density that gives cellulose an elastic modulus rivaling that of steel but makes the structure impervious to degradation whether chemical or enzymatic.

Within the plasma membrane, the structure synthesizing cellulose is called a "terminal complex" (Haigler and Brown, 1986; Kimura et al., 1999; Somerville, 2006). In land plants and related green algae, the terminal complex as seen in the electron microscope comprises six subunits with hexagonal symmetry and is called a "rosette" (Doblin et al., 2002). The primary constituents of the rosette are CELLULOSE SYNTHASE A (CESA) proteins. These proteins belong to processive glycosyltransferase family 2 and are thought to be the catalytic subunits for polymerizing the glucose chain. In angiosperms, CESAs usually comprise a small gene family with around ten members (Yin et al., 2009; Carroll and Specht, 2011).

Identification of CESA proteins and characterization of their function has greatly benefited from the facile genetics of *A. thaliana*. From this work, it emerged that certain CESA proteins synthesize the primary cell wall whereas others synthesize the secondary cell wall (Desprez et al., 2007; Persson et al., 2007). Furthermore, it appeared that a given cell must express three distinct CESA proteins to produce cellulose at optimal levels. For the secondary cell wall, a screen based on collapsed xylem cells led to the identification

of several *irregular xylem* (*irx*) lines, three of which, *irx5*, *irx3*, and *irx1*, harbor lesions in AtCESA4, 7, and 8, respectively (Taylor et al., 1999; Taylor et al., 2000; Taylor et al., 2003). Supporting the hypothesis of non-redundancy, these genes are expressed at similar levels in similar cell types, and the null mutants have indistinguishable phenotypes, including weak stems, collapsed xylem, and thin secondary cell walls that are deficient in cellulose.

Identification of a trio of CESA genes primarily responsible for synthesizing cellulose in secondary cell walls has been supported by work in other systems, including grasses. First, AtCESA4, 7, and 8 are usually represented in most other angiosperm species by a single sequence each, and orthologs are more closely related than homologs (i.e., CESA4s of various species resemble each other more closely than do CESA4, 7, and 8 of a single species) (Song et al., 2010; Carroll and Specht, 2011; Li et al., 2013). Second, in grasses, a mutant of barley, *fragile stem 2*, with brittle stems and low cellulose content in the mature plant was attributed to a lesion in the barley ortholog of AtCESA4 (Burton et al., 2010). In addition, "brittle culm" mutants in rice (*Oryza sativa*) have been mapped to the three orthologs of AtCEA4, 7 and 8 and again these mutants have similar phenotypes, including modest dwarfism, thinner and weaker culms, and reduced cellulose content (Tanaka et al., 2003). However, it is not understood why three distinct proteins are needed nor is it known which of the loss-of-function phenotypes result directly from the lost protein and which result as a consequence of cumulative effects.

Here we describe the *CESA* gene family in *B. distachyon* with the aid of gene expression profiling and phylogeny. Furthermore a detailed analysis of the candidate

secondary CESAs were performed by functionally characterizing mutants generated with specific artificial microRNA constructs.

2.2 Materials and Methods

2.2.1 Plant material and growth

Brachypodium distachyon (L.) line Bd 21-3 was used throughout. Seeds were imbibed in moist paper towels for seven days at 6˚C, planted on potting mix (#2; Conrad Fafard Inc. Agawam, MA), and grown in a growth chamber at 20˚C with 20 h light: 4 h dark, at a fluence rate of 220 $\mu mol \cdot m^{-2} \cdot s^{-1}$ and relative humidity of ~68%. For plate-grown plants, seeds were de-hulled and imbibed in water for 2 h with shaking. Then, seeds were treated with 70% ethanol for 20 s, rinsed with sterile water, and soaked in 1.3% NaClO for 4 min at room temperature while shaking. Seeds were subsequently rinsed three times with sterile water and stored in the dark at 4˚C for a minimum of 2 days in a sterile Petri dish with filter paper. Seedlings were grown for seven days on 0.5 strength MS medium adjusted to a pH of 5.8 with KOH and containing 0.7% agar (Difco "Bacto agar").

2.2.2 Identification of *CELLULOSE SYNTHASE A* genes

Complete, translated amino acid sequences of the ten *A. thaliana* CESA proteins and members of the cellulose synthase-like families were used as queries in BLASTP of Phytozome v8.0 and NCBI databases to identify homologous *B. distachyon* CESA genes and identified proteins were named according to the *A. thaliana* counterparts, where possible. Multiple sequence alignments were performed using ClustalW program and a phylogenetic trees were generated by the neighbor-joining method using MEGA5 software with 1000 bootstrap permutations (Tamura et al., 2011).

2.2.3 Measurements of transcript abundance

For microarray analysis, different growth regimes were used for leaves and stems, versus roots. For leaves and stems, approximately 30 days following germination and growth on soil, total leaf and stem were collected when the inflorescence began to emerge from the flag leaf. Leaves were separated from the stems with a curved-tip probe. Nodes and internodes from the second leaf junction to the internode below the inflorescence were frozen in liquid nitrogen. For roots, seven-day-old whole seedlings were flash frozen in liquid nitrogen and then the roots were snapped off into a sterile culture tube. For all organs, material was harvested six times during the day (2, 6, 10, 14, 18, and 22 h circadian time). Three plants were dissected for each time point and in triplicate for each tissue type. Samples were stored in liquid nitrogen or at -80˚C until RNA extraction. Tissue was ground with mortar and pestle in liquid nitrogen. RNA was extracted using a kit (Plant RNaeasy, Qiagen, Valencia, CA) according to the manufacturer's instructions. For hybridization, cDNA probes were synthesized using a kit (WT Ambion Santa Clara, CA).

The probes were applied to the *B. distachyon* BradiAR1b520742 whole genome tiling array (Affymetrix, Santa Clara, CA). The array contains ~6.5 million unique 25-mer oligonucleotide features, both the forward and reverse strand sequence. The complete genome sequence is tiled with an average of 30 bases between each array feature; 1.6 million features correspond to exons and introns and 4.9 million features between gene models (Todd Mockler, Donald Danforth Plant Science Center, personal communication). Approximately ~95% (~26,670) of the genes have at least five corresponding exon array features and from those a summary value was calculated for

each gene model Probeset values were calculated using gcRMA (Wu et al., 2004).

For obtaining RNA from the transgenics, stems were collected at the same developmental stage when the inflorescence was just visible from the flag leaf. First, second, and third nodes and internodes of the tallest stem were frozen in liquid nitrogen, homogenized, and RNA extracted as described above.

For RT-QPCR, on-column DNA digestion was performed using RNase-free DNase I (Qiagen). First strand cDNA was synthesized from 1 μg of total RNA using Superscript III reverse transcriptase with oligo dT primers (Invitrogen, Grand Island, NY). Samples were diluted three-fold with RNase free water (Qiagen) and 1 μL from each cDNA sample was used for RT-QPCR to check for genomic DNA contamination using GapDH primers. Triplicate quantitative PCR reactions were performed using 20 μL reaction volumes with 1 μL of diluted cDNA in each reaction with the QuantiFast SYBR Green PCR Kit (Qiagen). The reactions were conducted in an Eppendorf Realplex^2 Mastercycler using the following conditions: 95°C for 2 min, followed by 40 cycles of 95°C for 15 s, 60°C for 15 s and 68˚C for 20 s. As reference genes for normalization, *BdUBC18* (ubiquitin-conjugating enzyme 18) and *Bd5g25870* (Belonging to nuclear hormone receptor binding category) were used (Hong et al., 2008). QuantiPrime primer design tool was used for qPCR primer design (Arvidsson et al., 2008).

2.2.4 RNA *in situ* hybridization

RNA *in situ* hybridization was performed using methods described previously (Harding et al., 2002; Kao et al., 2002). Briefly, the 3’ end of *BdCESA4* and *BdCESA7* were cloned into the pGEM-T Easy vector and used as the template to generate labeled sense and anti-sense ribo probes using a kit (DIG labeling kit, Roche, Indianapolis, IN).

Three week old stem sections were frozen at -80˚C and fixed in 4% paraformaldehyde/ethanol:acetic acid (3:1) overnight at 4˚C (Harrington et al., 2007). Fixed tissue was encased in 4% agarose and sectioned using a Vibratome. Sections were removed from agarose using forceps and washed in phosphate-buffered saline (PBS; 33 mM Na_2HPO_4, 1.8 mM NaH_2PO_4 and 140 mM NaCl, pH 7.2) and post-fixed in PBS containing 3.7% (v/v) formaldehyde and 2 mg/mL glycine for 20 min at room temperature. Sections were washed in PBS and dehydrated in graded ethanol series and pre-hybridized for 1 h at 65˚C with hybridization solution comprising 20X SSC (3M NaCl, 0.3M Na citrate) containing 20% SDS, 3.7% formamide, and 10 mg/mL yeast tRNA made up in DEPC-treated water. Stems were then hybridized with DIG-labeled sense and anti-sense probes overnight at 65˚C and washed with a series of SSC and SDS containing solutions. Anti-DIG alkaline phosphatase fab fragments (Roche) were used at a dilution of 1:1000 and incubated over night at 4˚C. Alkaline phosphatase was detected using nitroblue tetrazolium and 5-bromo-4-chloro-3-indoyl-phosphate (Roche) in 0.1 M Tris (pH 9.5), 0.1 M NaCl and imaged with a PixeLINK camera attached to a Nikon eclipse 200 microscope.

2.2.5 Artificial microRNA constructs

Artificial microRNA sequences were designed on the Web MicroRNA Designer platform (http://wmd3.weigelworld.org) based on JGI *B. distachyon* genome annotation version 1.0 (Initiative, 2010). The *amiR-CESA4* construct targets AAGGGACCCATTCTTAAGCCA with hybridization energy of -37.74 kcal/mol and the *amiR-CESA7* construct targets ACGCCCACCATTGTCATCATC with hybridization energy of -37.37 kcal/mol. Constructs were engineered from the pNW55 plasmid to

replace the targeting regions of the native rice microRNA precursor *osaMIR528* (Warthmann et al., 2008). MicroRNA targets were PCR amplified (Table S2) according to Warthmann et al. (Warthmann et al., 2008) and cloned into pENTR/D-TOPO (Invitrogen). Sequence confirmed clones were recombined with a modified version of the destination vector pOL001 (Vogel et al., 2006), pOL001-ubigate-ori1 and transformed into *Agrobaterium tumefaciens* strain *AGL1* via electroporation.

2.2.6 Plant transformation

Transformation was carried out according to Vogel et al. (Vogel and Hill, 2008). Briefly, seeds were collected from six to seven week old plants and deglumed. Seeds were surface sterilized with a 1.3% NaClO solution containing 0.01% Triton-X100 for 4 min. Embryos were dissected and placed on callus initiation medium under aseptic conditions. Calli were propagated for seven weeks with two subsequent subcultures at four and six weeks following dissection. Seven-week-old calli were immersed in an *A. tumefaciens* suspension for 5 min and dried on filter paper. Next, they were co-cultivated on dry filter paper for three days at 22˚C in the dark. Following co-cultivation, calli were moved to selective plates containing 40 mg/L hygromycin and 200 mg/L timentin for four weeks in the dark at 28˚C. Following selection, calli were moved to Linsmaier and Skoog media for regeneration at 28˚C under constant light and next onto Murashige and Skoog media for root establishment under same conditions. Next they were transplanted to soil and grown as described above.

2.2.7 Genomic DNA extraction and genotyping

Genomic DNA was extracted from leaves according to Csaikl et al. (Csaikl et al., 1998) with slight modification. Briefly, leaves were frozen in liquid nitrogen and ground

using 3.2 mm diameter stainless steel metal balls (Biospec Products, Bartlesville, OK) in a ball mill (Mixer Mill, MM400, Retsch, Newtown, PA). Ground samples were incubated in DNA extraction buffer (100 mM NaCl, 50 mM Tris, 25 mM EDTA, 1% SDS, 10 mM 2-mercaptoethanol) for 10 min at 65°C. Next, they were mixed with 5 M potassium acetate and incubated on ice for 20 min and centrifuged for 10 min. DNA was pelleted by mixing the supernatant with isopropanol and centrifuging at maximum speed for 10 min followed by a 70% ethanol wash. Pelleted DNA was resuspended in 1X TE and the integrity of the samples was measured using a spectrophotometer (NanoDrop 1000, Themo Scientific, Waltham, MA). Genotyping was carried out by PCR for the Hgyromycin locus using the extracted genomic DNA as template with following cycler conditions; Initial denaturation at 98˚C for 2 min, flowed by 30 cycles of 98˚C for 30 s, 59˚C for 30 s, 72˚C for 55 s and a final extension at 72˚C for 7 min. PCR confirmed positive transformants were used in subsequent experiments.

2.2.8 Light microscopy

For histochemical analysis, stems were hand sectioned using a razor blade and stained with 0.002% Toluidine blue in water for 30 s. Stained sections were mounted with water and observed under an Eclipse E200MV R microscope (Nikon) and imaged using a PixeLINK 3 MP camera. Images were captured at 4 X magnification and stem area was measured by freehand tracing of a perimeter in ImageJ (http://rsb.info.nih.gov/ij/). Images captured at 20 X magnification were used for cell wall thickness measurements.

For polarized-light microscopy, internode segments were fixed in 2% glutaraldehyde in 50 mM Na_2PO_4 buffer (pH 7.2) for a minimum of 2 h at room

temperature and then handled and embedded in Spurr's resin by standard techniques. Semi-thin sections (0.5 µm) were cut on an ultra-microtome, mounted in immersion oil, and observed through a microscope equipped with the LC-PolScope (CRI Cambridge MA) as described previously (Baskin et al., 2004). Briefly, this instrument uses circularly polarized light and computer-controlled liquid crystal compensators to obtain four images with known compensator settings and from them calculates a fifth image in which the intensity of each pixel is proportional to birefringent retardation and a sixth image (not used here) in which the intensity of each pixel represents the orientation of the crystal's optical axis within the specimen plane (Oldenbourg and Mei, 1995). Several sections of each genotype were observed and the images shown are representative.

2.2.9 X-ray diffraction profiles and sum-frequency-generation vibration spectroscopy

X-ray diffraction and calculation of a crystallinity index were done as described by Ruland et al. (Ruland, 1961) with slight modifications. Fully senesced stems were ground as described above for genomic DNA extraction. Diffraction was analyzed on an X'Pert Pro powder X-ray diffractometer (PANanalytical BV, The Netherlands) operated at 45 kV and 45 mA using CuKα radiation at both Kα1 (λ = 1.5406 Å) and Kα2 (λ = 1.5444 Å). The diffraction profile was acquired from 5 to 50˚ in 0.0167 steps, with 66 s per step. The crystallinity index was calculated using the amorphous subtraction method, which determines crystallinity by subtracting the amorphous contribution from the diffraction profile obtained from xylan (Aldrich) measured in parallel. A scale factor was applied to the xylan profile to avoid negative values in the subtracted profiles. For each group, eight to twelve individuals were analyzed. Sum-frequency vibration spectroscopy was done as previously described (Barnette et al., 2011). Intact first internodes were

excised from completely senesced plants, and each group contained four to six individuals with ten measurements per stem. The amount of crystalline cellulose was estimated by comparing the 2944 cm^{-1} intensity of the sample with that of Avicel, as previously described (Barnette et al., 2012).

2.2.10 Statistical analysis

For each measurement, three to twelve independent plants were sampled from three or four different T_3 families for each transgene. Student's *t*-tests were performed in R (v 2.15.0). Significance was set a $P < 0.05$. No significant differences were observed among the different independent transgenic lines for a given construct and were thus pooled.

2.3 Results

2.3.1 *Brachypodium distachyon CESA* gene family

In *B. distachyon,* as in *A. thaliana* and rice, the CESA family comprises ten genes (Fig. 2.1). Their genomic sequences range from 3045 to 7601 bp, with 5 to 14 exons that form coding regions ranging from 2331 to 3279 bp (776 to 1092 amino acids). Amino acid sequence comparisons revealed extensive similarity among the BdCESA proteins, with conserved structural features that are characteristic of CESA protein families (Yin et al., 2009; Carroll and Specht, 2011). All ten BdCESA proteins contain eight transmembrane domains and two hyper-variable regions. With the exception of BdCESA10, they have the signature motif of glucosyl transferases, D,D,D,QxxRW, which is essential for binding UDP-Glucose. With the exception of BdCESA5 and 10, the proteins contain the RING-type zinc finger domain with eight cysteine residues spaced as

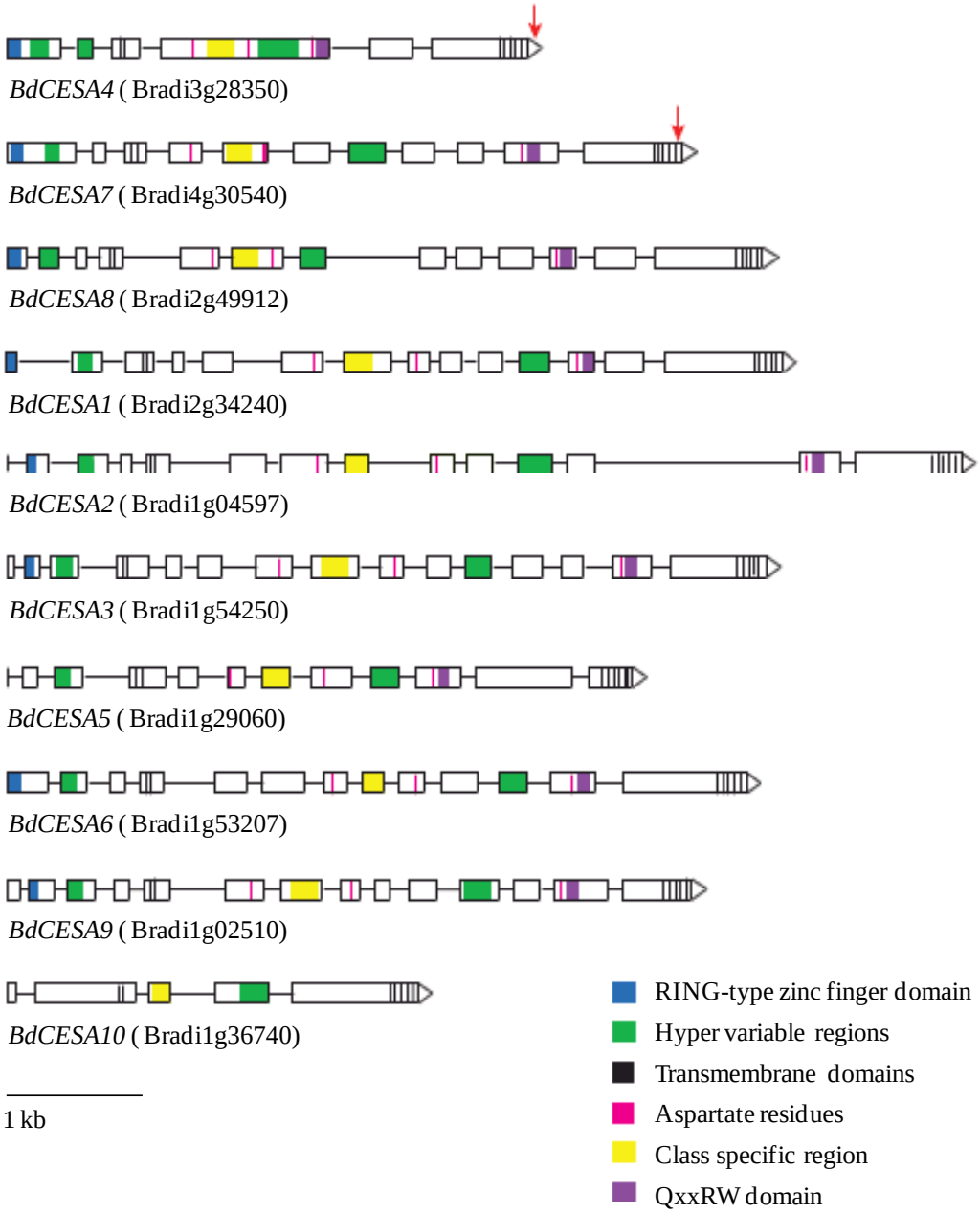

Figure 2.1. Models of *Brachypodium distachyon CESA* genes. Exons are indicated by boxes and introns by lines. Genes are drawn to scale; the bar in the lower left indicates 1 kb. Red arrows indicate regions used for artificial microRNA targeting.

characteristic for CESA proteins. Even though BdCESA10 is categorized as a CESA, it is short, and lacks the RING-type zinc finger motif, a portion of the first hyper-variable region, and two of the conserved aspartate residues of the QxxRW motif. This is also true of OsCESA11 and sorghum (*Sorghum bicolor*) Sb10g023430, both of which are nonetheless considered part of the CESA family (Wang et al., 2010).

Based on the amino acid sequence, the ten BdCESA proteins fall into reasonably well-established phylogenetic groups (Fig. 2.2). This is clearest for proteins associated with the secondary cell wall. *Brachypodium distachyon* has three sequences that are highly similar to those characterized in other species as secondary cell wall CESAs. We numbered these genes *BdCESA4*, *BdCESA7*, and *BdCESA8* based on their apparent orthologs in *A. thaliana*. Note that the published numbering differs for rice. While the three secondary CESA clades have a single member in each species, there is no complete one-to-one relationship for the CESAs associated with the primary cell wall. For example, *B. distachyon* has a single sequence in the CESA1 clade whereas *A. thaliana* has two. In contrast, *B. distachyon* has two sequences in the CESA3 clade whereas *A. thaliana* has only one. Interestingly, both *B. distachyon* and *A. thaliana* have several members of the CESA6 clade but these duplications appear to have formed in each lineage after the divergence of the two species. Based on sequence and conserved domains BdCESA10 seems to be the least similar to other CESA proteins. However a comprehensive phylogenetic analysis of the Cellulose synthase A proteins and Cellulose Synthase like proteins of *A. thaliana*, rice and *B. distachyon* revealed the distinct similarity of BdCESA10 to other CESA proteins (Fig. 2.3). Where possible, we numbered the *B. distachyon* gene after its closest relative in *A. thaliana*.

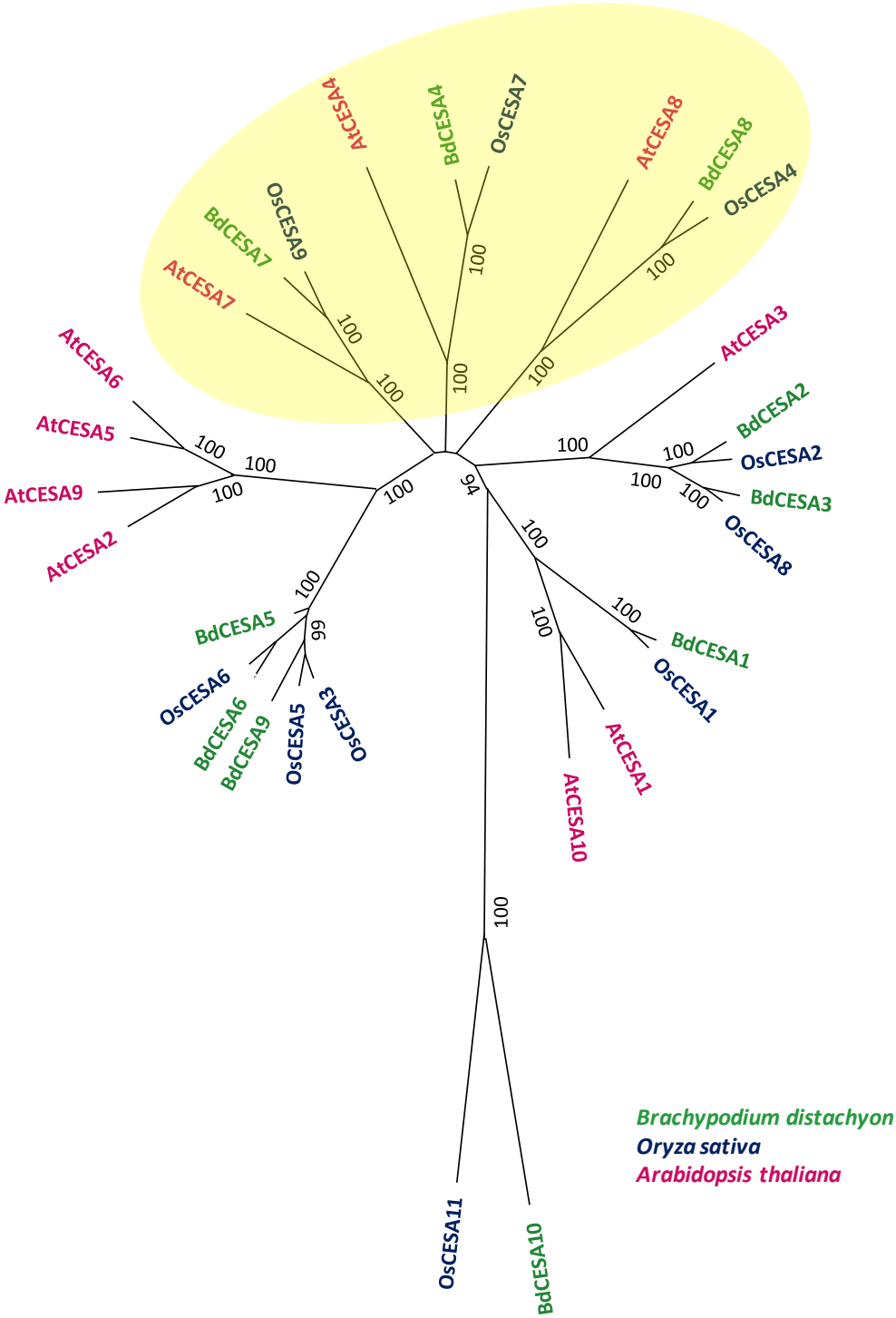

Figure 2.2. Phylogenetic analysis of CESA amino acid sequences. Numerical values on branches refer to neighbor-joining bootstrap support. Yellow oval denotes proteins associated with secondary cell walls.

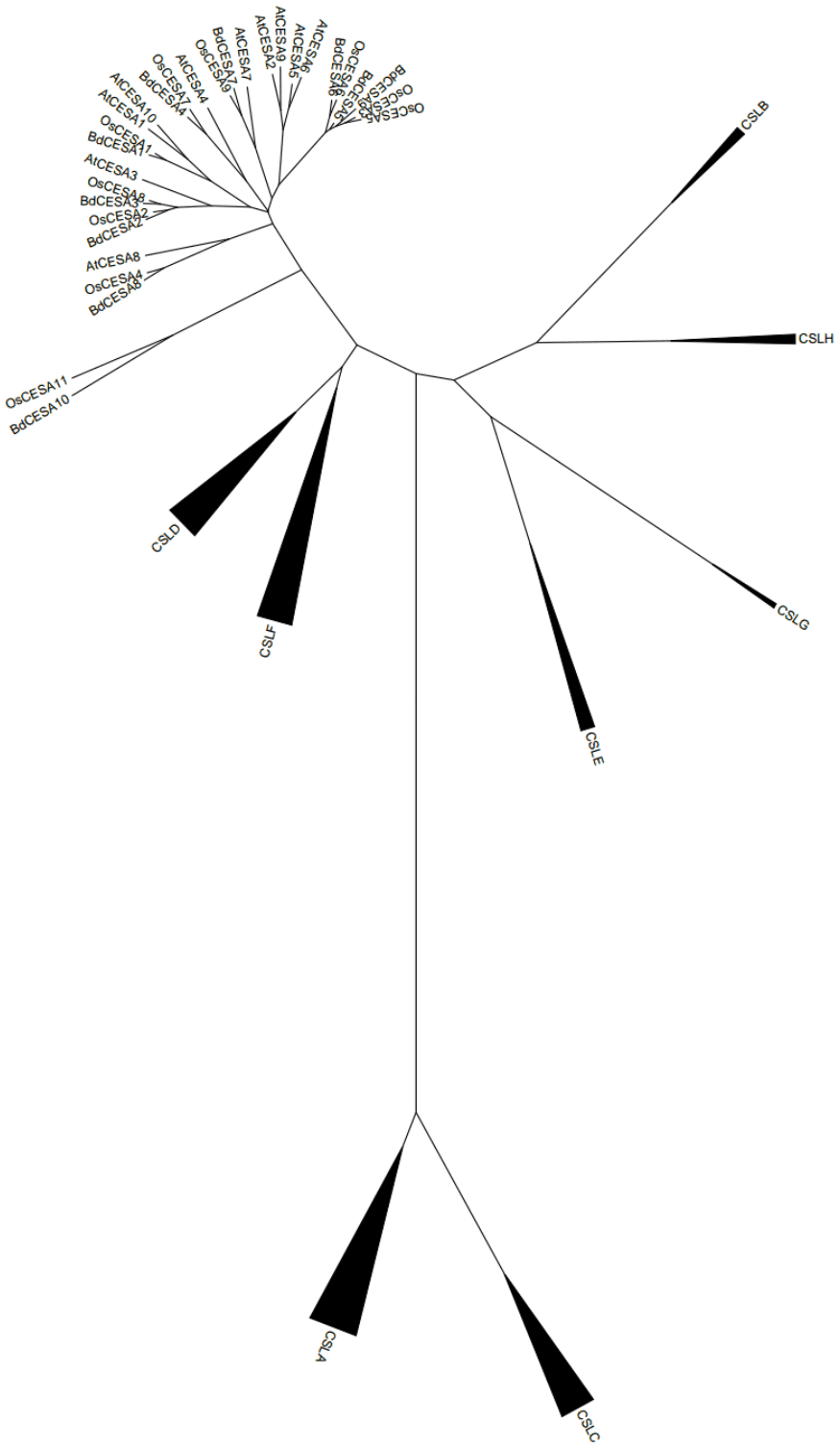

Figure 2.3. Phylogenetic analysis of *A. thaliana*, *B. distachyon* and rice CESA superfamily amino acid sequences. A consensus phylogeny was constructed with the neighbor-joining method with 1000 bootstrap permutations. The CESA clade is illustrates as an expanded sub-tree and the CSL clades are illustrated as condensed sub-trees.

