12-2013

Reducing cell wall recalcitrance and increasing biomass for enhanced ethanol production in switchgrass (Panicum virgatum L.)

Charleson Rajendran Poovaiah
University of Tennessee - Knoxville, cpoovaia@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Charleson Rajendran Poovaiah entitled "Reducing cell wall recalcitrance and increasing biomass for enhanced ethanol production in switchgrass (Panicum virgatum L.)." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Plants, Soils, and Insects.

C. Neal Stewart Jr., Major Professor

We have read this dissertation and recommend its acceptance:

Robert N. Trigiano, Feng Chen, Zong-Ming Cheng

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Reducing cell wall recalcitrance and increasing biomass for enhanced ethanol production in switchgrass (*Panicum virgatum* L.)

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Charleson Rajendran Poovaiah

December 2013

Dedication

This dissertation is dedicated to my lovely wife, Regina, and son, Marcus, my parents, G. Poovaiah and Retnavathy Poovaiah, my parents-in-law, Leela and Henry Iyathurai, and my aunt and uncle Mrs. & Dr. Augustine Peter whose love, patience, and prayers allowed for its completion.

Acknowledgements

I would like to offer my sincere appreciation to Dr. C. Neal Stewart, Jr. for taking on the role of my major professor in Plant Sciences. He has been more than an academic advisor for me during my graduate studies. His drive for excellence in science and sociable personality inspired me to try to be a better scientist. I would also sincerely thank my committee members for their guidance through my PhD studies.

I would like to thank Dr. Robert Trigiano for his continuous encouragement and time and Dr. Zong-Ming Cheng for his guidance and encouragement. I would like to thank Dr. Feng Chen for his advice and showing me his passion for science as a young scientist. I would like to thank Mitra Mazarei and David Mann for their advice and help throughout my PhD studies and previous lab members, Murali Rao and Hong Moon for their help and fun memories. And a special thanks to my fellow colleagues for their assistance and advice: Reggie Millwood, Jason Burris, Kellie Burris, Jonathan Willis, Jennifer Hinds and all other members of the Stewart lab.

I would like to thank Dr. Jonathan Mielenz, Dr. Kelsey Yee, Miguel Rodriguez and Choo Hamilton for their advice and assistance for experiments at Oak Ridge National Laboratory.

I would like to express my deepest gratitude to my lovely wife, Regina, for her love, patience, and encouragement during my PhD studies. I would like to thank my parents, G. Poovaiah and Retnavathy Poovaiah, my sister Pearlin Jacob for their love, prayers, patience, support and encouragement. I would also like to thank my parents-in-laws, Henry and Leela Iyathurai, for their love, prayers and support. I couldn't finish this work without their love, patience, and prayers. I specially thank my aunt and uncle Dr. Augustine Peter and Damayanthy Peter for their

motivation, inspiration, encouragement and moral support throughout my stay in US. Also I would like to my family both in US and India for their love and prayers.

Lastly, I would like to thank God for his endless and bottomless love for me.

Abstract

Lignocellulosic biomass is a potential large-scale biofuel feedstock for conversion to ethanol through saccharification and fermentation. The presence of lignin in lignocellulosic biomass impedes its degradation and subsequent fermentation. The removal of lignin by pretreatment is the most expensive step in the production of lignocellulosic biofuels. Manipulation of monolignol pathway is needed to reduce lignin and for the rational design of engineered cell walls of lignocellulosic feedstocks.

PvMYB4, a transcriptional repressor of lignin gene expression was identified and evaluated for its potential for improving switchgrass as a lignocellulosic feedstock. Ectopic overexpression of *PvMYB4* in transgenic switchgrass reduced lignin content and increased sugar release efficiency up to three-fold over that of the non-transgenic control. The transgenic plants yielded up to 2.6-fold more ethanol than controls. Detailed biomass characterization revealed alteration of lignin content, xylan/pectin and lignin linkages, lignin polymer size and internal linkages in lignin leading to reduced recalcitrance and increased ethanol yield. Genetically engineered *PvMYB4* switchgrass therefore can be used as potential germplasm for improvement of lignocellulosic feedstocks. It is currently being grown in field experiments.

Increasing biomass production is also important for biofuel crops. Sucrose synthase catalyzes the conversion of sucrose and uridine di-phosphate (UDP) into UDP-glucose which is used by cellulose synthase for cell wall biosynthesis. Bioinformatic and cluster analysis was used to identify four sucrose synthase genes in switchgrass. Transient subcellular localization revealed that *PvSUS1* localizes to the plasma membrane. Transgenic switchgrass plants overexpressing

PvSUS1 had increases in biomass and cellulose content. For switchgrass and other bioenergy feedstocks, the overexpression of *SUS1* genes might be a realistic strategy to increase both plant biomass and cellulose content to maximize biofuel production per land area cultivated.

Taken together *PvMYB4* and *PvSUS1* would be two excellent candidate genes to stack in switchgrass to provide improvements in two important aspects of feedstock improvement: greater amounts of less recalcitrant biomass. In addition, these genes could also be overexpressed in other crops such as maize. It could be envisage that a US corn crop overexpressing both these genes together could give greater corn grain yield and less recalcitrant stover for lignocellulosic biofuel production.

Table of contents

List of Tables

List of Figures

List of Abbreviations

4CL	4-Coumarate:CoA ligase
AG	Arabinogalactan
AIR	Alcohol insoluble residue
ARK1	ARBORKNOX
C3H	*p*-Coumarate 3-hydroxylase
C4H	Cinnamate 4-hydroxylase
CA	*p*-Coumaric acid
CAD	Cinnamyl alcohol dehydrogenase
CCoAOMT	Caffeoyl-CoA O-methyltransferase
CCR	Cinnamoyl-CoA reductase
Cg1	Corngrass1
CHS	Chalcone synthase
CHI	Chalcone isomerase
COMT	Caffeic acid 3-O-methyltransferase
CP/MAS NMR	Cross polarization/magic angle spinning nuclear magnetic resonance
CWR	Cell wall residues
DMSO	Dimethylsulfoxide
EDTA	Ethylenediaminetetraacetic acid
F5H	Ferulate 5-hydroxylase
FA	Ferulic acid
FLS F	Flavonol synthase

G	Guaiacyl lignin
GA20ox	Gibberellin 20-oxidase
GC-MS	Gas chromatography–mass spectrometry
GPC	Gel-permeation chromatography
H	*p*-hydroxyphenyl lignin
HCT,	*p*-hydroxylcinnamoyl CoA\:quinate/shikimate *p*-hydroxycinnamoyltransferase
HG	Homogalacturonan
HMF	5-hydroxymethylfurfural
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single quantum coherence
mAbs	Monoclonal antibodies
MESP	Minimum ethanol selling price
miRNAs,	microRNAs
MYB	Myeloblastosis family
NST1	NAC secondary wall thickening promoting factor
PAL	Phenylalanine ammonia-lyase
PvMYB4-OX	Switchgrass plants overexpressing the switchgrass MYB4 gene
RG	Rhamnogalacturonan
RNAi	Ribonucleic acid interference
S	Syringyl lignin
SAD	Sinapyl alcohol dehydrogenase

SND	Secondary wall-associated NAC domain protein1
SPL	SQUAMOSA PROMOTER BINDING LIKE
SSF	Simultaneous saccharification and fermentation
THF	Tetrahydrofuran
UBQ	ubiquitin

Chapter 1: Introduction

1.1 Lignocellulosic biomass as a source for biofuels

Plant cell walls are made up of three main components – cellulose, hemicellulose and lignin. Cellulose, a structural polymer is found in plants organized in to crystalline fibers in combination with other polysaccharides like hemicellulose and bound together by lignin mainly through hydrogen bonds. Together they form the complex lignocellulosic substrate. Lignocellulosic biomass provides an excellent source of mixed sugars for fermentation into biofuels and other materials. Sources of lignocellulosic biomass include solid municipal wastes, agriculture and forestry residues and dedicated bioenergy crops like poplar and switchgrass. Nearly 70% of plant biomass consists of 5- and 6-carbon sugars. These sugars are found in lignocellulosic biomass, which is comprised of mainly cellulose (a homologous polymer comprised of long chains of glucose); less so, hemicelluloses (heterologous polymer of 5- and 6-carbon sugars); and least of all lignin (a complex aromatic polymer). The major component cellulose is a homopolysaccharide comprised of glucose units, linked by β-(1→4) glycosidic bonds. Cellobiose is the smallest repeating unit of cellulose and can ultimately be converted into glucose. can potentially provide inputs for producing bioethanol equivalent to 30% of US gasoline consumption (Perlack et al., 2005). Reduction of recalcitrance can improve sugar release efficiency and increase the amount of ethanol produced. Additionally, reducing recalcitrance can also decrease the cost of pretreatment and reduce the amount of enzymes used to breakdown cellulosic biomass into simple sugars for fermentation (Chen & Dixon, 2007).

1.2 Switchgrass as a lignocellulosic biomass source

Switchgrass (*Panicum virgatum* L.), a perennial warm season C4 grass which has been identified as a potential crop for lignocellulosic ethanol production because of its faster growth rate, its ability to use nutrients efficiently and its widespread distribution throughout North America (McLaughlin & Walsh, 1998). Being a perennial grass switchgrass increases water quality, enhances carbon sequestration and reduces soil erosion and greenhouse gas emissions with an estimated reduction of 94% compared to greenhouse gas emissions from petroleum fuel (Schmer et al., 2008). Also in this study switchgrass yielded 540% more energy than the energy used to grow and process it.

Switchgrass has a deep root system which increases its capacity to utilize water and nutrients from deeper regions in the soil. The root system also stores nutrients and energy in the soil and increases carbon sequestration by increasing soil organic carbon levels. According to a study by Liebig et al., (2008) switchgrass fields had an average of seven tons more soil carbon per acre. Greater soil carbon under switchgrass was observed at all depths, but it was most seen at one to three feet - a depth in the soil profile where switchgrass has more root biomass.

Switchgrass is tolerant of wet and dry soils with the lowland varieties more suited to wet soil and upland varieties are suited to well-drained soils (Moser & Vogel, 1995). Switchgrass, being a native grass has a low risk of being invasive (Moser & Vogel, 1995) and it is easy to propagate with a quick stand establishment. Soil acidity does not limit the growth of switchgrass. In fact switchgrass seedlings are highly tolerant of acid soils and switchgrass have been reported to grow at pH 3.7 and switchgrass has been used in the reclamation of acid soils (Bona & Belesky,

1992, Stucky et al., 1980). Switchgrass, being a native grass has a low risk of being invasive (Moser & Vogel, 1995) and it is easy to propagate with a quick stand establishment. The above mentioned facts demonstrates that switchgrass is a good candidate as a bioenergy crop.

1.3 Genetic manipulation of switchgrass for biofuels

Value-added genes that cannot be transferred through crossing and selection can be incorporated into switchgrass through transgenic technology. *Agrobacterium*-mediated transformation efficiency of switchgrass has greatly improved with the identification of elite genotypes through selection. Incorporation of value added genes in to switchgrass should have a high impact on recalcitrance and biomass yield for biofuel purposes.

1.3.1 Lignin manipulation

Manipulation of monolignol biosynthetic pathway is one such trait as it reduces resistance of the cell wall to degradation by enzymes. Lignin is a complex phenolic heteropolymer associated with polysaccharides in the cell wall. Studies have shown that there is a strong negative correlation of lignin amount with digestibility (Jung & Vogel, 1986, Chen & Dixon, 2007). Enzymatic saccharification requires a thermochemical pretreatment step to increase the digestibility of the lignocellulosic feedstock. Manipulating lignin content and composition and cell wall characteristics through genetic engineering can reduce biomass conversion costs by developing crop varieties that have less lignin. Genetically engineered plants optimized for ethanol production will enable us to even skip the pretreatment step further reducing cost for ethanol production (Chen & Dixon, 2007). Optimizing plant biomass for more efficient cellulosic ethanol processing requires a better understanding of cell wall synthesis which can be

done by manipulating the genes in the cell wall synthesis pathway. The plant cell has evolved complex chemical and structural mechanisms to resist degradation by microbes and animals (Akin, 2007). These complex mechanisms must be broken down to release the sugars and these degradation processes must also be economically viable for commercial ethanol production. Cellulose, which can be fermented to ethanol, is tightly linked to lignin and pretreatment is necessary to effectively release the polysaccharides from the lignocellulose for saccharification and ethanol production (Mosier et al., 2005). Knowledge of genes affecting cell wall synthesis enables modification of the cell wall to reduce recalcitrance of lignocellulosic feedstocks and to produce economically-viable ethanol.

1.3.2 Increasing plant biomass

Genetic modification can also be used to increase biomass yield in switchgrass. Modification of plant growth regulators such as gibberellin biosynthetic genes has been used to increase plant biomass (Eriksson et al., 2000) and other potential ways to increase biomass yield have been discussed before (Sticklen, 2007, Rojas et al., 2010). Enzymes involved in sucrose catabolism have been demonstrated to increase in biomass yield in dicots (Coleman et al., 2010, Jiang et al., 2012). It is highly possible that many undiscovered genes that can increase plant growth are in existence.

The goal of this research was to determine if modifications to the monolignol biosynthetic pathway can enhance ethanol production and also to determine if modification of sucrose catabolism pathway can increase biomass yield in swtichgrass. Combination of transgenic

technology and traditional breeding will allow us to develop switchgrass with reduced recalcitrance and higher biomass yield for enhanced ethanol production.

1.4 References

Akin DE. 2007. Grass lignocellulose: strategies to overcome recalcitrance. *Applied Biochemistry and Biotechnology* 137-140: 3-15.

Bona L, Belesky DP. 1992. Evaluation of swtichgrass entries for acid soil tolerances. *Communications in Soil Science and Plant Analysis* 23: 1827-1841.

Chen F, Dixon RA. 2007. Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology* 25: 759-761.

Coleman H, Beamish L, Reid A, Park J-Y, Mansfield S. 2010. Altered sucrose metabolism impacts plant biomass production and flower development. *Transgenic Research* 19: 269-283.

Eriksson ME, Israelsson M, Olsson O, Moritz T. 2000. Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nature Biotechnology* 18: 784-788.

Jiang Y, Guo W, Zhu H, Ruan YL, Zhang T. 2012. Overexpression of GhSusA1 increases plant biomass and improves cotton fiber yield and quality. *Plant Biotechnology Journal* 10: 301-312.

Jung HG, Vogel KP. 1986. Influence of lignin on digestibility of forage cell wall material. *Journal of Animal Science* 62: 1703-1712.

Liebig MA, Schmer MR, Vogel KP, Mitchell RB. 2008. Soil carbon storage by switchgrass grown for bioenergy. *BioEnergy Research* 1: 215-222.

McLaughlin SB, Walsh ME. 1998. Evaluating environmental consequences of producing herbaceous crops for bioenergy. *Biomass & Bioenergy* 14: 317-324.

Moser LE, Vogel KP 1995. Switchgrass, Big Bluestem, and Indiangrass.In R.F. Barnes, D.A. Miller, Nelson CJ. *Forages Volume 1: An introduction to grassland agriculture*: Iowa State University Press. 409-420.

Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology* 96: 673-686.

Perlack RD, Wright LL, Turhollow AF, Graham RL, Stokes BJ, Erbach DC 2005. Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. OAK RIDGE NATIONAL LAB TN.

Rojas CA, Hemerly AS, Ferreira PCG. 2010. Genetically modified crops for biomass increase. Genes and strategies. *GM Crops* 1: 137-142.

Schmer MR, Vogel KP, Mitchell RB, Perrin RK. 2008. Net energy of cellulosic ethanol from switchgrass. *Proceedings of the National Academy of Sciences, USA* 105: 464-469.

Sticklen MB. 2007. Feedstock crop genetic engineering for alcohol fuels. *Crop Science* 47: 2238-2248.

Stucky D, Bauer J, Lindsey T. 1980. Restoration of acidic mine spoils with sewage sludge: I Revegetation. *Reclam. Rev.;(United States)* 3.

Chapter 2: Changing lignin to improve lignocellulosic biofuel feedstocks

(This chapter has been submitted to Plant Biotechnology Journal with the following authors Charleson R. Poovaiah, Madhugiri Nageswara-Rao, Jaya R. Soneji and C. Neal Stewart Jr.)

Charleson R Poovaiah is the principle author of the manuscript

2.1 Abstract

Lignocellulosic feedstocks can be converted to biofuels, which can conceivably replace a large fraction of fossil fuels currently used for transformation. However, lignin, a prominent constituent of secondary cell walls, is an impediment to the conversion of cell walls to fuel: the recalcitrance problem. Biomass pretreatment for removing lignin is the most expensive step in the production of lignocellulosic biofuels. Even though we have learned a great deal about the biosynthesis of lignin, we do not fully understand its role in plant biology, which is needed for the rational design of engineered cell walls for lignocellulosic feedstocks. This review will recapitulate our knowledge of lignin biosynthesis and discuss how lignin has been modified and the consequences for the host plant.

Keywords: Bioenergy, Fermentable sugars, Genetic manipulation, Lignin biosynthetic pathway, Transcription factors

2.2 Lignocellulosic biomass as a source for biofuel

Lignocellulosic biomass is a potential large-scale biofuel feedstock wherein entire aboveground biomass would be converted to fuels. The major candidates of lignocellulosic biomass are dedicated biofuel crops such as switchgrass, *Miscanthus*, and poplar, as well as "waste" products from forest biomass, crop residuals and household refuse. In the US alone, biofuels could sustainably replace nearly 30% of petroleum currently used in transportation fuels. Dedicated perennial biomass crops could have many environmental and resource advantages over non-renewable fossil fuels as well as annual crops. For example, a five-year field study led to projections that lignocellulosic biofuels derived from switchgrass would decrease greenhouse gas emissions and would act as a carbon sink compared with fossil fuels equivalents (Schmer et al., 2008).

Carbohydrate-rich plant cell walls are the major energy sink in plant biomass (Hisano et al., 2009). After growth and harvest of biomass, the major biofuel expense is pretreatment to loosen/breakdown cell walls, followed by saccharification and their conversion into biofuels (Rubin, 2008). Although lignocellulosic biofuels potentially provide benefits like reducing greenhouse gases, the high costs of pretreatment, saccharification and fermentation can render them economically non-viable when taken in context of other fuel opportunities (Hisano et al., 2009). Therefore, a major challenge is to decrease the effort and cost to convert these feedstocks into fuel. Fortunately, advances in our knowledge of the lignin biosynthetic pathway and plant biotechnology can be used to reduce the resistance of lignocellulosic biomass to breakdown in to simple sugars and increase biofuel production; we now have effective tools to address the recalcitrance problem.

2.3 Why lignin needs to be manipulated in lignocellulosic feedstocks

Cellulose is Earth's most prevalent polymer and the primary chemical feedstock of biofuels. It is, however, difficult to convert it into simple sugars and biofuels because of its crystallinity and occlusion by lignin; decreasing it in cell walls enhances saccharification (Saballos et al., 2012; Vanholme et al., 2010; Vanholme et al., 2008). It is thought that lignin occludes cellulose by physically shielding it from microbial or enzymatic decomposition (Sticklen, 2006; Vanholme et al., 2010; Weng et al., 2008b). Lignin can also adsorb hydrolytic enzymes to prevent them from breaking down cellulose into monosaccharides. In addition lignin degradation products inhibit the activity of the enzymes for fermentation of monosaccharides (Keating et al., 2006). Such breakdown products or derivatives of lignin can also negatively affect the growth of microorganisms that can be used in consolidated bioprocessing (Agbor et al., 2011). Thus, lignin has received significant attention in biofuels research with the goal of improving conversion efficiency.

The assumption has been made that altering lignin in the plant will be less expensive than eliminating it via pretreatment. Pretreatment methods, such as acid hydrolysis, alkaline hydrolysis, and enzymatic hydrolysis are currently being used to breakdown the lignocellulosic matrix and remove lignin (Kumar et al., 2009; Mosier et al., 2005). Given the cost of pretreatment, altering lignin has been viewed as a strategy to decrease, or even eliminate pretreatment. Reducing the cost of pretreatment and microbial/enzyme-based conversion is a major driver for altering lignin in feedstock.

2.4 Evolution and biosynthesis of lignin

It is widely accepted that the adaptation of vascular plants to terrestrial habitats was made possible by lignin (Figure 1), which provides structural support and affiliated traits required for an erect growth habit (Bonawitz and Chapple, 2010; Vanholme et al., 2010). Lignin, once believed to be limited to vascular plants, has recently been discovered in the bryophyte *Marchantia polymorpha* (Espineira et al., 2011). Gene network studies have revealed the earliest appearance of the monolignol biosynthetic genes families in the moss *Physcomitrella patens* but lignin and lignin monomers have not been found (Xu et al., 2009). Studies by Martone et al. (2009), however, have identified the existence of secondary walls and lignin in the marine red alga *Calliarthron cheilosporioides* through derivatization followed by reductive cleavage (DFRC), a β-O-4 linkage specific lignin analysis technique. The relevant pathways might have evolved more than one billion years ago prior to the divergence of red and green algae or have developed convergently in *C. cheilosporioides* and land plants (Martone et al., 2009). Syringyl lignin (S-lignin), which was thought to be a unique characteristic of angiosperms, has been found in some lycophytes (*Selaginella*), ferns (*Dennstaedtia bipinnata*) and gymnosperms (*Podocarpus macrophyllus* and *Tetraclinis articulate*). Thus, it might have evolved independently among plant taxa (Espineira et al., 2011; Martone et al., 2009; Weng et al., 2008a; Weng and Chapple, 2010). The discovery of a novel ferulate 5-hydroxylase (*F5H*) in *Selaginella*, which is structurally distinct from the angiosperm *F5H*, supports the hypothesis that the enzyme evolved independently at least twice (Vanholme et al., 2010; Weng and Chapple, 2010). It is interesting to note that *Ginkgo biloba* cell suspension cultures have the ability to synthesize S-lignin, a trait not observed in its woody tissues (Uzal et al., 2009).

Lignin is the second most abundant polymer in nature. Lignin functions as inter- and intra-molecular adhesive, providing rigidity to secondary cell walls and the plant body (Patzlaff et al., 2003). Lignin is formed via oxidative coupling of dihydroxycinnamyl alcohols (or monolignols), hydroxycinnamaldehydes (coniferaldehyde and sinapaldehyde), ferulates and acylated monolignols (Pomar et al., 2012). It is synthesized from *p*-coumaryl, coniferyl and sinapyl alcohols which produce *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units (Figure 2). The relative quantity of each component varies by species, organ, and growth stage (Hisano et al., 2009). G and S units are the main components of the dicot lignin, and a combination of G, S, and H units form the monocot lignin while gymnosperm lignin comprises of mostly G-units along with low levels of H-units (Boerjan et al., 2003). Biosynthesis of monolignols followed by their polymerization is the two main steps in lignin biosynthesis (Figure 2). Monolignol biosynthesis is thought to be highly conserved among angiosperm taxa, at least in its basic layout. The deamination of phenylalanine to produce cinnamic acid in the presence of phenylalanine ammonia lyase (*PAL*) followed by the conversion of cinnamic acid to *p*-coumaric acid by cinnamate 4-hydroxylase (*C4H*), are the initial steps in the monolignol biosynthesis pathway. Following the conversion to *p*-coumaric acid, the aromatic ring of cinnamic acid is *O*-methylated and its side chain, through reduction is converted from acid to alcohol, which leads to the formation of monolignols (Humphreys and Chapple, 2002; Xu et al., 2009) (Figure 2). Characterization of most of the enzymes in the monolignol biosynthesis pathway along with the identification of various intermediates and precursors has led to the discovery of pathways for further hydroxylation and methylation steps (Humphreys and Chapple, 2002; Ye et al., 1994). The quantity of various monolignols produced in the earlier steps of lignin biosynthesis is the

main determinant of the lignin polymer structure and is dependent on the plant species, cell type and developmental stage (Bonawitz and Chapple, 2010).

2.5 Cornucopia of choices for lignin modifications

The genes in the lignin biosynthesis pathway have been identified (Bonawitz and Chapple, 2010; Vanholme et al., 2010; Weng and Chapple, 2010). More than 10 enzymatic conversions are involved in monolignol synthesis. Polymerization of monolignols is catalyzed by peroxidases and laccases (Berthet et al., 2011; Herrero et al., 2013; Lee et al., 2013). The expression of monolignol biosynthesis genes are regulated by a series of transcription factors during secondary cell wall formation (Zhao and Dixon, 2011; Zhong et al., 2010a; Zhong et al., 2008). It is clear that there are many approaches that can be used to change lignin in cell walls using biotechnology given the large number of biosynthetic genes and transcription factors that are involved. Indeed, we know more about the identity of biosynthetic enzymes than the transcription factors that regulate the synthesis of lignin.

2.5.1 Manipulation of monolignol biosynthetic genes

Manipulation of genes in the monolignol biosynthesis pathway has a significant effect on total lignin and lignin composition. Depending on genes chosen for modification, parts or all of the metabolite flux can be restricted for monolignol synthesis (Vanholme et al., 2012). A number of genes have been overexpressed or downregulated in transgenic plants to understand their effect on lignin biosynthesis (Table 1). The enzyme at the top of the phenylpropanoid pathway is phenylalanine ammonia lyase (PAL) and caffeoyl-CoA *O*-methyltransferase (CCoAOMT) is the enzyme at the end of the general pathway. Monolignol-specific biosynthesis starts with cinnamoyl-CoA reductase (CCR) (Figure 2). Downregulation of genes upstream in the

phenylpropanoid pathway can reduce lignin content, but can also cause pleiotropic effects, such as stunted growth and collapsed vasculature (Bonawitz and Chapple, 2013). For example in *PAL* downregulated tobacco, the number and size of the vessels were reduced, which compromised vascular integrity of the transgenic plants (Korth et al., 2001). Allelic mutants of cinnamate 4-hydroxylase (C4H) had discrete effects on plant growth, lignin content and the S:G monolignol ratio. The weakest *Arabidopsis* mutant not only had less lignin content but also had the least yield penalty and was similar to wild type plants (Schilmiller et al., 2009).

One enzyme in the lignin biosynthesis pathway that has been studied a great deal, 4-coumarate:CoA ligase (4CL), catalyzes the conversion of *p*-coumaric and caffeic acids to their thio ester form in the phenylpropanoid pathway and takes part in the regulation of monolignol precursors thereby controlling lignin content and composition (Saballos et al., 2012). Biomass production in *4CL* downregulated plants varied from 0 to 75% reduced lignin depending on the gene expression level and host species. For example, when *4CL* was downregulated in switchgrass, there was no change in biomass yield (Xu et al., 2011), whereas *4CL* downregulation in poplar resulted in a positive relationship between *4CL* expression and biomass production (Voelker et al., 2010). When monolignol pools were reduced, flavonoids accumulated, which could be observed as reddish brown coloration in stems (Besseau et al., 2007; Voelker et al., 2010; Xu et al., 2011). Changes in lignin content and S:G ratios caused by downregulation of *4CL* greatly increased saccharification efficiency. For instance in sorghum brown rib2 mutants with mutations in *4CL* gene increased glucose release efficiency up to 25% and in switchgrass saccharification of dilute acid pretreated biomass of *4CL*-downregulated plants yielded 57% more sugars (Saballos et al., 2008; Xu et al., 2011).

Another series of conversions are catalyzed by *p*-coumarate 3-hydroxylase (C3H) and *p*-hydroxycinnamoyl-CoA:quinate/shikimate *p*-hydroxycinnamoyltransferase (HCT). C3H catalyzes the addition of hydroxyl groups at the meta-position of *p*-coumarate and is involved in the conversion of *p*-coumarate coA to caffeoyl coA. HCT catalyzes the reactions before and after the addition of the 3-hydroxyl group (Hoffmann et al., 2004). While there was reduction in height and lignin content in *HCT* downregulated plants, there was a large increase in H units besides reduction in S and G units (Shadle et al., 2007). The recent identification of a rare natural *HCT* mutant in poplar that has a C-terminal deletion and more H units showed no growth penalty. The G unit pool was significantly reduced in dormant, but not in actively growing tissues (Vanholme et al., 2013a).

The recent identification of an otherwise uncharacterized lignin biosynthesis gene is caffeoyl shikimate esterase (CSE), catalyzing the hydrolysis of caffeoyl shikimate to caffeate adds another step to the pathway that can be manipulated (Vanholme et al., 2013b). *Arabidopsis* cse T-DNA mutants have reduced lignin content and increased saccharification efficiency. Biomass yield was lower in one of the mutants compared to control and other growth abnormalities were not seen in the mutants. While manipulation of genes in the general phenylpropanoid pathway can cause severe growth defects besides reduction in lignin content, identification of natural mutants can be used to manipulate the phenylpropanoid pathway for reduced lignin content and composition without significant impact on biomass production. Also large increases in saccharification efficiency can compensate some losses in biomass production (Chen and Dixon, 2008).

Monolignol-specific pathway enzymes include CCR, cinnamyl alcohol dehydrogenase (CAD), ferulate 5-hydroxylase (F5H), and caffeic acid *O*-methyltransferase (COMT). Changes in

expression level of these genes directly affect the lignin composition and amount of lignin (Saballos et al., 2012). Digestibility of biomass was increased with downregulation of *CCR*, *COMT* and *CAD*. For instance, downregulation of *CCR* in poplar increased digestibility and doubled the amount of cell wall sugar release by *Clostridium cellulolyticum* relative to the non-transgenic parent (Boudet et al., 2003). Downregulation of *COMT* modified lignin content and, especially, composition with enhanced digestibility and no yield penalty (Fu et al., 2011a; Piquemal et al., 2002). The final step in the biosynthesis of monolignols is catalyzed by *CAD* whose downregulation directed the assimilation of hydroxycinnamaldehydes (Boerjan et al., 2003; Li et al., 2008). The involvement of *CAD* in lignin biosynthesis has been studied in poplar, eucalyptus, switchgrass, *Brachypodium* besides other species (Bouvier d'Yvoire et al., 2013; Chabannes et al., 2001; Fornale et al., 2012; Jackson et al., 2008; Lapierre et al., 2004; Ralph et al., 2001; Saathoff et al., 2011; Valerio et al., 2003). The downregulation of *CAD* in grasses reduced lignin content, modified lignin structure, and improved saccharification efficiency (Bouvier d'Yvoire et al., 2013; Dien et al., 2009; Fu et al., 2011b; Saathoff et al., 2011). In the plants studied for downregulation of *CAD*, biomass production was not significantly affected. These studies signify the importance of downregulation of monolignol biosynthetic genes and increase in saccharification efficiency. In view of the fact that *F5H* is the main enzyme responsible for S-lignin production, its overexpression led to an increase in S-lignin while its downregulation resulted in lignin mostly composed of G-units (Stewart et al., 2009). The S:G ratio could be further decreased by the downregulation of *HCT* or *C3H* genes (Coleman et al., 2008; Hoffmann et al., 2004; Ralph et al., 2006; Wagner et al., 2007).

COMT downregulation has proven to be a robust means of decreasing the S:G ratio and recalcitrance (Fu et al., 2011a; Jung et al 2012; 2013;Vanholme et al., 2008). However, increases

in the S:G ratio have also been shown to increase saccharification efficiency. A study by Li et al. (2010) in *Arabidopsis* has shown that digestibility of biomass rich in S-lignin increased when subjected to liquid hot water pretreatment indicating that higher S:G ratio might impact saccharification. But higher amounts of S-lignin in poplar had no effect on saccharification efficiency (Mansfield et al., 2012).

After monolignols are produced, they are oxidized and combined by coupling to form lignin polymers (Vanholme et al., 2008). To understand the role of laccases in lignification, Berthet et al. (2011) downregulated *LAC4* and *LAC17* in *Arabidopsis* and observed that the transgenic plants had decreased lignin, leading to enhanced saccharification efficiency. They also observed that the deposition of G-lignin could be influenced by a *LAC17* mutation. Interestingly, double mutants of *LAC4* and *LAC17* had reduced lignin levels and dwarfism was not observed under normal growth conditions. The double mutants also had increased saccharification, even though saccharification efficiency of *lac17* plants was similar to that of controls. Sugarcane laccase *SofLAC* was able to complement the *Arabidopsis lac17* mutation and restore lignin levels to those comparable to wild type plants (Cesarino et al., 2013).

2.5.2 Manipulation of transcription factors of lignin biosynthetic genes

Expression of genes in the monolignol biosynthetic pathway is regulated by R2R3-type transcription factors (TFs) which have the *MYB* DNA binding domain (Rogers and Campbell, 2004; Stracke et al., 2001). Manipulation of these transcription factors can have a dramatic effect on lignin content and composition as these TFs can regulate several genes at once. *MYB* TFs bind to AC elements in promoters, which are rich in adenosine and cytosine residues. Most monolignol biosynthetic gene promoters contain these AC elements, which allows for xylem

specific expression (Rogers and Campbell, 2004). One exception is *F5H,* which does not contain AC elements in its promoter; instead it is directly regulated by *NST3* (Zhao et al., 2010). Secondary cell wall biosynthesis is regulated by master switches early in the monolignol biosynthetic pathway (Figure 3). *SND1*, *SND2*, *NST3* are NAC domain proteins which act to initiate monolignol biosynthesis (Hussey et al., 2011; Zhong et al., 2010b; Zhong et al., 2006; Zhong et al., 2007). *Eucalyptus EgMYB2, Pinus PtMYB1, PtMYB, PtMYB8, Poplar PtrMYB3, PtrMYB20* act as transcriptional activators while *Eucalyptus EgMYB1,* maize *ZmMYB31 ZmMYB42,* Poplar *PttMYB21a* and switchgrass *PvMYB4* act as transcriptional repressors (Zhao and Dixon, 2011). Besides these biofuel crops, transcriptional activators and repressor have been identified in other species. For instance in *Antirrhinum AmMYB308* and *AmMYB330* repress the activity of *4CL*, *CAD* and *C4H* genes and Arabidopsis *AtMYB46*, *AtMYB83*, *AtMYB58*/*63*, *AtMYB85*, *AtMYB61* act as activators (Zhao and Dixon, 2011). Overexpression of repressors can increase saccharification efficiency, but can have negative consequences similar to that of monolignol gene downregulation. Greenhouse studies of switchgrass with overexpressed PvMYB4 resulted in transgenic plants with greatly reduced total lignin and plant stature with no change in the S:G ratio. There was up to a 250 % increase in saccharification efficiency and biofuel production compared with the non-transgenic parent (Shen et al., 2012; Shen et al., 2013).

Other known transcription factors involved in monolignol biosynthesis are basic helix-loop-helix (bHLH) and LIM TFs. Basic helix-loop-helix TFs such as *SbbHLH1* downregulates monolignol biosynthetic genes. Surprisingly, when the *MYB* TFs *MYB46*, *MYB83*, *MYB58* and *MYB63* were activated, lignin content was reduced. It is thought that the presence of *SbbHLH1* might repress monolignol biosynthetic genes, which, in turn, activates expression of *MYB* TF. *SbbHLH1* could

be combined with a *MYB* activator to repress lignin synthesis (Yan et al., 2013). LIM family of TF members have been known to regulate some genes in the monolignol pathway (Kawaoka et al., 2000). These genes have a consensus cysteine-rich protein sequence known as the LIM domain which binds to AC elements of monolignol biosynthetic gene promoters. Studies in tobacco by Kawaoka et al. (2000) have shown *NtLIM1* acts as weak activator of lignin genes to maintain a basal level of genes in the monolignol biosynthetic pathway. Suppression of *NtLIM1* did not produce any abnormal growth phenotype and even with a 27% decrease in lignin content, vessels did not collapse. One of the most desirable phenotypes for lignocellulosic biomass has been demonstrated in transgenic rice by the overexpression of *Arabidopsis* TF *SHINE*, which increased the cellulose content, decreased the lignin content and changed the lignin composition without decreasing the biomass of the transformed plants (Ambavaram et al., 2011). Though many studies have been conducted with constitutive promoters, TFs can be powerful tools to perturb the lignin biosynthetic pathway, but the TFs must be carefully regulated themselves, say, by the use of tissue-specific promoters, or they could lead to undesirable pleiotropic effects in transgenic plants.

2.5.3 Manipulation of regulatory factors that affect lignin biosynthesis

MicroRNAs (miRNAs) regulate various developmental programs of plants. MiRNAs regulate the expression of transcription factors that have been found to increase biomass yield, modify lignin content and composition, reduce recalcitrance and other such characteristics important for biofuel production (Fu et al., 2012; Rubinelli et al., 2013). A miRNA, such as the one encoded by the maize *Corngrass1* (*Cg1*) gene that belongs to the *miR156* class, targets the *SQUAMOSA PROMOTER BINDING LIKE* (*SPL*) family of transcription factors and supports the development of juvenile morphology with reduced lignification of cell walls (Rhoades et al.,

2002). The *Cg1* gene has been constitutively expressed in poplar and switchgrass (Chuck et al., 2011; Rubinelli et al., 2013) where it modified the content and composition of lignin and also had a severe effect on the plant structure when the miRNA was highly overexpressed. Similar results were also obtained in transgenic switchgrass when the rice *OsmiR156b* precursor was overexpressed (Fu et al., 2012). Three classes of transgenic plant phenotypes were obtained from the switchgrass study. High transgene overexpression resulted in severely dwarf plants. Medium expressers had delayed flowering, less recalcitrance and more biomass. Low expressers of *OsmiR156b* had increased tiller number, but were otherwise similar to the non-transgenic parent switchgrass. Recently, *miR397* targeting laccases has been identified in poplar. Twenty nine *PtrLACs* were identified to be the *PtrmiR397a*. Overexpression of *PtrmiR397a* reduced lignin content without affecting the expression level of monolignol biosynthetic genes, suggesting that PtrmiR397 is involved in lignin polymerization (Lu et al., 2013).

To date, bioenergy feedstock research targeting lignin has employed a single gene strategy whereby genes were expressed or silenced using constitutive promoters. This research has been valuable to down select gene targets that might ultimately provide desired phenotypes, but no single experiment has yielded perfect phenotypes. It is clear that spatial and temporal regulation of multiple transgenes at "sweet spot" expression levels will be needed to decrease or modify lignin in aboveground tissues for increased biofuel without negative pleiotropic effects. That said, constitutive promoter driven transgenes have been heuristic for our increased understanding of the consequences of modifying lignin with regards to recalcitrance and plant performance.

2.6 Consequences of lignin modification

As we see above, there are numerous genes targets that can change the S:G ratio and total lignin; when either of these is decreased, the saccharification efficiency increases, thereby yielding more biofuel per mass of feedstock. However, there can be off-effects associated with perturbing lignin in plants (Vermerris et al., 2010). The high overexpression of *PvMYB4* in transgenic switchgrass also led to shorter plant stature (Shen et al., 2012). Plants with reduced activity of *C4H*, *C3H*, *HCT*, or *CCR* are often dwarfed (Bonawitz and Chapple, 2010; Hoffmann et al., 2004; Reddy et al., 2005). Cell wall collapse in vasculature, which can lead to water conductance problems, might be one cause of dwarfism in lignin-altered plants. When linked with a significant enhancement in saccharification efficiency, moderate growth penalties might be tolerable, but extreme dwarfism or increased pest susceptibility would not be useful in agricultural fields.

The role of lignin in biotic stresses, such as responses to pathogens, is just beginning to be understood. *CAD* gene expression is induced in response to pathogen infection in *Arabidopsis* (Tronchet et al., 2010) indicating a possible role of lignin in disease resistance. On the other hand, the downregulation of the gene encoding *HCT* was associated with an increased tolerance to fungal infection and drought in alfalfa (Gallego-Giraldo et al., 2011).

Clearly, we need to gain a better understanding of lignin's role in whole plant physiology in field settings. The generation of transgenic plants with a wide variety of lignin-associated transgenes will allow for transgene stacking using plant hybridization to better understand synergistic transgene effects, as will crossing transgenic plants into a variety of genetic backgrounds. Tissue-specific promoters will likely be required to regulate lignin production in tissues crucial

for biofuel production, such as aboveground tissue in perennial grasses, while not perturbing lignin biosynthesis in roots. We also envisage the utilization of high biomass germplasm which can be hybridized with low lignin germplasm to decrease yield penalties associated with altering certain lignin biosynthesis genes. Certainly, biotechnology holds much promise to increase biofuel yield per hectare, but the fine-tuning of lignin biosynthesis will be needed to realize this promise.

2.7 Acknowledgements

This project was made possible through funding from USDA NIFA Biotechnology Risk Assessment Grants (BRAG) Program grant # 2010-39211-21699. C. Neal Stewart Jr. and Charleson Poovaiah also received support from the BioEnergy Science Center, a Bioenergy Research Center supported by the Office of Biological and Environmental Research in the US Department of Energy Office of Science.

2.8 References

Agbor VB, Cicek N, Sparling R, Berlin A, Levin DB. 2011. Biomass pretreatment: fundamentals toward application. *Biotechnology Advances* 29: 675-685.

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

Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Cezard L, Le Bris P, Borrega N, Herve J, Blondet E, Balzergue S, Lapierre C, Jouanin L. 2011. Disruption of *LACCASE4* and *17* results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell* 23: 1124-1137.

Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M. 2007. Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. *Plant Cell* 19: 148-162.

Boerjan W, Ralph J, Baucher M. 2003. Lignin biosynthesis. *Annual Review of Plant Biology* 54: 519-546.

Bonawitz ND, Chapple C. 2010. The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annual Review of Genetics* 44: 337-363.

Bonawitz ND, Chapple C. 2013. Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Current Opinion in Biotechnology* 24: 336-343.

Boudet AM, Kajita S, Grima-Pettenati J, Goffner D. 2003. Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends in Plant Science* 8: 576-581.

Bouvier d'Yvoire M, 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. 2013. Disrupting the cinnamyl alcohol dehydrogenase 1 gene (*BdCAD1*) leads to altered lignification and improved saccharification in *Brachypodium distachyon*. *Plant Journal* 73: 496-508.

Cesarino I, Araújo P, Mayer JLS, Vicentini R, Berthet S, Demedts B, Vanholme B, Boerjan W, Mazzafera P. 2013. Expression of *SofLAC*, a new laccase in sugarcane, restores lignin content but not S: G ratio of Arabidopsis lac17 mutant. *Journal of Experimental Botany* 64: 1769-1781.

Chabannes M, Barakate A, Lapierre C, Marita JM, Ralph J, Pean M, Danoun S, Halpin C, Grima-Pettenati J, Boudet AM. 2001. Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (*CCR*) and cinnamyl alcohol dehydrogenase (*CAD*) in tobacco plants. *Plant Journal* 28: 257-270.

Chen F, Dixon RA. 2008. Genetic manipulation of lignin biosynthesis to improve biomass characteristics for agro-industrial processes. *In Vitro Cellular & Developmental Biology-Animal* 44: S28-S29.

Chuck GS, Tobias C, Sun L, Kraemer F, Li CL, Dibble D, Arora R, Bragg JN, Vogel JP, Singh S, Simmons BA, Pauly M, Hake S. 2011. Overexpression of the maize *Corngrass1* microRNA prevents flowering, improves digestibility, and increases starch content of switchgrass. *Proceedings of the National Academy of Sciences, USA* 108: 17550-17555.

Coleman HD, Park J-Y, Nair R, Chapple C, Mansfield SD. 2008. RNAi-mediated suppression of p-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *Proceedings of the National Academy of Sciences, USA* 105: 4501-4506.

Dien BS, Sarath G, Pedersen JF, Sattler SE, Chen H, Funnell-Harris DL, Nichols NN, Cotta MA. 2009. Improved sugar conversion and ethanol yield for forage sorghum (*Sorghum bicolor* L. Moench) lines with reduced lignin contents. *BioEnergy Research* 2: 153-164.

Espineira JM, Novo Uzal E, Gomez Ros LV, Carrion JS, Merino F, Ros Barcelo A, Pomar F. 2011. Distribution of lignin monomers and the evolution of lignification among lower plants. *Plant Biology* (*Stuttgart*) 13: 59-68.

Fornale S, Capellades M, Encina A, Wang K, Irar S, Lapierre C, Ruel K, Joseleau JP, Berenguer J, Puigdomenech P, Rigau J, Caparros-Ruiz D. 2012. Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. *Molecular Plant* 5: 817-830.

Fu C, Mielenz JR, Xiao X, Ge Y, Hamilton CY, Rodriguez M, Jr., Chen F, Foston M, Ragauskas A, Bouton J, Dixon RA, Wang ZY. 2011. Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proceedings of the National Academy of Sciences, USA* 108: 3803-3808.

Fu C, Sunkar R, Zhou C, Shen H, Zhang J-Y, Matts J, Wolf J, Mann DGJ, Stewart CN, Tang Y, Wang Z-Y. 2012. Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotechnology Journal* 10: 443-452.

Fu C, Xiao X, Xi Y, Ge Y, Chen F, Bouton J, Dixon R, Wang Z-Y. 2011. Downregulation of cinnamyl alcohol dehydrogenase (*CAD*) Leads to improved saccharification efficiency in switchgrass. *BioEnergy Research* 4: 153-164.

Gallego-Giraldo L, Jikumaru Y, Kamiya Y, Tang Y, Dixon RA. 2011. Selective lignin downregulation leads to constitutive defense response expression in alfalfa (*Medicago sativa* L.). *New Phytologist* 190: 627-639.

Gui J, Shen J, Li L. 2011. Functional characterization of evolutionarily divergent 4-coumarate:coenzyme a ligases in rice. *Plant Physiology* 157: 574-586.

Guo D, Chen F, Inoue K, Blount JW, Dixon RA. 2001. Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl coA 3-O-methyltransferase in transgenic alfalfa: Impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell* 13: 73-88.

Herrero J, Esteban-Carrasco A, Zapata JM. 2013. Looking for *Arabidopsis thaliana* peroxidases involved in lignin biosynthesis. *Plant Physiology and Biochemistry* 67: 77-86.