In view of our interest in the secondary cell wall, we examined the *B. distachyon* secondary cell wall sequences at greater depth (Fig. 2.4). In the cysteine rich RING-type zinc finger domain, the sequences differ in the spacing between the second and third cysteine. BdCESA4 has the canonical 15 amino acids whereas BdCESA7 has a single amino acid insertion and BdCESA8 has an eight amino acid deletion. Such spacing variations have apparently not been previously reported for *A. thaliana* secondary CESAs. However, the rice secondary CESA homologous to BdCESA8 exhibits the same eight amino acid deletion (Tanaka et al., 2003). The RING-type zinc finger domain is the distinguishing feature of CESA proteins seen in the first portion of the N-terminus and is thought to be involved in CESA protein dimerization (Haigler and Brown, 1986; Kimura et al., 1999).

2.3.2 *Brachypodium distachyon* secondary cell wall *CESA* gene expression

To analyze *CESA* gene expression, we profiled transcripts with a whole genome tiling array focusing on organs expected to be enriched in secondary cell wall synthesis. To obtain RNA, leaves and stems were harvested when the inflorescence emerged from the flag leaf, whereas roots were harvested from seven-day-old seedlings. Additionally, to minimize changes in transcript abundance due to diurnal and circadian rhythms, RNA was pooled from material harvested at six different circadian time points over a 24-hour period. In all three organs, *BdCESA2* and *10* were expressed at essentially background levels and *BdCESA5* was slightly greater (Fig. 2.5). The four genes associated with the primary wall, *BdCESA1, 3, 6,* and *9,* were expressed at high levels in both root and stem. On the other hand, those associated with the secondary wall were expressed at much lower levels in roots than stems, having a ratio of expression roughly consistent with

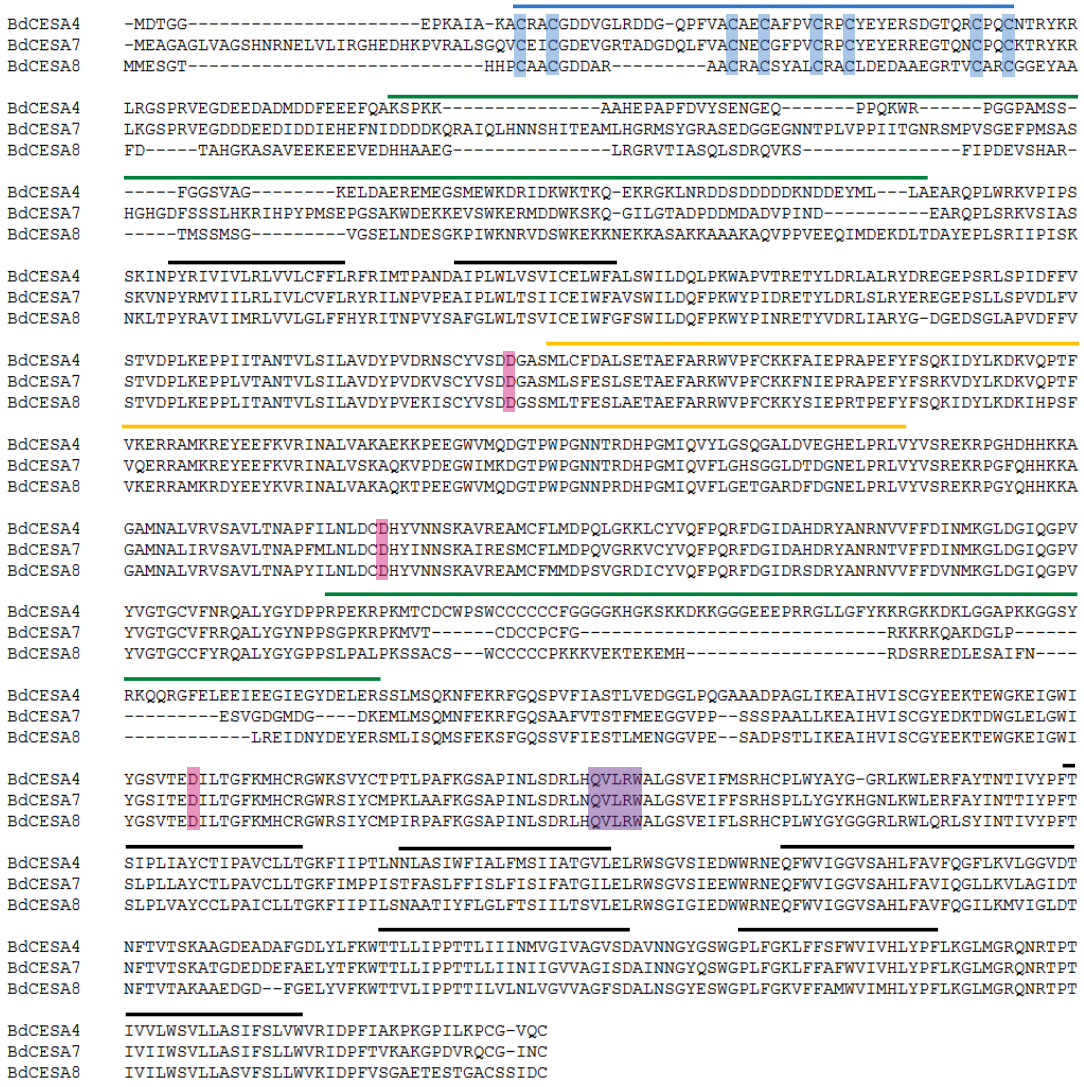

Figure 2.4. Alignment of the *B. distachyon* three secondary cell wall CESA amino acid sequences. The overhead blue line shows the RING-type zinc finger motif, with cysteine residues highlighted in blue. The overhead green lines show the two hypervariable regions. Pink highlights the three putatively catalytic aspartate residues and purple highlights the QxxRW motif. The overhead yellow line shows the plant-specific region and overhead black lines show the eight putative transmembrane domains.

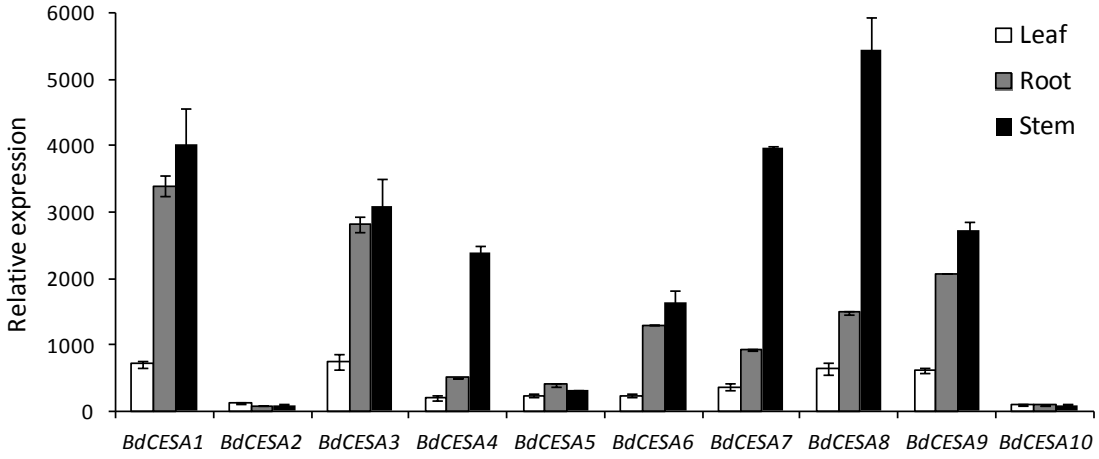

Figure 2.5. Relative abundance of *CESA* transcripts in different organs measured with a microarray. Bars plot mean ± standard deviation of three biological replicates.

Table 2.1. The ratios of relative transcripts abundance of the *CESA*s in root (R), leaf (L) and stem (S) tissue.

Gene	R/L	S/L	S/R
BdCESA1	4.8	5.6	1.2
BdCESA2	0.7	0.8	1.1
BdCESA3	3.8	4.2	1.1
BdCESA4	2.6	12.5	4.8
BdCESA5	1.7	1.4	0.8
BdCESA6	5.5	6.9	1.3
BdCESA7	2.6	11.0	4.3
BdCESA8	2.3	8.6	3.7
BdCESA9	3.4	4.4	1.3
BdCESA10	0.9	0.9	1.0

secondary cell wall content across the three organ types (Table 2.1).

2.3.3 Localization of putative secondary cell wall *CESA* transcripts

To localize transcripts, we performed RNA *in situ* hybridization on stems (Fig. 2.6). Stems were fixed, sectioned on a Vibratome, and hybridized with labeled sense and anti-sense probes. Upon color development, positive hybridization was detected mainly in the sections probed with the anti-sense probes. Consistent with a role in secondary cell wall synthesis, hybridization for each probe was strong in the vascular bundles and the surrounding mechanical cells including the sclerenchyma fibers and the epidermis. Hybridization was essentially undetectable in pith parenchyma, which undergoes limited cell wall thickening. These results strengthen the assignment of *BdCESA4* and *BdCESA7* as secondary cell wall related CESAs.

2.3.4 Artificial microRNAs targeting *BdCESA4* and *BdCESA7*

To examine the function of BdCESA4 and 7, we sought to reduce transcript levels by means of artificial microRNAs (*amiR*). The rice microRNA, *osaMIR528*, was modified to specifically target either *BdCESA4* or *BdCESA7* (Fig. 2.7 A, B). Each modified microRNA is predicted to target only a single gene. The *amiR-CESA4* construct specifically targets nucleotides 3106 to 3126, which are in the last exon immediately after the last transmembrane domain; in contrast, *amiR-CESA7* targets nucleotides 3028 to 3048, which are in the last exon in between the seventh and the eighth transmembrane domains (Fig. 2.1). To characterize the efficacy of these constructs, we measured mRNA levels in the stem of T_3 generation plants. For each construct, three to five individuals from families derived from three or four independent transformation events were analyzed by reverse transcriptase quantitative PCR (RT-QPCR). Stems were harvested

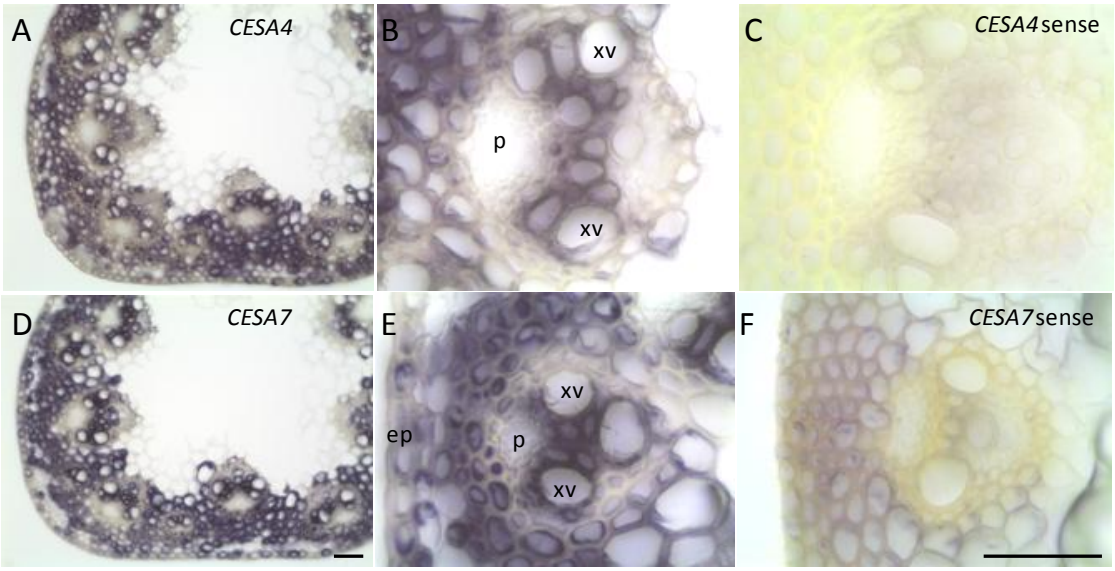

Figure 2.6. *BdCESA4* and *BdCESA7* are expressed in cells undergoing secondary wall deposition in the stem. Expression of *CESA4* (A, B) and *CESA7* (D, E) analyzed by *in situ* hybridization at three weeks of development when the inflorescence was just emerging from the flag leaf. Cross sections through the first internode were labeled with anti-sense probes, imaged at 10x (A, D) and 40x (B, E); and sense probes, imaged at 40x (C, F). xv, xylem vessel; p, phloem; ep, epidermis; Scale bar = 50 µm.

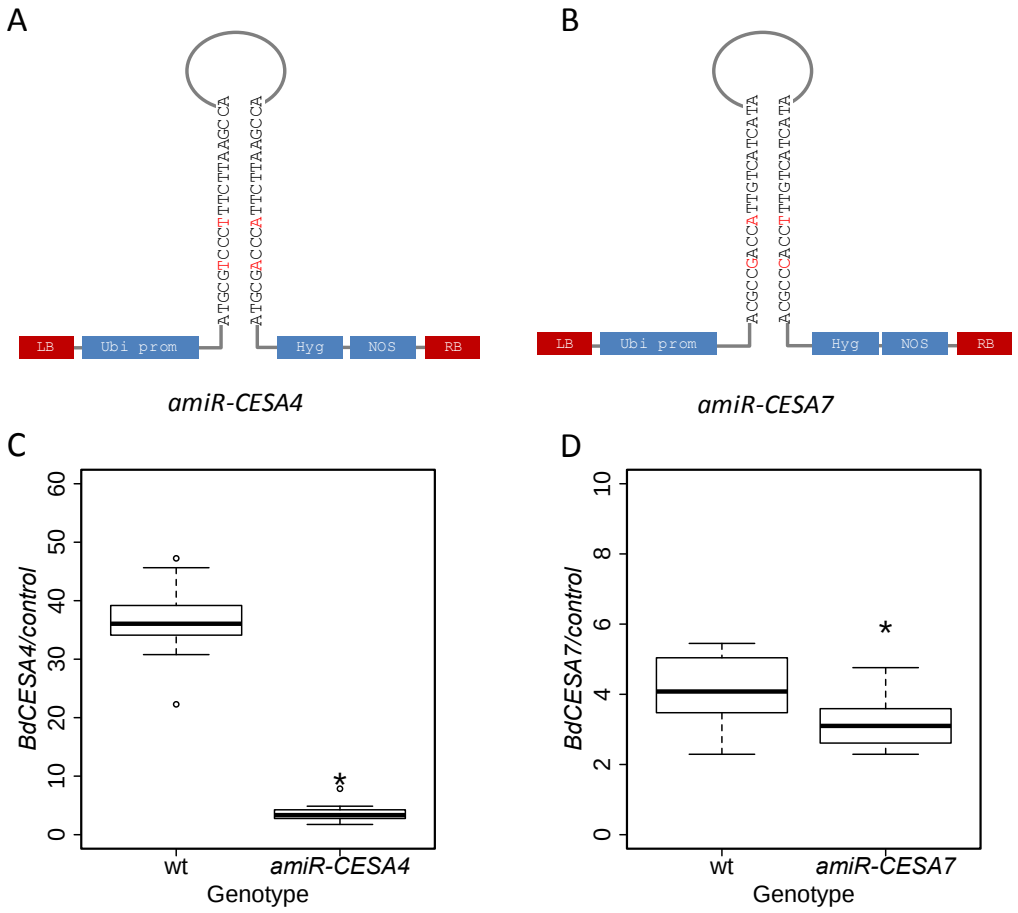

Figure 2.7. Targeting *CESA* expression by means of artificial microRNAs. (A, B) Schematic of the constructs used. The hairpin model illustrates the 21-mer sequence of each microRNA construct and red letters indicate the mismatch in each hairpin recognized by the DICER complex. LB, left border; Ubi prom, maize ubiquitin promoter; Hyg, hygromycin phosphotransferase gene; NOS, nopeline synthase terminator; RB, right border. (C, D) Relative levels of transcript measured by RT-QPCR. Reference genes are given in methods. Stem tissue was collected at the same development stage when inflorescence was just emerging from the flag leaf. Three to five individuals from three to four independent transgenic lines were analyzed for each construct. The boxes show interquartile range, the whiskers show the outer quartile edge, and the black line represents the median of each distribution. Open circles represent outliers, when present. * Denotes significance at the 5% level.

when the inflorescence was just emerging from the flag leaf, at developmentally equivalent time points. Both of the artificial microRNA constructs significantly reduced transcript abundance of the corresponding target (Fig. 2.7C, D). Specifically, *BdCESA4* was reduced 9.5 fold and *BdCESA7* was reduced 1.5 fold. Neither transgene detectably reduced the level of *BdCESA8* and *amiR-CESA7* caused no significant decrease in *BdCESA4*; however, *amiR-CESA4* modestly decreased the expression of *BdCESA7* (Fig. 2.8).

2.3.5 *BdCESA4* and *BdCESA7* knock-down lines and the structure of the stem

The *BdCESA* knock-down lines were modestly but significantly decreased in stature and delayed in inflorescence emergence (Fig. 2.9). To investigate anatomical changes, we examined first internode morphology of the same plants assayed for mRNA levels. Because the transgenic lines were grown at different times, each comparison included a wild-type control, and differences between these controls presumably reflect differences in growth conditions. Hand-cut transverse stem sections were stained with the polychromatic basic dye, toluidine blue, and imaged using a light microscope. Toluidine blue differentially stains cell wall polymers— polysaccharides purplish blue and lignified cell walls turquoise—allowing cell types to be distinguished. The artificial microRNA constructs had little if any effect on the overall shape and arrangement of the vascular bundles (Fig. 2.10A). Likewise, there was no significant difference between the genotypes in number of vascular bundles in either the inner or the outer ring (data not shown). While changes in anatomy were minor or absent, stem diameter appeared to be reduced, consistent with the decreased plant height. Measurements of stem area revealed

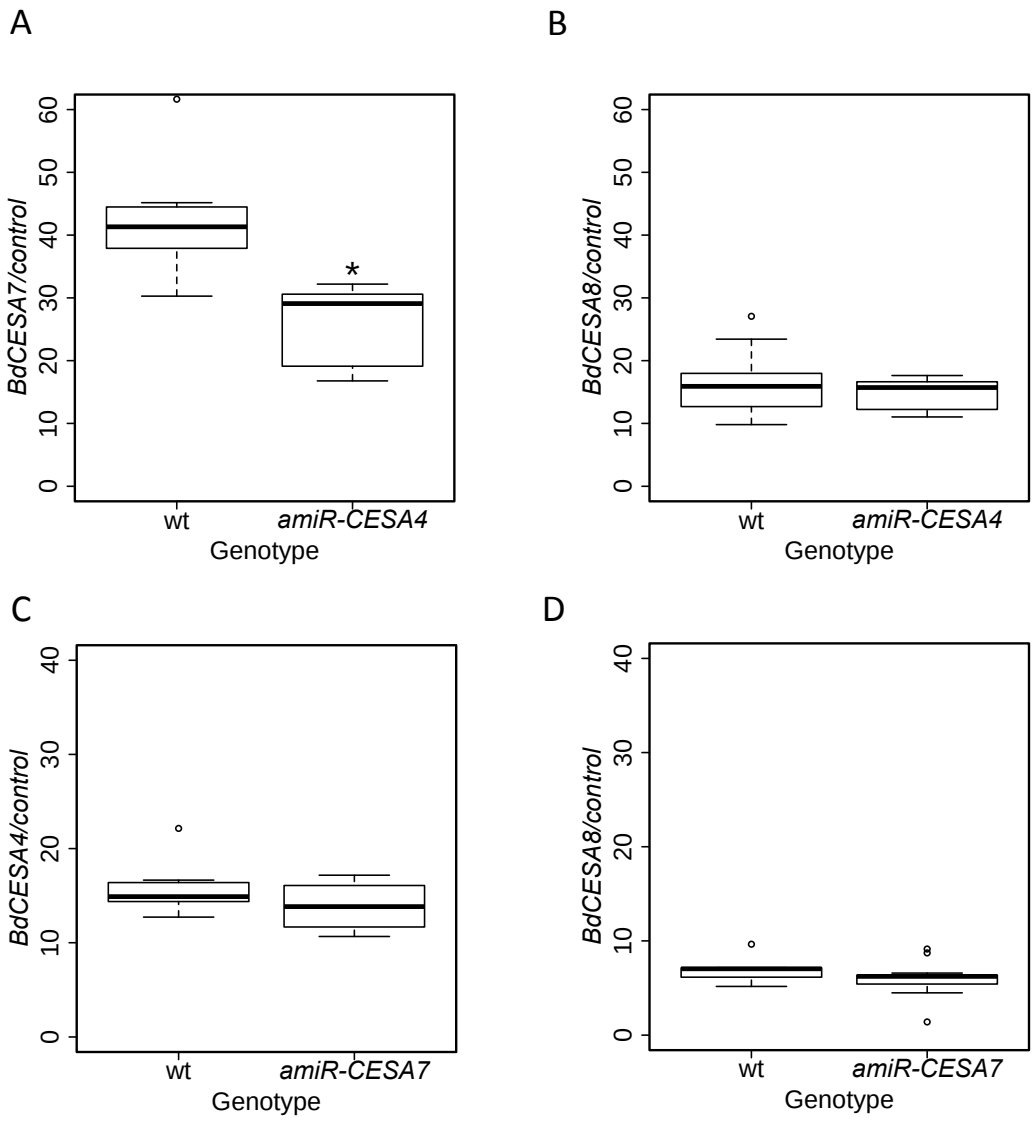

Figure 2.8. **Relative expression of selected non-targeted *BdCESA* genes.** Transcript abundance measured by RT-QPCR. The boxes comprise data from three to five individuals from three to four independent transgenic lines. Stem tissue was collected at the same development stage when inflorescence was just emerging from the flag leaf. Box plots and significance are as described for figure 2.7.

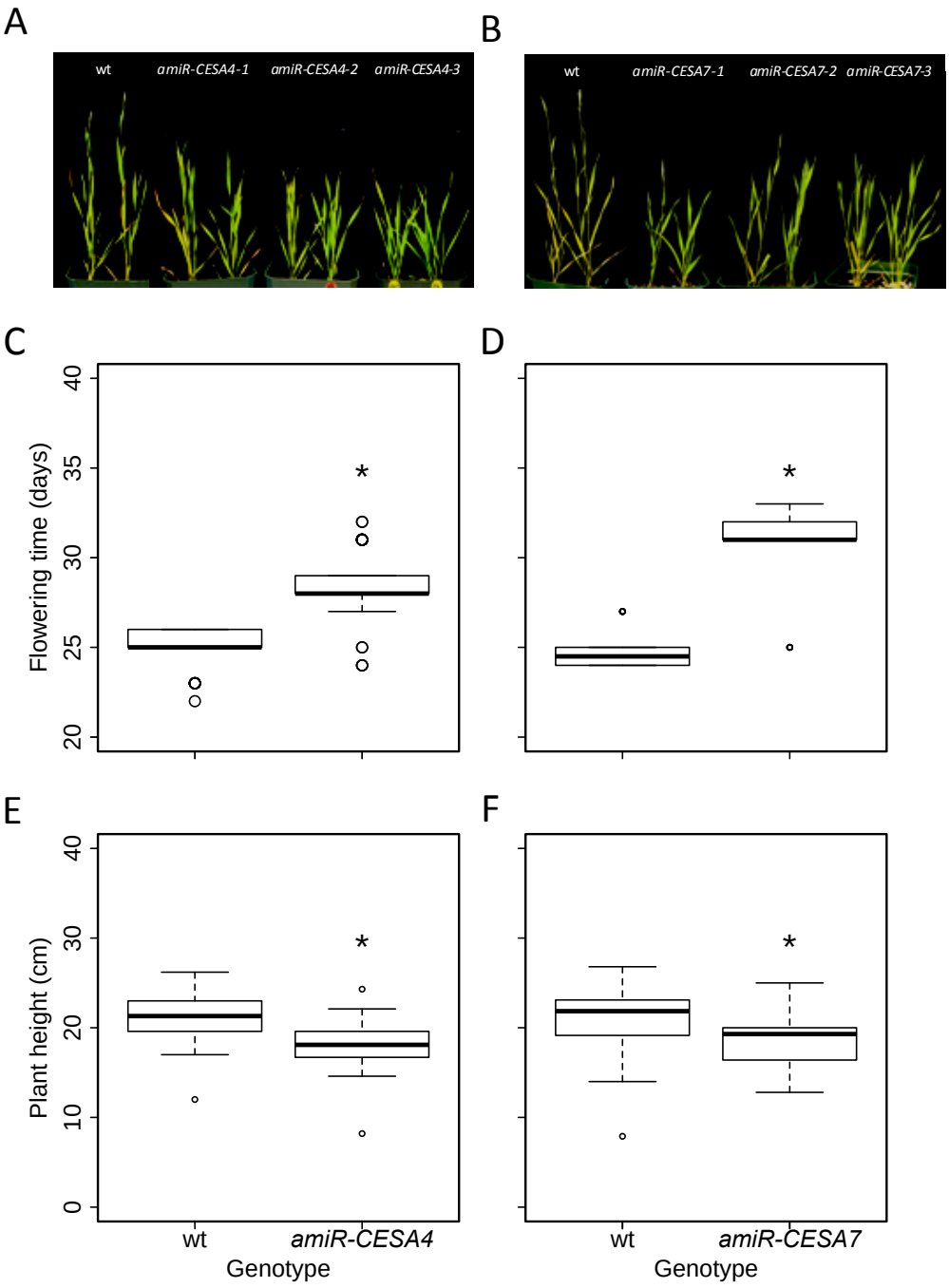

Figure 2.9. Whole plant phenotypes. (A, B) Plants at the time of wild-type inflorescence emergence. Representative plants of wild-type and of three independent lines used for each construct. (C, D) Days to inflorescence emergence. (E, F) Mature plant height. Twenty to thirty individuals from three independent lines were analyzed for each construct. Box plots and significance are as described for Figure 2.7.

a modest reduction, but one that was significant for both *amiR-CESA4* and *amiRCESA7* (Fig. 2.10B, C).

To examine the effects of the artificial microRNA constructs on cell wall structure, we measured cell wall thickness in the toluidine blue-stained sections. For both targets, the constructs reduced the thickness of cell walls modestly, but significantly (Fig. 2.10D-G). The reduction was similar for xylem as well as for interfascicular fibers. This observation is consistent with the phenotypes of secondary *CESA* mutants characterized in other grass species (Tanaka et al., 2003; Zhang et al., 2009; Burton et al., 2010).

2.3.6 Knock-down of *BdCESA4* and *BdCESA7* and crystalline cellulose content

Reduction in stem size along with the thinner cell walls indicated the possibility of a change in cell wall structure. Completely senesced and homogenized stem tissue was analyzed. First, crystalline cellulose content was assayed by X-ray powder diffraction, using the same individuals analyzed for figures 6 and 7, with the two wild-type samples pooled. The transgenic genotypes gave diffraction patterns with lower intensities at nearly all measured angles, although the effect in *amiR-CESA4* was stronger than in *amiR-CESA7* (Fig. 2.11A). To evaluate cellulose crystallinity, as described in Methods, we calculated a crystallinity index by means of the so-called "amorphous subtraction method" (Fig. 2.11B). The crystallinity index confirmed the visual impression of the diffraction patterns, namely a significant reduction in cellulose crystallinity for *amiR-CESA4*.

Second, we evaluated the amount of crystalline cellulose by sum-frequency-generation (SFG) vibration spectroscopy (Barnette et al., 2011). In this method, the sample is irradiated simultaneously by visible laser pulses at 532 nm and by infrared laser

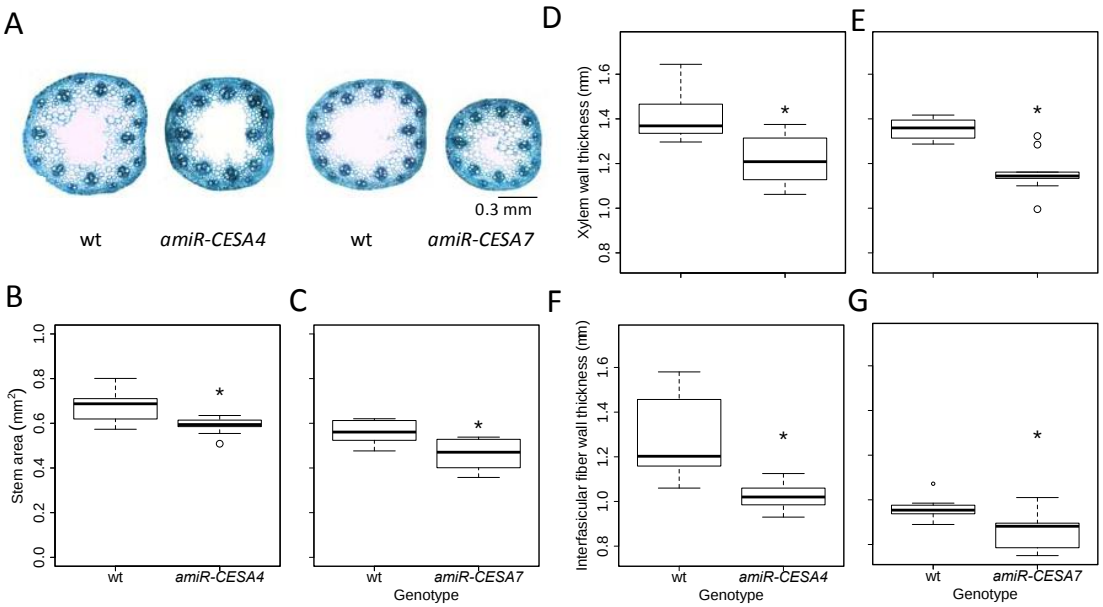

Figure 2.10. Stem anatomy. (A) Toluidine blue-stained transverse sections. (B, C) Stem area. Cell wall thickness of (D, E) metaxylem and (F, G) interfascicular fibers. First internode was collected and analyzed at inflorescence immergence from developmentally equivalent plants. Three to five individuals from three independent lines were analyzed. Box plots and significance are as described for figure 2.7.

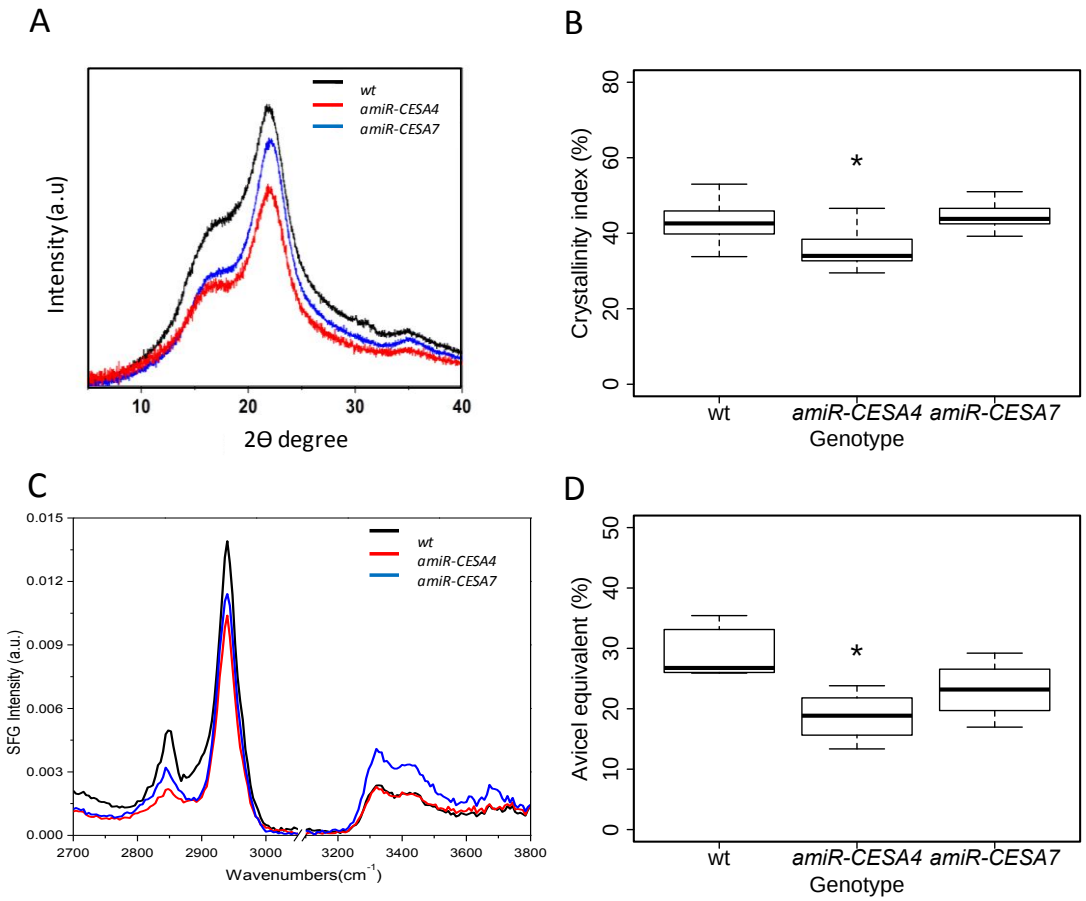

Figure 2.11. Spectroscopic analysis of cellulose crystallinity. (A) X-ray powder diffraction profiles. (B) Crystallinity index derived from the diffraction profiles using the amorphous cellulose subtraction method. (C) Sum-frequency-generation vibration spectra. (D) Cellulose crystallinity derived from the spectra based on comparison to Avicel. Eight to twelve individuals from three independent lines were analyzed for each transgene. Box plots and significance are as described for figure 2.7.

pulses with a tunable frequency. Among reflected and scattered lights, there are photons whose frequency is the sum of two input laser frequencies, which can be filtered and recorded separately. Due to the symmetry requirements, this frequency summation can be caused by crystalline cellulose but not by amorphous non-crystalline cell wall components (Barnette et al., 2011). For this analysis, lines were grown at the same time. In the spectra, there are three prominent peaks indicative of crystalline cellulose Iβ: namely 2850 cm^{-1} ascribed to symmetric CH_2 stretching, 2944 cm^{-1} ascribed to asymmetric CH_2 stretching, and 3320 cm^{-1} ascribed to the intra-chain hydrogen-bonded hydroxyl stretch (Fig. 2.8C) (Barnette et al., 2011; Barnette et al., 2012). Taking Avicel, a model cellulose Iβ with a high crystallinity, as a standard, the intensity at 2944 cm^{-1} can be used to estimate the crystalline cellulose amount by means of the previously determined calibration curve (Barnette et al., 2012). The comparison of the intensities recorded at 2944 cm^{-1} indicated that the Avicel-equivalent crystalline cellulose content tended to be reduced in *amiR-CESA7*and was significantly reduced in *amiR-CESA4* (Fig. 2.11D). These results are comparable to those from X-ray diffraction.

To examine cellulose crystallinity on a cellular scale, we used polarized-light microscopy on plants grown at one time and used also for the sum frequency generation spectroscopy. As described in Materials and Methods, our microscope is based on circularly polarized light, allowing contrast to be independent of crystal orientation within the plane perpendicular to the microscope's optical axis, and generating images in which intensity is proportional to birefringent retardance (Fig. 2.12). As expected for the ubiquitous presence of cellulose, all cell walls of the wild-type had retardance, with epidermis, metaxylem, and cells of the vascular sheath being particularly strong (Fig.

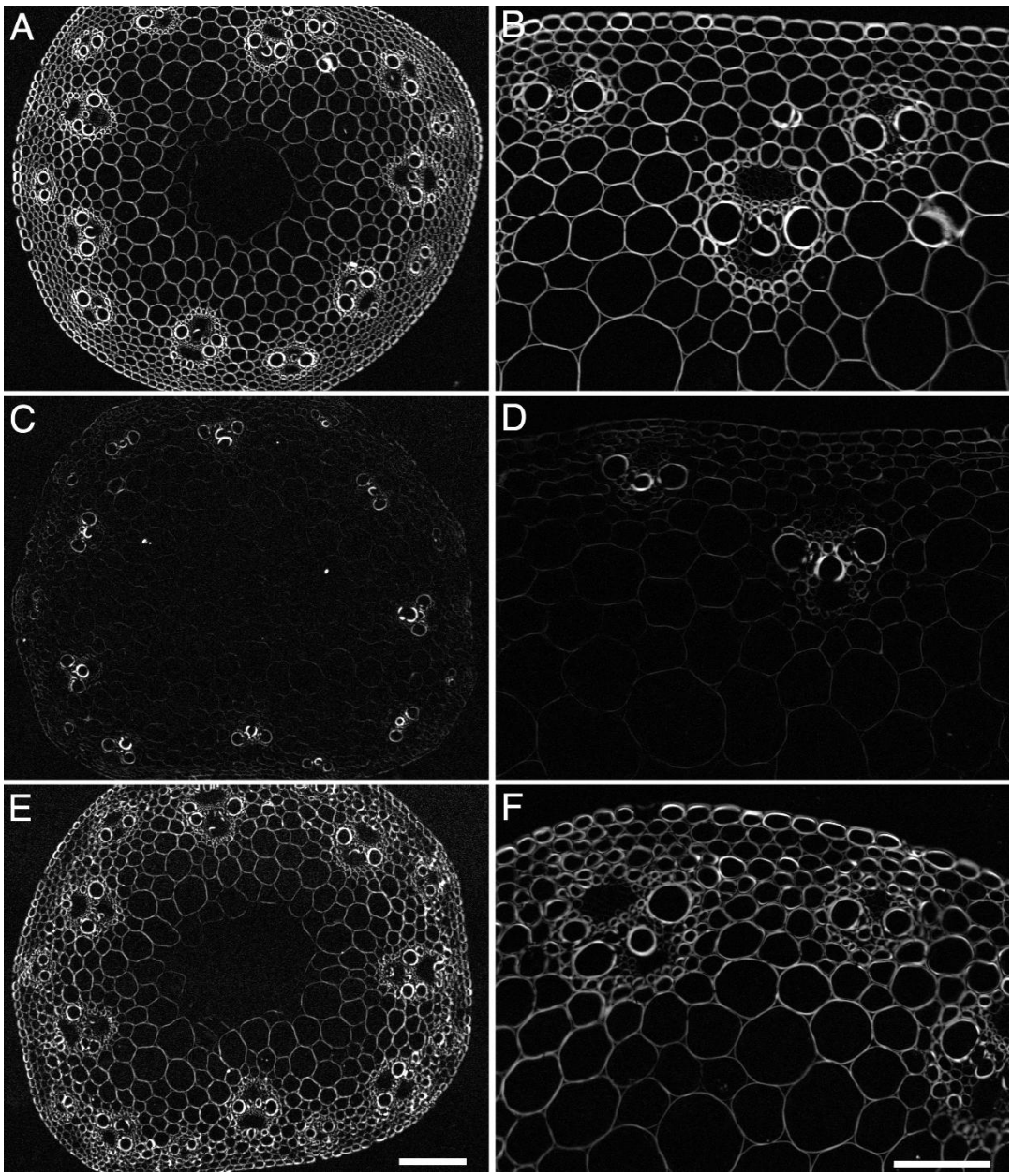

Figure 2.12. Polarized-light micrographs of stem internode transverse cross-sections. Representative images of (A, B) wild type; (C, D) *amiR-CESA4*; (E, F) *amiR-CESA7*. Left hand panels are observed through a 4x lens and a gray scale value of 255 indicates a retardance value of 5 nm; Right hand panels are observed through a 20x lens and a gray scale value of 255 indicates a retardance value of 13 nm. Bars = 50 µm (E) and 100 µm (F).

Table 2.2. List of oligonucleotides used in this study.

Oligo name	Sequence (5'- 3')
used for amiRNA construct development	
G-4368	CTG CAA GGC GAT TAA GTT GGG TAA C
G-4369	GCG GAT AAC AAT TTC ACA CAG GAA ACA G
Bd28350_I-miR-s1	AGTGGCTTAAGAATGGGTCGCATCAGGAGATTCAGTTTGA
Bd28350_II-miR-a1	TGATGCGACCCATTCTTAAGCCACTGCTGCTGCTACAGCC
Bd28350_III-miR*s1	CTATGCGTCCCTTTCTTAAGCCATTCCTGCTGCTAGGCTG
Bd28350_IV-miR*a1	AATGGCTTAAGAAAGGGACGCATAGAGAGGCAAAAGTGAA
Bd30540_I-miR-s2	AGTATGATGACAATGGTCGGCGTCAGGAGATTCAGTTTGA
Bd30540_II-miR-a2	TGACGCCGACCATTGTCATCATA CTGCTGCTGCTACAGCC
Bd30540_III-miR*s2	CT ACGCCCACCTTTGTCATCATATTCCTGCTGCTAGGCTG
Bd30540_IV-miR*a2	AATATGATGACAAAGGTGGGCGTAGAGAGGCAAAAGTGAA
used for qRT-PCR	
qPCRBdCESA4_F	GCGTTTCGCATACACCAACACC
qPCRBdCESA4_R	ACTCGCTAGGTTGTTCAGTGTGG
qPCRBdCESA7_F	GCGATTCGCCTACATCAACACC
qPCRBdCESA7_R	GGCTGGCAAATGTGCTAATCGG
Bd5g25870qPCR_F	TCAGCAGGGTGCTAATTCAGTTC
Bd5g25870qPCR_R	CGACAGAGTTTAGCGGTCTTAGC
BdUBC18_F	TCACCCGCAATGACTGTAAG
BdUBC18_R	ACCACCATCTGGTCTCCTTC
BdGAPDH_F	ATGGGCAAGATTAAGATCGGAATCAACGG
BdGAPDH_R	AGTGGTGCAGCTAGCATTTGAGACAAT
used for genotyping	
Hyg_F	AGAATCTCGTGCTTTCAGCTTCGA
Hyg_R	TCAAGACCAATGCGGAGCATATAC
used for *in situ* probe development	
CESA4_probe_aF	CAA GGG GCT CAT GGG AAG GCA
CESA4_probe_aR	GAA GCA GCC GAA TGC AGT AAC AA
CESA7_probe_aF	GCG CCT GAA GAA TTC TGA AGA TTG
CESA7_probe_aR	GCC TGT ATT CTT GAG GTT TGG TG

2.12A, B). In contrast, cell walls in the stem of *amiR-CESA4* had weak retardance (Fig. 2.12C, D). Although retardance would tend to decrease along with cell wall thickness, the decrease in retardance was much larger than that of cell wall thickness. Likewise, compared to the wild-type, cell walls of *amiR-CESA7* had less retardance (Fig. 2.12E, F). However, for this genotype the decrease was modest and within the range that might be attributable to the thinner cell walls. These results, taken together with those from diffraction and spectroscopy, confirm that the reduced expression of *BdCESA4* reduced the amount of crystalline cellulose present in the secondary cell wall.

2.4 Discussion

For *B. distachyon*, this is apparently the first study detailing the *CESA* gene family and functionally characterizing *BdCESA*s involved in secondary cell wall synthesis. Members of the CESA family are best characterized in *A. thaliana*, where they are assigned a role for cellulose synthesis in either primary or secondary cell wall. Other vascular plants have a similar gene family structure (Carroll and Specht, 2011). The *B. distachyon CESA* gene family includes the uncharacterized grass-specific CESA clade (*BdCESA10*), which has been identified in all grass genomes sequenced to date, and all the other previously described clades (Paterson et al., 2009; Wang et al., 2010; Carroll and Specht, 2011).

Using amino acid sequences, the ten BdCESA genes were categorized into two groups: primary and secondary. Each of the three secondary CESA clades contains a single *A. thaliana* protein and a single rice protein, which have all been functionally characterized (Taylor et al., 1999; Taylor et al., 2000; Tanaka et al., 2003). Fittingly, a

single BdCESA protein was present in each of these clades, suggesting they have not expanded since the time of eudicot and monocot divergence 140-150 million years ago (Chaw et al., 2004). This differs for the primary CESAs, which have differentially expanded between eudicots and monocots. There are different numbers of proteins in the CESA1 and CESA3 clades in *A. thaliana*, *B. distachyon*, and rice, and the CESA6 clade has diverged into separate eudicot and monocot clades, referred to as CESA6A and CESA6B, respectively (Carroll and Specht, 2011). Among the three secondary *B. distachyon* CESAs, *BdCESA8* was the most divergent, as was that clade member (*OsCESA4*) in rice (Tanaka et al., 2003). Moreover, *BdCESA8* and *OsCESA4* both lack eight amino acids in the first RING type zinc finger motif. However, *A. thaliana CESA8* neither lacks those eight amino acids nor is notably divergent.

Confirming the deduction from phylogeny, expression of the secondary *BdCESAs* was generally enriched in stems, which are abundant in secondary cell walls, and transcripts were specifically abundant in stem vascular tissue and the surrounding mechanical tissue including sub-epidermal cell layers, tissues that all make secondary cell walls. This is consistent with the tissue-specific patterns of secondary CESA expression observed in rice, maize, *A. thaliana*, and barley (Taylor et al., 2003; Appenzeller et al., 2004; Burton et al., 2004; Wang et al., 2010). Interestingly, in developing stems of barley, *HvCESA8* expression was two-fold greater than that of *HvCESA4* (Burton et al., 2004), which was also observed here for *BdCESA8* and *BdCESA4*. Taken together, we conclude that *BdCESA4*, *7*, and *8* encode cellulose synthase catalytic subunits for the secondary cell wall.

To our knowledge, this is the first report of artificial microRNA-mediated gene silencing in *B. distachyon*. We characterized the loss-of-function lines generated for *BdCESA4* or *BdCESA7* for changes in morphology and cellulose crystallinity. While these lines had a small but significant delay in inflorescence emergence and a reduction in stature, whole plant morphology was similar to wild-type plants. Similarly, delayed flowering and reduced stature was observed in the *A. thaliana irx3* mutant and in rice *brittle culm* mutants (Turner and Somerville, 1997; Zhang et al., 2009). Sections of the stem revealed that the knock-down lines had a small but significant reduction in stem internode transverse cross-sectional area, a reduction that resembles that observed in rice *brittle culm* mutants and barley *fragile stem 2* (Kokubo et al., 1991; Tanaka et al., 2003).

The loss-of-function lines for *BdCESA4* and *7* had cell walls in xylem and sclerenchyma that were modestly but significantly thinner than wild-type. Thinner cell walls are expected from a defect in secondary cell wall cellulose synthesis, and have been observed consistently in secondary cell wall *cesa* mutants (Tanaka et al., 2003; Zhang et al., 2009; Burton et al., 2010). In *A. thaliana*, the defining phenotype of secondary *CESA* mutants is irregular or collapsed xylem (Taylor et al., 1999; Taylor et al., 2000). However, xylem contours appeared to be regular here, as they do in the secondary *CESA* mutants of rice (Tanaka et al., 2003). These differences probably reflect distinctions between eudicots and monocots in xylem tension or cell wall composition rather than differences in protein function (Vogel, 2008; Handakumbura and Hazen, 2012).

Cellulose is a linear glucan polymer made of β (1-4) linked glucose molecules. Thirty-six such glucan chains are predicted to form a cellulose microfibril (Taylor, 2008). Hydrogen bonding between microfibrils crystallizes multiple cellulose chains together

providing the physical properties necessary to maintain rigid plant cell walls. Crystallinity also allows cellulose to be readily quantified. We took advantage of spectroscopic techniques (X-ray diffraction and sum frequency generation) and polarized-light microscopy to assess the status of cellulose in the loss-of-function lines.

For the *amiR-CESA7* mutants there was no significant decrease in crystallinity seen spectroscopically and the small decrease in birefringent retardance might reflect the thinner cell walls. Given that *BdCESA7* expression was reduced by less than a factor of two, a generally wild-type cellulose status is perhaps not unexpected. In contrast, the *amiR-CESA4* mutants had substantially decreased retardance and significantly decreased crystallinity from both diffraction and spectroscopic methods. The most parsimonious explanation for these data is that the cell walls of this line simply contained less cellulose than those of the wild-type. A nuanced explanation is that the cellulose synthesized in this line contained more defects, such as amorphous regions, or comprised microfibrils with fewer than the usual number (e.g., 36) of glucose chains. These alternatives are difficult to distinguish. Be that as it may, the significant decrease in crystalline cellulose amount or quality in this line matches the nearly 10-fold reduction in *BdCESA4* expression. This result implies that the function of *BdCESA4* is at least partially non-redundant and it is noteworthy that the expression of neither *BdCESA7* nor *8* was elevated in an attempt at compensation.

Interestingly, even though the two lines differed in the strength of the knock-down and in the consequent loss of cellulose, the morphological defects in the plants were similar and modest. This implies that small decreases in CESA activity are sufficient to impact the plant, possibly through a feedback system analogous to the one *A.*

thaliana invoking THESEUS (Hématy et al., 2007), but that progressively larger changes to the plant need not follow from larger decreases in CESA activity. In this connection, it will be interesting to observe the phenotype of a complete loss of function mutant for BdCESA4.

CHAPTER 3

BdMYB48 IS A GRASS SPECIFIC POSITIVE REGULATOR OF SECONDARY CELL WALL SYNTHESIS AND BIOFUEL FEEDSTOCK ATTRIBUTES IN *Brachypodium distachyon*

3.1 Introduction

The transcriptional network that governs the regulation of secondary cell wall biosynthesis in eudicots consists of several MYB family transcription factors (Zhong et al., 2010; Hussey et al., 2013; Schuetz et al., 2013). MYB transcription factor proteins are present in all eukaryotes and are generally encoded by large gene families (Rosinski and Atchley, 1998; Lipsick 1996; Wilkins et al., 2009). For example, *Arabidopsis thaliana* and rice (*Oryza sativa*) have an estimated 197 and 155 MYBs, respectively (Katiyaret al. 2012). In addition to overall sequence similarity, the family is categorized based on the number and type of DNA-binding domains in each protein. They vary from one to four amino acid sequence repeats and are named based on the number of repeats contained within each DNA-binding domain. Majority of plant MYB proteins, the R2R3-MYB group, harbor two repeats and are extensively described for *A. thaliana*, rice and maize (Stracke et al., 2001; Du et al., 2012; Prouse and Campbell, 2012). The MYB proteins that are known to regulate cell wall regulatory network include AtMYB46/83/58/63/85/20/52/54/69/103/4/32 and 75. Among these AtMYB46 and AtMYB83 are thought of as master regulators capable of activating downstream transcription factors and cell wall genes. Over-expression of AtMYB46/83 resulted in thicker secondary cell walls in the xylem vessels while dominant repression resulted in thinner walls (Zhong et al., 2007). Some of these proteins specifically regulate the

biosynthesis of a single component of the cell wall. For instance AtMYB58/63 directly binds the promoters of lignin genes and activates lignin biosynthesis. Dominant repression of these two proteins results in reduction of secondary wall thickening and lignin content. As a result these plants exhibit a pendent stem phenotype. On the other hand, over-expression leads to up regulation of the expression of lignin pathway genes and ectopic deposition of lignin in cells that are normally not lignified. However, over-expression of AtMYB58/63 does not result in thicker secondary cell walls (Zhou et al., 2009). Several other MYBs including AtMYB4/20/32/52/54/69/85/103 are abundant in stem tissue and facilitate cell wall gene expression and regulate the overall secondary cell wall biosynthesis process. Similar to AtMYB58/63, AtMYB85 is also shown to activate the lignin pathway genes. Dominant repression of AtMYB85 resulted in a significant reduction in secondary wall thickness of interfascicular and xylary fibers and over-expression leads to ectopic deposition of secondary cell walls in the epidermis and cortex of the stems (Zhong et al., 2008). Dominant repression of MYB103/52/54/69 also reduces secondary wall thickening in interfascicular and xylary fiber cells however, none of them had a significant impact on the vessels (Zhong et al., 2008). Moreover all these MYB proteins, MYB46/83/58/63/85/52/54 and 103 act as activators of cell wall regulation.

On the other hand MYB4/32 and 75 act as repressors of cell wall biosynthesis. MYB75 loss-of-function mutants possess thicker secondary cell walls in interfascicular and xylary fibers. In addition knocking out MYB75 up regulates a set of lignin pathway genes and affects the lignin content in these plants (Bhargava et al. 2010). These genes are activated by AtMYB46/83, which also directly bind the promoters of downstream cell

wall genes. Those that have been well characterized for DNA-binding affinity have been shown to interact with a *cis*-regulatory AC-element in the promoters of cell wall genes to facilitate transcription. The three variants of the AC element AC-I (ACCTACC), AC-II (ACCAACC), AC-III (ACCTAAC) are present in promoters of many lignin pathway genes and are thought to regulate lignin biosynthesis (Lois et al., 1989; Hatton et al., 1995; Raes et al., 2003).The AC element was further described to have four more variations by interchanging the T with a C at the last base position, ACC(T/A)A(A/C)(C/T) and named the secondary wall MYB responsive element (SMRE) or the MYB46 responsive *cis*-element (M46RE). This element was over represented in genes that were up regulated by over-expression of MYB46 under an inducible promoter. Among these, the representation was particularly high for the cellulose synthase genes, cellulose synthase like genes and lignin genes associated with MYB46 regulated secondary cell wall biosynthesis (Kim et al., 2012).

Unlike *A. thaliana* and poplar, the current transcriptional regulatory network for grass secondary cell wall biosynthesis is sparse (Handakumbura and Hazen, 2012). To date, the function of only a few MYB transcription factors have been demonstrated in a grass species. ZmMYB31 and PvMYB4 are two MYB repressors characterized in maize and switchgrass respectively. They are both shown to repress the lignin genes by binding to the AC elements found in the promoters of these genes and to subsequently regulate cell wall biosynthesis (Fornale et al., 2010; Shen et al., 2012). Rice (*Oryza sativa*) and maize (*Zea mays*) orthologs of AtMYB46/83 are the only MYB activators characterized to date. These proteins were capable of activating the entire secondary cell wall biosynthetic pathway when over-expressed in *A. thaliana* (McCarthy et al., 2009; Zhong

et al., 2011). However, no activators have been shown to directly activate the cell wall genes in grasses. Here we describe a member of a grass specific clade that directly activates cell wall genes and regulates cell wall biosynthesis. We identified *BdMYB48* as a candidate gene for the positive regulation of secondary cell wall biosynthesis based on expression profiling and phylogeny. Here we demonstrate the function of a novel grass specific MYB regulator, BdMYB48 in the model monocot *Brachypodium distachyon*.

3.2 Materials and Methods

3.2.1 Phylogenetic analysis

Amino acid sequences of *A. thaliana* and rice MYBs were downloaded from the PlnTFDB plant transcription factor database (Perez-Rodriguez et al., 2010) and *B. distachyon* MYB sequences Bradi4g06317 (BdMYB1), Bradi2g47590 (BdMYB48), Bradi2g17980 (BdMYB31), Bradi2g40620 (BdMYB44), Bradi2g11080 (BdMYB27), Bradi2g36730 (BdMYB41), Bradi5g21970 (BdMYB108), Bradi1g20250 (BdMYB4), Bradi1g61400 (BdMYB18), Bradi5g20130 (BdMYB104), Bradi3g52260 (BdMYB75) were downloaded from phytozome (http://www.phytozome.net/). DNA-binding domains were extracted from the above amino acid sequences and two separate neighbor-joining phylogenies were constructed with 1000 bootstrap permutations using MEGA 5.0 one with the full-length protein and one with the DNA-binding domains alone.

3.2.2 Yeast one-hybrid assay

Yeast one-hybrid screens were conducted as previously reported (Pruneda-Paz et al., 2009). Two synthetic promoters were generated for the AC-like element (ACCAAC) and the mutated AC-like element (TTTAAC) by fusing four copies of the sequence in

tandem. A spacer sequence of tttagatatcataa was included at the 3'end of the synthetic promoters. Sequence confirmed clones were recombined with pLacZi plasmid (Clonetech) containing the *LacZ* reporter gene and stably integrated into the YM4271 yeast strain. *BdMYB48* was fused in frame with *GAL4* activation domain in pDEST22 destination vector.

3.2.3 Plant material and growth conditions

Brachypodium distachyon (L.) line Bd21-3 was used as the genetic background. Wild type, control, and mutant seeds were imbibed on moist paper towels for seven days at 4˚C and planted in potting mix (#2; Conrad Fafard Inc. Agawam, MA). Plants were grown at control conditions in a growth chamber at 20˚C in 20h: 4h light: dark cycles at a fluence rate of 220 $\mu mol \cdot m^{-2} \cdot s^{-1}$ and relative humidity of ~68%.