Hisano H, Nandakumar R, Wang ZY. 2009. Genetic modification of lignin biosynthesis for improved biofuel production. *In Vitro Cellular & Developmental Biology-Plant* 45: 306-313.

Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, and, Legrand M. 2004. Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyl-transferase affects phenylpropanoid biosynthesis. *Plant Cell* 16: 1446-1465.

Humphreys JM, Chapple C. 2002. Rewriting the lignin roadmap. *Current Opinion in Plant Biology* 5: 224-229.

Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield SD. 2003. Significant increases in pulping efficiency in C4H-F5H-transformed poplars: improved chemical savings and reduced environmental toxins. *Journal of Agricultural and Food Chemistry* 51: 6178-6183.

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 Biology* 11: 173-173.

Jackson LA, Shadle GL, Zhou R, Nakashima J, Chen F, Dixon RA. 2008. Improving saccharification efficiency of alfalfa stems through modification of the terminal stages of monolignol biosynthesis. *BioEnergy Research* 1: 180-192.

Jouanin L, Goujon T, de Nadai V, Martin MT, Mila I, Vallet C, Pollet B, Yoshinaga A, Chabbert B, Petit-Conil M, Lapierre C. 2000. Lignification in transgenic poplars with extremely reduced caffeic acid O-methyltransferase activity. *Plant Physiology* 123: 1363-1374.

Kajita S, Hishiyama S, Tomimura Y, Katayama Y, Omori S. 1997. Structural characterization of modified lignin in transgenic tobacco plants in which the activity of 4-coumarate:coenzyme A ligase is depressed. *Plant Physiology* 114: 871-879.

Kawaoka A, Kaothien P, Yoshida K, Endo S, Yamada K, Ebinuma H. 2000. Functional analysis of tobacco LIM protein Ntlim1 involved in lignin biosynthesis. *Plant Journal* 22: 289-301.

Keating JD, Panganiban C, Mansfield SD. 2006. Tolerance and adaptation of ethanologenic yeasts to lignocellulosic inhibitory compounds. *Biotechnology and Bioengineering* 93: 1196-1206.

Korth KL, Blount JW, Chen F, Rasmussen S, Lamb C, Dixon RA. 2001. Changes in phenylpropanoid metabolites associated with homology-dependent silencing of phenylalanine ammonia-lyase and its somatic reversion in tobacco. *Physiologia Plantarum* 111: 137-143.

Kumar P, Barrett DM, Delwiche MJ, Stroeve P. 2009. Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Industrial & Engineering Chemistry Research* 48: 3713-3729.

Lapierre C, Pilate G, Pollet B, Mila I, Leple J-C, Jouanin L, Kim H, Ralph J. 2004. Signatures of cinnamyl alcohol dehydrogenase deficiency in poplar lignins. *Phytochemistry* 65: 313-321.

Lee D, Meyer K, Chapple C, Douglas CJ. 1997. Antisense suppression of 4-coumarate:coenzyme A ligase activity in *Arabidopsis* leads to altered lignin subunit composition. *Plant Cell* 9: 1985-1998.

Lee Y, Rubio Maria C, Alassimone J, Geldner N. 2013. A mechanism for localized lignin deposition in the endodermis. *Cell* 153: 402-412.

Li X, Weng J-K, Chapple C. 2008. Improvement of biomass through lignin modification. *Plant Journal* 54: 569-581.

Li X, Ximenes E, Kim Y, Slininger M, Meilan R, Ladisch M, Chapple C. 2010. Lignin monomer composition affects *Arabidopsis* cell-wall degradability after liquid hot water pretreatment. *Biotechnology for Biofuels* 3: 27.

Lu S, Li Q, Wei H, Chang MJ, Tunlaya-Anukit S, Kim H, Liu J, Song J, Sun YH, Yuan L, Yeh TF, Peszlen I, Ralph J, Sederoff RR, Chiang VL. 2013. Ptr-miR397a is a negative regulator of

laccase genes affecting lignin content in *Populus trichocarpa*. *Proceedings of the National Academy of Sciences, USA* 110: 10848-10853.

Mansfield SD, Kang K-Y, Chapple C. 2012. Designed for deconstruction – poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. *New Phytologist* 194: 91-101.

Martone PT, Estevez JM, Lu FC, Ruel K, Denny MW, Somerville C, Ralph J. 2009. Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. *Current Biology* 19: 169-175.

Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology* 96: 673-686.

Patzlaff A, McInnis S, Courtenay A, Surman C, Newman LJ, Smith C, Bevan MW, Mansfield S, Whetten RW, Sederoff RR, Campbell MM. 2003. Characterisation of a pine *MYB* that regulates lignification. *Plant Journal* 36: 743-754.

Piquemal J, Chamayou S, Nadaud I, Beckert M, Barriere Y, Mila I, Lapierre C, Rigau J, Puigdomenech P, Jauneau A, Digonnet C, Boudet A-M, Goffner D, Pichon M. 2002. Down-regulation of caffeic acid o-methyltransferase in maize revisited using a transgenic approach. *Plant Physiology* 130: 1675-1685.

Pomar F, Barceló AR, Novo-Uzal E, Ros LVG, Espiñeira MJ 2012. Evolutionary history of lignins. In: Jouanin L, Lapierre C eds. *Advances in Botanical Research*: Academic Press, 309–350.

Ralph J, Akiyama T, Coleman HD, Mansfield SD. 2012. Effects on lignin structure of coumarate 3-hydroxylase downregulation in Poplar. *BioEnergy Research* 5: 1009-1019.

Ralph J, Akiyama T, Kim H, Lu F, Schatz PF, Marita JM, Ralph SA, Reddy MSS, Chen F, Dixon RA. 2006. Effects of coumarate 3-hydroxylase down-regulation on lignin structure. *Journal of Biological Chemistry* 281: 8843-8853.

Ralph J, Lapierre C, Marita JM, Kim H, Lu FC, Hatfield RD, Ralph S, Chapple C, Franke R, Hemm MR, Van Doorsselaere J, Sederoff RR, O'Malley DM, Scott JT, MacKay JJ, Yahiaoui N, Boudet AM, Pean M, Pilate G, Jouanin L, Boerjan W. 2001. Elucidation of new structures in lignins of CAD- and COMT-deficient plants by NMR. *Phytochemistry* 57: 993-1003.

Reddy MSS, Chen F, Shadle G, Jackson L, Aljoe H, Dixon RA. 2005. Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.). *Proceedings of the National Academy of Sciences, USA* 102: 16573-16578.

Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. 2002. Prediction of plant microRNA targets. *Cell* 110: 513-520.

Rogers LA, Campbell MM. 2004. The genetic control of lignin deposition during plant growth and development. *New Phytologist* 164: 17-30.

Rubin EM. 2008. Genomics of cellulosic biofuels. *Nature* 454: 841-845.

Rubinelli PM, Chuck G, Li X, Meilan R. 2013. Constitutive expression of the *Corngrass1* microRNA in poplar affects plant architecture and stem lignin content and composition. *Biomass and Bioenergy* 54: 312-321.

Saathoff AJ, Sarath G, Chow EK, Dien BS, Tobias CM. 2011. Downregulation of cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. *PLOS One* 6: e16416.

Saballos A, Sattler SE, Sanchez E, Foster TP, Xin Z, Kang C, Pedersen JF, Vermerris W. 2012. Brown midrib2 (Bmr2) encodes the major 4-coumarate:coenzyme A ligase involved in lignin biosynthesis in sorghum (*Sorghum bicolor* (L.) Moench). *Plant Journal* 70: 818-830.

Saballos A, Vermerris W, Rivera L, Ejeta G. 2008. Allelic association, chemical characterization and saccharification properties of brown midrib mutants of sorghum (*Sorghum bicolor* (L.) Moench). *BioEnergy Research* 1: 193-204.

Schilmiller A, Stout J, Weng J, Humphreys J, Ruegger M, Chapple C. 2009. Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in *Arabidopsis*. *Plant Journal* 60: 771-782.

Schmer MR, Vogel KP, Mitchell RB, Perrin RK. 2008. Net energy of cellulosic ethanol from switchgrass. *Proceedings of the National Academy of Sciences, USA* 105: 464-469.

Sewalt VJH, Ni W, Blount JW, Jung HG, Masoud SA, Howles PA, Lamb C, Dixon RA. 1997. Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of L-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase. *Plant Physiology* 115: 41-50.

Shadle G, Chen F, Srinivasa Reddy MS, Jackson L, Nakashima J, Dixon RA. 2007. Down-regulation of hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry* 68: 1521-1529.

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

Shen H, Poovaiah CR, Ziebell A, Tschaplinski TJ, Pattathil S, Gjersing E, Engle NL, Katahira R, Pu Y, Sykes R, Chen F, Ragauskas AJ, Mielenz JR, Hahn MG, Davis M, Stewart CN, Jr., Dixon RA. 2013. Enhanced characteristics of genetically modified switchgrass (*Panicum virgatum* L.) for high biofuel production. *Biotechnol Biofuels* 6: 71.

Sibout R, Eudes A, Mouille G, Pollet B, Lapierre C, Jouanin L, Seguin A. 2005. Cinnamyl alcohol dehydrogenase-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of *Arabidopsis*. *Plant Cell* 17: 2059-2076.

Stewart JJ, Akiyama T, Chapple C, Ralph J, Mansfield SD. 2009. The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiology* 150: 621-635.

Sticklen M. 2006. Plant genetic engineering to improve biomass characteristics for biofuels. *Current Opinion in Biotechnology* 17: 315-319.

Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* 4: 447-456.

Tronchet M, Balague C, Kroj T, Jouanin L, Roby D. 2010. Cinnamyl alcohol dehydrogenases-C and D, key enzymes in lignin biosynthesis, play an essential role in disease resistance in *Arabidopsis*. *Molecular Plant Pathology* 11: 83-92.

Uzal EN, Ros LVG, Pomar F, Bernal MA, Paradela A, Albar JP, Barcelo AR. 2009. The presence of sinapyl lignin in *Ginkgo biloba* cell cultures changes our views of the evolution of lignin biosynthesis. *Physiologia Plantarum* 135: 196-213.

Vailhé MAB, Besle JM, Maillot MP, Cornu A, Halpin C, Knight M. 1998. Effect of down-regulation of cinnamyl alcohol dehydrogenase on cell wall composition and on degradability of tobacco stems. *Journal of the Science of Food and Agriculture* 76: 505-514.

Valerio L, Carter D, Rodrigues JC, Tournier V, Gominho J, Marque C, Boudet AM, Maunders M, Pereira H, Teulieres C. 2003. Down regulation of cinnamyl alcohol dehydrogenase, a lignification enzyme, in *Eucalyptus camaldulensis*. *Molecular Breeding* 12: 157-167.

Vanholme B, Cesarino I, Goeminne G, Kim H, Marroni F, Acker R, Vanholme R, Morreel K, Ivens B, Pinosio S. 2013. Breeding with rare defective alleles (BRDA): a natural *Populus nigra* HCT mutant with modified lignin as a case study. *New Phytologist* 198: 765–776.

Vanholme R, Cesarino I, Rataj K, Xiao Y, Sundin L, Goeminne G, Kim H, Cross J, Morreel K, Araujo P, Welsh L, Haustraete J, McClellan C, Vanholme B, Ralph J, Simpson GG, Halpin C, Boerjan W. 2013. Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis. *Science* 341: 1103-1106.

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

Vanholme R, Morreel K, Ralph J, Boerjan W. 2008. Lignin engineering. *Current Opinion in Plant Biology* 11: 278-285.

Vanholme R, Storme V, Vanholme B, Sundin L, Christensen JH, Goeminne G, Halpin C, Rohde A, Morreel K, Boerjan W. 2012. A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. *Plant Cell* 24: 3506-3529.

Vermerris W, Sherman DM, McIntyre LM. 2010. Phenotypic plasticity in cell walls of maize brown midrib mutants is limited by lignin composition. *Journal of Experimental Botany* 61: 2479-2490.

Voelker SL, Lachenbruch B, Meinzer FC, Jourdes M, Ki C, Patten AM, Davin LB, Lewis NG, Tuskan GA, Gunter L, Decker SR, Selig MJ, Sykes R, Himmel ME, Kitin P, Shevchenko O, Strauss SH. 2010. Antisense down-regulation of *4CL* expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiology* 154: 874-886.

Wagner A, Donaldson L, Kim H, Phillips L, Flint H, Steward D, Torr K, Koch G, Schmitt U, Ralph J. 2009. Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata*. *Plant Physiology* 149: 370-383.

Wagner A, Ralph J, Akiyama T, Flint H, Phillips L, Torr K, Nanayakkara B, Kiri LT. 2007. Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA : shikimate hydroxycinnamoyltransferase in *Pinus radiata*. *Proceedings of the National Academy of Sciences, USA* 104: 11856-11861.

Weng J-K, Banks JA, Chapple C. 2008. Parallels in lignin biosynthesis: A study in *Selaginella moellendorffii* reveals convergence across 400 million years of evolution. *Communication in Integrative Biology* 1: 20-22.

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

Weng JK, Li X, Bonawitz ND, Chapple C. 2008. Emerging strategies of lignin engineering and degradation for cellulosic biofuel production. *Current Opinion in Biotechnology* 19: 166-172.

Xu B, Escamilla-Trevino LL, Sathitsuksanoh N, Shen Z, Shen H, Zhang YHP, Dixon RA, Zhao B. 2011. Silencing of 4-coumarate:coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. *New Phytologist* 192: 611-625.

Xu ZY, Zhang DD, Hu J, Zhou X, Ye X, Reichel KL, Stewart NR, Syrenne RD, Yang XH, Gao P, Shi WB, Doeppke C, Sykes RW, Burris JN, Bozell JJ, Cheng ZM, Hayes DG, Labbe N, Davis M, Stewart CN, Yuan JS. 2009. Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. *BMC Bioinformatics* 10: S3

Yan L, Xu C, Kang Y, Gu T, Wang D, Zhao S, Xia G. 2013. The heterologous expression in Arabidopsis thaliana of sorghum transcription factor *SbbHLH1* downregulates lignin synthesis. *Journal of Experimental Botany* 64: 3021-3032.

Ye ZH, Kneusel RE, Matern U, Varner JE. 1994. An alternative methylation pathway in lignin biosynthesis in *Zinnia*. *Plant Cell* 6: 1427-1439.

Zhao Q, Dixon RA. 2011. Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends in Plant Science* 16: 227-233.

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. *Proceedings of the National Academy of Sciences, USA* 107: 14496-14501.

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

Zhong R, Lee C, Ye Z-H. 2010b. Functional characterization of poplar wood-associated NAC domain transcription factors. *Plant Physiology* 152: 1044-1055.

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

Zhong RQ, 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 RQ, Lee CH, Zhou JL, McCarthy RL, Ye ZH. 2008. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell* 20: 2763-2782.

Zhong RQ, Richardson EA, Ye ZH. 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.

2.9 Appendix

2.9.1 Figures and tables

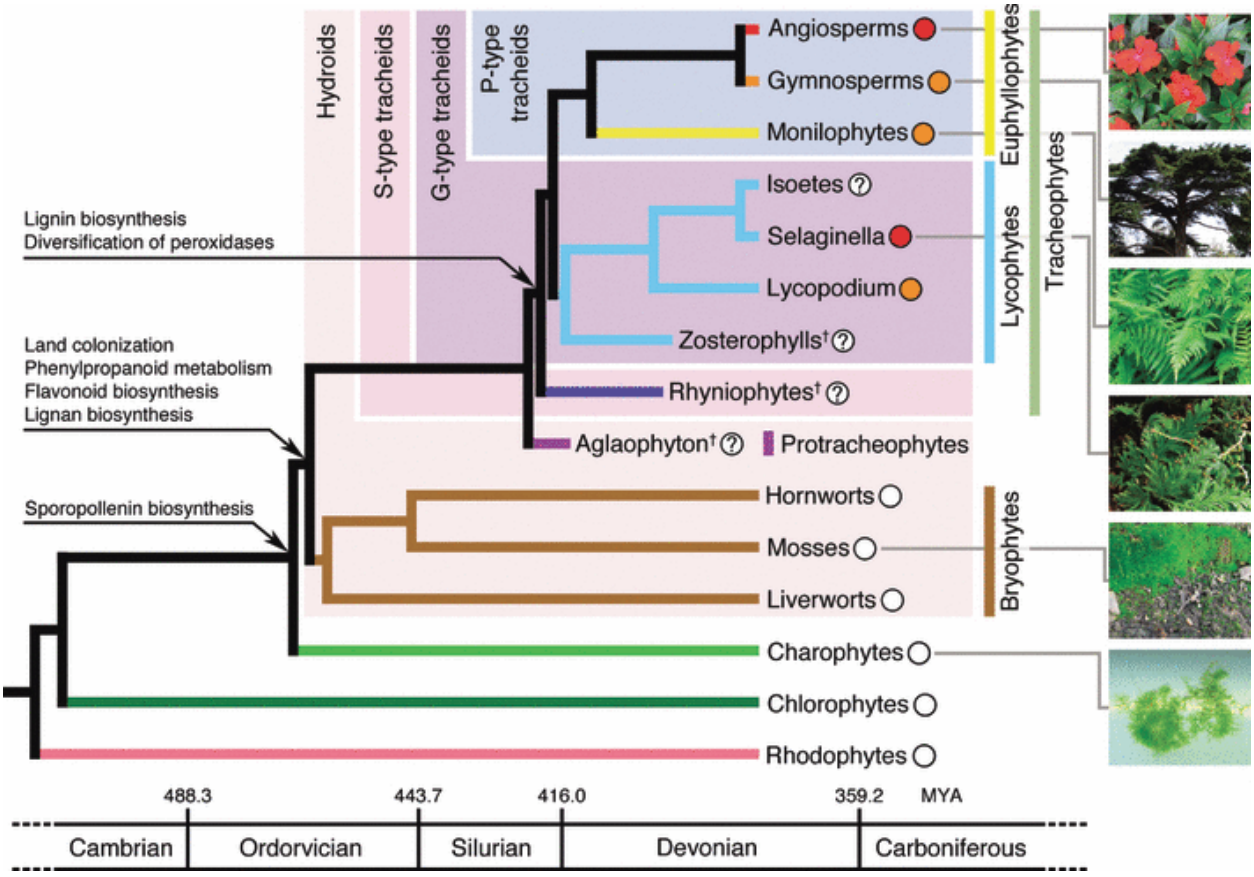


Figure 2-1 The evolution of lignin biosynthesis. Phylogenetic tree depicting the evolution of monolignols across major plant phyla

Hollow circles indicate absence of lignin, orange circles indicate plants with H and G lignin and red circles indicate plants with S, G and H lignin. The presence of red circles more than once in the evolutionary tree indicates convergent evolution of S lignin in *Selaginella* and angiosperms. Image used with permission from John Wiley & Sons, Inc.

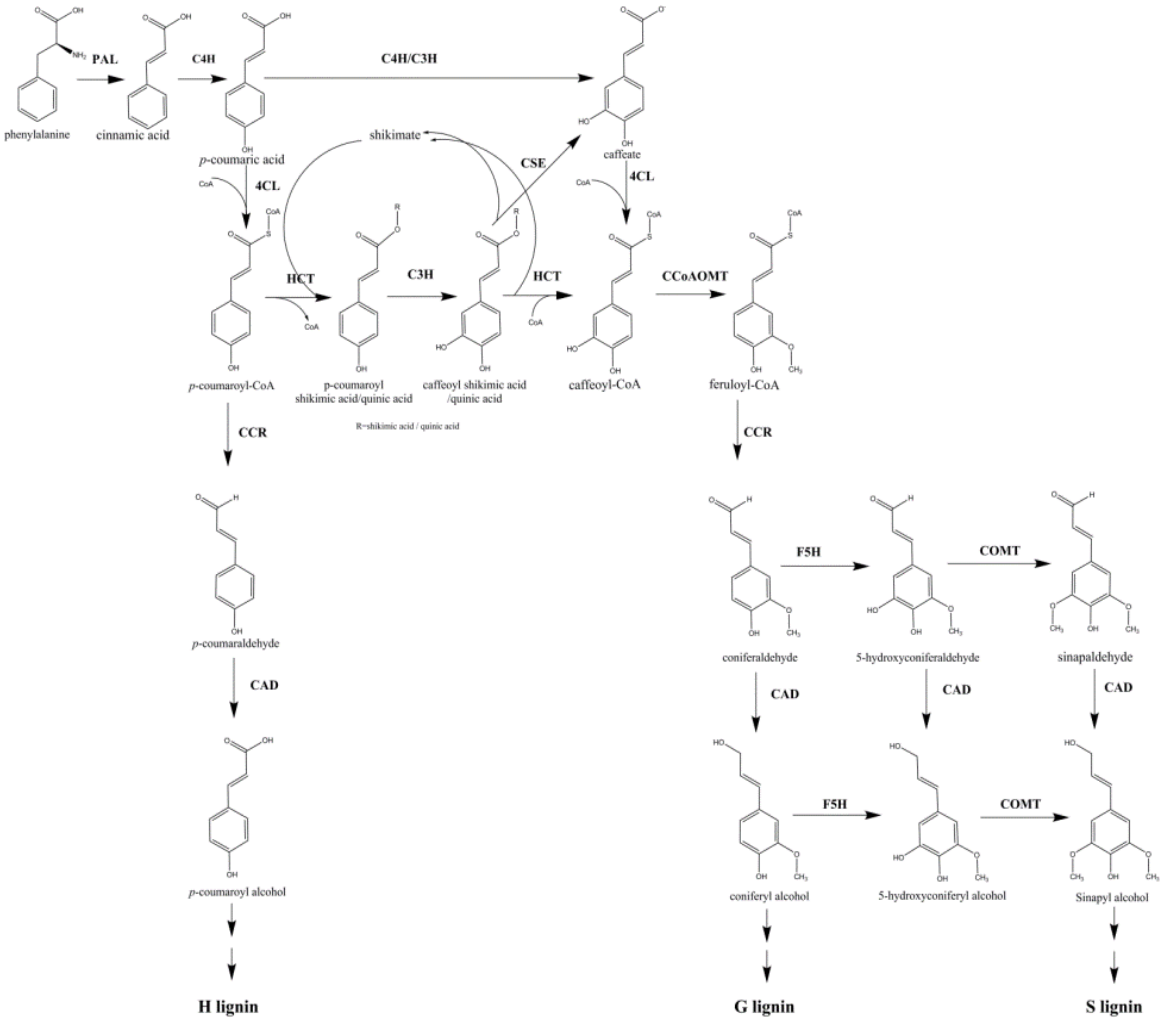


Figure 2-2 Lignin biosynthetic pathway

Diagram of the current accepted model of monolignol biosynthesis for *p*-coumaryl, coniferyl, and sinapyl alcohol in angiosperms including the newly identified enzyme caffeoyl shikimate esterase (Bonawitz and Chapple, 2010; Vanholme et al., 2013b). The abbreviations: *PAL*, phenylalanine ammonia-lyase; *C4H*, cinnamate 4-hydroxylase; *CSE*, caffeoyl shikimate esterase; *4CL*, 4-coumarate:CoA ligase; *C3H*, *p*-coumarate 3-hydroxylase; *HCT*, *p*-hydroxycinnamoyl-CoA:quinate/shikimate *p*-hydroxycinnamoyltransferase; *CCoAOMT*, caffeoyl-CoA *O*-methyltransferase; *CCR*, cinnamoyl-CoA reductase; *F5H*, ferulate 5-hydroxylase; *COMT*, caffeic acid *O*-methyltransferase; *CAD*, cinnamyl alcohol dehydrogenase.

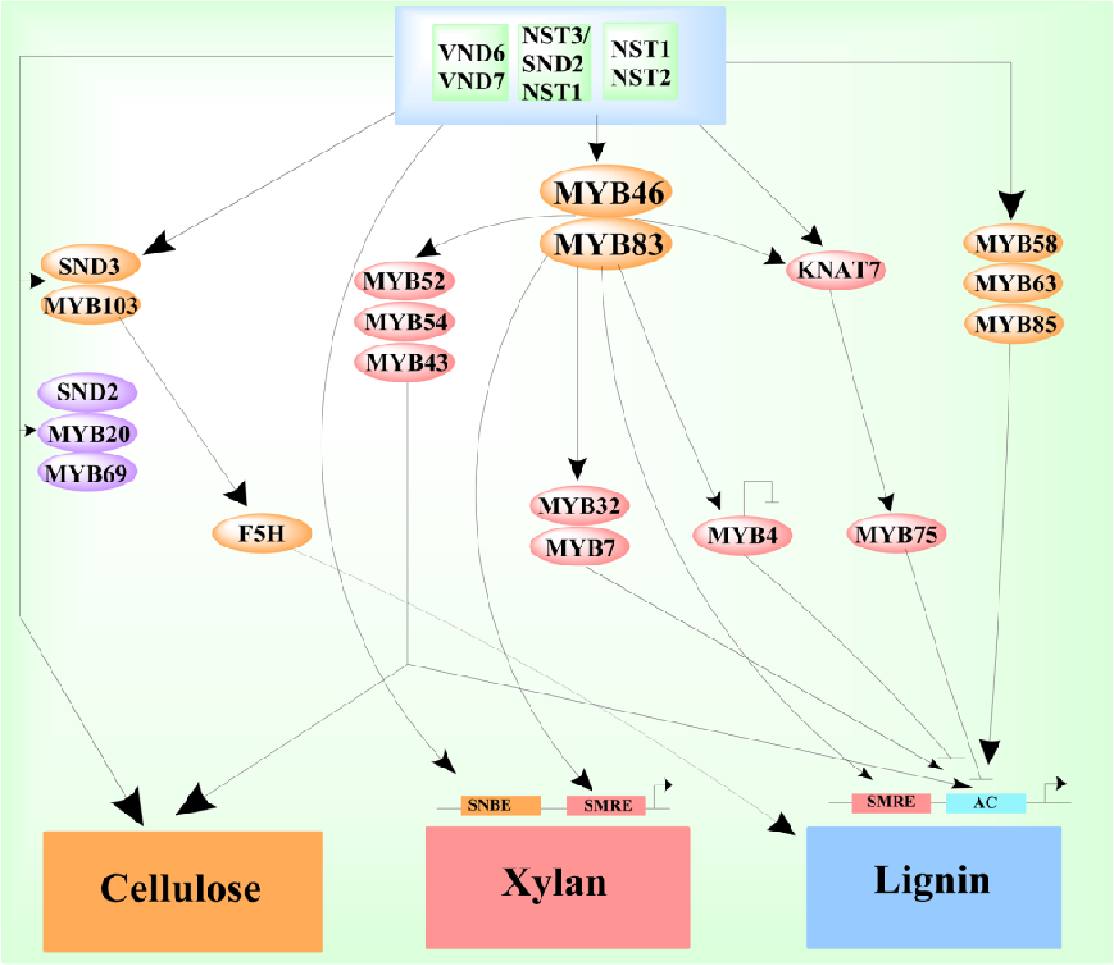


Figure 2-3 Lignin and secondary cell wall biosynthesis transcriptional regulatory network. *SND*, *VND*, *NST* and its homologs act as master switches of secondary cell wall biosynthesis in different tissues. *MYB* transcription factors bind to AC elements in lignin biosynthesis gene promoters to regulate lignin biosynthesis and also bind to SMRE elements. The secondary cell wall master switches also bind to SNBE elements to regulate xylan synthesis. Cellulose synthesis is regulated by the *SND* and *MYB* homologs. Abbreviations: *SND*, secondary wall associated NAC domain protein; *NST*, *NAC* secondary wall thickening promoting factor; *VND*, vascular-related *NAC*-domain; *F5H*, ferulate-5-hydroxylase; *SNBE*, secondary wall *NAC* binding element; *SMRE*, secondary wall *MYB* responsive element. This figure is an expanded version of a published model (Zhong and Ye, 2012).

Table 2-1 Genetically engineered plants in which the expression of lignin biosynthetic genes have been altered

Gene	Species	Method of manipulation[a]	Lignin content	S:G ratio	Method of lignin determination	Refs
PAL	Tobacco	DR	↓ 48%	↑	Klason	(Sewalt et al., 1997)
	Tobacco	S	↓ 70%	↑	Klason	(Korth et al., 2001)
	Tobacco	OE	↑ 45%	↓	Klason	(Korth et al., 2001)
C4H	Tobacco	AS, SS	↓ 20%	↓	Klason	(Sewalt et al., 1997)
	Alfalfa	AS	↓ 40%	↓	AcBr	(Reddy et al., 2005)
4CL	Tobacco	DR	↓ 12%	↓	Klason	(Kajita et al., 1997)
	Arabidopsis	AS	↓ 16%	↑	Thioglycolic acid	(Lee et al., 1997)
	Pinus	S	↓ 8-36%	H/G↑	AcBr	(Wagner et al., 2009)
	Populus	AS	↓ 10%	↑	AcBr	(Voelker et al., 2010)
	Rice	AS	↓ 21%	↓	Klason	(Gui et al., 2011)
	Switchgrass	S	↓ 17-32%	↑	72% H2SO4 + HPLC	(Xu et al., 2011)
HCT	Tobacco	VIGS	↓ 15%	↓	Klason	(Hoffmann et al., 2004)
	Arabidopsis	S	↓	NAb	Histochemical staining	(Hoffmann et al., 2004)
	Alfalfa	AS	↓ 35-50%	H/G↑	AcBr	(Shadle et al., 2007)

Table 2-1 continued

Gene	Species	Method of manipulation[a]	Lignin content	S:G ratio	Method of lignin determination	Refs
HCT	*Pinus radiata*	S	↓	H/G↑	AcBr	(Wagner et al., 2007)
C3H	Alfalfa	AS	↓ 30%	H/G↑	AcBr	(Reddy et al., 2005)
	Poplar	S	↓	H/G↑	72% H2SO4 + HPLC	(Coleman et al., 2008; Ralph et al., 2012)
CCoAMT	Alfalfa	AS	↓ 14%	↑	Klason	(Guo et al., 2001)
CCR	Alfalfa	AS	↓ 4-30%	↑	AcBr	(Jackson et al., 2008)
F5H	Poplar	OE	NCc	↑	72% H2SO4 + HPLC	(Huntley et al., 2003)
	Alfalfa	AS	↑ 13%	↓	AcBr	(Reddy et al., 2005)
COMT	Tobacco	DR	↓	NC	Klason	(Sewalt et al., 1997)
	Poplar	OE	↑ 5%	↓	Klason	(Jouanin et al., 2000)
	Alfalfa	S	↓ 25%	↓	Klason	(Guo et al., 2001)
	Maize	AS	↓ 36%	↓	Klason	(Piquemal et al., 2002)
	Switchgrass	DR	↓ 6-14%	↓	AcBr	(Fu et al., 2011a)
CAD	Tobacco	AS	↓ 18%	↓	AcBr	(Vailhé et al., 1998)

Table 2-1 continued

Gene	Species	Method of manipulation[a]	Lignin content	S:G ratio	Method of lignin determination	Refs
CAD	Alfalfa	AS	↓ 4-9%	↓	AcBr	(Jackson et al., 2008)
	Poplar	AS	↓ 10%	↓	Klason	(Lapierre et al., 2004)
	Eucalyptus	AS	NC	NC	FTIR	(Valerio et al., 2003)
	Arabidopsis	DM	↓ 39%	↓	Klason	(Sibout et al., 2005)
	Switchgrass	S	↓ 14-22%	↓	AcBr ANKOM ADLd	(Fu et al., 2011b; Saathoff et al., 2011)
	Brachypodium	S	↓ 69%	↓	Klason	(Bouvier d'Yvoire et al., 2013)
	Maize	S	NC in stems, ↓ 6.4% in midribs	↓	Klason	(Fornale et al., 2012)

[a]DR: down regulation; OE: over-expression; SS: sense suppression; AS: anti-sense expression; S: silencing using RNAi; VIGS: virus-induced gene silencing; DM: double mutant.

[b]NA: not available.

[c]NC: no change.

[d]ANKOM ADL: ANKOM technology-9/99, method for determining acid detergent lignin (ADL), ANKOM Technology Corp., Fairport, NY, USA.

Chapter 3: Functional characterization of the switchgrass (*Panicum virgatum* L.) R2R3-MYB transcription factor *PvMYB4* for improvement of lignocellulosic feedstocks

(This chapter was published in the New Phytologist 193: 121-36. with the following authors Hui Shen, Xianzhi He, Charleson R. Poovaiah, Wegi A. Wuddineh, Junying Ma, David G. J. Mann, Huanzhong Wang, Lisa Jackson, Yuhong Tang, C. Neal Stewart Jr, Fang Chen, and Richard A. Dixon)

Charleson R Poovaiah's contribution was generation of transgenic switchgrass plants, generation of figure 3-6 and contribution to methods section and revision the manuscript.

3.1 Abstract

The major obstacle for bioenergy production from switchgrass biomass is the low saccharification efficiency caused by cell wall recalcitrance. Saccharification efficiency is negatively correlated with both lignin content and cell wall ester-linked *p*-coumarate:ferulate (*p*-CA:FA) ratio. In this study, we cloned and functionally characterized an R2R3-MYB transcription factor from switchgrass and evaluated its potential for developing lignocellulosic feedstocks. The switchgrass PvMYB4 cDNAs were cloned and expressed in *Escherichia coli*, yeast, tobacco and switchgrass for functional characterization. Analyses included determination of phylogenetic relations, *in situ* hybridization, electrophoretic mobility shift assays to determine binding sites in target promoters, and protoplast transactivation assays to demonstrate domains active on target promoters. PvMYB4 binds to the AC-I, AC-II and AC-III elements of monolignol pathway genes and down-regulates these genes *in vivo*. Ectopic overexpression of PvMYB4 in transgenic switchgrass resulted in reduced lignin content and ester-linked *p*-CA : FA ratio, reduced plant stature, increased tillering and an approx. threefold increase in sugar release efficiency from cell wall residues.We describe an alternative strategy for reducing recalcitrance in switchgrass by manipulating the expression of a key transcription factor instead of a lignin biosynthetic gene. PvMYB4-OX transgenic switchgrass lines can be used as potential germplasm for improvement of lignocellulosic feedstocks and provide a platform for further understanding gene regulatory networks underlying switchgrass cell wall recalcitrance.

3.2 Introduction

Switchgrass is a C4 perennial forage grass native to most areas of the North American grasslands. Since switchgrass has high biomass yield and is well adapted to marginal lands, it has been selected as a dedicated lignocellulosic feedstock for bioenergy production in the United States (McLaughlin & Adams Kszos, 2005; Bouton, 2007; Schmer et al., 2008). The biomass yield of switchgrass varies according to precipitation during the growing season, annual temperature, nitrogen fertilization, and the type of cultivar (Fuentes & Taliaferro, 2002; Wullschleger et al., 2010). The estimated annual biomass yield of switchgrass in the United States is projected to average 12.9 metric tons per hectare for lowland ecotypes and 8.7 metric tons for upland ecotypes (Wullschleger et al., 2010).

The major barrier to efficient conversion of lignocellulose to liquid transportation fuels is the recalcitrant nature of the cell wall. The cellulose microfibrils are embedded in a matrix of hemicelluloses that are covalently linked with lignin and other aromatic phenolic compounds. These linkages, and the masking effect of lignin itself, increase the pretreatment costs and block the access of enzymes to the polysaccharide chains (Akin, 2007; Himmel et al., 2007; Pauly and Keegstra, 2008). Reduction of lignin content through transgenic approaches improves fermentable sugar yield and saccharification efficiency of alfalfa stems (Chen & Dixon, 2007; Jackson et al., 2008), and chemical analysis of field-grown switchgrass stems indicated that the saccharification efficiency is negatively correlated with lignin content and ester-linked p-CA : FA ratio (Shen et al., 2009a). Therefore, genetic modification of lignin and cell wall phenolic ester synthesis in switchgrass would be predicted to serve as an efficient method for reducing recalcitrance and thus improving bioenergy production efficiency (Hisano et al., 2009; Keshwani & Cheng, 2009).

Lignin is formed by the oxidative polymerization of three monolignols, p-coumaryl, coniferyl and sinapyl alcohols, to give the hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units of the mature lignin polymer. Although the monolignol biosynthetic pathway is now well understood (Figure 3-10), and a picture is beginning to emerge of its complex regulatory architecture, to date only a few of the lignin biosynthetic and regulatory genes from switchgrass have been functionally characterized or annotated by comparative bioinformatics analysis. This situation will soon be changed with the ongoing efforts in switchgrass genetics and genomics (Okada et al., 2010).

Common regulatory cis-elements have been identified within the promoters of the genes in the lignin biosynthesis pathway (Hatton et al., 1995; Raes et al., 2003; Legay et al., 2007). A detailed analysis of the bean phenylalanine ammonia-lyase 2 (PAL2) promoter identified three AC elements together with a G-box involved in xylem-specific expression (Hatton et al., 1995). Mutation of the AC-I (ACCTACC), AC-II (ACCAACC) and AC-III (ACCTAAC) elements resulted in a decrease in xylem-associated expression (Hatton et al., 1995). Bioinformatic analysis of the promoters of all lignin biosynthetic genes in Arabidopsis thaliana indicated that such AC elements are present in the majority of the genes except for the promoters of ferulate 5-hydroxylase (F5H) and caffeic acid 3-O-methyltransferase (COMT) (Raes et al., 2003). However, examination of 2 kb of a maize COMT gene promoter region identified a putative ACIII box (aACCTAAC) 200 bp upstream of the transcription start site (Fornaléet al., 2006).

The lignin biosynthesis pathway is coregulated with the secondary cell wall biosynthesis program through a master switch system which includes a group of NAC and R2R3-MYB transcription factors (TFs) (Mitsuda et al., 2007; Zhong et al., 2010). The Arabidopsis genome contains at least 114 NAC genes (Olsen et al., 2005; Shen et al., 2009b) and 126 members of the

R2R3-MYB family (Romero et al., 1998; Stracke et al., 2001). The R2R3-MYB family members are classified into at least 22 subfamilies based on their conserved C-terminal motifs (Romero et al., 1998; Stracke et al., 2001). A group of R2R3-MYB transcriptional activators have been shown to bind directly to AC elements in vitro. These activators include PtMYB1 and PtMYB4 from pine (Patzlaff et al., 2003; Bomal et al., 2008), EgMYB2 from Eucalyptus (Goicoechea et al., 2005), VvMYB5a from Vitis vinifera (Deluc et al., 2006), NtMYBJS1 from tobacco (Gális et al., 2006), PtrMYB21a from poplar (Karpinska et al., 2004; Bylesjo et al., 2008) and AtMYB61, AtMYB68 and AtMYB63 from Arabidopsis (Zhou et al., 2009; Wang et al., 2010). It has been suggested that AtMYB68, AtMYB63 and their close homolog PtrMYB28 from poplar are lignin-specific TFs, whereas PtMYB4 and EgMYB2 also regulate the biosynthesis of cellulose and xylan and therefore serve as secondary cell wall master switches (Zhong et al., 2010).

By contrast, members from subfamily 4 of the R2R3-MYB family have been shown to act as transcriptional repressors of monolignol biosynthetic genes. This was first demonstrated for AmMYB308 and AmMYB330 from Antirrhinum majus (Tamagnone et al., 1998). Overexpression of AmMYB308 and AmMYB330 in tobacco causes reduced plant growth and a white lesion phenotype on leaves (Tamagnone et al., 1998). Similar phenotypes were observed on overexpression of AtMYB4 in Arabidopsis (Jin et al., 2000). AtMYB32 (Preston et al., 2004), a close homolog of AtMYB4, Eucalyptus gunnii EgMYB1 (Legay et al., 2007) and maize ZmMYB31 and ZmMYB42 have also been identified as lignin repressors, and ZmMYB31 was proposed to be a good candidate for biotechnological applications (Fornaléet al., 2006, 2010; Sonbol et al., 2009). However, no genetic manipulation of lignin repressors in monocot plants has yet been reported.

In the present study, we have identified and characterized the lignin repressor PvMYB4 from switchgrass. Effects of overexpression of PvMYB4 in transgenic tobacco and switchgrass suggest that the gene is functionally orthologous to AtMYB4 and ZmMYB31. PvMYB4 binds directly to AC-I (preferred), AC-II and AC-III elements, both in vitro and in a yeast transcription system. Lignin biosynthetic genes are significantly down-regulated in PvMYB4-overexpressing plants, associated with reductions in lignin content and ester-linked p-CA : FA ratio. Overexpression of PvMYB4 in transgenic switchgrass increases saccharification efficiency threefold.

3.3 Materials and Methods

3.3.1 Plant materials and growth conditions

Tobacco (Nicotiana tabacum cv Xanthi NN) and switchgrass (*Panicum virgatum* L. cv Alamo) were grown in the glasshouse under standard conditions (temperature range 25–29°C with a 16 h day from 06:00 to 22:00 h facilitated by supplementary lighting (parabolic aluminized reflector, 125–55 μmol m−2) and relative humidity 77–22%, average 51%). Plants were watered two to three times wk–1, with fertilizer (Peters 20-10-20; J.R. Peters Inc., Allentown, PA, USA) added in the last watering. For tissue-specific gene expression pattern analysis, roots, leaves, leaf sheaths, internodes, and flowers were collected at the reproduction (R1) developmental stage (Moore et al., 1991). Samples were immediately frozen in liquid nitrogen and kept at −80°C for storage, or were ground to powder in a freezer mill (SPEX SamplePrep, Metuchen, NJ, USA) under liquid nitrogen for further RNA isolation and analysis of lignin and soluble and wall-bound phenolics.

3.3.2 Analysis of lignin, phenolic compounds and sugar release efficiency

Cell wall residues (CWR) were prepared by sequentially extracting tobacco stems and switchgrass whole tillers with chloroform/methanol (1 : 1), 100% methanol, 50% methanol, and water (three times each). Fifteen milligrams of lyophilized sample was used for lignin analysis. The acetyl bromide method was employed to determine total lignin content, and thioacidolysis followed by GC-MS was used to identify and quantify lignin-derived monomers. Soluble phenolics were extracted from 30.0 mg freeze-dried tissue powder with 1.5 ml 50% methanol plus 1.5% acetic acid for 12 h at room temperature. One hundred milligrams of extractive-free CWR were used for analysis of esterified cell wall-bound phenolics using low-temperature alkaline hydrolysis. Approx. 100 and 125 mg of CWR were used for determination of total sugar release and enzymatic saccharification without acid pretreatment, respectively. Full details of all analytical methods have been described previously (Shen et al., 2009a).

3.3.3 RNA isolation and quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Total RNA was isolated with an RNeasy Mini Kit (Qiagen). RNA quality was analyzed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). Two micrograms (tobacco) and 3 μg (switchgrass) of total RNA were treated with DNase (Applied Biosystems, Ambion, Austin, TX, USA) for 1 h to remove genomic DNA contamination and then used for reverse transcription with SuperScript® III Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. The cDNA samples were diluted 20-fold and 2 μl diluted cDNA samples were used as the qRT-PCR templates. qRT-PCR and data analysis were as described

previously (Karlen et al., 2007; Shen et al., 2009a). Primer pairs used for qRT-PCR are listed in Table 3-1.

3.3.4 Transcriptional repression and domain mapping in yeast

The pYES2 (Invitrogen) and pGBT-9 (Clontech Laboratories, Mountain View, CA, USA) vectors were used to construct the effector plasmids for transcriptional repression activity and motif mapping assays, respectively. The pLacZi-based vectors (3AC-I and 3AC-II) were a gift from Dr Malcolm Campbell (University of Toronto) (Patzlaff et al., 2003). The six GAL4 DNA-binding motif was cloned from Addgene plasmid 14590 (Zhang et al., 2003) and inserted into pLacZi vector in front of the Mini-pCYC1 promoter. The VP16 motif was cloned from Addgene plasmid 14594 (Kurosu & Peterlin, 2004) into pGBT-9 vector to fuse with the GAL4 binding domain and PvMYB4 C-terminal deletions. The reporter plasmids were first integrated into the genome of yeast strain YM4721 purchased from ATCC (American Type Culture Collection, Manassas, VA, USA) as described by Rose et al. (1990) to make the reporter strains. Approx. 1 μg of effector plasmid was transformed into the yeast reporter strains with the EZ-Yeast Transformation Kit (MP Biomedicals, Solon, OH, USA). β-Galactosidase assays were performed as described in the yeast protocols handbook (Clontech).

3.3.5 In situ hybridization

Primers spanning 757 bp of the PvMYB4 open reading frame were designed using PrimerQuest (http://www.idtdna.com/Scitools/Applications/Primerquest/). After the cDNA template was obtained, separate reactions were performed for making the sense and antisense probes. The T7 promoter sequence was added in front of the reverse primer and the PCR reaction was conducted with the forward primers without T7 to make an antisense probe. In the same way, T7 promoter

was added to the front of the forward primer in combination with the reverse primer without T7 to make the sense control probe. The specific primers used for PvMYB4 are listed in Table 3-1. The probe was synthesized by in vitro transcription with a MAXIscript Kit (AM1308–AM1326) with 0.4–0.6 μg of template DNA obtained as described earlier. Digoxigenin-11-uridine-5′-triphosphate (DIG-11-UTP; Roche) was used for the labeling. The quality and quantity of the probes were checked with a Bioanalyzer.

The tissue preparation including fixation, dehydration, and paraffin embedding were as described previously (Jackson, 1991). In brief, the switchgrass stem sections were rehydrated with an ethanol to water series after the paraffin had been removed by two 10 min incubations in Histo-clear (National Diagnostics Inc., Charlotte, NC). After brief equilibration in 0.1 M triethanolamine, the tissue was acetylated with 0.25% acetic anhydride in 0.1 M triethanolamine for 10 min. Sections were then washed twice with 1× PBS buffer for 5 min before and after triethanolamine treatment. Prehybridization, in situ hybridization and imaging methods were as described previously (Zhou et al., 2010).