3.2.4 Plasmid construction and plant transformation

The full-length coding region of *BdMYB48* (Bradi2g47590) was PCR amplified from Bd21-3 stem cDNA to include the stop codon using Phusion high-fidelity DNA polymerase (New England bio labs) and cloned into pENTR/D-TOPO vector (Invitrogen). Sequence confirmed plasmid was recombined with pOL001 ubigate ori1 destination vector (modified from pOL001, described in (Vogel et al., 2006) to generate the MYB48 gain-of-function construct (*MYB48-OE*). A 39 nucleotide dominant repressor (CRES) sequence was synthesized using overlapping oligonucleotides with *att*B2 and *att*B5 flanking sites and cloned into pDONR 221 P5-P2. Full-length coding region of *BdMYB48* without the stop codon (MYB48NS) was cloned into pDONR 221 P1-P5r. CRES and MYB48NS entry clones were recombined with pOL001 ubigate ori1 destination vector to generate the MYB48 dominant repressor construct (*MYB48-DR*). All constructs were

transformed into *Agrobaterium tumefaciens* strain *AGL1* via electroporation for calli transformations. *Brachypodium distachyon* calli transformations were carried out with minor modifications as previously described (Handakumbura et al., 2013). Primary transgenic were PCR confirmed for the hygromycin resistance gene and propagated for three subsequent generations and the resulting T_4 progeny were PCR confirmed for presence of the hygromycin phosphotransferase II gene using a kit (Thermo scientific Phire plant direct PCR kit) according to manufacturer's specifications. PCR confirmed transgenics were used for subsequent experiments.

3.2.5 Measurements of transcript abundance and localization

Total RNA was extracted using a kit (Plant RNaeasy, Qiagen, Valencia, CA) according to the manufacturer's instructions. First, second, and third nodes and internodes of the tallest stem were frozen in liquid nitrogen from developmentally comparable individuals at inflorescence immergence, stage 51 on the BBCH-scale for cereals (Lancashire et al. 1991). On-column DNA digestions were performed using RNase-free DNase I (Qiagen). First strand cDNA was synthesized using oligo dT primers (Invitrogen) and QRT-PCR reactions were performed in triplicate as previously described (Handakumbura et al., 2013). Values were normalized against two housekeeping genes, *BdUBC18* (ubiquitin-conjugating enzyme 18) and *BdGapDH* (Hong et al., 2008). Primers were designed using QuantiPrime primer design tool (Arvidsson et al., 2008).

RNA in situ hybridization was performed as previously described (Handakumbura et al., 2013) using stem cross sections of the first internode of the tallest stem at flowering stage.

3.2.6 Microscopy

First internode of the tallest stem at complete senescence was used for histochemical analysis. Hand-cut stem cross sections were stained with phloroglucinol-HCl and observed under an Eclipse E200MV R microscope (Nikon) and imaged using a PixeLINK 3 MP camera. Images captured at 4X magnification were used for area measurements by freehand tracing of a perimeter in ImageJ (http://rsb.info.nih.gov/ij/). First internode of the tallest stem of mutant and vector control plants were excised when the inflorescence was first visible from the flag leaf and fixed in 2% glutaraldehyde in 50mM phosphate buffer at room temperature for two hours. Next samples were post-fixed in phosphate buffered Osmium tetraoxide under same conditions. Samples were rinsed thrice with water and dehydrated in a graded ethanol series. Fixed tissue was infiltrated with 30%, 50%, 70% and 100% spurr's resin for an hour each. Samples were infiltrated overnight in 100% resin, embedded in fresh resin and allowed to solidify in an oven. Embedded samples were sectioned using an ultra-cut microtome; post-fixed with uranyl acetate and lead citrate, and observed with an electron microscope.

3.2.7 Acetyl bromide soluble lignin measurements

Acetyl bromide soluble lignin (ABSL) content was measured as previously described (Foster et al., 2010). Briefly, 1.5 mg of senesced ground stem tissue was incubated with 100 µl of freshly made acetyl bromide solution (25% v/v acetyl bromide in glacial acetic acid) at 50 ˚C for 2 h followed by an additional hour of incubation with vortexing every 15 min. Next, samples were cooled on ice to room temperature, and mixed with 400 µl of 2 M sodium hydroxide and 70 µl of freshly prepared 0.5 M hydroxylamine hydrochloride and 1.43 ml glacial acetic acid was added to the samples to make

the total volume up to 2 ml. 200 µl of each sample was pipetted into a UV-specific 96-well plate and read in a plate reader at 280 nm. Six to sixteen individuals from three events were analyzed for each transgene and ABSL% was calculated as previously described (Foster et al., 2010).

3.2.8 Cell wall digestibility measurements

Fully senesced stems were washed with 70% ethanol at 70 ˚C for 1 h to remove soluble cell wall material and air dried overnight. Next they were ground using a ball mill and weighed into individual wells in 20 mg duplicated and fermented with *Clostridium phytophementans* as previously described (Lee et al., 2012; Lee et al., 2012). Nine individuals from three independent events were analyzed in duplicate for each transgene.

3.2.9 Chromatin immunoprecipitation

About 2 g of whole stems tissue was harvested from three week old plants and treated for 15 min under vacuum with cross-linking buffer (10 mM Tris, pH 8.0, 1 mM EDTA, 250 mM sucrose, 1 mM PMSF and 1% formaldehyde). Cross-linking was quenched using 125 mM glycine, pH 8.0, under vacuum for 5 min, followed by three washing steps in double-distilled water. Next the tissue was rapidly frozen in liquid nitrogen, ground to a fine power using a mortar and pestle and stored at -80 ˚C. Chromatin was extracted using a kit (Zymo-spin) according to the manufacturer specification. Cross-linked samples were washed with the provided Nuclei prep buffer, resuspended in chromatin shearing buffer and sonicated on ice for four cycles using 40% amptitude.100 µl of the sheared chromatin was incubated with the anti-GFP antibody overnight at 4˚C. Samples were recovered using ZymoMag Protein A beads,

washed thrice with the chromatin wash buffers and eluted using the chromatin elusion buffer. Eluted DNA was treated with 5M NaCl at 75 ˚C for 5 min and incubated with Proteinase K at 65 ˚C for 30 min, ChIP DNA was purified using the provided columns and eluted using 8 µl of elution buffer. Triplicate QRT-PCR reactions were performed for three biological replicates using Quantifast SYBR Green PCR Kit (QIAGEN), with 2 ng of DNA with the following cycler conditions: 2 min at 95 ˚C, followed by 40 cycles of 15 s at 95 ˚C, 20s at 55 ˚C and 20 s at 68 ˚C. Results were normalized to the input DNA, using the following equation:$100 \times 2^{(\text{Ct input-3.32--Ct ChIP})}$

3.2.10 Statistical analysis

For each measurement three different T_4 families were analyzed for each transgene. Student's *t*-tests were performed in R v2.15.0. Significance was set a $P < 0.05$.

3.3 Results

3.3.1 *BdMYB48* is highly expressed in stem and is localized to the interfascicular fibers

Regulators of secondary cell wall biosynthesis are commonly highly expressed in tissues abundant in secondary cell walls and the genes they activate mimic their expression behavior. To investigate *BdMYB48* transcript abundance in leaf, stem and root, a *B. distachyon* microarray data set was utilized (Handakumbura et al., 2013). *BdMYB48* transcript abundance in stems was approximately six and three-fold greater relative to leaf and root, respectively (Fig. 3.1A). Grass stems are substantially enriched for secondary cell walls (Matos et al., 2013) and account for a larger proportion of above ground plant biomass. However, not all cell types in the stem undergo secondary wall

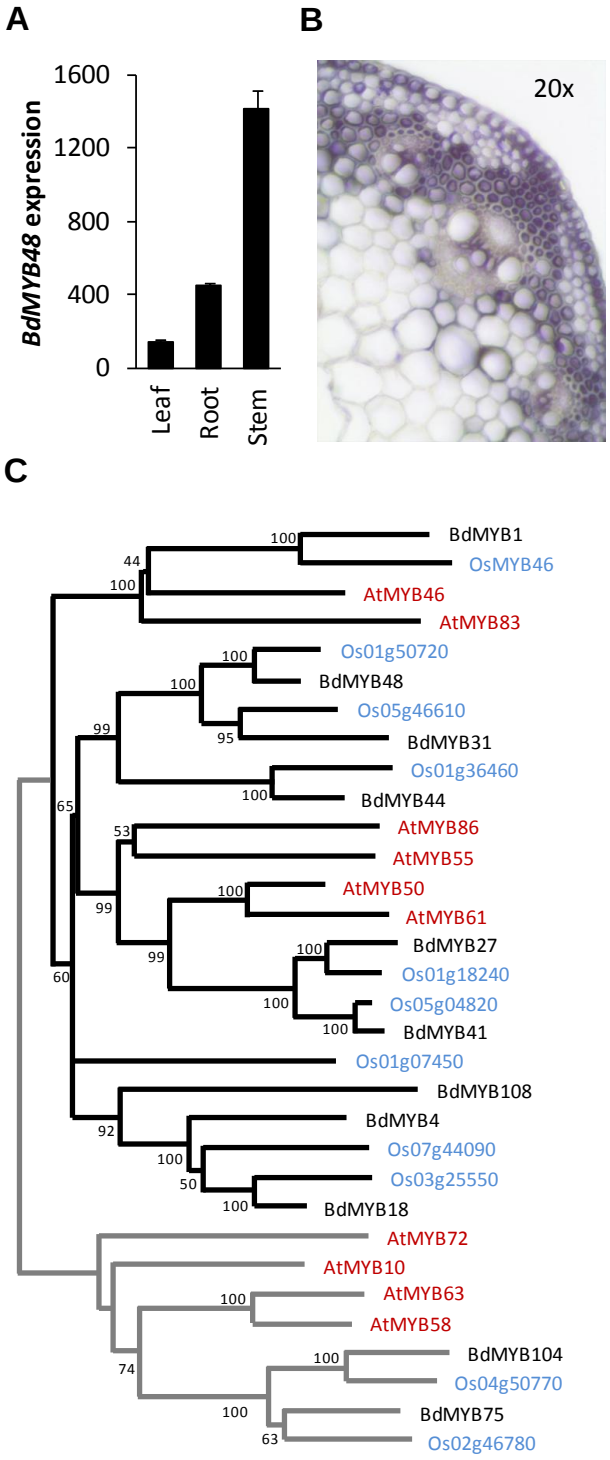

Figure 3.1. BdMYB48 is a grass-specific MYB transcription factor highly expressed in stem. Relative transcript abundance of *BdMY48* in leaf, root and stem tissue measured with a microarray. Mean ± standard deviation of three biological replicates (A). RNA *in situ* hybridization illustrating *BdMYB48* transcript localization in stem cross sections six weeks post germination (B). A subclade of the MYB family illustrating amino acid sequence similarity between *Arabidopsis thaliana* (red), rice (blue) and *Brachypodium distachyon* (black) (C). A rooted neighbor-joining phylogeny was constructed using MEGA 5 with 1000 bootstrap permutations. Numbers on each branch indicate bootstrap support. Branches indicated in gray were used as an out group.

development. To confirm that *BdMYB48* gene expression is associated with those cells undergoing secondary wall thickening, RNA *in situ* hybridization was carried out on stem cross sections at six weeks post germination (Fig. 3.1B). *BdMYB48* transcripts were localized to cell types with thickened secondary walls at this stage of development. They were mainly localized to the epidermis, sclerenchyma fibers and xylem cells in both small and large vascular bundles. No hybridization was detected in the parenchyma cells and phloem fibers. Hybridization with control *BdMYB48* sense probe showed no labeling as expected.

3.3.2 BdMYB48 belongs to a grass specific MYB clade

MYB proteins are involved in a variety of processes including growth and development, cell wall thickening, cell cycle regulation and defense responses. Some of the *A. thaliana* MYB proteins such as AtMYB46 and AtMYB83 directly regulate cell wall biosynthesis (Zhong et al., 2008). Amino acid sequence comparison revealed extensive similarity between BdMYB48 and the large R2R3-MYB family of *A. thaliana* and rice proteins, which have two repeats of the MYB DNA-binding domain (Stracke et al., 2001). To determine the protein similarity between BdMYB48 and the functionally characterized MYBs, a phylogeny was constructed between the MYB protein sequences from *A. thaliana*, *B. distachyon* and rice (Fig. 3.1C). An outgroup of phylogenetically distinct MYB proteins was also included. Additionally, a separate phylogeny was constructed using the DNA-binding domains alone (Fig. 3.2). These two phylogenies revealed a similar relationship. However, the phylogeny created with the full-length proteins (Fig. 1C) had greater bootstrap support. The closest ortholog to the well-characterized *A. thaliana* MYBs, AtMYB46 and AtMYB83 is a gene described by

Valdivia et al. (2103)(Valdivia et al., 2013) as BdMYB1. The BdMYB48 clade includes

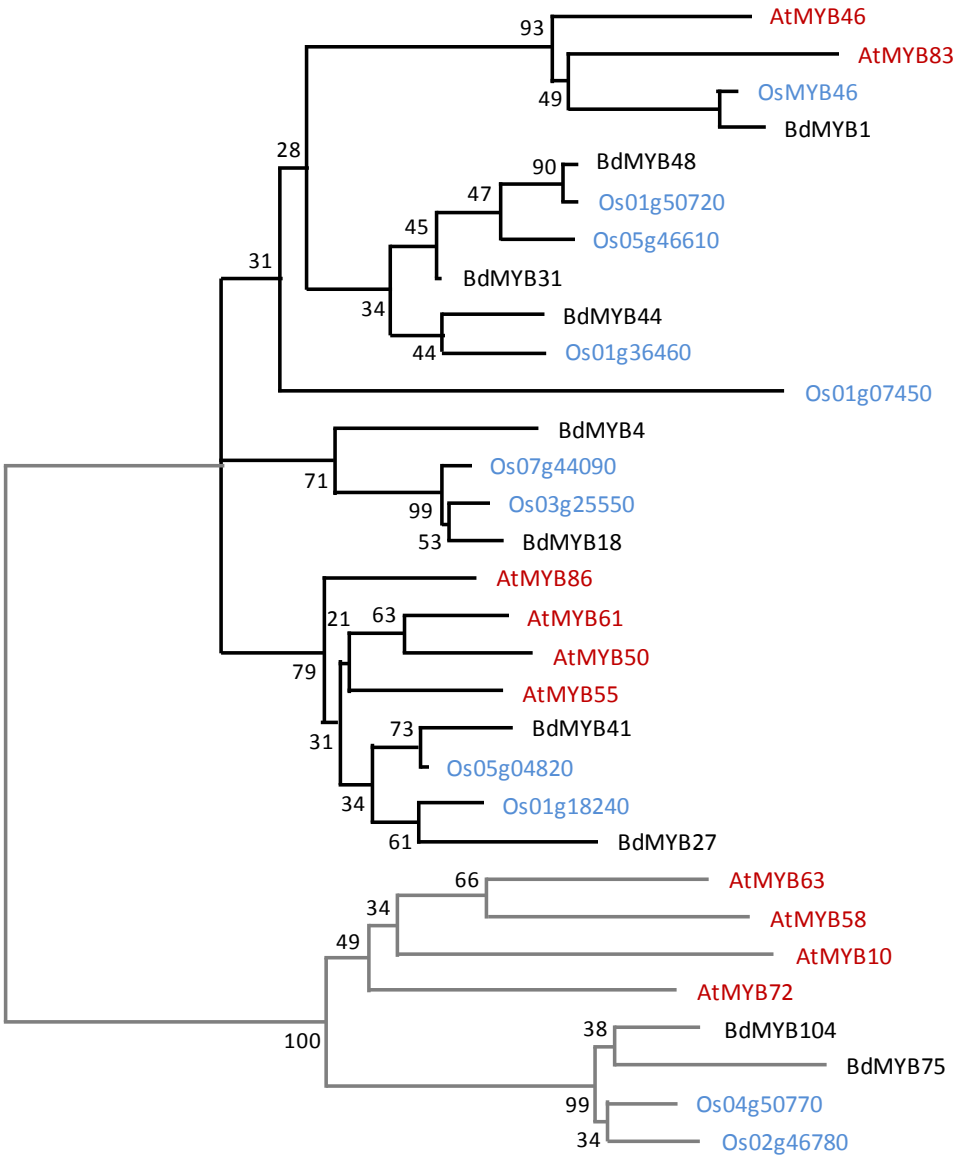

Figure 3.2. The DNA-binding domain of BdMYB48 does not have an Arabidopsis ortholog. Phylogenetic relationship of the DNA-binding domains of a sub set of MYB proteins from *Arabidopsis thaliana* (red), *Brachypodium distachyon* (black), and rice (blue). DNA-binding domains were extracted from the MYB proteins analyzed in Fig. 3.1C and a neighbor-joining phylogeny was constructed using MEGA 5 with 1000 bootstrap permutations. Numbers on each branch indicates the bootstrap support. Branches indicated in gray were used as an out-group.

BdMYB31 and BdMYB44, each with a corresponding rice protein, but none from *A. thaliana*. This group is distinct from AtMYB46 and AtMYB83 by overall amino acid sequence similarity and a neighboring group that includes the *A. thaliana* wall associated regulators AtMYB86 and AtMYB50. Therefore, BdMYB48 may have a cell wall biosynthesis regulatory role unique to grasses.

3.3.3 Gain-of-function and dominant repression of *BdMYB48* results in reciprocal whole plant phenotypes

To investigate the function of *BdMYB48*, gain-of-function lines were developed by over-expressing the full-length coding region under the maize ubiquitin promoter (*MYB48-OE*). Similarly, dominant repressor lines were generated by over-expressing the full-length coding region fused to a 39 base pair dominant repressor sequence (*MYB48-DR*). Multiple independent events were generated and tested for each transgene. In general, the *BdMYB48* gain-of-function and dominant repressor plants exhibited reciprocal phenotypes. Two weeks after germination, both lines had similar stature, but phenotypically diverged when stem internodes began to elongate (Fig. 3.3A). When the inflorescence emerged from the flag leaf, *MYB48-OE* lines were slightly taller than the control lines and the dominant repressor lines were significantly shorter. These phenotypes persisted throughout development until plants completely senesced (Fig. 3.3B). Inflorescence emergence was significantly delayed by about 5 days in both lines compared to the controls (Fig. 3.3C). Moreover, they differed significantly in above ground biomass. At complete senescence, above ground biomass yield was significantly greater for *MYB48-OE* plants and reduced for *MYB48-DR* plants (Fig. 3.3D). While there was no significant difference in plant height between control and *MYB48-OE* lines,

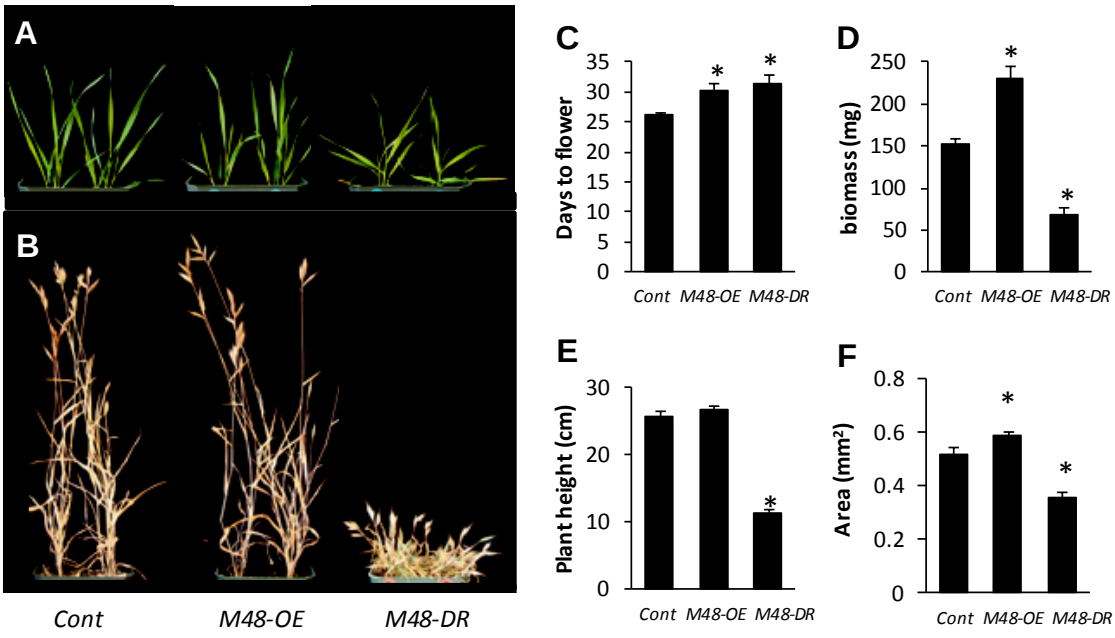

Figure 3.3. BdMYB48 is an activator of plant above ground biomass accumulation. Vector control (left), *M48-OE* (center), and *M48-DR* (right) plants were planted at the same time. All three lines appeared similar two weeks following germination (A). When completely senesced, *M48-DR* lines remained dwarf (B). Number of days before the first inflorescence was visible (C). Total above ground biomass (D), plant height (E), and stem cross section area of the first internode (F) at complete senescence. Twelve to sixteen individuals from three independent events were analyzed for each trait. * $p < 0.05$.

MYB48-DR lines were dramatically shorter (Fig. 3.3E). Leaf, stem node and internode count were unchanged; however, the internodes were not fully elongated in *MYB48-DR* plants. Possible explanations for a short internode include fewer cells, shorter cells, or both.

To test these possibilities we next examined propidium iodide treated longitudinal sections of the first internode using confocal microscopy. An equivalent number of long cells were observed among the three lines, but the *MYB48-DR* long cells were not as elongated (Fig. 3.4). In addition, the transverse stem cross section area differed among the three lines at senescence. Relative to control, *MYB48-OE* plants had a larger stem cross section area and *MYB48-DR* plants had a significantly smaller area (Fig. 3.3F). This observation may account for some of the differences observed in above ground biomass.

3.3.4 BdMYB48 regulates secondary cell wall lignification and biofuel conversion efficiency

To further investigate the function of BdMYB48, stem cross sections were analyzed using bright field light microscopy for changes in vascular patterning and composition. Vascular bundle shape and arrangement appeared similar in all three transgenic lines; however, a striking difference was observed in the cells between the vascular bundles when stained with a lignin-indicator dye, phloroglucinol-HCl (Fig. 3.5A-C). The interfascicular fiber region of *MYB48-OE* sections were bright red and the *MYB48-DR* sections were yellow relative to the control sections indicating the presence of very little lignin. However, the color intensity of the vascular bundles was similar among the three transgenic lines. The striking change in histochemical staining led us to investigate the lignin content in these lines. Fully senesced pulverized stem tissue was

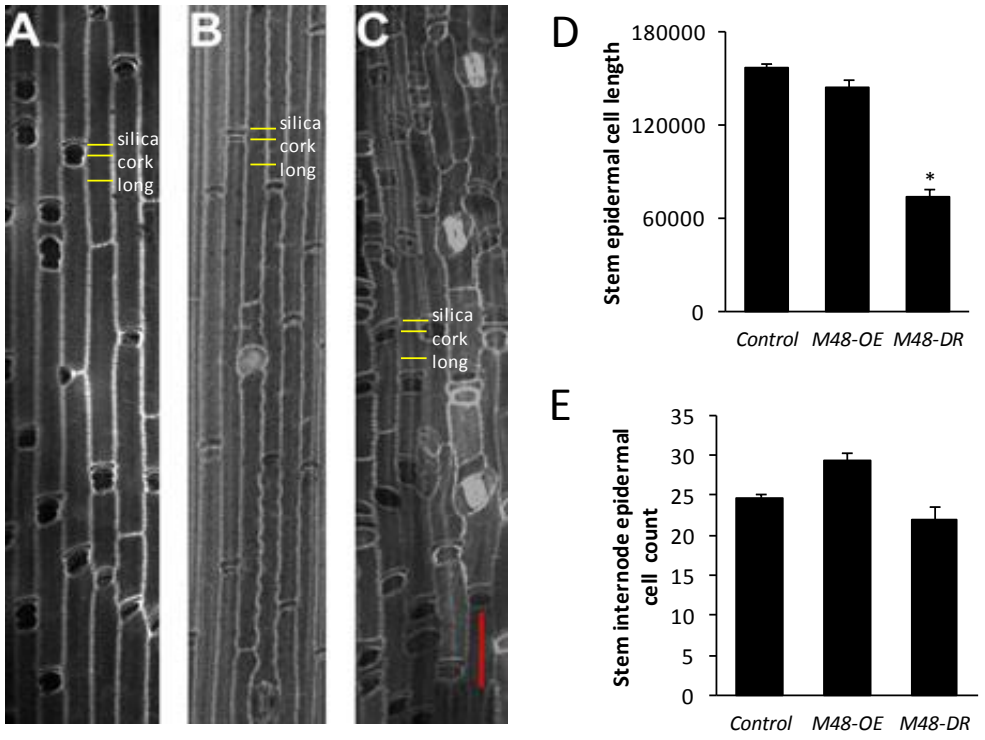

Figure 3.4. *MYB48-DR* plants are dwarf due to non-elongated stem cells. Longitudinal stem sections illustrating the cell length of control (A), *M48-OE* (B), and *M48-DR* (C). Confocal images of propidium iodide stained longitudinal stem sections of the first internode of flowering stems. Scale bar = 50 µm. Stem epidermal cell length (D), and stem internode epidermal cell count (E) of the first internode of flowering stems.

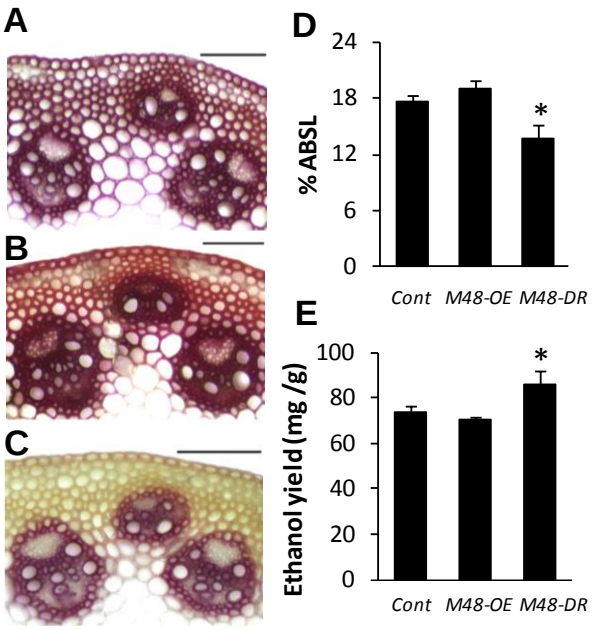

Figure 3.5. BdMYB48 is an activator of stem lignin accumulation. First internodes of fully senesced plants were hand sectioned and stained with phloroglucinol-HCl and representative images are illustrated (A-C). Compared to the control (A) *M48-OE* stem sections (B) stained a dark orange red color and *M48-DR* stem sections (C) were stained yellow in the interfascicular region and a less intense red color in the vascular bindles. Scale bar = 0.1 mm. Acetyl bromide soluble lignin content (D) and Ethanol yield (E) of completely senesced stem tissue. Pulverized stem tissue from six to sixteen individuals from three independent events were analyzed for each line.* $p < 0.05$.

assayed for acetyl bromide soluble lignin (ABSL) content. There was a slight increase in lignin content in *MYB48-OE* stems and a significant decrease in *MYB48-DR* stems (Fig. 3.5D). Considering lignin content is generally inversely correlated with bioconversion efficiency phenotypes, we measured ethanol yield after culturing senesced stems with *Clostridium phytofermentans*. As expected, a decrease in ethanol yield was observed for *MYB48-OE* lines and conversely a significant increase for *MYB48-DR* lines (Fig. 3.5E).

3.3.5 BdMYB48 is an activator of secondary cell wall biosynthesis and a regulator of cell wall thickening

The interfascicular fiber walls were examined in greater detail using transmission electron microscopy due to the striking changes observed in lignin staining and overall stem area. The first internode of the tallest stem when the inflorescence had just emerged from the flag leaf was fixed and sectioned using an ultra-cut microtome (Fig. 3.6). In comparison to the cell wall thickness of the control samples, *MYB48-OE* walls were thicker and *MYB48-DR* walls noticeably thinner. These results suggest that changes in secondary wall thickness may account for the overall differences observed in stem area and the above ground biomass (Fig. 3.4 D-F). Based on these results it is evident that *BdMYB48* plays an important role in activating secondary wall thickening.