3.3.6 Dual luciferase assay

Arabidopsis protoplasts were isolated according to a previously published protocol with minor modifications (Sheen, 2001; Asai et al., 2002). In brief, leaves from healthy 30-d-old Arabidopsis were cut into 0.5–1 mm strips with fresh razor blades. The leaf strips were put into an enzyme solution composed of cellulase and macroenzyme, vacuum-infiltrated for 20 min and then digested for 3 h without shaking in the dark. After filtration, the protoplasts were collected and transformed by polyethylene glycol transfection. To make the effector constructs, coding sequences of PvMYB4 were inserted after the 35S promoter of the Gateway overexpression

vector P2GW7 (http://gateway.psb.ugent.be/). Reporter constructs were prepared as reported (Wang et al., 2010). Promoter activities were represented by Firefly Luciferase (LUC/Renilla) LUC activities, and normalized to the value obtained from protoplasts transformed with empty vector.

3.3.7 Cell imaging and histochemical staining

Switchgrass internode samples and tobacco leaf petioles were collected in the glasshouse and immediately frozen in liquid nitrogen. These samples were then cut with a Leica CM 1850 cryostat (Leica Microsystems Inc., Buffalo Grove, IL, USA) at −20°C and prepared for microscopy as described previously (Nakashima et al., 2008). Phloroglucinol-HCl staining and Mäule staining were carried out as previously described (Fu et al., 2011). Ultraviolet absorption microspectrophotometry was performed as previously described (Nakashima et al., 2008). Photographs were taken using a Nikon DXM 1200 color camera attached to a Nikon microphot-FX microscope system with ACT-1 software (Nikon Instruments Inc., Laguna Hills, CA, USA). Tobacco leaf infiltration was performed as described previously (Sparkes et al., 2006). Imaging of GFP fluorescence by confocal microscopy was performed as described previously (Wang et al., 2008).

3.3.8 Tissue culture and transformation

Tobacco (Nicotiana tabacum cv Xanthi NN) was transformed by the leaf-disc method (Horsch et al., 1985). Leaf discs from a tobacco plant grown in a Magenta box were incubated with *Agrobacterium tumefaciens* (AGL1) harboring the MYB construct for 20 min. The leaf discs were then blotted dry on filter paper and plated on co-cultivation medium for 4 days' co-cultivation in the dark. Leaf discs were transferred to regeneration medium after co-cultivation.

Leaf discs were subcultured every 2 wk in regeneration medium until plant regeneration. Transgenic plants were maintained in the glasshouse for analysis.

Agrobacterium-mediated switchgrass transformation was performed based on a previously published protocol (Xi et al., 2009). The pANIC vector was used for switchgrass transformation (Mann et al., 2011). Briefly, the binary vector contained the attR1-CmR-ccdB-attR2 Gateway-compatible cassette cloned downstream under the control of the ZmUbi1 promoter. The PvMYB4 coding sequence was cloned into the pCR8/GW/TOPO backbone, sequence-verified, and recombined into the expression vector using Gateway® LR Clonase® II enzyme mix (Invitrogen). The switchgrass ST2 line was used for stable transformation. The ST2 line was specifically selected for its high tissue culture response and provided by Dr Zeng-Yu Wang, Noble Foundation. Since the ST2 line is vegetatively propagated by tillers, all the lines are in the same genetic background, an important point for a highly heterozygous outcrossing species.

3.3.9 Electrophoretic mobility shift assays

PvMYB4a was cloned into the pDEST17 expression vector by the Gateway cloning method (Invitrogen). After sequencing, the plasmid was transformed into BL21 DE3 Escherichia coli competent cells. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce expression of PvMYB4. The recombinant protein was purified with the MagneHis™ Protein Purification System (Promega) then concentrated with an Amicon Ultra-15 Centrifugal Filter (Millipore) by washing with Tris buffer (20 mM Tris–HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA and 10 mM β-mercaptoethanol (β-ME)) to remove carried-over HEPES buffer and imidazole. Approx. 700–1000 ng of purified PvMYB4 protein was used for electrophoretic mobility shift assay (EMSA) in each reaction. The probes were labeled by annealing biotin-labeled olignucleotides.

Oligonucleotides used for EMSA are listed in Table 3-1. Binding conditions were 12 mM Tris–HCl (pH 7.5), 20 mM NaCl, 50 mM KCl, 2.5 mM MgCl2, 1 mM DL-dithiothreitol (DTT), 0.2 mM EDTA, 2.5% glycerol, 1 mM β-ME, 0.05% NP-40 and 2 fmol biotin-labeled probes. The samples were loaded and run in a 6% DNA retardant gel (Invitrogen) in the cold room after the reactions had been incubated at 4°C for 30 min. The DNA was transferred onto nylon membranes and signal detected with a LightShift® Chemiluminescent EMSA Kit (Thermo Fisher Scientific Inc., Rockford, IL, USA) using standard protocols.

3.3.10 Statistical analysis

Different letters on the bars of figures indicate significant differences of values at the $P \leq 0.05$ levels. The multiple comparisons were done by Duncan grouping with SAS software (SAS Institute Inc., Cary, NC, USA). Asterisks on top bars indicate values that were determined by the Student's t-test (Microsoft Office Excel 2007) to be significantly different from the wild-type or its equivalent control ($P < 0.05$).

3.4 Results

3.4.1 Switchgrass PvMYB4 belongs to the R2-R3 MYB subfamily 4

To clone orthologs of the MYB subfamily 4 genes in switchgrass, we used the AtMYB4 amino acid sequence to blast against the switchgrass expressed sequence tag (EST) database (Tobias et al., 2005). After identifying the most homologous gene sequence, we designed the primer pairs HS010 and HS011 (Table 3-1) to clone the open reading frame (ORF) from switchgrass (cv Alamo) using RT-PCR. An 836 bp full-length cDNA was cloned into the pENTR_D topo vector for sequencing. The sequence alignment of PvMYB4 with other MYB subfamily 4 members

shows that PvMYB4 has a highly conserved R2–R3 domain at the N-terminal region, while the C-terminal domain is more divergent, both in sequence and in length (Figure 3-1a). However, three typical protein motifs of the MYB subgroup 4 were identified at the C-terminal; 'LlsrGIDPxT/SHRxI/L', 'pdLNLD/ELxiG/S' and 'CX1–2CX7–12CX2C' (Stracke et al., 2001; Fornaléet al., 2006). We also identified a 'FLGLX4–7V/LLD/GF/YR/SX1LEMK' motif using the ClustalW sequence alignment tool (http://www.ebi.ac.uk/Tools/msa/clustalw2/) (Figure 3-1a). These motifs were termed as the C1, C2, Zf and C4 motifs, respectively (Figure 3-1b; Legay et al., 2007).

We performed a phylogenetic analysis using the PhyML method and tools available at Phylogeny.fr: (http://www.phylogeny.fr/) to identify the most closely related R2R3-MYB factors within the subfamily 4 group (Dereeper et al., 2008). Phylogenetic analysis also shows that PvMYB4 is more closely related to ZmMYB42, ZmMYB38 and ZmMYB31 than to the well-known dicot phenylpropanoid/lignin biosynthesis repressors such as AmMYB308, AtMYB4, and AtMYB32 (Figure 3-1c).

Sequencing of the PvMYB4 clones indicated the presence of at least five gene variants, namely PvMYB4a, 4b, 4c, 4d and 4e, in the tetraploid out-crossing switchgrass genome (Figure 3-11). A similar phenomenon has been reported for the switchgrass lignin biosynthesis gene PvCCR1 (Escamilla-Trevino et al., 2009). All the constructs used in the present work were made based on the PvMYB4a sequence.

3.4.2 PvMYB4 is expressed in the vascular bundles of switchgrass

To test the tissue-specific expression pattern of PvMYB4 in switchgrass, we isolated total RNA from different tissues at the R1 developmental stage and performed qRT-PCR analysis (Moore et

al., 1991; Sarath et al., 2007). The PvMYB4 gene was expressed in all the tissues tested, with highest expression in leaf and leaf sheath (Figure 3-2a). In-situ hybridization analysis of young stem tissue sections indicated that PvMYB4 is expressed throughout the stem, including the parenchyma cells, vascular tissues and epidermis. The highest expression level appeared to be in vascular bundles (Figure 3-2b). To test whether PvMYB4 is localized to the nucleus, we performed infiltration assays of tobacco epidermal cells with a construct for expression of a PvMYB4: eGFP fusion protein. GFP fluorescence was shown to localize to the nucleus by confocal microscopy (Figure 3-2c).

3.4.3 PvMYB4 is a transcriptional repressor and binds to the AC elements in the genes involved in the lignin biosynthesis pathway

We tested for transcriptional repression activity using a dual Luciferase reporter assay in Arabidopsis protoplasts. When PvMYB4 was co-expressed with the Arabidopsis PAL4 or Arabidopsis CCoAOMT (caffeoyl CoA 3-O-methyltransferase) promoters driving a Luciferase reporter gene, Luciferase expression was repressed by almost 80–90% compared with the values obtained from these lignin pathway promoters in the absence of PvMYB4 expression (Figure 3-3a). Several R2-R3 MYB proteins from subfamily 4 have been reported to bind the AC elements in monolignol pathway gene promoters (Romero et al., 1998; Legay et al., 2007; Zhao et al., 2007). Use of EMSA led to the conclusion that these proteins might bind to 'GKTWGGTR' elements in vitro (Romero et al., 1998). The reverse complement sequence of 'GKTWGGTR' is 'YACCWAMC' with Y as T or C, W as T or A, and M as A or C. The possible elements can therefore be designated as 'T/CACCT/AAA/CC'. AtMYB4 can bind to the AC-I (ACCTACC), AC-II (ACCAACC) and AC-III (ACCTAAC) elements (Zhao et al., 2007). It also binds to the 'ACCGCCC' elements (MYB elements) found in its own promoter

region (Zhao et al., 2007). EgMYB1 binds to a region of the promoter of the EgCAD (cinnamyl alcohol dehydrogenase) gene with the sequence 'TACCAAGACCCACCTACCGCATGTCGA' containing both the 'ACCCACC' and 'ACCTACC' elements (AC-I) (Legay et al., 2007).
To test whether PvMYB4 can bind to similar AC-rich elements, three repeats of the AC-I, AC-II, AC-III and AC-IV (a hypothetical AC-rich element 'ACCAAAC') elements (Figure 3-3b) were cloned into the pLACZi vector immediately upstream of a pCYC1 minimal promoter driving the LacZ reporter gene, for expression in yeast (Figure 3-3c). Since PvMYB4 has transcriptional repression activity, we cloned PvMYB4 as a fusion with the GAL4 activation domain (AD) to generate a chimeric effector (Figure 3-3c). If PvMYB4 binds to the test AC-rich elements, it will block the expression of the LacZ reporter gene activated by the GAL4 AD. A R2R3-MYB transcriptional activator PtMYB4 (Bomal et al., 2008) and a bZip transcription factor HY5 (ELONGATED HYPOCOTYL 5) (Oyama et al., 1997) were used as positive and negative controls, respectively, along with the vector control containing only the GAL4-AD (Figure 3-3, Figure 3-12). The positive control PtMYB4 increased the lacz reporter gene expression by c. 50–100-fold compared with the vector and negative controls, indicating that the system is functional in yeast (Figure 3-12). By contrast, PvMYB4 repressed the expression from the three AC-I, AC-II and AC-III elements fused to the Mini-pCyC1 promoter driving the lacZ reporter gene by c. 37–82% compared with the vector and negative control (HY5) under the same conditions (Figure 3-3d). However, PvMYB4 did not repress transcription driven by the three AC-IV elements in yeast (Figure 3-3d), suggesting that it may not bind to the AC-IV element, or may do so only at very low affinity.

PvMYB4 was cloned into the pDEST17 vector with a His tag fusion at the N-terminal, and the purified protein from E. coli extracts was shown to bind to the AC-I, AC-II, AC-III and AC-IV

elements in EMSA assays (Figure 3-3e). Salmon sperm DNA was added as a nonspecific competitor (200×); it could not compete with the AC-I element but exhibited weak competition activity with the AC-II and AC-III elements, and stronger competition activity with the AC-IV element. The results indicate that PvMYB4 has the highest binding affinity to the AC-I element and the lowest affinity to the AC-IV element in vitro, consistent with the results of the repression activity assay in yeast (Figure 3-3d).

It has been reported that the C1 motif (also called GIDP motif) has putative activation activity (Matsui et al., 2008) and that C2 motifs in AtMYB4 function as the repression motifs (Jin et al., 2000). To investigate the functions of the C-terminal motifs in PvMYB4 in more detail, we fused different deletions of PvMYB4 between the VP16 activation motif from herpes simplex virus protein VP16 (Sadowski et al., 1988) and the GAL4 binding domain (Figure 3-3f). The reporter construct was made by placing six GAL4 binding motifs in front of the pCYC1 minimal promoter driving the LacZ reporter gene. After expression of reporter and effector constructs in yeast, β-galactosidase assays indicated that the C2 and C4 motifs all have transcriptional repression activities, because the deletion of these motifs restored the transcriptional activation effects contributed to PvMYB4 by the VP16 activation motif. When the C1 motif was further deleted, the transcriptional activation was reduced slightly, consistent with a previous suggestion of transcriptional activation activity for the C1 motif (Matsui et al., 2008). By contrast, the Zf motif (Figure 3-1b) did not appear to have repression activity (Figure 3-3g).

A region with a high percentage of proline (P) and glutamate (Q) residues is located between the C1 and C2 motifs of PvMYB4, and most of the sequence variations were found in this area (Figure 3-1b, a). It has been reported that P-Q-rich and acidic-blob-type regions are associated with transcriptional activation domains (Ruden et al., 1991). To test this hypothesis, we created a

fusion of the PQ motif (green box,) with the GAL4 binding domain as an effector plasmid, and expressed it in yeast with the 6GAL4-Mini-pCYC1: lacZ reporter construct; no transcriptional activation of β-galactosidase was detected (Figure 3-13), indicating that the P-Q-rich region has no transcriptional activation activity.

3.4.4 Expression of PvMYB4 in transgenic tobacco

To investigate the biological activity of PvMYB4 in vivo, we first overexpressed the gene under the control of the constitutive 35S promoter in tobacco. Several independent lines, such as #6, #9 and #16, showed high expression of PvMYB4 transcripts (Figure 3-4a). Overexpression of PvMYB4 gave rise to the same phenotypes as observed previously for AmMYB308 and AtMYB4 in transgenic tobacco (Tamagnone et al., 1998; Jin et al., 2000), namely a reduction in plant stature and appearance of numerous whitish lesions on the mature leaves (Figure 3-4b). The latter phenotype has been linked to reduced concentrations of hydroxycinnamic acid derivatives (Elkind et al., 1990; Tamagnone et al., 1998).

The most abundant soluble phenolic compound in methanolic extracts of tobacco stems is chlorogenic acid (CGA, Rt 14.87 min), which exists with a minor isomer (Rt = 14.37 min) (Figure 3-5a). Overexpression of PvMYB4 in tobacco caused an approx. 80% reduction in soluble CGA concentrations (Figure 3-5a, Figure 3-14a). The most abundant ester-linked wall-bound phenolic compounds released from low-temperature hydrolysis of tobacco stem cell walls are vanillin (Rt = 19.63), syringyl aldehyde (Rt = 22.59) and p-coumaric acid (p-CA, Rt = 23.78). Overexpression of PvMYB4 significantly reduced the content of vanillin and p-CA compared with the wild-type, but syringyl aldehyde concentrations were not significantly changed (Figure 3-5b, Figure 3-14b).

The lignin content and composition of PvMYB4-OX transgenic tobacco plants were also significantly changed. Leaf petiole sections of PvMYB4-OX plants showed reduced staining with phloroglucinol-HCl (Figure 3-14c), suggesting an overall reduced concentration of lignin, confirmed by the reduction in acetyl bromide lignin content (Figure 3-5c). Mäule staining showed a decrease in the red coloration of the xylem vessels, consistent with a reduction in S lignin as was confirmed by thioacidolysis analysis of lignin monomer yield (Figure 3-5d). The latter indicated a reduction in lignin content of from 40 to 60% in the PvMYB4-OX overexpressing lines compared with controls (Figure 3-5d). However, overall the G units were reduced more than the S units, leading to an increased S/G ratio in the MYB4 overexpressors (Figure 3-5e).

3.4.5 Overexpression of PvMYB4 in switchgrass

To generate switchgrass PvMYB4-OX lines, the full-length PvMYB4a ORF was cloned into a binary vector (pANIC vector, Mann et al., 2011) fused with an AcV5 epitope tag (Monsma & Blissard, 1995) at the C-terminal and driven by the maize ZmUbi promoter. The hygromycin B (hph) resistance gene was used as a selectable marker for transformation (Figure 3-15a). Embryonic callus generated from immature inflorescences of switchgrass cv Alamo line ST2 was transformed by *Agrobacterium*-mediated transformation. The transformation process, based on modifications of previously published protocols (Somleva et al., 2002; Xi et al., 2009), is illustrated in Figure 3-6(a).

Genomic DNA PCR analysis showed that transgenic lines 2A and 2B could serve as appropriate controls because only the hygromycin B resistance gene portion of the T-DNA was integrated into the genome, and the lines were phenotypically indistinguishable from untransformed ST2 lines (Figure 3-15b).

To confirm overexpression of the PvMYB4 gene in the transgenic plants, two primer pairs, HS543 and HS544, which cover the PvMYB4 ORF and AcV5 tag sequence, were used for qRT-PCR analysis (Table 3-1). The selected PvMYB4-OX lines exhibited a 10–12-fold increase in PvMYB4 expression (Figure 3-6b). The HS360 and HS361 primer pairs were designed from the 3′-UTR region of the PvMYB4 sequence to determine expression of the endogenous PvMYB4 genes in switchgrass (Figure 3-15c). The endogenous PvMYB4 gene was repressed in the transgenic plants, indicating a similar self-repression feedback regulatory network as described for AtMYB4 in Arabidopsis (Zhao et al., 2007) (Figure 3-15c).

Overexpression of PvMYB4 in switchgrass had dramatic effects on plant morphology. The PvMYB4-OX transgenic plants showed a reduction in plant height (by c. 40% on average) but more tillers (up to a 2.5-fold increase) (Figure 3-7a, b). By contrast to the PvMYB4-OX tobacco plants, no white lesions were observed on the leaves.

PvMYB4 overexpression reduces lignin content, alters cell wall phenolic content, and reduces recalcitrance in transgenic switchgrass

The lignin content in the middle parts of the different internodes of E4 stage stems of PvMYB4-OX and control switchgrass plants were first evaluated by staining with phloroglucinol-HCl and Mäule reagent. Clearly, total lignin content (phloroglucinol-HCl staining, red color) and S lignin (Mäule staining, reddish-brown color) were reduced in the stems of the transgenic plants, especially the mature I1 and I2 internodes (Figure 3-7c). Furthermore, although the structure of the vascular bundles remained the same, their size was significantly reduced in the transgenic plants based on the measurement of vascular bundles from internode 2 (Figure 3-7c, Figure 3-17a). We also found that the tillers of PvMYB4-OX lines were thinner than those of the controls (Figure 3-17b).

Total lignin content in whole stems of E4 stage plants was reduced by c. 40–50% (determined as AcBr lignin) or 60–70% (determined by thioacidolysis) as a result of overexpression of PvMYB4 (Figure 3-8a,b). However, in contrast to the PvMYB4-OX tobacco lines, the S/G ratio remains unchanged (Figure 3-16). The ester-linked p-CA : FA ratio was reduced by c. 50% compared with the control lines (Figure 3-8d) and this was mainly caused by the significant reduction of the ester-linked p-CA content in the cell wall residue (Figure 3-8c). Since both the lignin content and the ester-linked p-CA : FA ratio are reduced in the PvMYB4-OX transgenic switchgrass lines, we predicted that these lines would exhibit significantly higher saccharification efficiency than the control lines based on our previous observations of natural variation (Shen et al., 2009a). To test this, we measured sugar release from CWR by enzymatic saccharification without acid pretreatment. The total sugar released from the PvMYB4-OX transgenic switchgrass lines as a function of total available cell wall sugar content was approx. threefold higher from the PvMYB4 overexpressing lines than from the controls (Figure 3-8e,f). We also estimated the dry biomass production under glasshouse conditions; the transgenic PvMYB4-OX lines had less (1A and 1B lines), similar (line 1C and 1E lines) or c. 20% increased (line 1D) biomass depending on the line (Figure 3-18).

3.4.6 Downstream target genes of PvMYB4

Given that PvMYB4 represses phenylpropanoid metabolism and lignin content in both switchgrass and tobacco, and represses transcription of the Arabidopsis PAL4 and CCoAOMT7 promoters in vitro, we used the transgenic PvMYB4-OX lines to gain a broader picture of the downstream gene targets within the phenylpropanoid and monolignol biosynthesis pathways.

Total RNA was extracted from the leaves of tobacco lines and young seedlings of switchgrass lines. First-strand cDNA was synthesized and qRT-PCR primer pairs (Table 3-1) were designed

for each of the genes encoding the enzymes of phenylpropanoid and monolignol biosynthesis. In the transgenic control lines, F5H (ferulate 5-hydroxylase), CCoAOMT and HCT (hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase) genes showed overall lower transcript abundance than other genes (Figure 3-9a, b). The transcript abundance of PAL, C4H, C3′H (coumaroyl shikimate 3′-hydroxylase), 4CL (4-coumarate: CoA ligase), COMT (caffeic acid 3-O-methyltransferase), CCR (cinnamoyl CoA reductase) and CAD genes was reduced by at least 70% in the PvMYB4-OX tobacco lines and 50% in the switchgrass lines (Figure 3-9a, b). Because of potential cross-talk and/or spillover between lignin and flavonoid pathways in plants down-regulated in lignin biosynthesis (Besseau et al., 2007), we also analyzed genes involved in flavonoid biosynthesis (chalcone synthase, chalcone isomerase and flavonol synthase); none of these was affected by overexpression of PvMYB4 (Figure 3-19). Combining the results from biochemical and target gene expression analyses, we conclude that PvMYB4 functions as a repressor of the phenylpropanoid and lignin pathways.

3.5 Discussion

3.5.1 PvMYB4 is a lignin biosynthetic pathway repressor

R2R3-MYB4 subfamily 4 members have been shown to function as repressors of lignin biosynthesis in several species (Tamagnone et al., 1998; Jin et al., 2000; Fornaléet al., 2006, 2010; Legay et al., 2007; Zhao & Dixon, 2011). The PvMYB4 gene is expressed in all tissues of switchgrass including the roots, stems, leaves, leaf sheaths and inflorescences at the R1 development stages. In situ hybridization of young stem sections showed that PvMYB4 is highly expressed in the vascular bundles, consistent with its binding activity for the xylem expression-associated AC elements. The strong expression level of PvMYB4 in the vascular bundles and

parenchyma cells at earlier developmental stages is consistent with lower degrees of lignification in young compared with mature stems.

Protein motif analysis showed that PvMYB4 has four conserved motifs, C1, C2, Zf and C4, in the C-terminal regulatory region. The C1 motif may function as an activation domain, confirming a previous report (Matsui et al., 2008), and the C2 motif functions as the repression domain, consistent with results from similar analysis of AtMYB4 (Jin et al., 2000). The C4 motif also acts as a repression motif although AmMYB308, another lignin repressor, does not have this motif. The coexistence of the activation and repression motifs at the C-terminal of PvMYB4 reflects the complexity of the protein structures and potential gene regulatory network. The function of the Zf motif within the C-terminal regulatory region remains unknown for the R2R3-MYB subfamily 4 proteins.

Five PvMYB4 gene variants were identified by EST cloning in switchgrass. Allelic variation in a polyploid genome has been suggested as being responsible for heterosis (Springer & Stupar, 2007). The lignin biosynthetic gene PvCCR has at least four gene variants (Escamilla-Trevino et al., 2009). These natural gene variants provide a basis for molecular marker discovery. It will be interesting to determine whether these variations in lignin pathway synthetic and regulatory genes and alleles are linked to recalcitrance phenotypes using marker-assisted approaches.

Overexpression of transcriptional repressors of the lignin synthetic pathway reduces lignin content in a range of dicot plants (Tamagnone et al., 1998; Jin et al., 2000; Fornaléet al., 2006, 2010; Legay et al., 2007). However, this reduction had not yet been shown in monocots. The basal internodes of switchgrass have high lignin content during development (Sarath et al., 2007; Shen et al., 2009a). Chemical staining and wet chemistry analysis of the stems of PvMYB4-OX

transgenic plants showed that lignin content was significantly reduced. The amounts of both S and G subunits were reduced in both tobacco and switchgrass overexpressing PvMYB4; the reduction in S and G units was similar in switchgrass, whereas the greater reduction in G units resulted in an increased S/G ratio in tobacco. Phylogenetic analysis shows that PvMYB4 is closely related to ZmMYB42 and ZmMYB31, and overexpression of ZmMYB42 in transgenic Arabidopsis decreased S/G ratio through the relative enrichment of H and G subunits and reduction of S subunits (Sonbol et al., 2009). However, overexpression of ZmMYB31 in transgenic Arabidopsis reduced lignin content but led to an increased amount of H subunits with no effect on the S/G ratio (Fornalé et al., 2010). Thus, qualitative regulation of monolignol composition by MYB repressors appears to be both repressor- and species-specific, and not simply a result of whether the MYB gene or target species is monocot or dicot.

Although all the phenylpropanoid/monoligol biosynthetic genes were repressed by PvMYB4 in transgenic tobacco and switchgrass, whether this effect is directly or indirectly regulated through the binding of PvMYB4 to AC elements is not clear. ZmMYB31 can bind directly to the ZmF5H gene promoter region containing an ACII-like element (tCCAACC), as shown by quantitative reverse transcription chromatin immunoprecipitation (qRT-ChIP) assays in maize (Fornalé et al., 2010). The potential presence of AC elements in (all) the switchgrass monolignol pathway genes, and direct binding of PvMYB4 to these putative elements, can be evaluated once the switchgrass genome sequence becomes available.

3.5.2 Biotechnological applications of PvMYB4 for development of switchgrass as a dedicated bioenergy crop

Bioenergy is commonly defined as the energy generated from renewable biological products (such as plant biomass) and is regarded as a clean and sustainable alternative to fossil sources of energy due to the potential contribution to reduced carbon dioxide emissions (McLaughlin et al., 2002; McLaughlin & Adams Kszos, 2005; Williams et al., 2009). Switchgrass feedstock is highly lignified plant biomass, and the high lignin content dramatically increases the energy cost for releasing the polysaccharide components from cell walls during biomass to biofuel conversion.

The molecular, genomic and biotechnological resources for switchgrass are limited but increasing exponentially (Tobias et al., 2005, 2008; Okada et al., 2010). Reliable genetic transformation of switchgrass using *Agrobacterium*-mediated methods has recently been developed (Burris et al., 2009; Xi et al., 2009), paving the way for increased biotechnological applications for biofuel and bio product improvements. For example, successful engineering of a functional metabolic pathway for the production of polyhydroxybutyrate (PHB) in transgenic switchgrass has been reported (Somleva et al., 2008), suggesting that complex traits can be engineered in this dedicated bioenergy crop. One obvious strategy to overcome the cell wall recalcitrance of switchgrass biomass is to reduce the lignin content (Hisano et al., 2009), based on the effectiveness of this approach in alfalfa and other species (Bell et al., 2004; Chen & Dixon, 2007). This has recently been achieved in switchgrass through down-regulation of the monolignol biosynthetic genes CAD (Saathoff et al., 2011) and COMT (Fu et al., 2011), leading to increased sugar release and, in the latter case, improved ethanol yield at lower cellulase loadings.

The present report describes an alternative strategy for reducing recalcitrance in switchgrass by manipulating the expression of a key transcription factor. The lignin content, wall-bound ester-linked p-CA content and p-CA : FA ratio were reduced in the transgenic switchgrass, both features that could result in significantly improved saccharification efficiency based on our previous studies of natural variation in these cell wall parameters (Shen et al., 2009a) and the phenotypes of the PvCOMT down-regulated transgenic switchgrass (Fu et al., 2011). Indeed, the PvMYB4-OX transgenic switchgrass lines have approx. threefold increased saccharification efficiency compared with the control lines. Down-regulation of lignin pathway genes can result in negative growth impacts, possibly as a result of metabolic spillover (Besseau et al., 2007; Li et al., 2010; Gallego-Giraldo et al., 2011). The PvCOMT down-regulated switchgrass plants have a normal growth phenotype under glasshouse conditions, whereas the PvMYB4-OX plants have reduced stature but increased tillering. More detailed analysis of gene expression in these plants utilizing the recently available switchgrass Affymetrix Gene Chip may help explain the basis of this growth phenotype, which is reminiscent of the reduced stature but increased branching observed in alfalfa with down-regulated HCT expression (Gallego-Giraldo et al., 2011). In these alfalfa plants, alterations in cytokinin and gibberellin concentrations and/or signaling may contribute to the phenotype (Gallego-Giraldo et al., 2011).

Based on an estimation of dry biomass production under glasshouse conditions, some of the transgenic PvMYB4-OX lines have similar (line 1C) or c. 20% increased (line 1D) biomass production. Evaluation of these lines is therefore necessary under field conditions, with the potential for selection of lines retaining both improved biomass characteristics and good growth phenotypes.

3.6 Acknowledgements

We thank Dr Jin Nakashima for assistance with cell imaging, Dr Mohamed Bedair for GC-MS analysis, Dr Guifen Li, Jason N. Burris, and Ellen Reeves for assistance with switchgrass tissue culture and transformation, and Drs Qiao Zhao and Cheol-Min Yoo for critical reading of the manuscript. This work was supported by the BioEnergy Science Center, a US Department of Energy Bioenergy Research Center, through the Office of Biological and Environmental Research in the DOE Office of Science.

3.7 References

Akin DE. 2007. Grass lignocellulose: strategies to overcome recalcitrance. *Applied Biochemistry and Biotechnology* 137: 3–15.

Asai T, Tena G, Plotnikova J, Willmann MR, Chiu WL, Gomez-Gomez L, Boller T, Ausubel FM, Sheen J. 2002. Map kinase signalling cascade in *Arabidopsis* innate immunity. *Nature* 415: 977–983.

Bell J, Lehmann D, Wang ZY, Chen L, Auh CK, Dowling P. 2004. Transgenic down-regulation of caffeic acid o-methyltransferase (COMT) led to improved digestibility in tall fescue (*Festuca arundinacea*). *Functional Plant Biology* 31: 235–245.

Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M. 2007. Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects Auxin transport and plant growth. *Plant Cell* 19: 148–162.

Bomal C, Bedon F, Caron S, Mansfield SD, Levasseur C, Cooke JEK, Blais S, Tremblay L, Morency MJ, Pavy N. 2008. Involvement of *Pinus taeda* MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis. *Journal of Experimental Botany* 59: 3925–3939.

Bouton JH. 2007. Molecular breeding of switchgrass for use as a biofuel crop. *Current Opinion in Genetics & Development* 17: 553–558.

Bylesjo M, Nilsson R, Srivastava V, Gronlund A, Johansson AI, Jansson S, Karlsson J, Moritz T, Wingsle G, Trygg J. 2008. Integrated analysis of transcript, protein and metabolite data to study lignin biosynthesis in hybrid aspen. *Journal of Proteome Research* 8: 199–210.

Chen F, Dixon RA. 2007. Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology* 25: 759–761.

Deluc L, Barrieu F, Marchive C, Lauvergeat V, Decendit A, Richard T, Carde J-P, Merillon J-M, Hamdi S. 2006. Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. *Plant Physiology* 140: 499–511.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M. 2008. Phylogeny. Fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research* 36: W465–W469.

Elkind Y, Edwards R, Mavandad M, Hedrick SA, Ribak O, Dixon RA, Lamb CJ. 1990. Abnormal plant development and down regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene. *Proceedings of the National Academy of Sciences, USA* 87: 9057–9061.

Escamilla-Trevino LL, Shen H, Uppalapati SR, Ray T, Tang Y, Hernandez T, Yin Y, Xu Y, Dixon RA. 2009. Switchgrass (*Panicum virgatum*) possesses a divergent family of cinnamoyl CoA reductases with distinct biochemical properties. *New Phytologist* 185: 143–155.

Fornalé S, Shi X, Chai C, Encina A, Irar S, Capellades M, Fuguet E, Torres JL, Rovira P, Puigdomènech P. 2010. ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant Journal* 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 Molecular Biology* 62: 809–

Fu C, Mielenz JR, Xiao X, Ge Y, Hamilton CY, Chen F, Bouton J, Foston M, Dixon RA, Wang Z-Y. 2011. Genetic manipulation of lignin biosynthesis in switchgrass significantly reduces recalcitrance and improves biomass ethanol production. *Proceedings of the National Academy of Sciences, USA* 108: 3803–3808.

Fuentes RG, Taliaferro CM. 2002. Biomass yield stability of switchgrass cultivars. In: Janick J, Whipkey A, eds. *Trends in new crops and new uses*. Alexandria, VA, USA: ASHS Press, 276–282.

Gális I, Šimek P, Narisawa T, Sasaki M, Horiguchi T, Fukuda H, Matsuoka K. 2006. A novel R2R3 MYB transcription factor NtMYBJS1 is a methyl jasmonate-dependent regulator of phenylpropanoid-conjugate biosynthesis in tobacco. *Plant Journal* 46: 573–592.

Gallego-Giraldo L, Jikumaru Y, Kamiya Y, Tang Y, Dixon RA. 2011. Selective lignin down-regulation leads to constitutive defense response expression in alfalfa (*Medicago sativa* L.). *New Phytologist* 190: 627–639.

Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N et al. 2005. EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant Journal* 43: 553–567.

Hatton D, Sablowski R, Yung MH, 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 Journal* 7*:* 859–876.

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315:

Hisano H, Nandakumar R, Wang Z-Y. 2009. Genetic modification of lignin biosynthesis for improved biofuel production. *In Vitro Cellular & Developmental Biology – Plant* 45: 306–313.

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT. 1985. A simple and general method for transferring genes into plants. *Science* 227: 1229–1231.

Jackson D. 1991. In-situ hybridization in plants. In: Bowles DJ, Gurr SJ, McPherson M, eds. *Molecular plant pathology: a practical approach*. New York, USA: Oxford University Press, 163–174.

Jackson LA, Shadle GL, Zhou R, Nakashima J, Chen F, Dixon RA. 2008. Improving saccharification efficiency of alfalfa stems through modification of the terminal stages of monolignol biosynthesis. *BioEnergy Research* 1*:* 180–192.

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*. *The EMBO Journal* 19: 6150–6161.

Karlen Y, McNair A, Perseguers S, Mazza C, Mermod N. 2007. Statistical significance of quantitative PCR. *BMC Bioinformatics* 8: 131.

Karpinska B, Karlsson M, Srivastava M, Stenberg A, Schrader J, Sterky F, Bhalerao R, Wingsle G. 2004. MYB transcription factors are differentially expressed and regulated during secondary vascular tissue development in hybrid aspen. *Plant Molecular Biology* 56: 255–270.

Keshwani DR, Cheng JJ. 2009. Switchgrass for bioethanol and other value-added applications: a review. *Bioresource Technology* 100: 1515–1523.

Kurosu T, Peterlin BM. 2004. VP16 and ubiquitin; binding of P-TEFb via its activation domain and ubiquitin facilitates elongation of transcription of target genes. *Current Biology* 14: 1112–1116.

Legay S, Lacombe E, Goicoechea M, Briere C, Séguin A, Mackay J, Grima-Pettenati J. 2007. Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Science* 173: 542–549.

Li X, Bonawitz ND, Weng JK, Chapple C. 2010. The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. *Plant Cell* 22: 1620–1632.

Mann DGJ, LaFayette PR, Abercrombie LL, King ZR, Mazarei M, Halter MC, Poovaiah CR, Baxter H, Shen H, Dixon RA et al. 2011. Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass and other monocot species. *Plant Biotechnology Journal*. 10(2): 226-236

Matsui K, Umemura Y, Ohme-Takagi M. 2008. Atmybl2, a protein with a single myb domain, acts as a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *Plant Journal* 55: 954–967.

McLaughlin SB, Adams Kszos L. 2005. Development of switchgrass (*Panicum virgatum*) as a bioenergy feedstock in the United States. *Biomass and Bioenergy* 28: 515–535.

McLaughlin SB, De La Torre Ugarte DG, Garten CT Jr, Lynd LR, Sanderson MA, Tolbert VR, Wolf DD. 2002. High-value renewable energy from prairie grasses. *Environmental Science & Technology* 36*:* 2122–2129.

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.

Monsma SA, Blissard GW. 1995. Identification of a membrane fusion domain and an oligomerization domain in the baculovirus GP64 envelope fusion protein. *Journal of Virology* 69: 2583.

Moore KJ, Moser LE, Vogel KP, Waller SS, Johnson BE, Pedersen JF. 1991. Describing and quantifying growth stages of perennial forage grasses. *Agronomy Journal* 83: 1073–1077.

Nakashima J, Chen F, Jackson L, Shadle G, Dixon RA. 2008. Multi-site genetic modification of monolignol bioynthesis in alfalfa (*Medicago sativa* L.)- effects on lignin composition in specific cell types. *New Phytologist* 179: 738–750.

Okada M, Lanzatella C, Saha MC, Bouton J, Wu R, Tobias CM. 2010. Complete switchgrass genetic maps reveal subgenome collinearity, preferential pairing, and multilocus interactions. *Genetics* 185: 745–760.

Olsen AN, Ernst HA, Lo Leggio L, Skriver K. 2005. NAC transcription factors: structurally distinct, functionally diverse. *Trends in Plant Science* 10: 79–87.

Oyama T, Shimura Y, Okada K. 1997. The *Arabidopsis* hy5 gene encodes a bZip protein that regulates stimulus-induced development of root and hypocotyl. *Genes & Development* 11: 2983.

Patzlaff A, McInnis S, Courtenay A, Surman C, Newman LJ, Smith C, Bevan MW, Mansfield S, Whetten RW, Sederoff RR et al. 2003. Characterisation of a pine MYB that regulates lignification. *Plant Journal* 36: 743–754.

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

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

Raes J, Rohde A, Christensen JH, Van de Peer Y, Boerjan W. 2003. Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiology* 133: 1051–1071.

Romero I, Fuertes A, Benito MJ, Malpica JM, Leyva A, Paz-Ares J. 1998. More than 80 R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. *Plant Journal* 14: 273–284.

Rose MD, Winston FM, Heiter P. 1990. *Methods in yeast genetics: a laboratory course manual*, Vol. 19. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press.

Ruden DM, Ma J, Li Y, Wood K, Ptashne M. 1991. Generating yeast transcriptional activators containing no yeast protein sequences. *Nature* 350: 250–252.

Saathoff AJ, Sarath G, Chow EK, Dien BS, Tobias CM, Herrera-Estrella A. 2011. Downregulation of cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. *PLoS ONE* 6: 804–807.

Sadowski I, Ma J, Triezenberg S, Ptashne M. 1988. Gal4-VP16 is an unusually potent transcriptional activator. *Nature* 335: 563–564.

Sarath G, Baird LM, Vogel KP, Mitchell RB. 2007. Internode structure and cell wall composition in maturing tillers of switchgrass (*Panicum virgatum*. L). *Bioresource Technology* 98: 2985–2992.

Schmer MR, Vogel KP, Mitchell RB, Perrin RK. 2008. Net energy of cellulosic ethanol from switchgrass. *Proceedings of the National Academy of Sciences, USA* 105: 464–469.

Sheen J. 2001. Signal transduction in maize and *Arabidopsis* mesophyll protoplasts. *Plant Physiology* 127: 1466–1475.

Shen H, Fu C, Xiao X, Ray T, Tang Y, Wang Z, Chen F. 2009a. Developmental control of lignification in stems of lowland switchgrass variety Alamo and the effects on saccharification efficiency. *BioEnergy Research* 2: 233–245.

Shen H, Yin Y, Chen F, Xu Y, Dixon R. 2009b. A bioinformatic analysis of NAC genes for plant cell wall development in relation to lignocellulosic bioenergy production. *BioEnergy Research* 2: 217–232.

Somleva MN, Snell KD, Beaulieu JJ, Peoples OP, Garrison BR, Patterson NA. 2008. Production of polyhydroxybutyrate in switchgrass, a value-added co-product in an important lignocellulosic biomass crop. *Plant Biotechnology Journal* 6: 663–678.

Somleva MN, Tomaszewski Z, Conger BV. 2002. *Agrobacterium*-mediated genetic transformation of switchgrass. *Crop Science* 42: 2080–2087.

Sonbol F-M, Fornalé S, Capellades M, Encina A, Touriño S, Torres J-L, Rovira P, Ruel K, Puigdomènech P, Rigau J et al. 2009. The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in Arabidopsis thaliana. *Plant Molecular Biology* 70: 283–296.

Sparkes IA, Runions J, Kearns A, Hawes C. 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols* 1: 2019–2025.

Springer NM, Stupar RM. 2007. Allelic variation and heterosis in maize: how do two halves make more than a whole? *Genome Research* 17: 264.

Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* 4: 447–456.

Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Macia FA, Roberts K, Martin C. 1998. The AmMYB308 and AmMYB330 transcription factors from antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10: 135–154.

Tobias CM, Sarath G, Twigg P, Lindquist E, Pangilinan J, Penning BW, Barry K, McCann MC, Carpita NC, Lazo GR. 2008. Comparative genomics in switchgrass using 61,585 high-quality expressed sequence tags. *The Plant Genome* 1: 111–124.

Tobias CM, Twigg P, Hayden DM, Vogel KP, Mitchell RM, Lazo GR, Chow EK, Sarath G. 2005. Analysis of expressed sequence tags and the identification of associated short tandem repeats in switchgrass. *Theoretical and Applied Genetics* 111: 956–964.

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. *Proceedings of the National Academy of Sciences, USA* 107: 22338–22343.

Wang YS, Yoo CM, Blancaflor EB. 2008. Improved imaging of actin filaments in transgenic *Arabidopsis* plants expressing a green fluorescent protein fusion to the C and N termini of the fimbrin actin binding domain 2. *New Phytologist* 177: 525–536.

Williams PRD, Inman D, Aden A, Heath GA. 2009. Environmental and sustainability factors associated with next-generation biofuels in the US: what do we really know*? Environmental Science & Technology* 43: 4763–4775.

Wullschleger SD, Davis EB, Borsuk ME, Gunderson CA, Lynd LR. 2010. Biomass production in switchgrass across the United States: database description and determinants of yield. *Agronomy Journal* 102: 1158–1168.

Xi Y, Ge Y, Wang ZY. 2009. Genetic transformation of switchgrass. *Methods in Molecular Biology* 581: 53–59.

Zhang F, Barboric M, Blackwell TK, Peterlin BM. 2003. A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb. Genes & Development 17: 748–758.

Zhao Q, Dixon RA. 2011. Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends in Plant Science* 16: 227–233.

Zhao J, Zhang W, Zhao Y, Gong X, Guo L, Zhu G, Wang X, Gong Z, Schumaker KS, Guo Y. 2007. SAD2, an importin -like protein, is required for UV-B response in *Arabidopsis* by mediating MYB4 nuclear trafficking. *Plant Cell* 19: 3805–3818.

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.

Zhou J, Lee C, Zhong R, Ye ZH. 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell* 21: 248–266.

Zhou R, Jackson L, Shadle G, Nakashima J, Temple S, Chen F, Dixon RA. 2010. Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proceedings of the National Academy of Sciences, USA* 107: 17803–17808.

3.8 Appendix

3.8.1 Figures and tables

Figure 3-1 PvMYB4 belongs to the R2R3-MYB subfamily 4.

(a) ClustalW alignment of the amino acid sequences of PvMYB4 and other R2R3-MYB subfamily 4 proteins. The R2, R3 MYB domains are underlined. The boxed sequences are the potential functional motifs. (b) Structure of PvMYB4 protein domains and motifs. The boxed sequences are the potential functional motifs. Different types of lines indicate different motifs as shown in (b). PPI, protein–protein interaction motif for interaction with bHLH protein; C1, C2, Zf and C4, C-terminal motifs. Accession numbers for subgroup four sequences: PvMYB4a (JF299185), PvMYB4b (JF299186), PvMYB4c (JF299187), PvMYB4d (JF299188), PvMYB4e (JF299189), AtMYB32 (NM_119665), EgMYB1 (CAE09058), AmMYB308 (JQ0960), AtMYB4 (AY519615), AmMYB330 (JQ0957), ZmMYB42 (CAJ42204) and ZmMYB31 (CAJ42202). (c) Phylogenetic analysis of PvMYB4 with other subfamily 4 protein sequences by Phylogeny.fr program (http://www.phylogeny.fr/).