3.3.6 *BdMYB48* regulates cellulose and lignin associated gene expression

To investigate the transcriptional function of *BdMYB48*, I measured gene expression in *BdMYB48* transgenic mutants. The tallest stem from developmentally equivalent plants was collected and flash frozen when the inflorescence was just visible from the flag leaf. Quantitative real time PCR (QRT-PCR) was utilized to examine transcript abundance of the transgenes and cell wall genes. Presence of the *MYB48-OE*

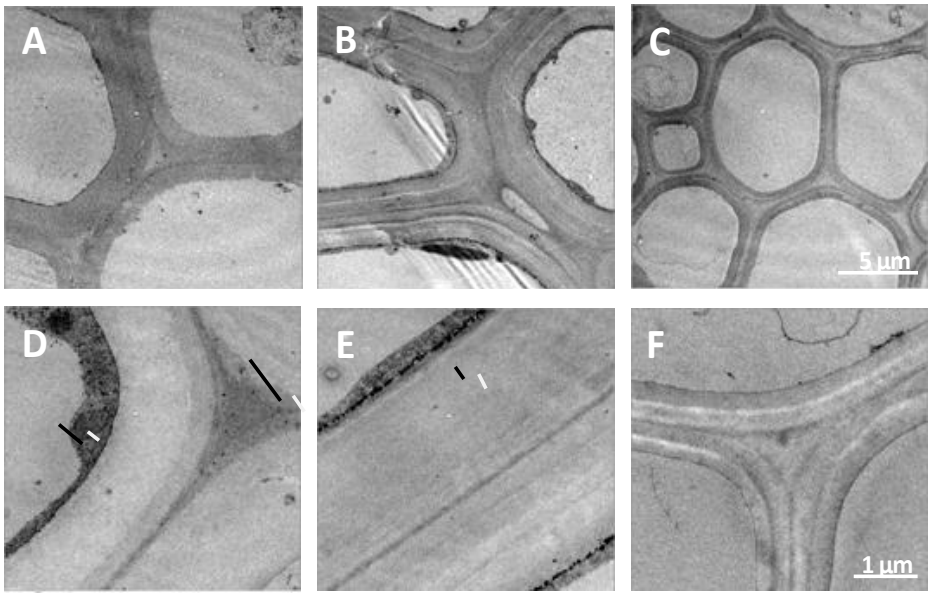

Figure 3.6. BdMYB48 is an activator of interfascicular fibers secondary cell wall thickening. Representative transmission electron micrographs illustrating the cell wall thickness of cells in the interfascicular fiber region of the first internode of inflorescence stems of control (A,D), *M48-OE* (B,E) and *M48-DR* (C,F). D-F are 5X magnified images of A-C. Black and white lines indicate the thickness of secondary and primary walls, respectively. Scale bars =5 µm (A-C) and 1 µm (D-F).

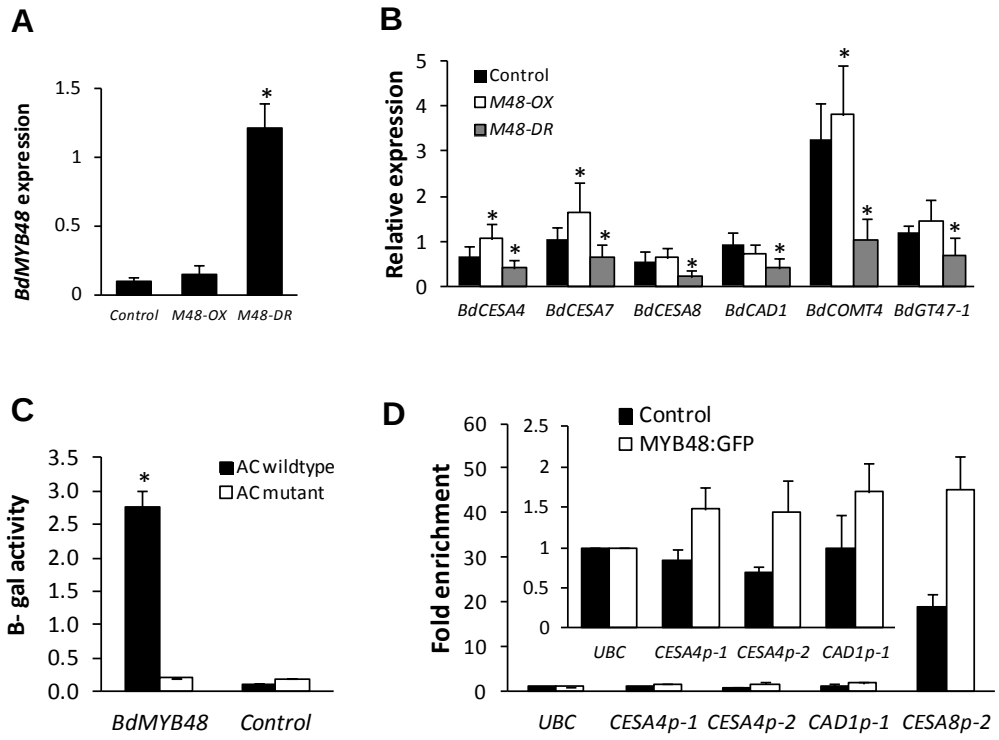

Figure 3.7. BdMYB48 directly interacts with an AC-like sequence motif to activate cell wall gene expression. Relative expression of *BdMYB48* (A) and secondary cell wall genes (B) in control, *M48-OE* and *M48-DR* stems. The tallest stem was collected from developmentally equivalent plants when the inflorescence was first visible from the flag leaf. Nine individuals from three independent events were analyzed in triplicate using QRT-PCR and normalized relative to two housekeeping genes, *BdUBC18* and *BdGapDH*. β-galactosidase activity illustrating a preferential interaction between the AC–II element and BdMYB48 protein in yeast (C). Values are mean of three independent yeast transformations. Relative fold enrichment of cell wall promoter fragments (D). Three biological samples were analyzed for control and MYB48:GFP in triplicate using QRT-PCR and normalized to input DNA. * $p < 0.05$.

transgene resulted in a slight increase in the abundance of the *BdMYB48* transcript and the *MYB48-DR* transgene resulted in a significant increase in the total *BdMYB48* transcript (Fig. 3.7A). Three cellulose genes, two lignin genes and one hemicellulose gene were analyzed using QRT-PCR. Among the three *CELLULOSE SYNTHASE A* genes, *BdCESA4* and 7 were significantly up regulated in *MYB48-OE* plants. All three *CESA* genes were significantly down regulated in *MYB48-DR* plants (Fig. 3.7B). A similar expression pattern was observed for the lignin gene, *BdCOMT4*, a significant up regulation in *MYB48-OE* plants and a significant down regulation in *MYB48-DR* plants. The second lignin gene *BdCAD1* exhibited no significant change in the transcript level in *MYB48-OE* plants however it was significantly down regulated in *MYB48-DR* plants (Fig. 3.7B). The transcript levels of the hemicellulose gene *BdGT47-1* was moderately increased in the *MYB48-OE* plants and significantly decreased in *MYB48-DR* plants. As demonstrated above, *BdMYB48* influences the expression of cellulose, hemicellulose and lignin gene expression. Changes in *BdCESA4/7/8*, *BdGT47-1*, *BdCOMT4* and *BdCAD1* transcript levels along with the changes observed in cell wall thickness and composition implies that BdMYB48 activates the transcription of secondary cell wall biosynthetic genes. It is reasonable to predict there will be transcriptional changes associated will other cell wall genes that are functional in a similar capacity in cellulose and lignin biosynthesis. We next searched the *cis*-regulatory regions of *BdCESA4/7/8*, *BdCOMT4* and *BdCAD1* for known regulatory elements. A sequence similar to the AC-ll element (Hatton et al., 1995) (ACCAAC) is upstream of the lignin genes *BdCAD1* and *BdCOMT4* and the cellulose genes *BdCESA4/7/8* (Fig. 3.8). A heterologous system was used to test for an affinity between this common motif and BdMYB48 protein. Four adjacent copies

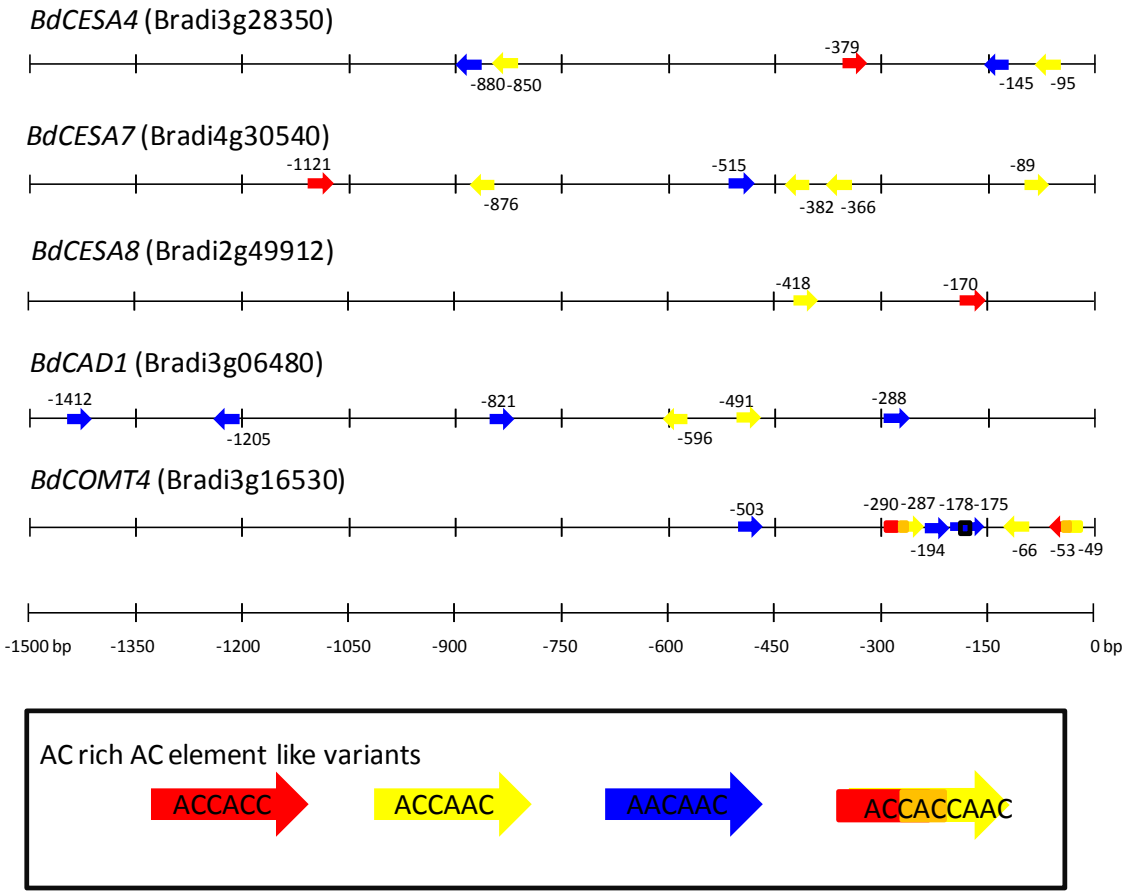

Figure 3.8. Variants of the AC element are present in *Brachypodium distachyon* cellulose and lignin gene promoters. 1500 bp upstream of the transcription start site was analyzed for the presence of the AC-like elements.

of the identified AC-II were fused to the LacZ reporter gene and stably integrated into the yeast genome. A similar strain was developed using a mutated motif (TTTAAC) as a negative control. BdMYB48 preferentially activated the promoter containing the AC-II element, but not the mutated version (Fig. 3.7C).

3.3.7 BdMYB48 directly interacts with secondary cell wall biosynthetic genes *in vivo*

Altogether these results support the notion that BdMYB48 directly activates cellulose and lignin gene expression by direct interaction with their *cis*-regulatory regions. To investigate these interactions *in vivo*, transgenic lines were generated with a GFP tagged BdMYB48 (*MYB48-GFP-OE*). These lines were used for chromatin immunoprecipitation along with control plants (*GFP-OE*). Immunoprecipitated chromatin was used as template for QRT-PCR to assay for enrichment of specific cell wall gene *cis*-regulatory regions. The results revealed a BdMYB48 specific enrichment of *BdCESA4*, *BdCESA8*, and *BdCAD1* promoter fragments (Fig. 3.7D). Moreover, all the enriched fragments contained an AC-like element (Fig 3.8). This provides evidence for direct binding of BdMYB48 to cell wall gene promoters.

3.4 Discussion

I focused on the MYB family of transcription factors and identified *BdMYB48* as a potential candidate for cell wall regulation based on its expression profile. *BdMYB48* transcript was abundant in stem, mirroring the expression profile of characterized *A. thaliana* cell wall regulators involved in promoting secondary cell wall thickening (Zhong et al., 2007; McCarthy et al., 2009). *BdMYB48* transcript was mainly localized in the interfascicular fibers, xylem and somewhat to the epidermis. BdMYB48 protein

shares close homology to *AtMYB46/83*, which are mainly expressed in xylem fibers (Zhong et al., 2007; McCarthy et al., 2009), but another protein, BdMYB1 is their closest ortholog. All the characterized MYB proteins are localized to the xylem and fibers and none have been reported with interfascicular fiber specific localization. Interestingly BdMYB48 has a strong localization pattern in the interfascicular fiber cells. The difference observed in tissue specific localization among BdMYB48 and AtMYB46/83 could be attributed to the aforementioned phylogenetic separation.

BdMYB48 over-expression and dominant repression transgenes were used for *in planta* functional characterization. The timing of inflorescence emergence was the only trait that exhibited a similar and significant effect. The whole plant phenotypes were dramatically different; *MYB48-OE* resulted in a larger plant with greater biomass. This has not been reported for the over-expression of *A. thaliana* MYB46/83/58 or 63 (Zhong et al., 2007; McCarthy et al., 2009; Zhou et al., 2009). The increase in above ground biomass appears to be in part the product of larger stems with thicker secondary cell walls. Thicker secondary cell walls were also observed in *A. thaliana* MYB46/83 over-expressers (Zhong et al., 2007; McCarthy et al., 2009). On the other hand, *MYB48-DR* resulted in severely dwarf plants with significantly diminished above ground biomass. Dwarfism of *MYB48-DR* was mainly attributed to the internodes cell length but not the cell count. *MYB48-DR* stem cells were significantly shorter compared to the control stem cells resulting in a significantly dwarf plant. These plants also had significantly smaller stems with thinner interfascicular fiber cell walls. AtMYB46/83 dominant repressors were also shown to have significantly thinner secondary cell walls in vessels and fibers. As a result these plants demonstrated a pendant phenotype (Zhong et al., 2007; McCarthy

et al., 2009). However unlike BdMYB48, dominant repression of AtMYB46/83 did not cause dwarfism. The difference seen between BdMYB48 and AtMYB46/83 plant height phenotypes under dominant repression could be the result of a functional difference between these MYB proteins.

Another striking observation was the difference in lignification in stem cross sections. Senesced plants were used for histo-chemical analysis and lignin composition analysis. Over-expression of *BdMYB48* resulted in a moderate increase in the lignin content in whole stems and dominant repression resulted in a significant decrease. Based on phloroglucinol-HCl staining it is evident that BdMYB48 has a profound impact on the interfascicular fiber cell wall lignification. BdMYB48 transcripts were also localized to the same cell types in stem cross sections mainly to the interfascicular fibers and somewhat to the xylem and epidermis. Transcript localization and histo-chemical analysis of stems further supports an important role for MYB48 in interfascicular fiber cell wall regulation. Reciprocal lignification patterns have been observed in histo-chemical studies performed on AtMYB46/83 stem sections, where over expression of the genes resulted in ectopic lignification in parenchymatous cells and dominant repression results in the absence of secondary cell wall lignification (Zhong et al., 2007; McCarthy et al., 2009). Even though we observed reciprocal phenotypes for total stem lignin, the prominent effect of BdMYB48 dominant repression on interfascicular fiber cells is a unique observation. As expected, over expression of *BdMYB48* upregulated secondary cell wall biosynthesis genes. Moreover, dominant repression resulted in a significant down regulation of the same cell wall genes. These findings argue that BdMYB48 is an activator of secondary cell wall biosynthesis.

Some MYB proteins are known to interact with AC-like elements and either activate or repress the transcription of downstream targets (Lois et al., 1989; Hatton et al., 1995; Raes et al., 2003, Zhong and Ye, 2012, Kim et al., 2013). Next I investigated the promoters of genes involved in secondary cell wall biosynthesis for over represented *cis*-regulatory elements. As suspected *BdCESA4/7/8*, *BdCAD1* and *BdCOMT4* genes all harbor AC-like elements in their promoter regions. In yeast, BdMYB48 preferentially bound the AC rich element relative to a mutated version. Therefore, it is plausible that BdMYB48 activates the downstream secondary cell wall genes via a protein-DNA interaction facilitated through above AC elements found in their promoter regions. This was further supported by the enrichment of cell wall gene promoters such as CESA4, CESA8 and CAD1 in BdMYB48 immuno precipitated chromatin. This provides *in planta* evidence for a direct interaction between BdMYB48 protein and the cell wall gene promoters. AtMYB46 has also been shown to interact with the AtCESA4/7/8 promoters *in planta* (kim et al., 2013). Therefore, based on the in silico predictions, synthetic AC element analysis and chromatin immunoprecipitation it is evident BdMYB48 is a direct activator of secondary cell wall biosynthesis in *B. distachyon*.

As might be expected for plants with a significant reduction in total lignin, a moderate increase in ethanol yield was observed following incubation of pulverized *MYB48-DR* stems with *Clostridium phytofermentans.* On the other hand *MYB48-OE* resulted in larger plants with a greater above ground bioamass but not more recalcitrant making BdMYB48 an obvious candidate for energy crop improvement. In conclusion we have demonstrated that BdMYB48 is a grass specific activator capable of regulating secondary cell wall biosynthesis. To our knowledge this is the first grass specific

activator characterized in a grass species. This protein has a greater impact on the interfascicular fiber cell walls which is ideal in manipulating energy crops without causing drastic effects to the vascular system.

Table 3.1. List of primers used in this study

Primer Name	Sequence
attB1MYB48_F	ACAAGTTTGTACAAAAAAGCAGGCTCTATGGGGCGGCACGCGGGCACT
attB2MYB48_R	ACCACTTTGTACAAGAAAGCTGGGTATCAAAAGTACTCGAGGTTGAAG
attB5rMYB48_NSR	ACAACTTTTGTATACAAAGTTGTAAAGTACTCGAGGTTGAAGTC
attB5CRES_F	ggggACAACTTTGTATACAAAAGTTGCTGTTGATCTTGATCTTGAATTGAGATTGGGT
attB2CRES_R	ggggACCACTTTGTACAAGAAAGCTGGGTATCAAGCAAAACCCAATCT
attB5GFP_F	ggggACAACTTTGTATACAAAAGTTGCTATG GTG AGC AAG GGC GAG GAG
attB2GFP_R	ggggACCACTTTGTACAAGAAAGCTGGGTATCA CTT GTA CAG CTC GTC CAT GCC
attB1GFP_F	ggggACAAGTTTGTACAAAAAAGCAGGCTCTATGGTGAGCAAGGGCGAGG
qPCRGT47-1_F	AATATAGCGCGCTGCATGTCCTC
qPCRGT47-1_R	AATATAGCGCGCTGCATGTCCTC
qPCRMYB48CDS_F	AGGAAACAGGTGGTCGCAGATTG
qPCRMYB48CDS_R	GCTTCTTCTTGAGGCAGCTGTTCC
qPCRUBC18_F	GGAGGCACCTCAGGTCATTT
qPCRUBC18_R	ATAGCGGTCATTGTCTTGCG
qPCRGAPDH_F	TTGCTCTCCAGAGCGATGAC
qPCRGAPDH_R	CTCCACGACATAATCGGCAC
qPCRCAD1_F	AGGATAGAATGGGCAGCATCGC
qPCRCAD1_R	ATCTTCAGGGCCTGTCTTCCTGAG
qPCRCOMT4_F	TGGAGAGCTGGTACTACCTGAAG
qPCRCOMT4_R	CGACATCCCGTATGCCTTGTTG
qPCRCESA4_F	GCGTTTCGCATACACCAACACC
qPCRCESA4_R	ACTCGCTAGGTTGTTCAGTGTGG
qPCRCESA7_F	GCGATTCGCCTACATCAACACC
qPCRCESA7_R	GGCTGGCAAATGTGCTAATCGG
qPCRCESA8_F	CAAAGCACAAGTTCCGCCTGTG
qPCRCESA8_R	TGGCTCGTATGCATCTGTCAAATC
qUBC18p_F	AAGGCTTGAACATGACAGCA
qUBC18p_R	ATGAAATGGGCACCTGAAAA
qCESA4p-1_F	TGCAAAAGGCCTCAGCTAAT
qCESA4p-1_R	TGGTGGCATACAAAACCTCA
qCESA4p-2_F	CTTCACGCTCACTCACCATC
qCESA4p-2_R	CGGAAGACCAAGAATGAAGC
qCESA8p-2_F	CTTGCTCTCACCGTCCTGA
qCESA8p-2_R	GGTTTCGAAGCGAAGGTGAC
qCADp-1_F	TTCCTATTGCAAGTACATCATGC
qCADp-1_R	TATCGTGTGCTGCCCATCTA
qCADp-3_F	AAACTGTTTGAAAATCAAATCTGC
qCADp-3_R	GGAAGTTGTCGTGGGATCAG
qCOMTp-2_F	TCGAGAAATAATGGTTCAGACG
qCOMTp-2_R	AGATATACTTGTTGTCGCGAAG
Hpt_F	AGAATCTCGTGCTTTCAGCTTCGA
Hpt_R	TCAAGACCAATGCGGAGCATATAC
Myb48_probe_aF	GCATGGCGCATTTTGACTTCAACC
Myb48_probe_aR	CTA CAC AAT GTT CAC ATT CCT ATA CC

CHAPTER 4

GRASS NAC REPRESSOR OF FLOWERING SUPPRESSES FLORAL TRANSITION AND SECONDARY WALL SYNTHESIS IN *Brachypodium distachyon*

4.1 Introduction

The large NAC (NAM, ATAF1/2, and CUC2) transcription factor family is comprised of plant-specific proteins. Several are well characterized with respect to the regulation of secondary cell wall biosynthesis and are thought of as key regulators in this process. *SECONDARY WALL ASSOCIATED NAC DOMAIN PROTEIN1* (*SND1*; also known as *NST3*) is a fiber specific *Arabidopsis thaliana* secondary cell wall regulator. *SND1* is the most extensively characterized among these NAC proteins that activate the expression of cellulose and lignin genes. Over-expression of *SND1* results in ectopic deposition of secondary cell walls in parenchymatous cells. Conversely, dominant repression results in a significant reduction in secondary cell wall deposition in fibers (Zhong et al., 2006). *NAC SECONDARY WALL THICKENING PROMOTING FACTOR1* and *2* (*NST1*/*2*) regulate secondary cell wall deposition in *A. thaliana* anther endothecium (Mitsuda et al., 2005). *NST1* and *SND1* are thought to play redundant regulatory roles and the double mutant causes a drastic reduction in the expression of secondary cell wall genes and a significant reduction in cellulose, xylan and lignin biosynthesis (Mitsuda et al., 2007; Zhong et al., 2007). *VASCULAR RELATED NAC DOMAIN6* and *7* (*VND6*/*7*) specifically expressed in protoxylem and metaxylem are required for vessel development in *A. thaliana* roots (Kubo et al., 2005; Demura and Fukuda, 2007). Moreover, over-expression of *VND6*/*7* results in ectopic vessel development and dominant repression

results in the loss of protoxylem and metaxylem development (Kubo et al., 2005). Collectively these five NAC proteins are considered master regulators capable of activating the entire secondary cell wall biosynthesis process by interacting and activating downstream transcription factors including *SND2*, *SND3*, *MYB46*, *MYB54*, *MYB42*, *MYB103*, *MYB85*, *MYB52*, *MYB69* and *KNAT7* (Zhong et al., 2007; Zhong et al., 2008).

To date, very little functional characterization has been done for the two wall biosynthesis associated NAC transcription factors *SND2* and *SND3*. These two proteins are distantly related to the previously described proteins when considering DNA-binding and transcriptional activation domain sequences (Hu et al., 2010; Wang et al., 2011). Even though SND1, SND2, and SND3 share the same identifier, SND1 is closely related to NST1 and NST2 and distantly related to SND2 and SND3. In a recent study NAC protein from 19 plant species were analyzed to understand the evolutionary relationship among NST, SND and VND sub groups. In this comparison, SND2 orthologs were found in both dicot and monocot species. However, SND3 was only found within dicot species suggesting it is a dicot specific regulator (Yao et al., 2012). Both SND2 and SND3 are highly expressed in *A. thaliana* stems mirroring a similar expression profile of many characterized cell wall regulators (Yao et al., 2012). Over-expression of *SND2* increases the secondary wall thickness in interfascicular and xylary fibers in *A. thaliana* (Zhong et al., 2008). Conversely, dominant repression results in a drastic reduction of secondary wall thickness in fibers. Moreover, *SND2* has been shown to regulate cellulose, hemicellulose, and lignin biosynthesis genes (Hussey et al., 2011). *SND2* orthologous proteins have also been characterized in poplar. Over-expression of the *SND2* ortholog

PopNAC154 in poplar reduced plant height; however, unlike in *A. thaliana*, over-expression had no effect on secondary cell wall thickness (Grant et al., 2010). On the other hand, dominant repression of the *SND2* ortholog (*PtSND2*) resulted in a significant reduction of secondary cell wall thickness in xylary fibers (Wang et al., 2013). These diverging results between *A. thaliana* and poplar suggest differing regulatory functions by the various SND2 orthologs in different plant species. This contrasts with SND1, where the rice and maize counterparts were able to complement the *snd1/nst1* double mutant and cause similar over-expression phenotypes in *A. thaliana* (Zhong et al., 2011; Yoshida et al., 2013).

Secondary cell walls, by weight, represent the majority of biomass in stem tissue. While grasses can produce abundant stem before flowering, *A. thaliana* grows as a rosette of leaves and only elongates stem following floral induction. The transition from vegetative to reproductive growth is an important phenomenon that is achieved through the integration of developmental and environmental cues such as day length and temperature. The molecular mechanisms of flowering have been well characterized in the model dicot *A. thaliana*. The mobile signal responsible for this transition is florigen, which is encoded by *FLOWERING LOCUS T* (*FT*). It moves from leaves to the shoot apical meristem to activate meristem identity genes, including *APETALA1* (*AP1*), *CAULIFLOWER* (*CAL*) and *FRUITFUL* (*FUL*), to transition the shoot apical meristem into a flower bud (Fig. 4.1) (Ream et al., 2014). *FT* is expressed following exposure to cold through epigenetic repression of repressors of flowering: *FLOWERING LOCUS C* (*FLC*) and *SHORT VEGETATIVE PHASE* (*SVP*) (Michaels and Amasino, 1999; Alexandre and Hennig, 2008; Gu et al., 2010). Long day photoperiod can also induce

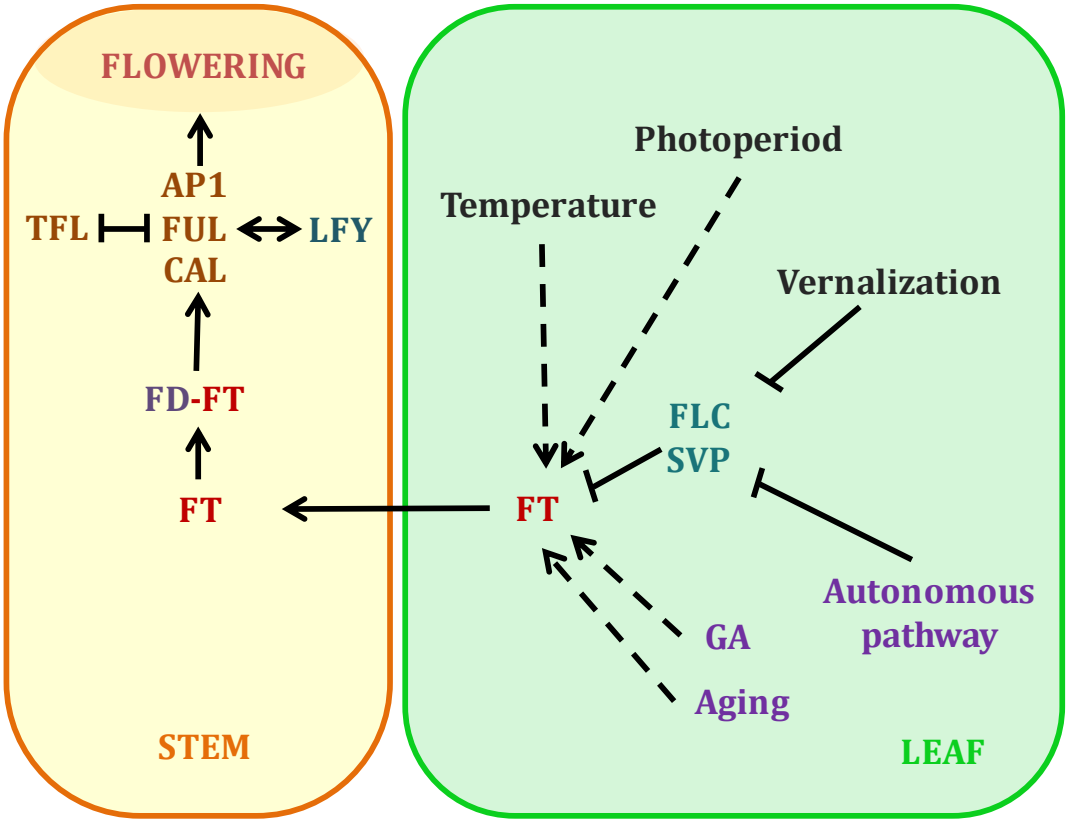

Figure 4.1. A simplified model for flowering in *Arabidopsis thaliana*. *FLOWERING LOCUS T* (*FT*) is activated by temperature, photoperiod, clock, hormones and aging. *FLOWERING LOCUS C* (*FLC*) and *SHORT VEGETATIVE PHASE* (*SVP*) are the main repressors of *FT*. *FLC* is activated by *FRIGIDA* (*FRI*) and is repressed by the autonomous pathway and cold. When the timing is suitable for flowering FT protein moves through the phloem to the stem and interacts with FD. FT-FD complex is responsible for activating the floral integrator gene *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*) and the floral meristem identity genes *APETALA 1* (*AP1*), *CAULIFLOWER* (*CAL*), *FRUITFUL* (*FUL*) and *LEAFY* (*LFY*) which results in the vegetative to reproductive transition.

flowering through CONSTANS induction of *FT* expression (Imaizumi and Kay, 2006). Once activated, the mobile signal can complex with the bZIP transcription factor FD to form an FT-FD complex in the stem (Jaeger and Wigge, 2007). This complex is responsible for initiating the transcription of meristem identity genes.