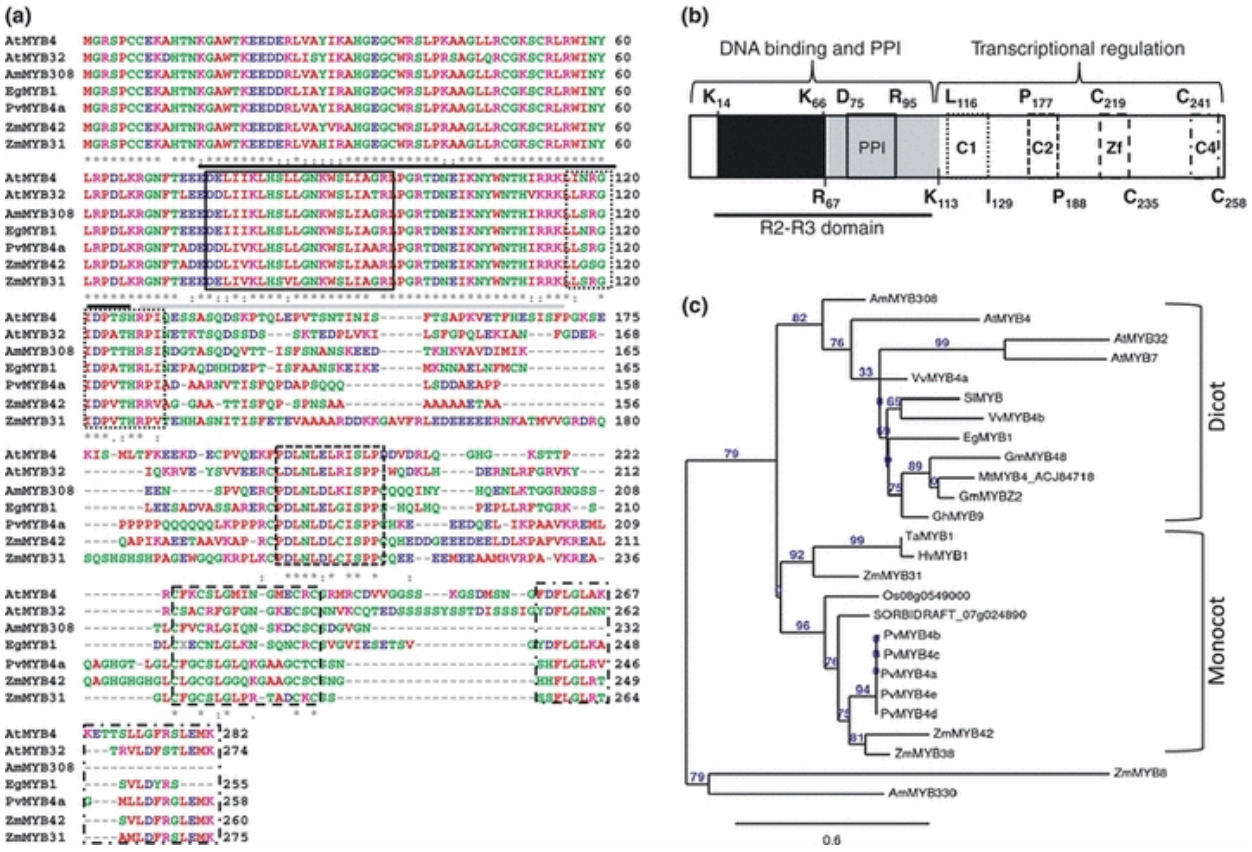

(a)
AtMYB4
AtMYB32
AmMYB308
EgMYB1
PvMYB4a
ZmMYB42
ZmMYB31
(b)
DNA binding and PPI
Transcriptional regulation
K14
K66
D75
R95
L116
P177
C219
C241
PPI
C1
C2
Zf
C4
R67
K113
I129
P188
C235
C258
R2-R3 domain
(c)
AmMYB308
AtMYB4
AtMYB32
AtMYB7
VvMYB4a
SlMYB
VvMYB4b
EgMYB1
GmMYB48
MtMYB4_ACJ84718
GmMYBZ2
GhMYB9
TaMYB1
HvMYB1
ZmMYB31
Os08g0549000
SORBIDRAFT_07g024890
PvMYB4b
PvMYB4c
PvMYB4a
PvMYB4e
PvMYB4d
ZmMYB42
ZmMYB38
ZmMYB8
AmMYB330
Dicot
Monocot
0.6

Figure 3-2 Expression patterns of PvMYB4.

(a) Quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis of tissue expression pattern of PvMYB4 in switchgrass (*Panicum virgatum*). Different letters on the bars of figures indicate significant differences at the $P \leq 0.05$ level as tested by Duncan grouping with SAS software (SAS Institute Inc.). UBQ, ubiquitin. (b) *In situ* hybridization of *PvMYB4* in young stem sections. VT, vascular tissue; P, parenchyma; E, epidermis. Bar, 500 μm. (c) Confocal images of nuclear localization of PvMYB4:eGFP protein in tobacco (*Nicotiana tabacum*) epidermal cells. Bars, 50 μm (upper panel); 10 μm (lower panel).

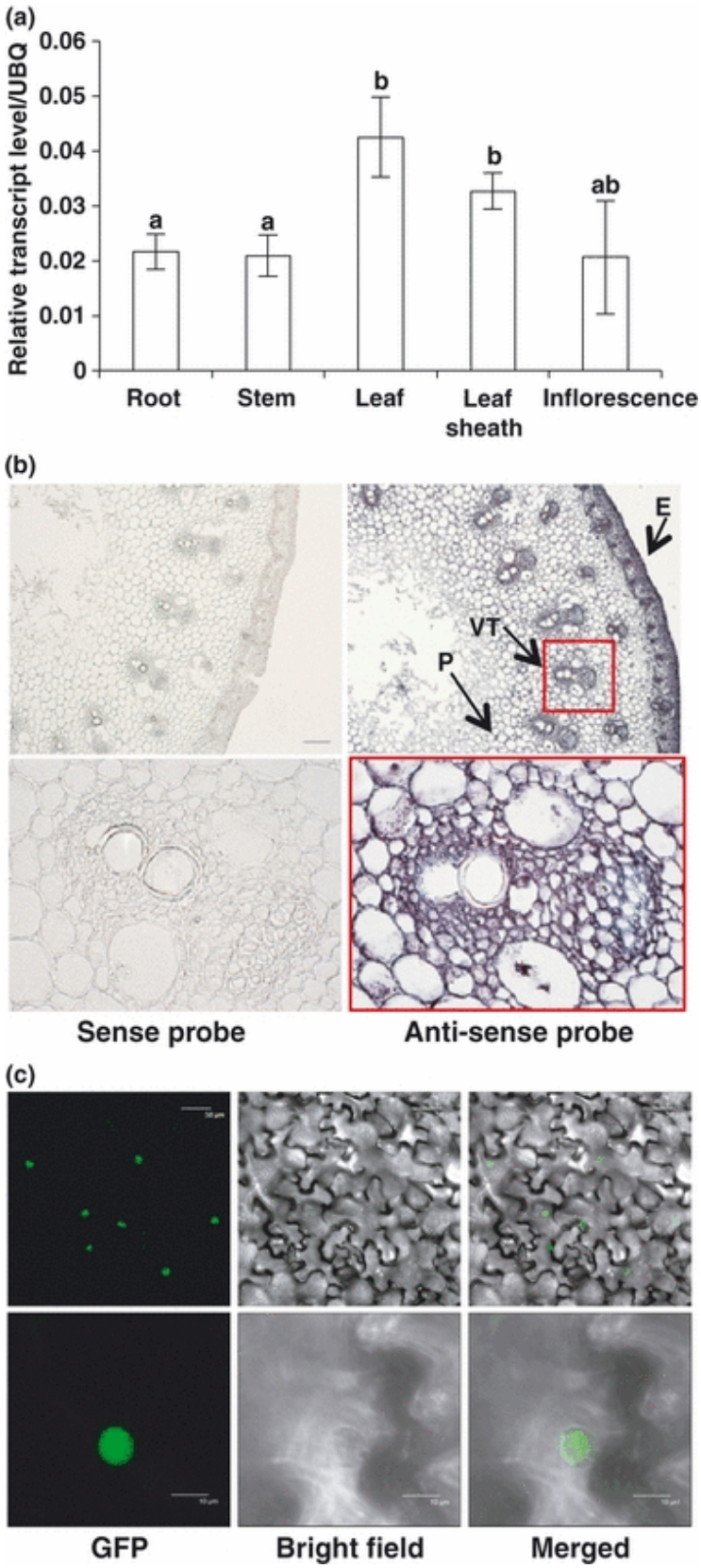
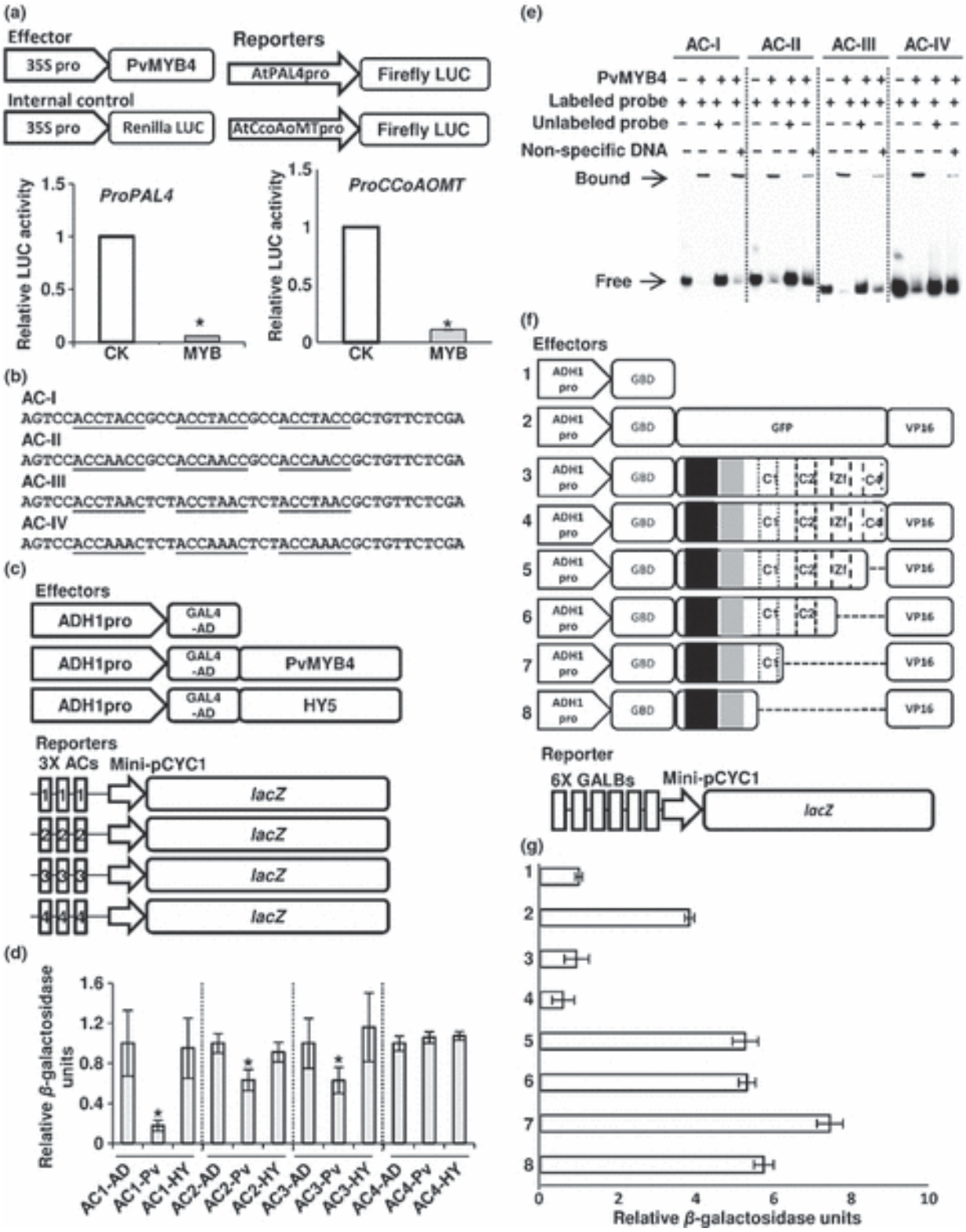

(a)
Relative transcript level/UBQ
0.06
0.05
0.04
0.03
0.02
0.01
0
a
a
b
b
ab
Root
Stem
Leaf
Leaf sheath
Inflorescence
(b)
E
VT
P
Sense probe
Anti-sense probe
(c)
GFP
Bright field
Merged

Figure 3-3 PvMYB4 is a transcriptional repressor and binds to AC elements.

(a) Dual *luciferase* (LUC) assay of 35S:PvMYB4 effector with firefly *luciferase* reporter constructs in *Arabidopsis thaliana* protoplasts. CK, vector control plasmid was used as effector; MYB, *35S:PvMYB4* plasmid was used as effector. (b) Synthetic AC elements tested in electrophoretic mobility shift assay (EMSA) and yeast assays. (c) Effector and reporter plasmids for testing the binding of PvMYB4 to AC elements in yeast. The transcription factor HY5 was used as negative control. (d) β-Galactosidase assays for lacZ expression driven by PvMYB4 binding to AC elements in yeast. AC(1,2,3,4)-AD, AC(1,2,3,4)-Pv and AC(1,2,3,4)-HY are the reporters containing three consecutive copies of the AC I, II, III, or IV elements fused to the pCYC1 minimal promoter and *lacZ* with either GAL4AD, GAL4AD-PvMYB4 or GAL4AD-HY5 as effectors. Data are means ± SE ($n = 4$). (e) EMSAs of PvMYB4 with AC elements. Free, free probe; bound, shifted band. Salmon sperm DNA (200×) was used as nonspecific competitor. (f) Effector and reporter constructs for testing the PvMYB4 repression motif in yeast. GBD, GAL4 DNA binding domain; VP16, activation motif of the VP16 protein; GALBs, GAL4 protein binding sites. (g) β-Galactosidase assays showing activation of lacZ expression from the reporter 6GALBs-*LacZ* by the effectors with different PvMYB4 C-terminal deletions shown in Figure 3-3(f). Data were normalized to the vector control pGBT-9, and are means ± SE ($n = 4$). Asterisks on top of the bars in (a) and (d) indicate values that were determined by Student's *t*-test to be significantly different from their equivalent control ($P < 0.05$).


(a)
Effector
35S pro
PvMYB4
Internal control
35S pro
Renilla LUC
Reporters
AtPAL4pro
Firefly LUC
AtCcoAoMTpro
Firefly LUC
ProPAL4
Relative LUC activity
CK
MYB
ProCCoAOMT
Relative LUC activity
CK
MYB
(b)
AC-I
AGTCCACCTACCGCCACCTACCGCCACCTACCGCTGTTCTCGA
AC-II
AGTCCACCAACCGCCACCAACCGCCACCAACCGCTGTTCTCGA
AC-III
AGTCCACCTAACTCTACCTAACTCTACCTAACGCTGTTCTCGA
AC-IV
AGTCCACCAAACTCTACCAAACTCTACCAAACGCTGTTCTCGA
(c)
Effectors
ADH1pro
GAL4 -AD
ADH1pro
GAL4 -AD
PvMYB4
ADH1pro
GAL4 -AD
HY5
Reporters
3X ACs
Mini-pCYC1
lacZ
lacZ
lacZ
lacZ
(d)
Relative β-galactosidase units
AC1-AD
AC1-Pv
AC1-HY
AC2-AD
AC2-Pv
AC2-HY
AC3-AD
AC3-Pv
AC3-HY
AC4-AD
AC4-Pv
AC4-HY
(e)
AC-I
AC-II
AC-III
AC-IV
PvMYB4
Labeled probe
Unlabeled probe
Non-specific DNA
Bound
Free
(f)
Effectors
ADH1 pro
GBD
GFP
VP16
C1
C2
IZI
C4
Reporter
6X GALBs
Mini-pCYC1
lacZ
(g)
Relative β-galactosidase units

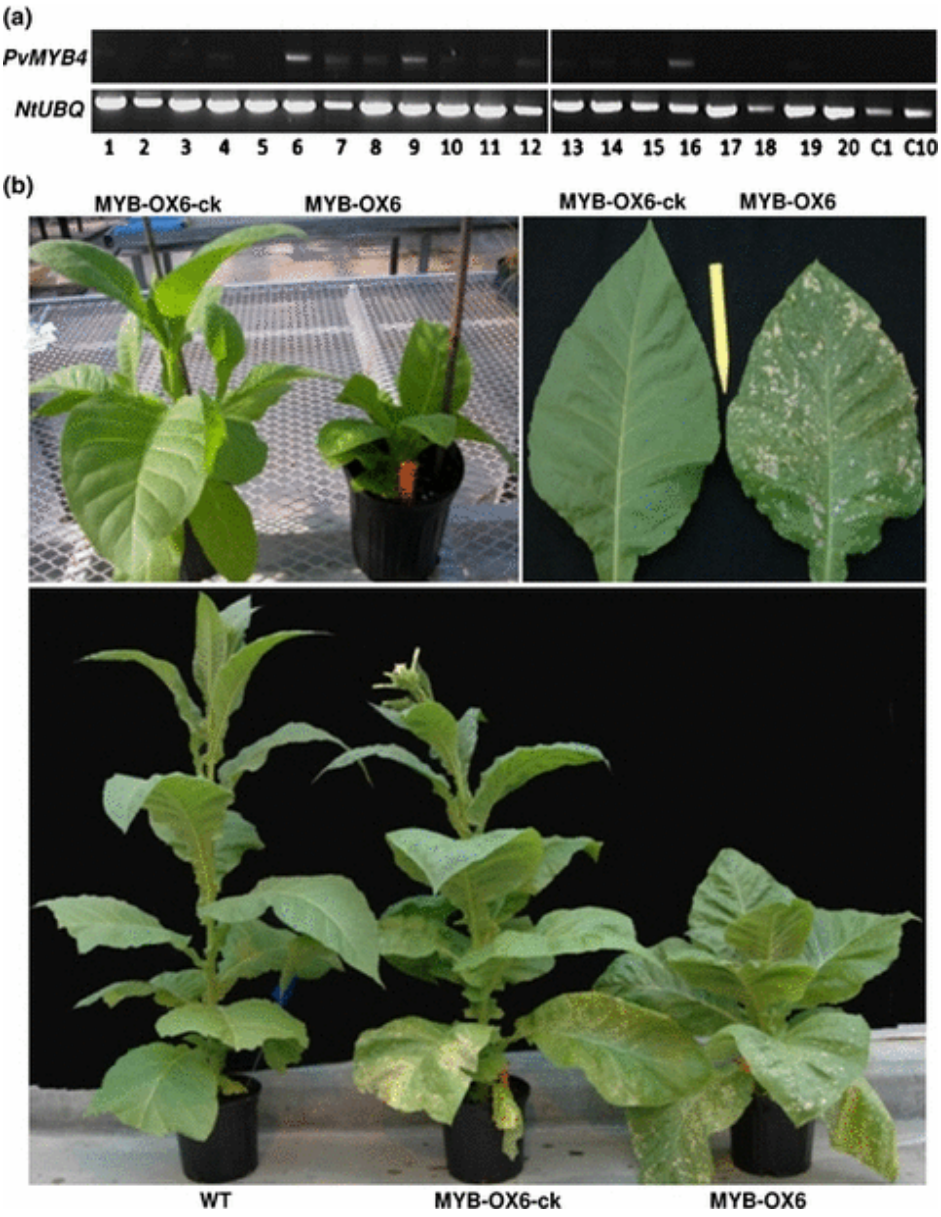
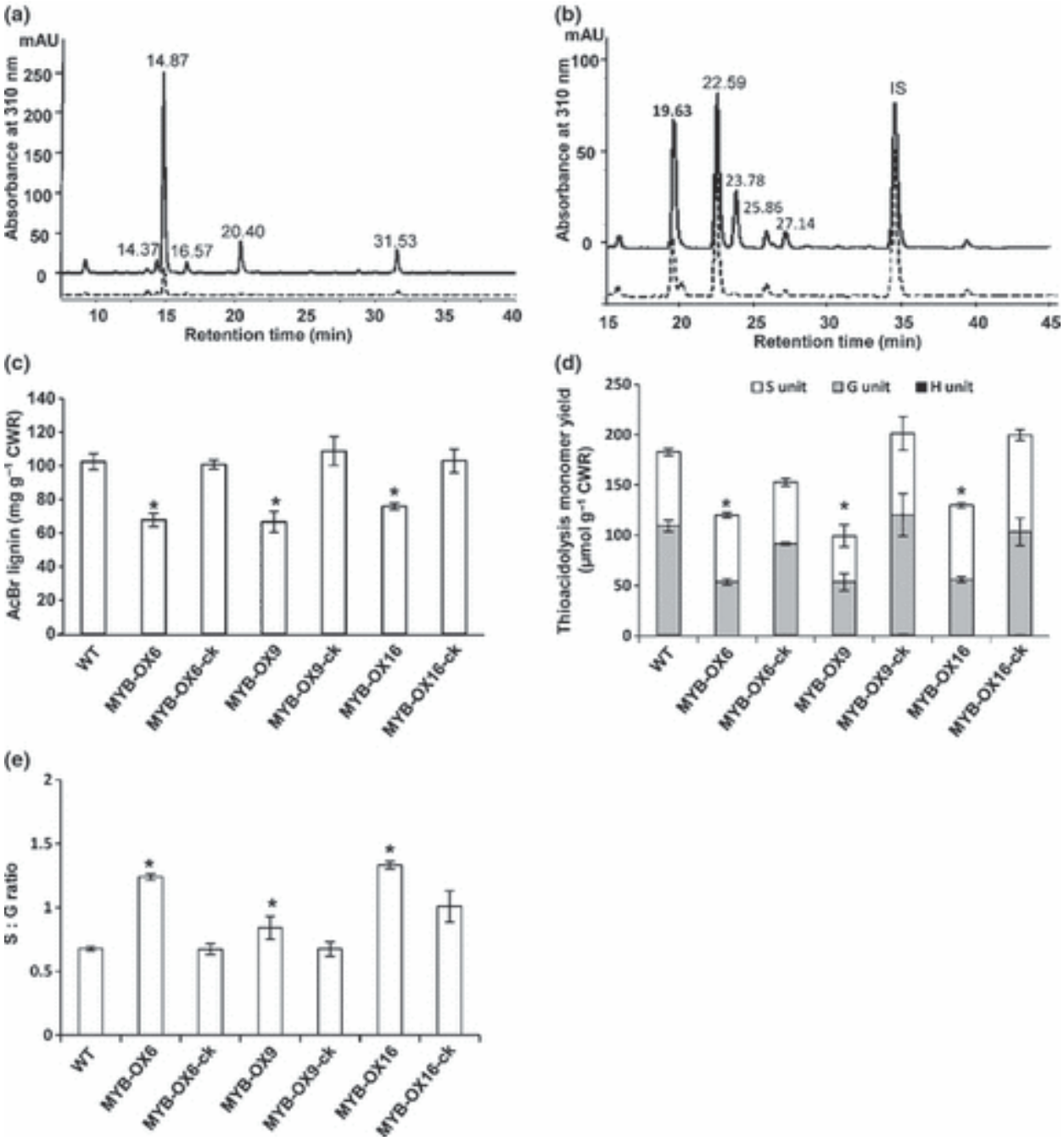


Figure 3-4 Transgenic tobacco (*Nicotiana tabacum*) plants overexpressing PvMYB4 show reduced growth and lesion mimic phenotypes.

(a) Semiquantitative reverse transcription polymerase chain reaction analysis of PvMYB4-OX transgenic tobacco plants. NtUBQ was used as reference gene.(b) Growth and leaf lesion phenotypes of PvMYB4-OX6 transgenic tobacco plants. MYB-OX6-ck indicates the T1 null-segregating line.

Figure 3-5 Overexpression of PvMYB4 alters phenylpropanoid metabolism in transgenic tobacco (*Nicotiana tabacum*) plants.