Many of these same components have been identified as having a role in the control of flowering in grasses. Similar to *A. thaliana*, *FT* (also known as *VERNALIZATION3* in barley and wheat) is the main floral integrator of developmental and environmental cues and turning on floral meristem identity genes including *VRN1* (Greenup et al., 2010; Wu et al., 2013; Lv et al., 2014; Ream et al., 2014). In temperate cereals, *FT* expression is induced under long days and *VRN2* is repressed in response to vernalization (Greenup et al., 2010). *VRN1* is the cereal *AP1* ortholog, but contrary to *A. thaliana* AP1, *VRN1* is expressed in both leaves and in the floral meristem (Trevaskis et al.; Alonso-Peral et al., 2011). *VRN1* is activated by low temperatures and as a result represses *VRN2* during winter allowing flowering in spring (Greenup et al., 2009). *PHOTOPERIOD1-H1* (*PPD*-H)1 is another component responsible for activating *FT* under reduced levels of VRN2. Once activated, FT protein binds *FLOWERING LOCUS D LIKE2* (FDL2), the FD ortholog found in cereals and initiates VRN1 transcription in the meristem (Li and Dubcovsky, 2008) triggering the transition to flowering (Fig. 4.2). Regardless of the similarities and distinctions for the signal transduction pathways that govern floral induction between *A. thaliana* and grasses, we expect fundamental differences in how that regulation relates to stem formation. Here we describe the function of a NAC transcription factor, *GRASS NAC REPRESSOR OF FLOWERING* (*GNRF*), using over-expression and loss-of-function mutant lines.

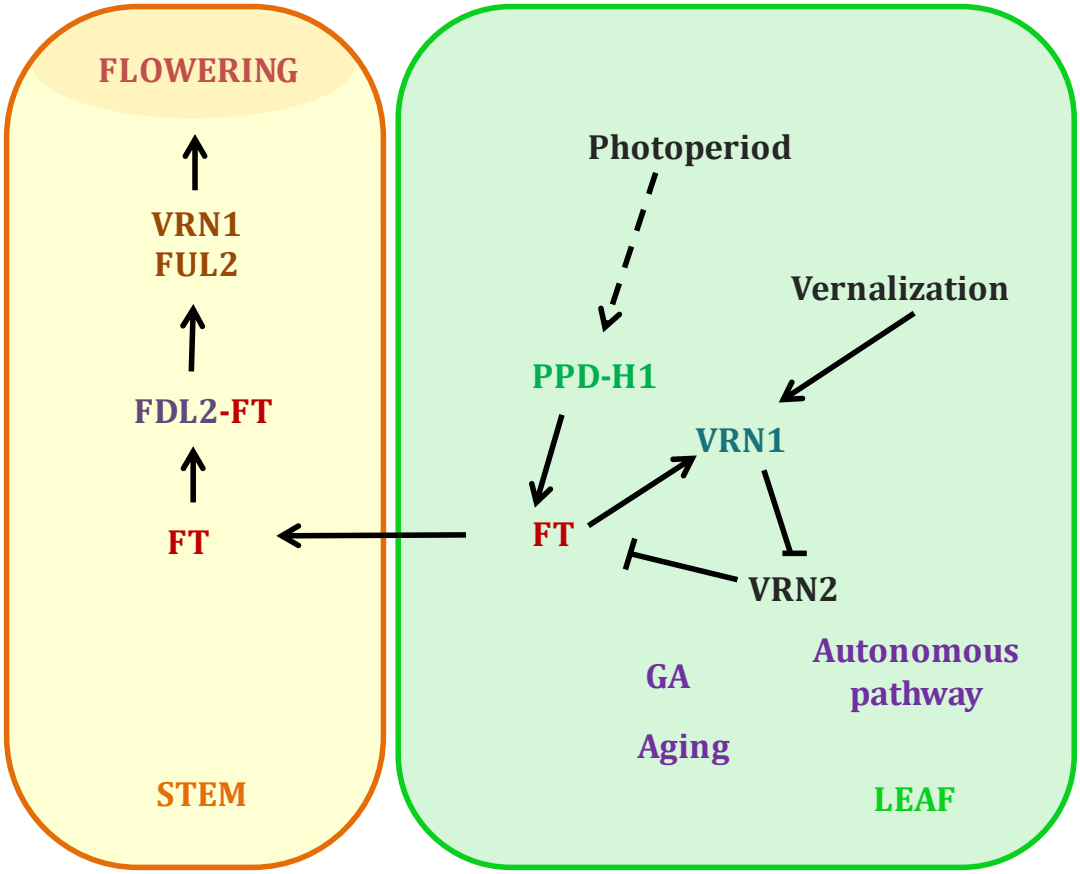

Figure 4.2. Proposed model for flowering in grasses. *FLOWERING LOCUS T* (*FT*) is activated by long days via *PHOTOPERIOD1-H1* (*PPD-H1*) and is repressed by *VERNALIZATION 2* (*VRN2*). During winter *VRN1* is activated and represses *VRN2* to initiate flowering in spring. Once activated FT moves through the phloem into the stem where it complexes with *FLOWERING LOCUS D LIKE2* (*FDL2*) to activate the floral meristem identity gene *VRN1*.

4.2 Material and Methods

4.2.1 Phylogenetic analysis

Amino acid sequences of *A. thaliana* NAC proteins SND2 and SND3 and the orthologous proteins from *O. sativa*, *Z. maize*, *P. trichocarpa* and *Sorghum biocolor* were downloaded from plant transcription factor data base (PlnTFDB; http://plntfdb.bio.uni-potsdam.de/v3.0). *B. distachyon* NAC sequences were downloaded from phytozome (http://www.phytozome.net/). A neighbor-joining phylogeny was constructed with 1000 bootstrap permutations using MEGA 5.0. An outgroup including *A. thaliana* NST1, NST2 and SND1 were also included in this analysis.

4.2.2 *GNRF-OE* plasmid construction

Full-length coding region of Bradi2g46197 (GNRF) was PCR amplified using Bd21-3 cDNA with the stop codon using Phusion high-fidelity DNA polymerase (New England Bio Labs) and cloned into pENTR/D-TOPO vector (Invitrogen). Sequence-confirmed entry clone was recombined with pOL001 ubigate ori1 destination vector (modified from pOL001, described in (Vogel et al., 2006) to generate the GNRF gain-of-function construct (*GNRF-OE*). Above construct was transformed into *Agrobacterium tumefaciens* strain *AGL1* via electroporation for *B. distachyon* calli transformations.

4.2.3 Plant material, growth conditions and calli transformation

Brachypodium distachyon (L.) line Bd21-3 was used as the genetic background. Seeds were imbibed a week and planted as previously described by Handakumbura et al., 2013. *Brachypodium distachyon* calli transformation was carried out as described by (Vogel and Hill, 2008) with minor modifications described in Handakumbura et al., 2013. Once shoots and roots were established primary transgenics were transplanted to soil.

Primary transgenic were PCR confirmed by the amplification of the hygromycin resistance gene and were propagated for two subsequent generations to obtain T_3 individuals.

4.2.4 *gnrf-1* mutant allele isolation

A homozygous *gnrf-1* mutant allele was isolated from the TILLING (Targeting Induced Local Lesions in Genomes) mutant population generated at INRA-Versailles by restriction enzyme digestions. Leaf genomic DNA was isolated from each individual using the method described by Handakumbura et al., 2013. A 614 bp fragment spanning the region with the mutations was PCR amplified using Taq DNA polymerase and the purified amplicons were digested with BaeGI at 37 °C. The digestion results in two fragments of 141 bp and 473 bp for the wild-type allele and a single undigested fragment for the mutant allele.

4.2.5 Genotyping and phenotyping

Genomic DNA was extracted from T_3 generation *GNRF-OE* plants and M2 generation *gnrf-1* plants as previously described (Handakumbura et al., 2013). *GNRF-OE* plants were PCR confirmed for the hygromycin selectable marker gene and were used in subsequent experiments. Flowering was induced in about five percent of the *GNRF-OE* plants with excessive fertilizer treatment with a N-P-K 10:30:20 fertilizer. Homozygous *gnrf-1* mutants were confirmed by restriction enzyme digestion as described previously. Phenotypic data such as flowering time, plant height, above ground biomass at senescence were manually recorded.

4.2.6 RNA extraction and QRT-PCR

Total RNA was extracted from the first, second, and third internodes of the tallest stem at BBCH stage 5 and from the fourth leaf from the base of the tallest stem at BBCH stage 4.5 (Hong et al.) (Plant RNaeasy, Qiagen, Valencia, CA) as previously described by Handakumbura et al., 2013. First strand cDNA was synthesized from 1 μg of DNase (Qiagen) treated total RNA using Superscript III reverse transcriptase with oligo dT primers (Invitrogen, Grand Island, NY). Triplicate quantitative PCR reactions were performed as described in Chapter 3. *BdUBC18* (ubiquitin-conjugating enzyme 18) and *GapDH* were used as reference genes for normalization (Hong et al., 2008). QuantiPrime primer design tool was used for qPCR primer design (Arvidsson et al., 2008).

4.2.7 Microarray analysis

For the microarray analysis, three stem samples at BBCH stage 5 (Hong et al., 2011) were pooled for each biological replicate and hybridized in triplicate for each line. RNA was extracted using a kit (Plant RNaeasy, Qiagen, Valencia, CA) as described above. Samples were hybridized to a *B. distachyon* BradiAR1b520742 whole genome tiling array (Affymetrix, Santa Clara, CA). Based on the hybridization signals a significantly differentially expressed gene list was generated using a modified *t*-test (Tusher et al., 2001).

4.2.8 Histo-chemical analysis of stem lignification

Histo-chemical assays were performed on the first internode of tallest stem at complete maturity. Hand cut sections were stained with phloroglucinol-HCl as described in Chapter 3 for total lignin. Stained sections were observed under an Eclipse E200MV R microscope (Nikon) and imaged using a PixeLINK 3 MP camera.

4.2.9 Acetyl bromide soluble lignin measurement

Fully senesced pulverized stem tissue was used for acetyl bromide soluble lignin (ABSL) measurements as previously described (Foster et al., 2010) and briefly specified in Chapter 3. Three to six individuals were analyzed in triplicate for each line.

4.3 Results

4.3.1 Bradi2g46197 transcript is abundant in stem tissue

The secondary cell wall regulators characterized to date exhibit the signature expression profile of being highly expressed in tissues abundant in secondary cell walls relative to other tissues. To investigate Bradi2g46197 transcript abundance in leaf, stem, and root, a *B. distachyon* microarray data set was utilized (Handakumbura et al., 2013). Bradi2g46197 transcript abundance was approximately nine and three-fold greater in stem, relative to leaf and root, respectively (Fig. 4.3A). Grass stems are substantially enriched for secondary cell walls (Matos et al., 2013). Therefore, based on annotation and expression analysis, Bradi2g46197 is a potential candidate regulator of grass secondary cell wall biosynthesis.

4.3.2 Bradi2g46197 is the closest ortholog to SND2

Bradi2g46197 belongs to the plant specific NAC transcription factor family that is involved in a variety of processes. Some of the *A. thaliana* NAC proteins such as NST1/2, SND1/2/3 are specifically involved in the regulation of cell wall biosynthesis. To determine the protein similarity between Bradi2g46197 and the functionally characterized NACs, a phylogeny was constructed between the NAC protein sequences from *A. thaliana*, *B. distachyon*, *O. sativa*, *Z. maize*, *P. trichocarpa* and *Sorghum bicolor* using the MEGA 5.0 software. NST1, NST2 and SND1 were used as an out-group in order to better understand the relationship of Bradi2g46197 to the other dicot and

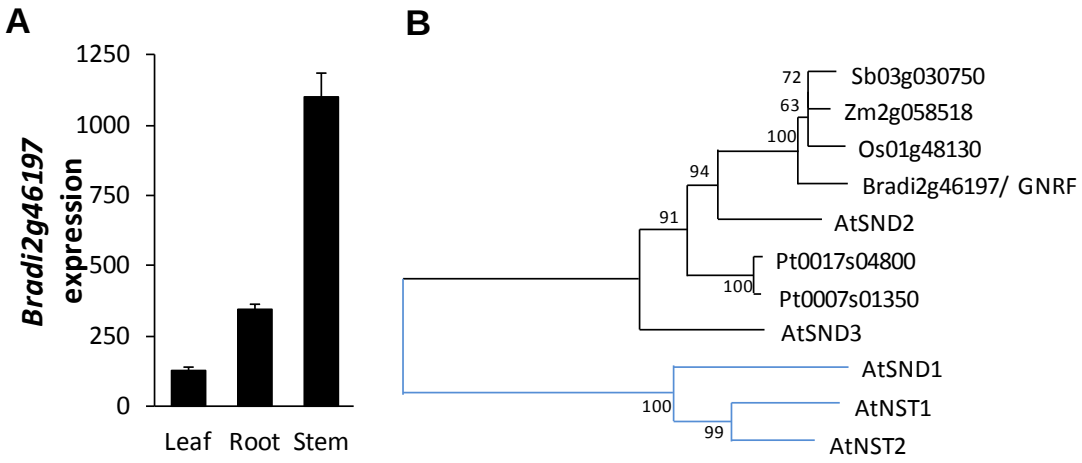

Figure 4.3. Relative expression and phylogeny of *Bradi2g46197*. (A) Relative transcript abundance of *Bradi2g46197* in leaf, root and stem tissue measured with a microarray. Mean ± standard deviation of three biological replicates. (B) A sub clade of the NAC phylogeny illustrating amino acid sequence similarity between *Arabidopsis thaliana*, *Brachypodium distachyon* ,*Oryza sativa*, *Zea maize*, *Populus trichocarpa and Sorghum biocolor*. A rooted neighbor-joining phylogeny was constructed using MEGA 5 with 1000 bootstrap permutations. Numbers on each branch indicate bootstrap support. Branches indicated in blue were used as an out group.

monocot NAC orthologs (Fig. 4.3B). Based on protein similarity, Bradi2g46197 is a grass ortholog to both SND2 and SND3. Both dicots included in this analysis, *A. thaliana* and poplar, have two orthologous proteins in this clade while the grasses have only one.

4.3.3 Over-expression of Bradi2g46197 results in persistent vegetative growth

To investigate the function of Bradi2g46197, gain-of-function lines were developed by over-expressing the full-length coding region under the maize Ubiqutin promoter. Multiple independent events were generated for this construct and at least three events were used in the subsequent experiments. A homozygous line with a nonsynonymous point mutation that modifies the fiftieth amino acid from a proline (P) to a leucine (L) was isolated from a TILLING (Targeting Induced Local Lesions in Genomes, (Dalmais et al., 2013) population to investigate the loss-of-function phenotypes. The aforementioned point mutation lies within the DNA binding domain in the N-terminus and predicted to alter protein function. Surprisingly, over-expression construct harboring transgenics remained vegetative until senescence (Fig. 4.4). These plants exhibited a branched and bushy phenotype and did not transition from vegetative to reproductive growth. We therefore named the Bradi2g46197 locus *GRASS NAC REPRESSOR OF FLOWERING* (*GNRF*). From here on, the gain-of-function mutant will be referred to as *GNRF-OE* and the loss-of-function mutant *gnrf-1*. Unlike *GNRF-OE* plants, which did not flower and were considerably shorter than control plants, *gnrf-1* plants flowered significantly earlier than the control plants by approximately 1.5 d (Fig. 4.5A) and were similar to control plants in stature (Fig. 4.5B). Moreover, the above ground biomass was significantly greater for *GNRF-OE* plants where as no significant difference was observed between *gnrf-1* plants and the controls (Fig. 4.5C).

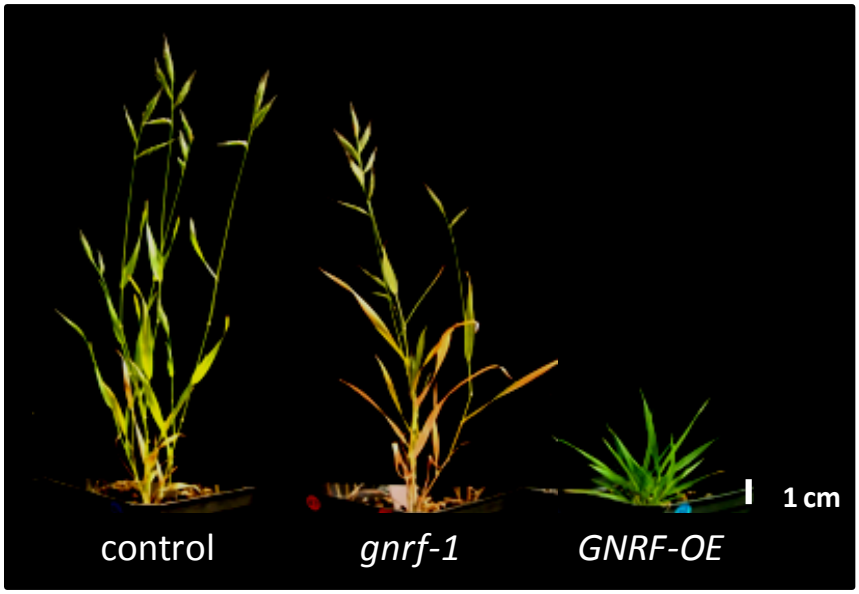

Figure 4.4. Plant phenotypes of control (left), *gnrf-1* (center) and *GNRF-OE* (right) at flowering stage. Images were captured at the time of control plant flowering. Scale bar = 1 cm.

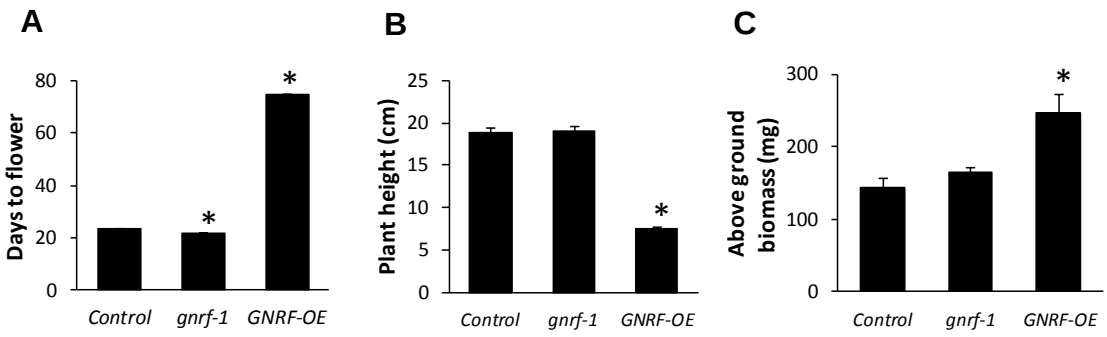

Figure 4.5. Whole plant phenotypes. (A) Days taken for the first flower to be visible from the flag leaf. Fifteen to forty eight individuals were analyzed for each line. Note that *GNRF-OE* remained vegetative even after 75 days, (B) Plant height at complete maturity, (C) Total above ground biomass at complete senescence. Fifteen to twenty one individuals were analyzed for each line. * $p < 0.05$.

4.3.4GNRF is a repressor of genes associated with cell wall, transport, and flowering

Considering the dramatic pleiotropic effects of the gain-of-function lines, a microarray experiment was performed using RNA from pooled stems from control, *GNRF-OE* and *gnrf-1* plants in order to isolate the potential effected pathways. Samples were hybridized to the *B. distachyon* BradiAR1b520742 whole genome tiling arrays and differentially expressed genes were identified using a modified *t*-test. No significant differences in gene expression were detected between control and *gnrf-1*. Some rather dramatic changes were observed within the *GNRF-OE* plants. While only 23 genes were up-regulated by *GNRF* over-expression, 372 genes were significantly down-regulated. One of the largest categories of repressed genes was transporters of numerous substrates: amino acids, ammonium, arsenite, carboxylate, lipids, peptides, potassium, silicon, sucrose, and sulfate (Table 4.1). Polysaccharide synthesis genes, lignin and lignin related genes, and lipid transfer genes were also abundant among the repressed genes. The most striking observation was the >50 fold repression of two floral meristem identity associated genes *VRN1*/*FUL1*/*MADS33* (Bradi1g08340) and *FUL2*/*MADS10* (Bradi1g59250). These two genes are the two closest orthologs to *A. thaliana AP1* and *FUL*. Apart from the above two genes, several other MADS-box, bHLH, WRKY, HB, NAC and MYB transcription factors were repressed in *GNRF-OE* stems. Another interesting observation is the 3-fold reduction of *BdMYB48* (Bradi2g47590). As I have previously described (Chapter 3) *BdMYB48* is a direct activator of secondary cell wall biosynthesis genes. The repression of the cellulose genes *CESA4*, *CESA7*, and *COMT4*, *4CL1*, *4CL3*, *CCoAMT* lignin genes in the *GNRF-OE* stems could be an indirect effect due to the repression of *BdMYB48*. As a result of the repression of many cellulose and

hemicellulose synthesis genes, *GNRF-OE* stems should show a significant reduction in the polysaccharide composition. Apart from lignin genes, many laccasses and peroxidases were also repressed in these mutants. These genes are necessary for the polymerization of lignin monomers and their subsequent polymerization within the cell wall (Boerjan et al., 2003). Moreover, several copper ion transporter genes were repressed. Copper ions act as a catalyst in the lignin polymerization process (Boerjan et al., 2003). Down regulation of many lignin and lignin associated genes along with the necessary catalysts should result in a significant reduction in cell wall lignification in *GNRF-OE* mutants.

4.3.5 *GNRF* represses meristem identity genes and floral integrators

Considering that *GNRF-OE* plants lacked the ability to transition from vegetative to reproductive growth, and many flowering associated genes were repressed as measured by microarray, flowering pathway genes were further investigated by QRT-PCR. I initially sought to validate the microarray results by analyzing the expression of *FUL1*, *FUL2* and *CENTRORADIALIS-LIKE1 HOMOLOGOUS TO TFL1* (*RCN2*) (Fig. 4.6). The primary stem was collected and flash frozen from developmentally equivalent plants at BBCH stage 5 (Hong et al., 2011) when the inflorescence had just begun to emerge from the flag leaf sheath. Eight to ten plants were individually analyzed for transcript abundance. In agreement with the microarray data, *FUL1*, *FUL2* and *RCN2* were significantly down-regulated in *GNRF-OE* stems. The same genes were modestly, but not significantly up-regulated in the *gnrf-1* stems.

Next we investigated relative transcript abundance of the mobile flowering signal. *FLOWERING LOCUST*, or florigen, is an activator of flowering and one of the terminal

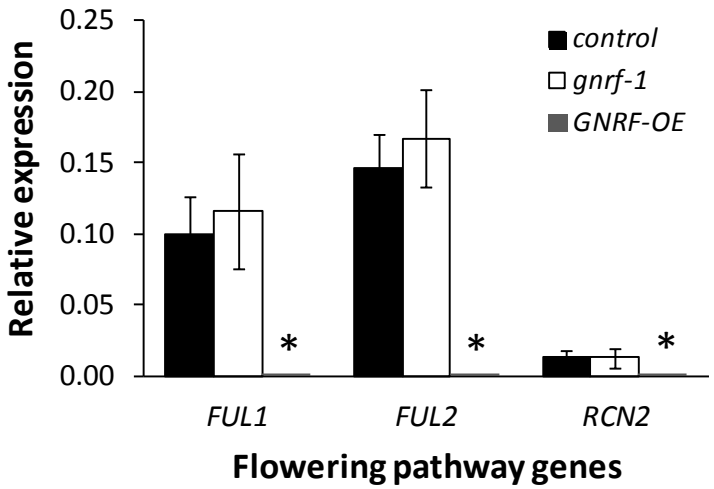

Figure 4.6. Transcript abundance of flowering pathway genes in stems. Relative expression of *FUL1*, *FUL2*, and *RCN2* genes in control, *gnrf-1* and *GNRF-OE* stems. Tallest stem was collected from developmentally equivalent plants when the inflorescence was first visible from the flag leaf. Eight to ten individuals from each line were analyzed in triplicate using QRT-PCR and normalized against two housekeeping genes. * $p < 0.05$.

genes in this pathway (Higgins et al., 2010). The protein functions as a mobile signal that moves from the leaf to the shoot apex to initiate flowering. Accordingly, leaf tissue was used to investigate the relative abundance of *FT* transcripts. As might be expected, *BdFT* was undetectable in *GNRF-OE* leaves and significantly up-regulated in *gnrf-1* leaves (Fig. 4.7A). In order to better understand this dynamic, I investigated several putative upstream repressors of *FT*. *FLOWERING LOCUS C* (*FLC*) and *SHORT VEGETATIVE PHASE* (*SVP*) act as immediate repressors of *FT* in *A. thaliana* and served as initial candidates in *B. distachyon* (Ruelens et al., 2013). While *SVP* expression was not altered in *GNRF-OE* (data not shown), expression of other FLC-like genes was altered (Fig. 4.7B-D). *BdMADS37*, *BdODDSOC1* and *BdODDSOC2* are *B. distachyon* orthologs of *A. thaliana FLC* and are collectively referred to as the FLC-like genes. *BdMADS37* was down-regulated in *GNRF-OE* leaves and *BbODDSOC1* and *BdODDSOC2* were significantly up-regulated in *GNRF-OE* leaves. The expression of *BdMADS37*, *BDODDSOC*, and *BdODDSOC2* were at similar levels to controls in *gnrf-1* leaves.

4.3.6 *GNRF* regulates genes associated with cellulose, xylan and lignin biosynthesis in stem tissue

Since NACs play a key regulatory role in secondary cell wall biosynthesis, I measured the gene expression of several cell wall genes in *GNRF-OE* and *gnrf-1* plants to investigate the transcriptional function of *GNRF*. Quantitative real time PCR was utilized to examine transcript abundance of the transgenes and candidate cell wall genes using the same cDNA samples used for detecting flowering pathway genes. As expected, *GNRF* was significantly up-regulated in *GNRF-OE* stems (Fig. 4.8). Three genes involved in cellulose biosynthesis namely *BdCESA4/7/8* and two genes associated with lignin

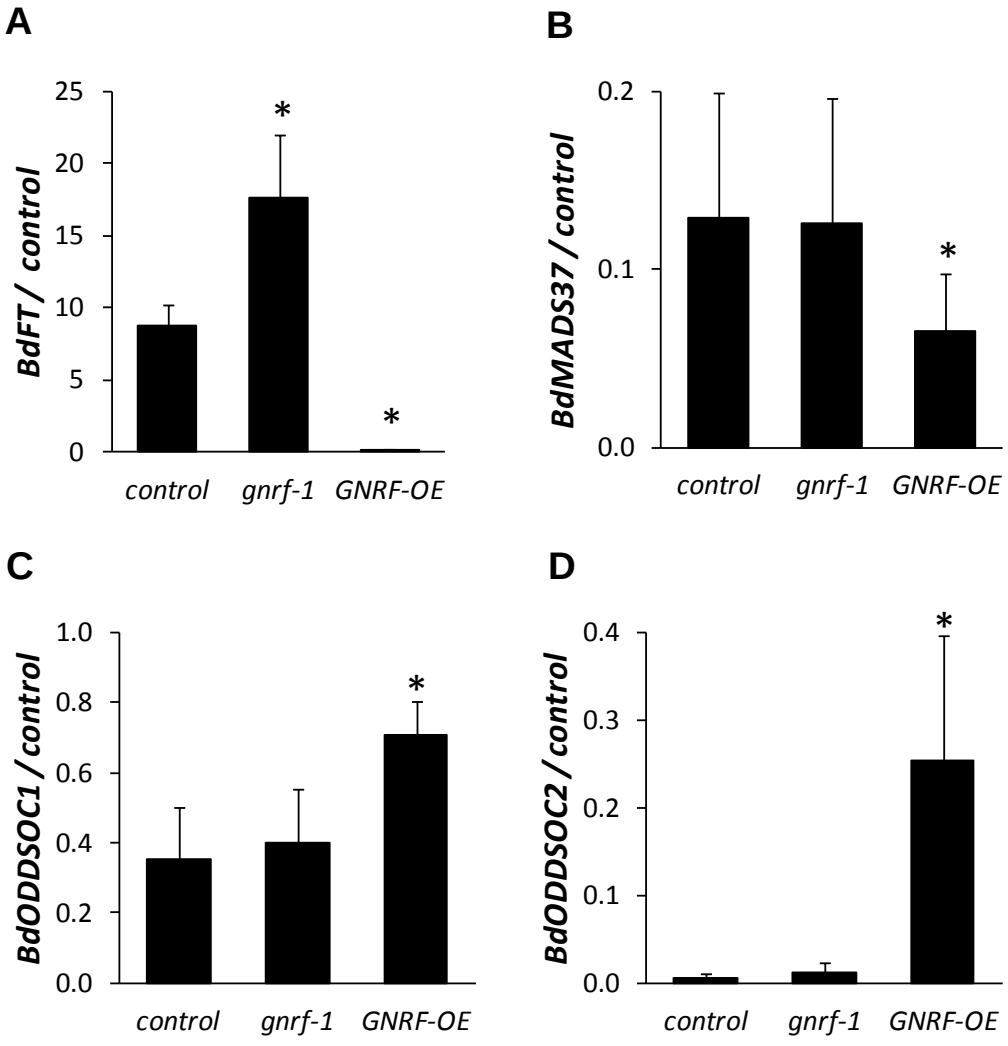

Figure 4.7. Relative expression of flowering pathway genes in leaves. (A*) FT*, (B) *MADS37*, (C) *ODDSOC1*, (D) *ODDSOC2*. Fourth leaf from the base of the tallest stem was collected from developmentally equivalent plants when the inflorescence was first visible from the flag leaf. Seven to nine individuals were analyzed in triplicate using QRT-PCR and normalized against two housekeeping genes. * $p < 0.05$

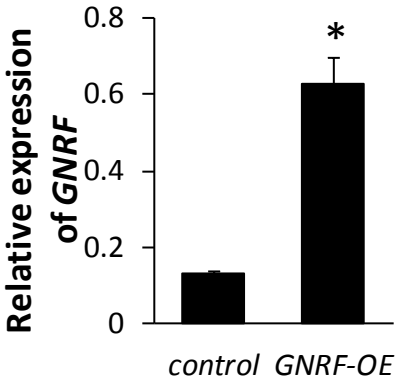

Figure 4.8. Transcript abundance of *GNRF* in control and *GNRF-OE* stems. Relative expression of *GNRF* was measured in the stem tissue of the tallest stem when the inflorescence was just visible from the flag leaf of the control plants. Ten individuals were analyzed in triplicate using QRT-PCR and normalized against two housekeeping genes. * $p < 0.05$.

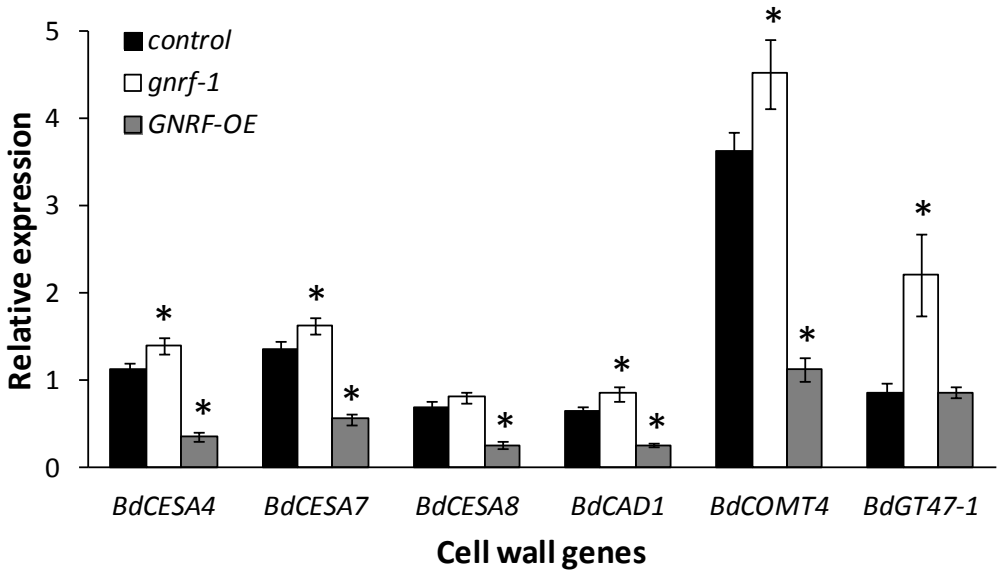

Figure 4.9. Target cell wall gene expression in stems. Relative expression of secondary cell wall genes in control, *gnrf-1* and *GNRF-OE* stems. Tallest stem was collected from developmentally equivalent plants when the inflorescence was first visible from the flag leaf. Eight to ten individuals were analyzed in triplicate using QRT-PCR and normalized against two housekeeping genes. * $p < 0.05$.

biosynthesis namely, *BdCAD1* and *BdCOMT4* and a gene with a predicted role in xylan biosynthesis *BdGT47-1* were analyzed for changes in transcript abundance (Fig. 4.9). All five cellulose and lignin genes were significantly down-regulated in *GNRF-OE* stems. The same genes were also found to be repressed in the microarray data set (Table 4.1). Conversely, the expression of *BdCESA4/7*, *BdCAD1*, and *BdCOMT4* was significantly up-regulated in *gnrf-1*stems. While the putative xylan gene, *BdGT47-1*, was significantly up-regulated in the *gnrf-1*lines, no change was observed in *GNRF-OE* lines. Overall, these data suggest that *GNRF* is a repressor of cellulose and lignin gene expression.

4.3.7 *GNRF* influences cell wall composition

In order to correlate gene expression with levels of cellulose and lignin, histochemical assays were preformed on stem tissue. First internodes of fully senesced stems were sectioned and stained with phloroglucinol-HCl to investigate possible changes in lignin content. While no visible changes in staining intensities were observed between control and *gnrf-1* stem cross sections, *GNRF-OE* sclerenchyma fibers and vascular bundles exhibited a lighter shade of orange (Fig. 4.10A). Since phloroglucinol-HCl stains *O*-4-linked coniferyl and sinapyl aldehydes in lignified tissue in a concentration indicative manner, the lighter staining pattern observed in *GNRF-OE* is likely due to a decrease in total lignin. Fully senesced pulverized tissue was used to measure the acetyl bromide soluble lignin (ABSL) content to further investigate the function of GNRF on cell wall composition and lignification (Fig. 4.10B). As expected based on the phloroglucinol-HCL staining, the *GNRF-OE* stem was significantly reduced in ABSL content compared to the controls whereas no significant change was observed between the *gnrf-1* and control samples. As many lignin and lignin related genes were

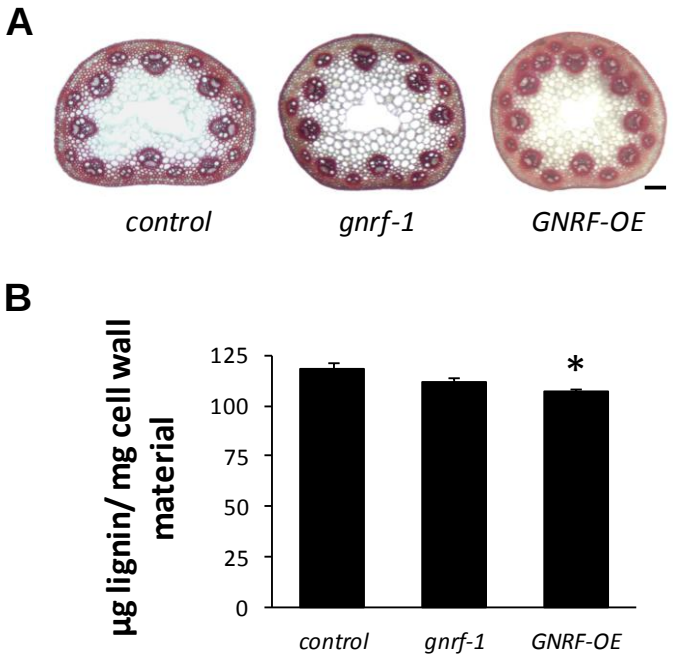

Figure 4.10. Histo-chemical and compositional analysis of stem lignification. (A) Lignin staining of stem cross sections. First internodes of fully senesced plants were hand sectioned and stained with phloroglucinol-HCl to visualize lignin. Representative images are illustrated. Scale bar = 50 µm. (B) Acetyl bromide soluble lignin content in senesced stems. Three to six individuals from each line were analyzed for the ASBL lignin content. * $p < 0.05$.

repressed in the *GNRF-OE* stems I expected to observe a reduction in lignin composition which is further demonstrated by the significant reduction in ABSL.

4.4 Discussion

Several NAC transcription factors have been shown to play crucial roles in secondary cell wall biosynthesis and overall plant growth and development. The *B. distachyon* NAC family is estimated to comprise 99 NAC proteins among which only one has been functionally characterized to date (International Brachypodium Initiative, 2010; Valdivia et al., 2013). GNRF was selected for functional characterization as it is highly expressed in stems mirroring an expression profile similar to that of characterized secondary cell wall regulators such as NST1, SND1 and SND2 (Mitsuda et al., 2005; Zhong et al., 2007). Moreover it is co-regulated with cell wall biosynthesis genes.

An over-expression construct was used to develop gain-of-function mutants with constitutive over-expression of GNRF. Simultaneously, a homozygous mutant allele harboring a nonsynonymous point mutation was isolated from a TILLING mutant collection to investigate the *gnrf-1* loss-of-function phenotypes. Surprisingly, over-expression mutants demonstrated a persistent vegetative phenotype. *GNRF-OE* plants failed to flower and could not transition into the reproductive stage. Conversely the *gnrf-1* mutants flowered significantly earlier than control plants. However apart from our observations, there are no known reports associating SND2 orthologs with flowering or floral transition. Over-expression of SND2 in *A. thaliana* had no adverse effects on flowering (Zhong et al., 2008). Additionally, no change in flowering time was observed with *PtSND2* over-expression mutants (Wang et al., 2013). These observations warranted

further investigation of the flowering pathway in *B. distachyon*. I first confirmed the over-expression of the GNRF trasngene and as expected GNRF transcripts were significantly up-regulated in both leaf and stem tissue in the *GNRF-OE* plants. FT is the key signal in the vegetative to reproductive transition, during which FT accumulates in *A. thaliana* leaves prior to flowering (Gu et al., 2010; Wu et al., 2013; Lv et al., 2014). In order to investigate the influence of GNRF on *FT* we analyzed *FT* expression in mutant leaves. As expected *FT* was not detectable in *GNRF-OE* leaves whereas it was significantly up-regulated in *gnrf-1* leaves, in agreement with the flowering time phenotypes observed for these mutants. Unavailability of the mobile signal FT in *GNRF-OE* leaves is a possible cause for the persistant vegetative phenotype. Recently, FT was also shown to bind phospholipids to accelerate flowering (Nakamura et al., 2014). Another possibility is the repression of numerous lipid transporter genes, which in turn influence the availability of these phospholipids thus delaying the flowering process.

However the direct cause of the persistent vegetative growth of *GNRF-OE* is most likely due to the >50 fold reduction of two MADS box transcription factors, *FUL1* and *FUL2*. *VRN1* is the closest ortholog to *A. thaliana* meristem identity gene *AP1* in temperate cereals and is responsible for transitioning the meristem into a flower bud. Moreover *FUL* is another meristem identity gene homologous to *AP1* (Ream et al., 2014). Even though BdMADS33/FUL1 and MADS10/FUL2 have not been functionally characterized, they are the closest orthologous proteins to AP1 and FUL, respectively, and therefore most likely regulate floral meristem identity in *B. distchyon*. It is interesting to note that similar to the *GNRF-OE* phenotype, mutations in *A. thaliana FUL* gene along with mutations in *AP1*/*CAL* genes result in non-flowering leafy phenotypes (Ferrandiz et

al., 2000). This provides strong support for the function of FUL1 and FUL2 as floral identity genes in *B. distachyon*. *AtAP1* is negatively regulated by the meristem identity gene *TERMINAL FLOWER* and the organ identity gene *AGAMOUS*. It is also regulated by the floral homeotic gene *PISTILLATA* and its interacting partner *APETALA3* (Sundström et al., 2006). However, to date there is no indication of any involvement of a NAC transcription factor in the regulation of these redundant floral homeotic genes, *AP1*/*CAL*/*FUL*. Furthermore, there are no known NAC transcription factors associated with dicot or monocot flowering pathways (Higgins et al., 2010). GNRF is the first NAC protein to be associated with the flowering pathway in *B. distachyon*. The architecture of flowering pathways in *B. distachyon* and *A. thaliana* are similar in some aspects, for instance vegetative to reproductive transition is signaled by leaf localized FT expression (Corbesier et al.; Wu et al., 2013; Lv et al., 2014). In both species, activation of orthologous floral meristem identity genes results in flowering. Interestingly, the vernalization responses between these two species are significantly different.

GNRF is the closest ortholog to SND2. Very little is known about the function of SND2 in any plant system. To date it has only been shown to activate cell wall biosynthesis in *A. thaliana* and poplar (Zhong et al., 2008; Grant et al., 2010; Wang et al., 2013). Based on the protein similarity and the expression profile, GNRF likely has a similar function to SND2. In support of that, we have shown that GNRF is a regulator of cellulose, lignin and hemicellulose genes. Over-expression of GNRF resulted in a significant reduction of these genes in the *GNRF-OE* stems. Therefore, unlike SND2, GNRF appears to act as a repressor of cell wall biosynthesis. However, analysis of GNRF sequence revealed no apparent repression domains. Based on the microarray results, a

handful of lignin genes, peroxidases, lacasses and copper ion transporters were repressed by *GNRF*. Since these components are essential for polymerization of lignin monomers, cell wall lignification should be effected in the *GNRF-OE* mutants. Complementing the expression analysis, histo-chemical analysis revealed a qualitative reduction in stem lignin in the *GNRF-OE* stems. This observation was further validated by the significantly lower levels of acetyl bromide soluble lignin content measured in *GNRF-OE* stems.

Based on the expression profiling and cell wall composition analysis it is evident that GNRF is involved in cell wall regulation. Unlike other classical cell wall regulators, GNRF has profound pleiotropic effects. Based on transcription profiling, GNRF is associated with cell wall, floral transition and transporters of numerous substrates. However using microarray expression analysis, direct and indirect regulation via GNRF over-expression is hard to decipher. Many different genes may be directly regulated by GNRF, including *FUL1* and *FUL2*. Further experiments will be required to determine whether these genes are under direct regulation by GNRF as well as if these genes share common GNRF binding sites.

Table 4.1. Genes repressed by GNRF in *GNRF-OE* stems

Gene name	Gene description	Fold change
Transcription factors		
Bradi1g08340	MADS33/VRN1/FUL1	53.8
Bradi1g59250	MADS10/FUL2	51.3
Bradi1g08326	MADS1	33.6
Bradi2g00280	WRKY47	7.7
Bradi3g26690	BEL-like Homeobox	7.0
Bradi1g48520	MADS7	6.1
Bradi1g69890	MADS11	5.7
Bradi1g26720	SBP family	4.6
Bradi1g57607	KNOX6 Homeobox	4.5
Bradi3g56290	bZip family	4.7
Bradi5g10640	NAC family, XND1-like	3.6
Bradi1g63690	MYB-like family	3.7
Bradi1g51960	MYB15	4.5
Bradi2g48690	MADS19	4.3
Bradi3g21480	Homeobox family	3.8
Bradi4g27720	bZIP70	4.4
Bradi3g34567	WRKY13	2.5
Bradi2g05700	NAC28	2.9
Bradi2g47590	MYB48/SWAM1	3.0
Bradi3g51800	MADS28	2.8
Bradi5g11270	MADS36	2.7
Bradi1g12690	KNOX4 Homeobox	3.0
Bradi1g10047	KNOX2 Homeobox	3.2
Bradi2g09720	C2C2-Dof family	3.4
Bradi1g73710	C2C2-Dof family	2.6
Bradi1g12780	bHLH family	3.4
Bradi1g71990	bHLH family	3.2
Bradi3g15440	bHLH family	2.7
Bradi3g16515	MYB59/LHY	2.9
Bradi3g03407	ARF family	3.1
Bradi4g33370	WRKY60	2.2
Bradi2g59110	SBP family	2.3
Bradi5g22920	bHLH family	2.2
Bradi2g23530	BEL-like Homeobox	2.3
Bradi5g17640	AP2 family	2.2
Cell wall polysaccharide synthesis		
Bradi2g60557	Glycosylhydrolase, GH17, β-1,3-glucanase 1	12.4
Bradi2g34650	Fasciclin-like arabinogalactan protein	8.1

Bradi1g64560	Glycosyltransferase, GT34 family, xylosyltransferase	7.6
Bradi2g08310	Glycosyltransferase, GT1 family, UDP-glucosyl transferase	7.1
Bradi1g12290	Glycosyltransferase, GT47 family	6.8
Bradi4g13697	Glycosyltransferase, GT37 family fucosyltransferase 1	6.2
Bradi2g61230	Glycosyltransferase, GT61 family	5.5
Bradi1g25117	Cellulose synthase-like F	3.8
Bradi1g33827	Glycosylhydrolase, GH6, xyloglucan endotransglycosylase 6	4.9
Bradi1g21990	Glycosyltransferase, GT75 family	4.8
Bradi1g12710	Glycosyl hydrolase, GH10 family	4.8
Bradi2g00220	Fasciclin-like arabinoogalactan protein	4.6
Bradi1g22030	COBRA	4.5
Bradi3g04080	Glycosylhydrolase, GH9 family glycosyl hydrolase 9B8	4.0
Bradi5g04120	Glycoside hydrolase, α-Expansin	3.9
Bradi2g53580	Glycoside hydrolase, α-Expansin	3.8
Bradi4g33490	Fasciclin-like arabinoogalactan protein	3.8
Bradi4g29640	Glycosylhydrolase, GH9 family	3.8
Bradi4g21240	Pfam:04669 Polysaccharide biosynthesis	3.7
Bradi2g02320	Glycosylhydrolase, GH10 family	3.7
Bradi2g59410	Glycosyltransferase, GT47 family	3.6
Bradi1g59880	COBRA-like	3.6
Bradi3g33130	Glycoside hydrolase, β-Expansin	3.6
Bradi3g28350	Glycosyltransferase, GT2 family, CESA4	3.0
Bradi4g30540	Glycosyltransferase, GT2 family, CESA7	2.8
Bradi3g33140	Glycosylhydrolase, β-Expansin	2.7
Bradi3g19087	Glycosyltransferase, GT2 family, CSLC	2.6
Bradi4g28260	Glycosyltransferase, GT77 family, Extensin	2.5
Bradi1g10347	Glycosylhydrolase, GH17 family	2.3
Lignin synthesis		
Bradi2g23370	Laccase	12.6
Bradi1g66720	Laccase	12.3
Bradi3g58560	Copper ion binding	11.6
Bradi1g27910	Peroxidase	11.3
Bradi2g09690	Peroxidase	6.9
Bradi2g20840	Peroxidase	5.9
Bradi1g27920	Peroxidase	5.3
Bradi3g30590	Ferulic acid 5-hydroxylase 1 (FAH1)	5.1
Bradi2g54680	Laccase	4.7
Bradi4g44810	Laccase	4.3

Bradi3g38540	Copper ion binding	4.5
Bradi1g24880	Laccase	4.3
Bradi4g11850	Laccase	4.3
Bradi4g28920	Copper transporter protein ATOX1-related	4.3
Bradi1g43680	Peroxidase	4.0
Bradi1g38297	Peroxidase	3.2
Bradi3g05750	4-coumarate--CoA ligase 3 (4CL3)	3.1
Bradi1g32870	Peroxidase	3.1
Bradi1g45790	Copper ion binding protein	2.9
Bradi1g68900	Peroxidase	2.9
Bradi3g39420	Caffeoyl CoA 3-O-methyltransferase (CCoAOMT)	2.8
Bradi3g16530	Caffeic acid O-methyltransferase (COMT4)	2.8
Bradi1g31320	4-coumarate--CoA ligase 1 (4CL1)	2.4
Bradi3g09240	Copper binding protein	2.3
Lipid transfer		
Bradi2g17550	Lipid transfer protein	8.5
Bradi2g17530	Lipid transfer protein	7.8
Bradi2g30490	Lipid transfer protein	3.8
Bradi1g19470	Lipid transfer protein	6.1
Bradi2g54970	Lipid transfer protein	4.9
Bradi2g32950	Lipid transfer protein	4.7
Bradi2g17540	Lipid transfer protein	4.4
Bradi5g17930	Lipid transfer protein	4.2
Transport		
Bradi3g39800	Dicarboxylate transporter	7.8
Bradi1g78100	Arsenite transport	6.5
Bradi1g45190	Amino acid transporter	6.3
Bradi3g05570	Potassium ion transporter	6.2
Bradi1g03500	Proton-dependent oligopeptide transporter	5.3
Bradi3g48950	Ammonium transporter	5.0
Bradi4g21790	Proton-dependent oligopeptide transporter	5.0
Bradi3g37850	Potassium ion transporter	4.7
Bradi1g21800	Sugar transporter	4.6
Bradi3g51280	Major facilitator superfamily	3.7
Bradi3g28920	UDP-glucuronic acid transporter	3.6
Bradi2g07830	Aquaporin transporter	3.6
Bradi1g69770	Aluminum activated citrate transporter	3.4
Bradi1g34140	ATPase-like zinc transporter	3.3
Bradi5g17990	ATP dependent copper transporter	3.2
Bradi1g25937	EamA-like transporter	3.1
Bradi4g28000	Sugar transporter	2.7
Bradi3g16130	ABC transporter	2.7

Bradi1g73170	Sucrose transporter	2.7
Bradi1g17830	Potassium transporter	2.6
Bradi3g39077	Oligopeptide transporter	2.6
Bradi4g34510	PINFORMED-Like auxin efflux carrier	2.5
Bradi2g24910	Amino acid transporter	2.5
Bradi3g32390	Tetracycline transporter	2.5
Bradi2g34560	ZIP Zinc transporter	2.5
Bradi5g24170	Sulfate transporter	2.4
Bradi3g51250	Mechanosensitive ion channel	2.4
Bradi1g59830	Amino acid transporter	2.3
Bradi1g34210	Cation transmembrane transporter	2.3
Bradi4g08130	ABC transporter	2.1

Table 4.2 Primers used in this study

Primer Name	Sequence
attB1NAC38_F	ggggACAAGTTTGTACAAAAAAGCAGGCTCTATGAC ATGGTGCAATAGCTTC
attB2NAC38_R	ggggACCACTTTGTACAAGAAAGCTGGGTATCAGGG GCCAAAGCCTGTCCC
qPCRNAC38CDS_F	CAAGAAGCAGCAGCAACAGCAAC
qPCRNAC38CDS_R	CGCAGCCTGCAACTGTTCATAC
qPCRUBC18_F	TCACCCGCAATGACTGTAAG
qPCRUBC18_R	ACCACCATCTGGTCTCCTTC
qPCRGAPDH_F	TCACCCGCAATGACTGTAAG
qPCRGAPDH_R	ACCACCATCTGGTCTCCTTC
qPCRCAD1_F	AGGATAGAATGGGCAGCATCGC
qPCRCAD1_R	ATCTTCAGGGCCTGTCTTCCTGAG
qPCRCOMT4_F	TGGAGAGCTGGTACTACCTGAAG
qPCRCOMT4_R	CGACATCCCGTATGCCTTGTTG
qPCRCESA4_F	GCGTTTCGCATACACCAACACC
qPCRCESA4_R	ACTCGCTAGGTTGTTCAGTGTGG
qPCRCESA7_F	GCGATTCGCCTACATCAACACC
qPCRCESA7_R	GGCTGGCAAATGTGCTAATCGG
qPCRCESA8_F	CAAAGCACAAGTTCCGCCTGTG
qPCRCESA8_R	TGGCTCGTATGCATCTGTCAAATC
qPCRGT47-1_F	AGGGTGGTTACTATGCAAGAGGTG
qPCRGT47-1_R	AATATAGCGCGCTGCATGTCCTC
qBd2g59187_F (ODDSOC2)	AAATCCAAGATATTGGCAAAACG
qBd2g59187_R	CCTTAGGCTCACTGGAGTTCTCA
qBd2g59120_F (ODDSOC1)	CCGGCAAGCTCTACGAGTACTC
qBd2g59120_R	GCTCCCGCAAATTGCTGAT
qBd3g41297_F (MADS37)	CAATCTGAGGATGAAGGTGTCACA
qBd3g41297_R	GCTTGACAAGTTGTTCGCTTTCT
qBd1g72150_F (SVP)	AACTCAAGGCTGAAGGAGCAAC
qBd1g72150_R	AATCAGCGGCAACCTGCATC
qBd1g48830_F (FT)	AACCAACCTGAGGGTGAGCTTC
qBd1g48830_R	AGCATCTGGGTCTACCATCACGAG
qBd1g08340_F (FUL1)	TTCGCCACCGACTCATGTATGG
qBd1g08340_R	TCTGCATAGGAGTAGCGCTCATAG
qBd1g59250_F (FUL2)	AAACTGAAGGCCAAGATTGAGACG
qBd1g59250_R	ATCCTCTCCCATGAGGTGCTTG
qBd3g44860_F (RCN2)	TTGGGAGGGAGATGGTGAGCTATG
qBd3g44860_R	TGAACCTGTGGATGCCGATGTTTG
Hpt_F	AGAATCTCGTGCTTTCAGCTTCGA
Hpt_R	TCAAGACCAATGCGGAGCATATAC

461_N1F2	ACATGGTGCAATAGCTTCAACGACG
461_N1R2	CCAGTCCTAATCGATCCGGGATC
461_N2F2	CACGACGTTGTAAAACGACCGGCAGCGGCCAAGAAGC
461_N2R2	GGATAACAATTTCACACAGGGGCCGACGCAATGCAAGCGC

CHAPTER 5

CONCLUSIONS

Transcriptional regulation of grass secondary cell wall biosynthesis is poorly understood. To date only two MYB transcription factors and one NAC transcription factor associated with cell walls have been characterized in a grass species. Extensive research conducted on the model eudicot *A. thaliana* has revealed a complex yet still evolving cell wall regulatory network. However due to the fundamental differences between dicots and grasses the grass cell wall regulatory network may possess unique regulators. My research objective was to uncover grass transcriptional regulators responsible for the regulation of secondary cell wall biosynthesis.

I established *B. distachyon* as a working model system in the Hazen laboratory for genetic manipulations. I used a reverse genetics approach to functionally characterize two *CELLULOSE SYNTHASE A* genes responsible for synthesizing cellulose in the secondary cell wall. Artificial microRNA mediated down regulation of *BdCESA4* or *BdCESA7* caused cell wall defects and a reduction in crystalline cellulose in these mutants. Next, using a yeast one-hybrid approach, I identified a number of candidate cell wall regulators interacting with these CESA gene promoters. Based on gene expression profiles and protein homology I selected *BdMYB48* and *GNRF* for functional characterization. According to phylogenetic analysis it was evident that BdMYB48 is part of a grass specific group within the MYB family. I generated gain-of-function and dominant repressor lines for *BdMYB48*. Based on the reciprocal phenotypes observed for plant height, above ground biomass, cell wall thickness and relative expression of cell wall genes it was clear that BdMYB48 is a direct activator of secondary cell wall biosynthesis.

More importantly BdMYB48 is capable of regulating biomass accumulation in the model grass *B. distachyon*. This work led to the functional characterization of the first grass specific transcriptional activator.

The second candidate, *GNRF*, is a NAC transcription factor abundantly expressed in stems. Gain-of-function mutants were generated for GNRF and a homozygous mutant allele was isolated from a TILLING population to investigate the loss-of-function phenotypes. Using these mutant lines I was able to demonstrate GNRF was capable of regulating cellulose and lignin gene expression. However, its function was distinct from the characterized orthologous proteins as GNRF was a repressor of cell wall biosynthesis. Surprisingly, unlike any other characterized cell wall regulator GNRF had a strong flowering phenotype. This led me to investigate the flowering pathway in these mutants. A microarray experiment revealed GNRF represses floral integrators and floral meristem identity genes and as a likely result the gain-of-function lines remained vegetative. Accordingly GNRF was also found to be a repressor of transporter functions. Further characterization of this pleiotropic regulator is needed to confirm its direct targets. In summary, I identified a grass specific transcriptional activator and a pleiotropic transcriptional repressor responsible for regulating cell wall biosynthesis in the model grass *B. distachyon* (Fig. 5.1). These findings will contribute to the current understanding of the grass cell wall regulation.

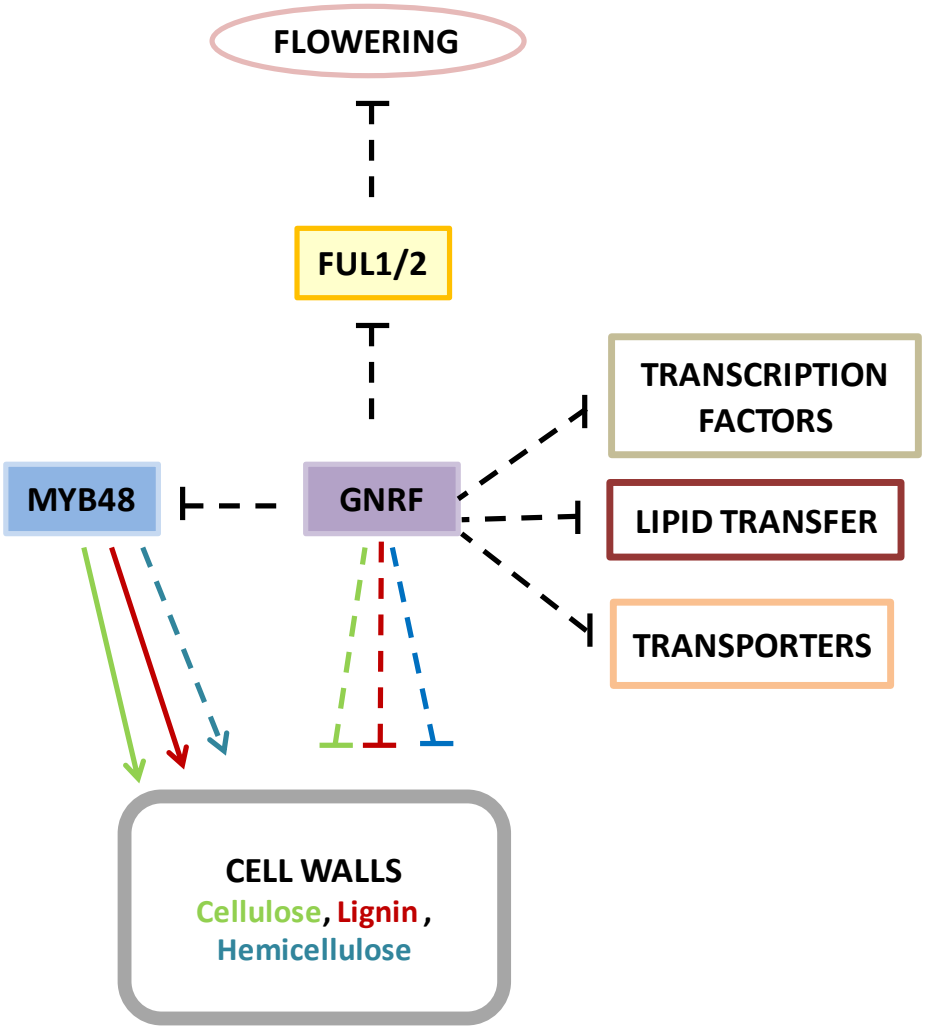

Figure 5.1 A model for the regulatory roles of *MYB48* and *GNRF*. MYB48 activates cellulose, hemicellulose, and lignin genes to regulate cell wall biosynthesis. GNRF represses cellulose, hemicellulose, and lignin genes, floral integrator genes such as *FRUITFUL1* and *2* (*FUL1/2*), other transcription factors , lipid transfer and transporters affecting cell wall biosynthesis, flowering and other developmental processes. Solid arrows illustrate direct interactions, dashed lines illustrate direct or indirect interactions.

BIBLIOGRAPHY

Alexandre CM, Hennig L (2008) FLC or not FLC: the other side of vernalization. J Exp Bot **59:** 1127-1135

Alonso-Peral MM, Oliver SN, Casao MC, Greenup AA, Trevaskis B (2011) The promoter of the cereal *VERNALIZATION1* gene Is sufficient for transcriptional induction by prolonged cold. PLoS ONE **6:** e29456

Ambavaram MMR, Krishnan A, Trijatmiko KR, Pereira A (2011) Coordinated activation of cellulose and repression of lignin biosynthesis pathways in rice. Plant Physiol **155:** 916-931

Appenzeller L, Doblin M, Barreiro R, Wang H, Niu X, Kollipara K, Carrigan L, Tomes D, Chapman M, Dhugga KS (2004) Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (*CesA*) gene family. Cellulose **11:** 287-299

Arvidsson S, Kwasniewski M, Riano-Pachon D, Mueller-Roeber B (2008) QuantPrime - a flexible tool for reliable high-throughput primer design for quantitative PCR. BMC Bioinformatics **9:** 465

Barnette AL, Bradley LC, Veres BD, Schreiner EP, Park YB, Park J, Park S, Kim SH (2011) Selective detection of crystalline cellulose in plant cell walls with sum-frequency-generation (SFG) vibration spectroscopy. Biomacromolecules **12:** 2434-2439

Barnette AL, Lee C, Bradley LC, Schreiner EP, Park YB, Shin H, Cosgrove DJ, Park S, Kim SH (2012) Quantification of crystalline cellulose in lignocellulosic biomass using sum frequency generation (SFG) vibration spectroscopy and comparison with other analytical methods. Carbohydr Polym **89:** 802-809

Baskin TI, Beemster GTS, Judy-March JE, Marga F (2004) Disorganization of cortical microtubules stimulates tangential expansion and reduces the uniformity of cellulose microfibril alignment among cells in the root of Arabidopsis. Plant Physiol **135:** 2279

Bevan MW, Garvin DF, Vogel JP (2010) Brachypodium distachyon genomics for sustainable food and fuel production. Curr Opin Biotechnol **21:** 211-217

Bhargava A, Mansfield SD, Hall HC, Douglas CJ, Ellis BE (2010) MYB75 functions in regulation of secondary cell wall formation in the Arabidopsis inflorescence stem. Plant Physiol **154:** 1428-1438

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol **54:** 519-546

Bosch M, Mayer C-D, Cookson A, Donnison IS (2011) Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. J Exp Bot **62:** 3545-3561

Bragg JN, Wu J, Gordon SP, Guttman ME, Thilmony R, Lazo GR, Gu YQ, Vogel JP (2012) Generation and characterization of the Western Regional Research Center Brachypodium T-DNA insertional mutant collection. PLoS ONE **7:** e41916

Brkljacic J, Grotewold E, Scholl R, Mockler T, Garvin DF, Vain P, Brutnell T, Sibout R, Bevan M, Budak H, Caicedo AL, Gao C, Gu Y, Hazen SP, Holt BF, Hong S-Y, Jordan M, Manzaneda AJ, Mitchell-Olds T, Mochida K, Mur LAJ, Park C-M, Sedbrook J, Watt M, Zheng SJ, Vogel JP (2011) Brachypodium as a model for the grasses: today and the future. Plant Physiol **157:** 3-13

Burton RA, Ma G, Baumann U, Harvey AJ, Shirley NJ, Taylor J, Pettolino F, Bacic A, Beatty M, Simmons CR, Dhugga KS, Rafalski JA, Tingey SV, Fincher GB (2010) A customized gene expression microarray reveals that the brittle stem phenotype *fs2* of barley Is attributable to a retroelement in the *HvCesA4* cellulose synthase gene. Plant Physiol **153:** 1716-1728

Burton RA, Shirley NJ, King BJ, Harvey AJ, Fincher GB (2004) The CesA gene family of barley (*Hordeum vulgare*): quantitative analysis of transcripts reveals two groups of co-expressed genes. Plant Physiol **134:** 224-236

Carroll A, Specht C (2011) Understanding plant cellulose synthases through a comprehensive investigation of the cellulose synthase family sequences. Front Plant Sci **2**

Chaw S-M, Chang C-C, Chen H-L, Li W-H (2004) Dating the monocot–dicot divergence and the origin of core eudicots using whole chloroplast genomes. J Mol Evol **58:** 424-441

Christensen U, Alonso-Simon A, Scheller HV, Willats WGT, Harholt J (2010) Characterization of the primary cell walls of seedlings of *Brachypodium distachyon* - potential model plant for temperate grasses. Phytochemistry **71:** 62-69

Christiansen M, Holm P, Gregersen P (2011) Characterization of barley (*Hordeum vulgare* L.) NAC transcription factors suggests conserved functions compared to both monocots and dicots. BMC Res Notes **4:** 302

Corbesier L, Vincent C Fau - Jang S, Jang S Fau - Fornara F, Fornara F Fau - Fan Q, Fan Q Fau - Searle I, Searle I Fau - Giakountis A, Giakountis A Fau - Farrona S, Farrona S Fau - Gissot L, Gissot L Fau - Turnbull C, Turnbull C Fau - Coupland G, Coupland G FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis.

Csaikl UM, Bastian H, Brettschneider R, Gauch S, Meir A, Schauerte M, Scholz F, Sperisen C, Vornam B, Ziegenhagen B (1998) Comparative analysis of different DNA extraction protocols: a fast, universal maxi-preparation of high quality plant DNA for genetic evaluation and phylogenetic studies. Plant Mol Biol Rep **16:** 69-86

d'Yvoire MB, Bouchabke-Coussa O, Voorend W, Antelme S, Cézard L, Legée F, Lebris P, Legay S, Whitehead C, McQueen-Mason SJ, Gomez LD, Jouanin L, Lapierre C, Sibout R (2012) Disrupting the *cinnamyl alcohol dehydrogenase 1* gene (*BdCAD1*) leads to altered lignification and improved saccharification in *Brachypodium distachyon*. Plant J **73:** 496-508

Dalmais M, Antelme S, Ho-Yue-Kuang S, Wang Y, Darracq O, d'Yvoire MB, Cezard L, Legee F, Blondet E, Oria N, Troadec C, Brunaud V, Jouanin L, Hofte H, Bendahmane A, Lapierre C, Sibout R (2013) A TILLING platform for functional genomics in *Brachypodium distachyon*. PLoS One **8:** e65503

Demura T, Fukuda H (2007) Transcriptional regulation in wood formation. Trends Plant Sci **12:** 64-70

Desprez T, Juraniec M, Crowell EF, Jouy H, Pochylova Z, Parcy F, Hofte H, Gonneau M, Vernhettes S (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. Proc Natl Acad Sci USA **104:** 15572-15577

Doblin MS, Kurek I, Jacob-Wilk D, Delmer DP (2002) Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell Physiol **43:** 1407

Ehlting J, Mattheus N, Aeschliman DS, Li E, Hamberger B, Cullis IF, Zhuang J, Kaneda M, Mansfield SD, Samuels L, Ritland K, Ellis BE, Bohlmann J, Douglas CJ (2005) Global transcript profiling of primary stems from Arabidopsis thaliana identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. Plant J **42:** 618-640

Ferrandiz C, Gu Q, Martienssen R, Yanofsky MF (2000) Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. Development **127:** 725-734

Fornalé S, Shi X, Chai C, Encina A, Irar S (2010) ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. Plant J **64:** 633-644

Fornalé S, Sonbol F-M, Maes T, Capellades M, Puigdomènech P, Rigau J, Caparrós-Ruiz D (2006) Down-regulation of the maize and *Arabidopsis thaliana* Caffeic acid *O*-methyl-transferase genes by two new maize R2R3-MYB transcription factors. Plant Mol Biol **62:** 809-823

Foster CE, Martin TM, Pauly M (2010) Comprehensive compositional analysis of plant cell walls (lignocellulosic biomass) Part I: lignin. J Vis Exp**:** e1745

Gomez L, Bristow J, Statham E, McQueen-Mason S (2008) Analysis of saccharification in *Brachypodium distachyon* stems under mild conditions of hydrolysis. Biotechnol Biofuels **1:** 15

Grant E, Fujino T, Beers E, Brunner A (2010) Characterization of NAC domain transcription factors implicated in control of vascular cell differentiation in Arabidopsis and *Populus*. Planta **232:** 337-352

Greenup A, Peacock WJ, Dennis ES, Trevaskis B (2009) The molecular biology of seasonal flowering-responses in Arabidopsis and the cereals. Ann Bot **103:** 1165-1172

Greenup AG, Sasani S, Oliver SN, Talbot MJ, Dennis ES, Hemming MN, Trevaskis B (2010) *ODDSOC2* Is a MADS box floral repressor that Is down-regulated by vernalization in temperate cereals. Plant Physiol **153:** 1062-1073

Gu Y, Kaplinsky N, Bringmann M, Cobb A, Carroll A, Sampathkumar A, Baskin TI, Persson S, Somerville CR (2010) Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. Proc Nat Acad Sci **107:** 12866-12871

Guillon F, Bouchet B, Jamme F, Robert P, Quemener B, Barron C, Larre C, Dumas P, Saulnier L (2011) *Brachypodium distachyon* grain: characterization of endosperm cell walls. J Exp Bot **62:** 1001-1015

Haigler CH, Brown RM (1986) Transport of rosettes from the Golgi apparatus to the plasma membrane in isolated mesophyll cells of *Zinnia elegans* during differentiation to tracheary elements in suspension culture. Protoplasma **134:** 111

Handakumbura P, Matos D, Osmont K, Harrington M, Heo K, Kafle K, Kim S, Baskin T, Hazen S (2013) Perturbation of *Brachypodium distachyon CELLULOSE SYNTHASE A4* or *7* results in abnormal cell walls. BMC Plant Biol **13:** 131

Handakumbura PP, Hazen SP (2012) Transcriptional regulation of grass secondary cell wall biosynthesis: playing catch-up with *Arabidopsis thaliana*. Front Plant Sci **3:** 74

Harding SA, Leshkevich J, Chiang VL, Tsai C-J (2002) Differential substrate inhibition couples kinetically distinct 4-Coumarate:coenzyme A Ligases with spatially distinct metabolic roles in quaking aspen. Plant Physiol **128:** 428-438

Harrington M, Hong E, Fasanmi O, Brewster R (2007) Cadherin-mediated adhesion regulates posterior body formation. BMC Dev Biol **7:** 130

Hatton D, Sablowski R, Yung M-H, Smith C, Schuch W, Bevan M (1995) Two classes of *cis* sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. Plant J **7:** 859-876

Heaton EA, Flavell RB, Mascia PN, Thomas SR, Dohleman FG, Long SP (2008) Herbaceous energy crop development: recent progress and future prospects. Curr. Opin. Biotechnol. **19:** 202

Hématy K, Sado P-E, Van Tuinen A, Rochange S, Desnos T, Balzergue S, Pelletier S, Renou J-P, Höfte H (2007) A receptor-like kinase mediates the response of *Arabidopsis* cells to the inhibition of cellulose synthesis. Curr Biol **17:** 922-931

Higgins JA, Bailey PC, Laurie DA (2010) Comparative genomics of flowering time pathways using *Brachypodium distachyon* as a model for the temperate grasses. PLoS ONE **5:** e10065

Hong S-Y, Seo P, Yang M-S, Xiang F, Park C-M (2008) Exploring valid reference genes for gene expression studies in Brachypodium distachyon by real-time PCR. BMC Plant Biol **8:** 112

Hong SY, Park JH, Cho SH, Yang MS, Park CM (2011) Phenological growth stages of *Brachypodium distachyon*: codification and description. Weed Res **51:** 612-620

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

Hussey SG, Mizrachi E, Creux NM, Myburg AA (2013) Navigating the transcriptional roadmap regulating plant secondary cell wall deposition. Front Plant Sci **4**

Hussey SG, Mizrachi E, Spokevicius AV, Bossinger G, Berger DK, Myburg AA (2011) SND2, a NAC transcription factor gene, regulates genes involved in secondary cell wall development in Arabidopsis fibres and increases fibre cell area in Eucalyptus. BMC Plant Biol **11:** 173

Imaizumi T, Kay S (2006) Photoperiodic control of flowering: not only by coincidence. Trends Plant Sci **11:** 550 - 558

Initiative TIB (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature **463:** 763-768

International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature **463:** 763-768

Jaeger KE, Wigge PA (2007) FT protein acts as a long-range signal in *Arabidopsis*. Curr Biol **17:** 1050-1054

Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. EMBO J **19:** 6150-6161

Kao Y-Y, Harding SA, Tsai C-J (2002) Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen. Plant Physiol **130:** 796-807

Kellogg E (2001) Evolutionary history of the grasses. Plant Physiol **125:** 1198 - 1205

Kiesselbach TA (1949) The structure and reporduction of corn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

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

Kimura S, Laosinchai W, Itoh T, Cui XJ, Linder CR, Brown RM (1999) Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *Vigna angularis*. Plant Cell **11:** 2075

Ko J-H, Kim W-C, Han K-H (2009) Ectopic expression of *MYB46* identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. Plant J **60:** 649-665

Kokubo A, Sakurai N, Kuraishi S, Takeda K (1991) Culm brittleness of barley (*Hordeum vulgare* L.) mutants is caused by smaller number of cellulose molecules in cell wall. Plant Physiol **97:** 509-514

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 Dev **19:** 1855-1860

Lee SJ, Warnick TA, Leschine SB, Hazen SP (2012) A high-throughput biological conversion assay for determining lignocellulosic quality. Methods Mol Biol **918:** 341-349

Lee SJ, Warnick TA, Pattathil S, Alvelo-Maurosa JG, Serapiglia MJ, McCormick H, Brown V, Young NF, Schnell DJ, Smart LB, Hahn MG, Pedersen JF, Leschine SB, Hazen SP (2012) Biological conversion assay using *Clostridium phytofermentans* to estimate plant feedstock quality. Biotechnol Biofuels **5:** 5

Li A, Xia T, Xu W, Chen T, Li X, Fan J, Wang R, Feng S, Wang Y, Wang B, Peng L (2013) An integrative analysis of four CESA isoforms specific for fiber cellulose production between *Gossypium hirsutum* and *Gossypium barbadense*. Planta **237:** 1585-1597

Li C, Dubcovsky J (2008) Wheat FT protein regulates *VRN1* transcription through interactions with FDL2. Plant J **55:** 543-554

Li E, Wang S, Liu Y, Chen J-G, Douglas CJ (2011) OVATE FAMILY PROTEIN4 (OFP4) interaction with KNAT7 regulates secondary cell wall formation in Arabidopsis thaliana. Plant J **67:** 328-341

Lois R, Dietrich A, Schulz W (1989) A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive cis-acting elements. EMBO J **8:** 1641-1648

Lv B, Nitcher R, Han X, Wang S, Ni F, Li K, Pearce S, Wu J, Dubcovsky J, Fu D (2014) Characterization of *FLOWERING LOCUS T1*(*FT1*) gene in *Brachypodium* and wheat. PLoS ONE **9:** e94171

Matos DA, Whitney IP, Harrington MJ, Hazen SP (2013) Cell walls and the developmental anatomy of the *Brachypodium distachyon* stem internode. PLoS ONE **8:** e80640

McCarthy RL, 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:** 1950-1964

Michaels SD, Amasino RM (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell **11:** 949-956

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 regulate secondary wall thickenings and are required for anther dehiscence. Plant Cell **17:** 2993-3006

Nakamura Y, Andrés F, Kanehara K, Liu Y-c, Dörmann P, Coupland G (2014) Arabidopsis florigen FT binds to diurnally oscillating phospholipids that accelerate flowering. Nat Commun **5**

Oh S, Park S, Han K-H (2003) Transcriptional regulation of secondary growth in *Arabidopsis thaliana*. J Exp Bot **54:** 2709-2722

Ohashi-Ito K, Oda Y, Fukuda H (2010) Arabidopsis VASCULAR-RELATED NAC-DOMAIN6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. Plant Cell **22:** 3461-3473

Oldenbourg R, Mei G (1995) New polarized light microscope with precision universal compensator. J Microsc **180:** 140-147

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob ur R, Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS (2009) The Sorghum bicolor genome and the diversification of grasses. Nature **457:** 551-556

Perez-Rodriguez P, Riano-Pachon DM, Correa LGG, Rensing SA, Kersten B, Mueller-Roeber B (2010) PlnTFDB: updated content and new features of the plant transcription factor database. Nucleic Acids Res **38:** D822-D827

Persson S, Paredez A, Carroll A, Palsdottir H, Doblin M, Poindexter P, Khitrov N, Auer M, Somerville CR (2007) Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in Arabidopsis. Proc Natl Acad Sci USA **104:** 15566-15571

Preston J, Wheeler J, Heazlewood J, Li SF, Parish RW (2004) AtMYB32 is required for normal pollen development in *Arabidopsis thaliana*. Plant J **40:** 979-995

Pruneda-Paz JL, Breton G, Para A, Kay SA (2009) A functional genomics approach reveals CHE as a component of the Arabidopsis circadian clock. Science **323:** 1481-1485

Pyo H, Demura T, Fukuda H (2007) TERE; a novel cis-element responsible for a coordinated expression of genes related to programmed cell death and secondary wall formation during differentiation of tracheary elements. Plant J **51:** 955-965

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

Rancour D, Marita J, Hatfield RD (2012) Cell wall composition throughout development for the model grass *Brachypodium distanchyon*. Front Plant Sci **3:** 266

Ream TS, Woods DP, Schwartz CJ, Sanabria CP, Mahoy JA, Walters EM, Kaeppler HF, Amasino RM (2014) Interaction of photoperiod and vernalization determines flowering time of *Brachypodium distachyon*. Plant Physiol **164:** 694-709

Rooney WL, Blumenthal Jr, Bean B, Mullet JE (2007) Designing sorghum as a dedicated bioenergy feedstock. Biofpr **1:** 147-157

Ruelens P, de Maagd RA, Proost S, Theiben G, Geuten K, Kaufmann K (2013) FLOWERING LOCUS C in monocots and the tandem origin of angiosperm-specific MADS-box genes. Nat Commun **4:** 2280

Ruland W (1961) X-ray determination of crystallinity and diffuse disorder scattering. Acta Cryst **14:** 1180-1185

Ruprecht C, Mutwil M, Saxe F, Eder M, Nikoloski Z, Persson S (2011) Large-scale co-expression approach to dissect secondary cell wall formation across plant species. Front Plant Sci **2:** e23

Schmer MR, Vogel KP, Mitchell RB, Perrin RK (2008) Net energy of cellulosic ethanol from switchgrass. Proc Natl Acad Sci USA **105:** 464-469

Schuetz M, Smith R, Ellis B (2013) Xylem tissue specification, patterning, and differentiation mechanisms. J Exp Bot **64:** 11-31

Shen H, He X, Poovaiah CR, Wuddineh WA, Ma J, Mann DGJ, Wang H, Jackson L, Tang Y, Neal Stewart C, Chen F, Dixon RA (2012) Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. New Phytol **193:** 121-136

Shiu S-H, Shih M-C, Li W-H (2005) Transcription factor families have much higher expansion rates in plants than in animals. Plant Physiol **139:** 18-26

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

Somerville C, Koornneef M (2002) A fortunate choice: the history of Arabidopsis as a model plant. Nat Rev Genet **3:** 883-889

Sonbol F-M, Fornalé S, Capellades M, Encina A, Touriño S, Torres J-L, Rovira P, Ruel K, Puigdomènech P, Rigau J, Caparrós-Ruiz D (2009) The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in *Arabidopsis thaliana*. Plant Mol Biol **70:** 283-296

Song D, Shen J, Li L (2010) Characterization of cellulose synthase complexes in Populus xylem differentiation. New Phytol **187:** 777-790

Soyano T, Thitamadee S, Machida Y, Chua N-H (2008) ASYMMETRIC LEAVES2-LIKE19/LATERAL ORGAN BOUNDARIES DOMAIN30 and ASL20/LBD18 regulate tracheary element differentiation in Arabidopsis. Plant Cell **20:** 3359-3373

Sundström JF, Nakayama N, Glimelius K, Irish VF (2006) Direct regulation of the floral homeotic *APETALA1* gene by *APETALA3* and *PISTILLATA* in Arabidopsis. Plant J **46:** 593-600

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:** 2731-2739

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

Taylor NG (2008) Cellulose biosynthesis and deposition in higher plants. New Phytologist **178:** 239-252

Taylor NG, Howells RM, Huttly AK, Vickers K, Turner SR (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. Proc Natl Acad Sci USA **100:** 1450-1455

Taylor NG, Laurie S, Turner SR (2000) Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in Arabidopsis. Plant Cell **12:** 2529-2540

Taylor NG, Scheible WR, Cutler S, Somerville CR, Turner SR (1999) The irregular xylem3 locus of Arabidopsis encodes a cellulose synthase required for secondary cell wall synthesis. Plant Cell **11:** 769

Trevaskis B, Bagnall DJ, Ellis MH, Peacock WJ, Dennis ES (2003) MADS box genes control vernalization-induced flowering in cereals. Proc Nat Acad Sci **100:** 13099-13104

Turner SR, Somerville CR (1997) Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. Plant Cell **9:** 689-701

Tusher V, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA **98:** 5116 - 5121

Valdivia ER, Herrera MT, Gianzo C, Fidalgo J, Revilla G, Zarra I, Sampedro J (2013) Regulation of secondary wall synthesis and cell death by NAC transcription factors in the monocot Brachypodium distachyon. J Exp Bot **64:** 1333-1343

Vignols F, Rigau J, Torres M, Capellades M, Puigdomenech P (1995) The *brown midrib3* (*bm3*) mutation in maize occurs in the gene encoding caffeic acid O-methyltransferase. Plant Cell **7:** 407 - 416

Vogel J (2008) Unique aspects of the grass cell wall. Curr Opin Plant Biol **11:** 301-307

Vogel J, Garvin D, Leong O, Hayden D (2006) *Agrobacterium*-mediated transformation and inbred line development in the model grass *Brachypodium distachyon*. Plant Cell Tiss Org **84:** 100179-100191

Vogel J, Hill T (2008) High-efficiency *Agrobacterium*-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3. Plant Cell Rep **27:** 471-478

Wang H, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA (2010) Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. Proc Natl Acad Sci USA **107:** 22338-22343

Wang H, Zhao Q, Chen F, Wang M, Dixon RA (2011) NAC domain function and transcriptional control of a secondary cell wall master switch. Plant J **68:** 1104-1114

Wang HH, Tang RJ, Liu H, Chen HY, Liu JY, Jiang XN, Zhang HX (2013) Chimeric repressor of PtSND2 severely affects wood formation in transgenic Populus. Tree Physiol **33:** 878-886

Wang L, Guo K, Li Y, Tu Y, Hu H, Wang B, Cui X, Peng L (2010) Expression profiling and integrative analysis of the *CESA*/*CSL* superfamily in rice. BMC Plant Biol **10:** 282

Warthmann N, Chen H, Ossowski S, Weigel D, Hervre P (2008) Highly specific gene silencing by artificial miRNAs in rice. PLoS ONE **3:** e1829

Wu L, Liu D, Wu J, Zhang R, Qin Z, Liu D, Li A, Fu D, Zhai W, Mao L (2013) Regulation of *FLOWERING LOCUS T* by a MicroRNA in *Brachypodium distachyon*. Plant Cell **25:** 4363-4377

Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F (2004) A model-based background adjustment for oligonucleotide expression arrays. J Am Statis Assoc **99:** 909-917

Yamaguchi M, Ohtani M, Mitsuda N, Kubo M, Ohme-Takagi M, Fukuda H, Demura T (2010) VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in Arabidopsis. Plant Cell **22:** 1249-1263

Yang C, Xu Z, Song J, Conner K, Vizcay Barrena G, Wilson ZA (2007) Arabidopsis MYB26/MALE STERILE35 regulates secondary thickening in the endothecium and Is essential for anther dehiscence. Plant Cell **19:** 534-548

Yao D, Wei Q, Xu W, Syrenne R, Yuan J, Su Z (2012) Comparative genomic analysis of NAC transcriptional factors to dissect the regulatory mechanisms for cell wall biosynthesis. BMC Bioinformatics **13:** S10

Yin YB, Huang JL, Xu Y (2009) The cellulose synthase superfamily in fully sequenced plants and algae. BMC Plant Biol **9:** 99-113

Yoshida K, Sakamoto S, Kawai T, Kobayashi Y, Sato K, Ichinose Y, Yaoi K, Akiyoshi-Endo M, Sato H, Takamizo T, Ohme-Takagi M, Mitsuda N (2013) Engineering the *Oryza sativa* cell wall with rice NAC transcription factors regulating secondary wall formation. Front Plant Sci **4:** 383

Zhang B, Deng L, Qian Q, Xiong G, Zeng D, Li R, Guo L, Li J, Zhou Y (2009) A missense mutation in the transmembrane domain of CESA4 affects protein abundance in the plasma membrane and results in abnormal cell wall biosynthesis in rice. Plant Mol Biol **71:** 509-524

Zhao C, Craig JC, Petzold HE, Dickerman AW, Beers EP (2005) The xylem and phloem transcriptomes from secondary tissues of the Arabidopsis root-hypocotyl. Plant Physiol **138:** 803-818

Zhao Q, Wang H, Yin Y, Xu Y, Chen F, Dixon RA (2010) Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. Proc Natl Acad Sci USA **107:** 14496-14501

Zhong R, Demura T, 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, Demura T, Ye ZH (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, McCarthy RL, Reeves CK, Jones EG, Ye Z-H (2011) Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. Plant Cell Physiol **52:** 1856-1871

Zhong R, Lee C, Ye Z-H (2010) Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. Trends in Plant Science **15:** 625-632

Zhong R, Lee C, Ye ZH (2010) 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 RL, 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, Richardson E, Ye Z-H (2007) 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, Richardson EA, Ye ZH (2007) The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. Plant Cell **19:** 2776-2792

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

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