(a, b) Typical high-performance liquid chromatography (HPLC) chromatograms of methanol/water-soluble fractions (a) and ester-linked wall-bound phenolics from stems of wild-type (CK) and PvMYB4-ox tobacco lines (b). CK, solid line; MYB-OX, dashed line. The major soluble phenolic compound is chlorogenic acid (Rt = 14.87 min). Wall-bound ester-linked phenolic compounds: vanillin (Rt = 19.63); syringyl aldehyde (Rt = 22.59); *p*-coumaric acid (*p*-CA, Rt = 23.78) and ferulic acid (FA, Rt = 27.14). IS, internal standard (c) AcBr lignin content of stems of CK and PvMYB4-ox transgenic tobacco. MYB-OX-ck indicates the T1 null-segregating line. (d) Lignin composition of stems of CK and PvMYB4-ox transgenic tobacco as determined by thioacidolysis. S, syringyl unit (open bars); G, guaiacyl unit (gray bars). (e) S : G ratio (from thioacidolysis) of stems of CK and PvMYB4-ox transgenic tobacco. CWR, cell wall residue. All data are means ± SE ($n = 6$). Asterisks on top of the bars indicate values that were determined by the Student's *t*-test to be significantly different from their equivalent control ($P < 0.05$).


(a)
mAU
Absorbance at 310 nm
14.87
14.37
16.57
20.40
31.53
Retention time (min)
(b)
mAU
Absorbance at 310 nm
19.63
22.59
23.78
25.86
27.14
IS
Retention time (min)
(c)
AcBr lignin (mg g⁻¹ CWR)
WT
MYB-OX6
MYB-OX6-ck
MYB-OX9
MYB-OX9-ck
MYB-OX16
MYB-OX16-ck
(d)
S unit
G unit
H unit
Thioacidolysis monomer yield (μmol g⁻¹ CWR)
WT
MYB-OX6
MYB-OX6-ck
MYB-OX9
MYB-OX9-ck
MYB-OX16
MYB-OX16-ck
(e)
S : G ratio
WT
MYB-OX6
MYB-OX6-ck
MYB-OX9
MYB-OX9-ck
MYB-OX16
MYB-OX16-ck

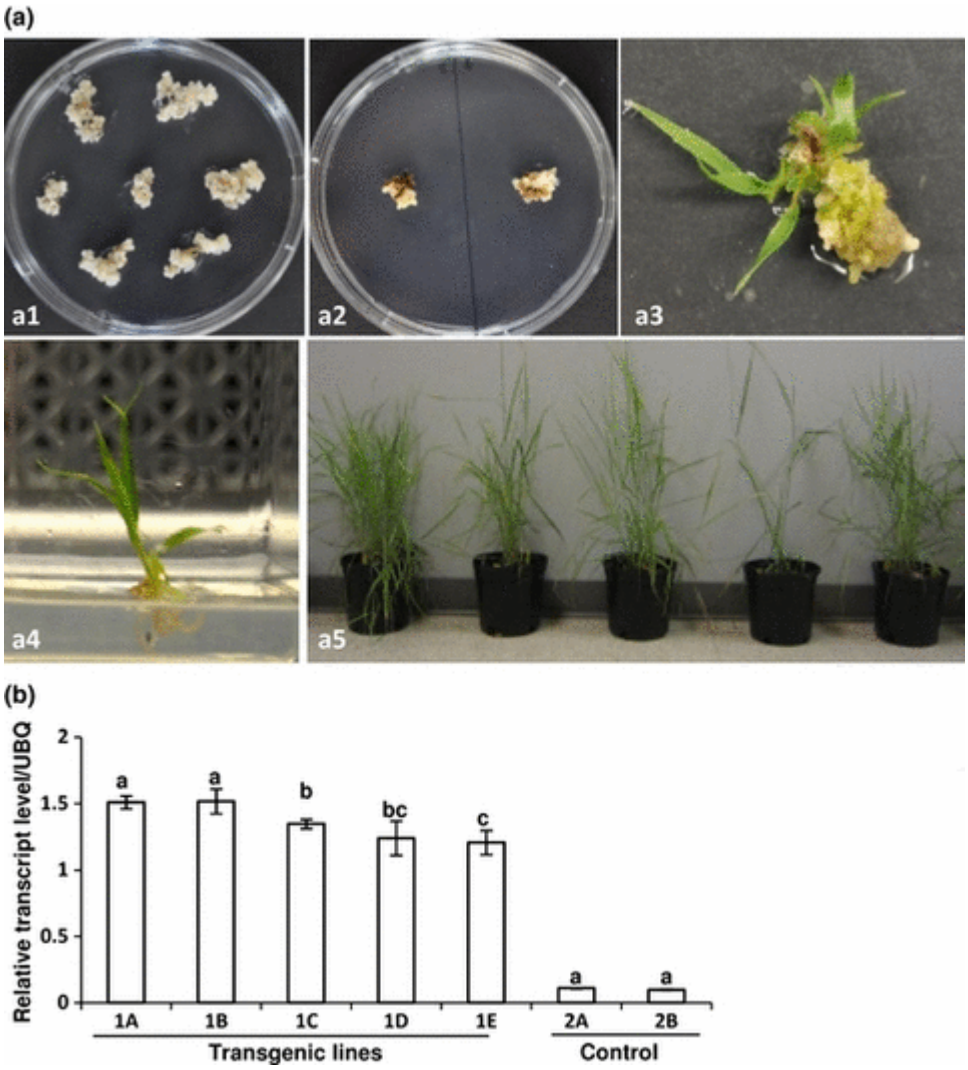
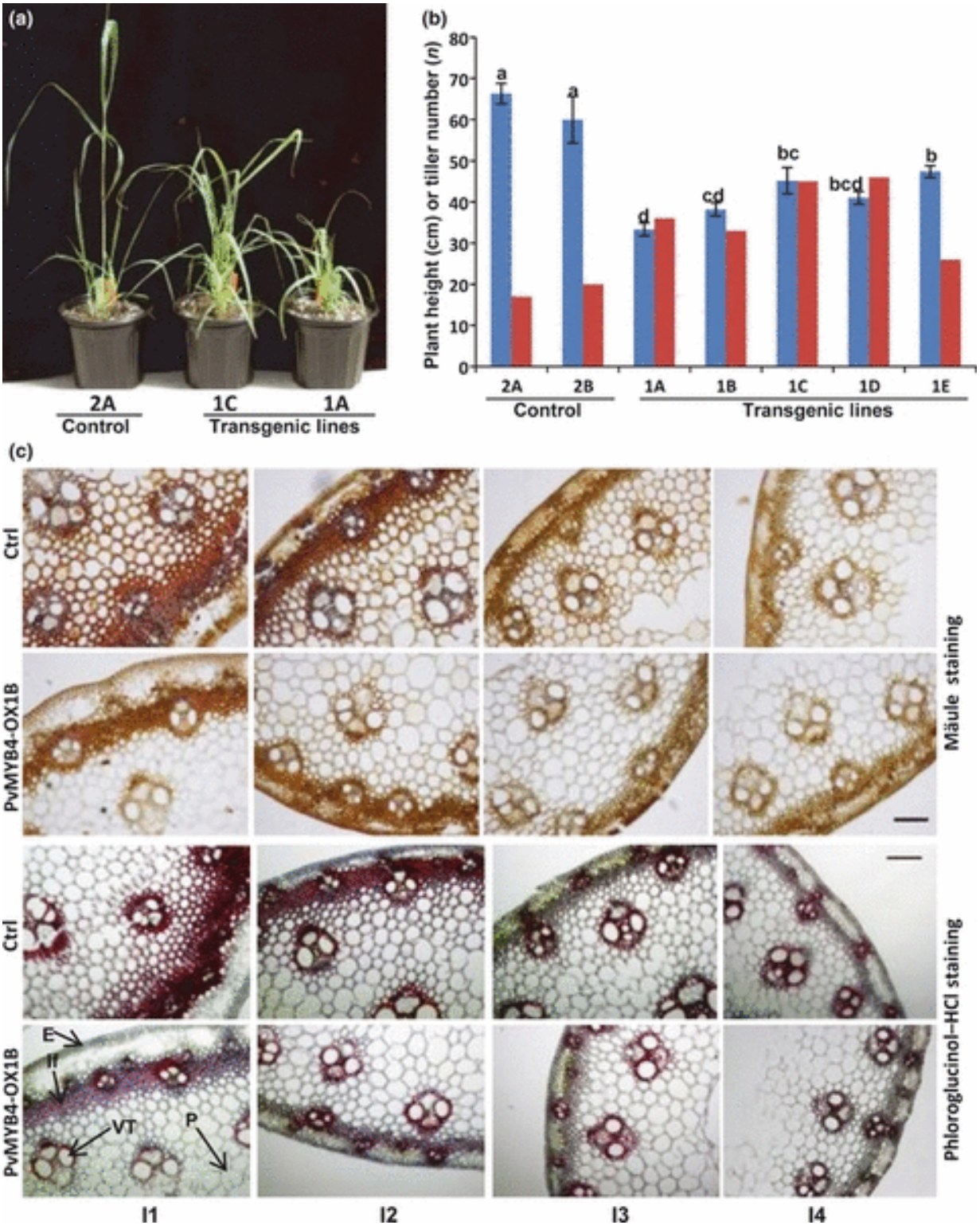


Figure 3-6 Overexpression of PvMYB4 in transgenic switchgrass (*Panicum virgatum*).

(a) Representative pictures of the switchgrass transformation procedure. a1, embryonic callus used for transformation; a2, transformed callus on selection medium; a3, regenerated seedlings on regeneration medium; a4, regenerated seedling on rooting medium; a5, regenerated PvMYB4-OX plants in the glasshouse. (b) Quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis of *PvMYB4* transcripts in control and PvMYB4-OX transgenic switchgrass. Data are means ± SE (n = 3).

Figure 3-7 PvMYB4-OX transgenic switchgrass (*Panicum virgatum*) has reduced lignin and altered growth morphology.

(a) Visible phenotypes of control (2A) and PvMYB4-OX transgenic switchgrass (1A, C). The overexpressing lines have reduced stature but increased tillering. (b) Plant height (blue bars) and tiller number (red bars) for control and PvMYB4-OX transgenic switchgrass. Data are means ± SE ($n \geq 4$). The letters indicates significant differences at the $P \leq 0.05$ level. (c) Phloroglucinol-HCl staining (lower panel) and Mäule staining (upper panel) of cross-sections of different internodes of PvMYB4-OX transgenic switchgrass. I1–I4, internode 1 (lower) to internode 4 (upper) of E4 tillers (Shen et al., 2009a); VT, vascular tissue; P, parenchyma, E, epidermis; If, interfascicular fiber. All the pictures were taken at 100× magnification. Bars, 100 μm.


(a)
2A
1C
1A
Control
Transgenic lines
(b)
Plant height (cm) or tiller number (n)
80
70
60
50
40
30
20
10
0
a
a
d
cd
bc
bcd
b
2A
2B
1A
1B
1C
1D
1E
Control
Transgenic lines
(c)
Ctrl
PvMYB4-OX1B
Mäule staining
Ctrl
PvMYB4-OX1B
Phloroglucinol–HCl staining
E
VT
P
I1
I2
I3
I4

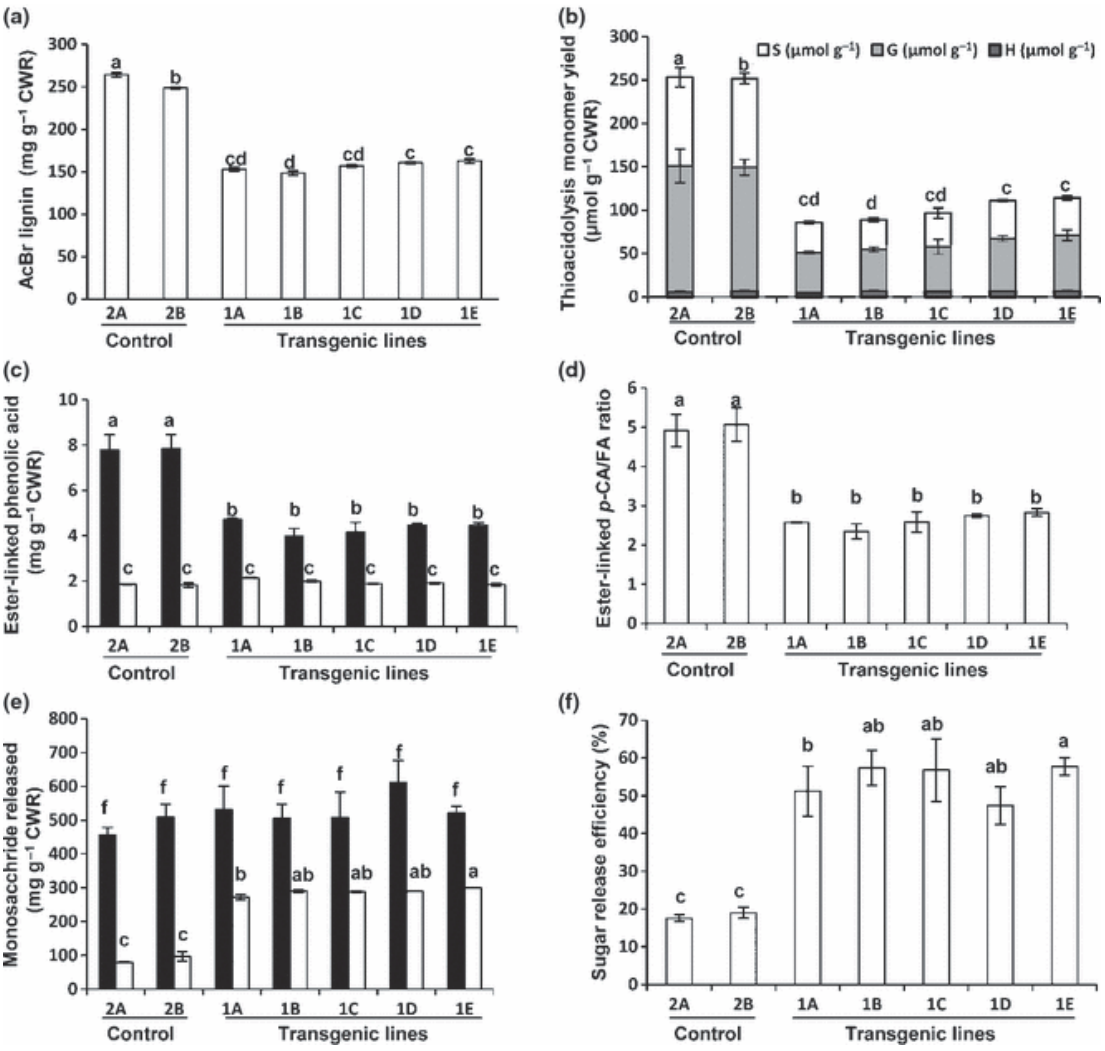


Figure 3-8 Overexpression of PvMYB4 alters phenylpropanoid metabolism in transgenic switchgrass (*Panicum virgatum*) plants.

(a) AcBr lignin content of whole E4 stems. (b) Lignin composition determined by thioacidolysis. (c) The content of ester-linked *p*-coumaric acid (*p*-CA, closed bars) and ferulic acid (FA, open bars) in switchgrass stems. CWR, cell wall residue. (d) *p*-CA : FA ratios of the ester-linked wall-bound phenolics. (e) Total monosaccharide released (closed bars) from cell wall residues and monosaccharide released from enzymatic saccharification without acid pretreatment (open bars). (f) Sugar release efficiency of enzymatic saccharification without acid pretreatment. Data are means ± SE (n = 3).

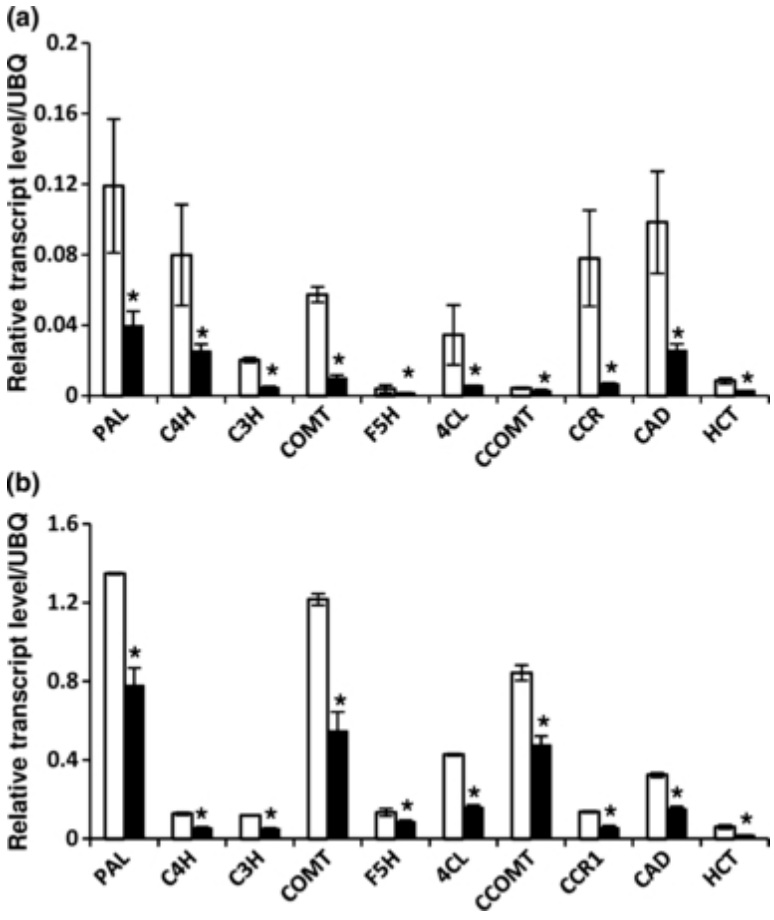


Figure 3-9 Identification of genes down-regulated by PvMYB4 in tobacco (*Nicotiana tabacum*) and switchgrass (*Panicum virgatum*)

(a) Quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis of the expression levels of lignin synthetic genes in PvMYB4-OX transgenic tobacco (closed bars). (b) qRT-PCR analysis of the expression levels of lignin synthetic genes in PvMYB4-OX transgenic switchgrass (closed bars). The CK value (open bars) represents the combined/average result from both control (2A and 2B) lines.

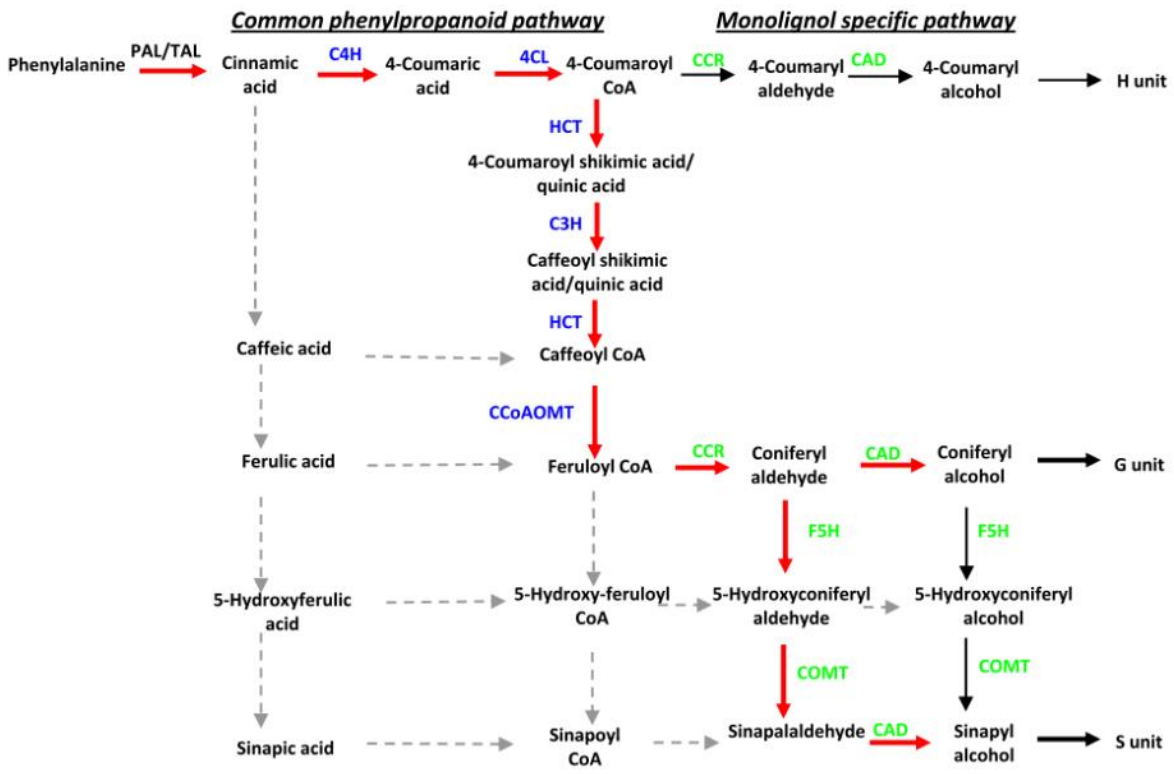


Figure 3-10 The phenylpropanoid and monolignol biosynthesis pathways.

Solid red and black arrows show reactions that are likely to occur in vivo, gray-colored arrows show hypothetical reactions, which are only detected by in vitro biochemical enzyme assay.

Figure 3-11 PvMYB4 gene variants in switchgrass

(a) ClustalW alignment of the amino acid sequences of PvMYB4a, 4b, 4c, 4d and 4e. Green box indicates the P-Q-rich region. (b) Different EST clones sequenced. The EST clone MYB_6497_S3 was used for making all the constructs reported in this work.

A

```
PvMYB4a MGRSPCCEKAHTNKGAWTKEEDDRLVAYIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60
PvMYB4c MGRSPFCEKAHTNKGAWTKEEDDRLVAYIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60
PvMYB4e MGRSPCCEKAHTNKGAWTKEEDDRLVAYIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60
PvMYB4d MGRSPCCEKAHTNKGAWTKEEDDRLVAYIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60
PvMYB4b MGRSPCCEKAHTNKCAWTKEEDDRLVAYIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60
        ***** ******** ******************************************

PvMYB4a LRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKNYWNTHIKRKLLSRG 120
PvMYB4c LRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKNYWNTHIKRKLLSRG 120
PvMYB4e LRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKNYWNTHIKRKLLSRG 120
PvMYB4d LRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKNYWNTHIKRKLLSRG 120
PvMYB4b LRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKNYWNTHIKRKLLSRG 120
        ************************************************************

PvMYB4a IDPVTHRPIADAARNVTISFQPDAPSQQQLSDDAEAPPPPPPPQQQQQQLKPPPHCPDLN 180
PvMYB4c IDPVTHRPIADAARNVTISFQPDAPSQQQLSDDAEAPPPPPPPQQQQQQLKPPPHCPDLN 180
PvMYB4e IDPVTHRPIADAARNVTISFQPDAPSQQQLSDDAEAPPPPP--QQQQQQLKPPPHCPDLN 178
PvMYB4d IDPVTHRPIADAARNVTISFQPDAPSQQQLSDDAEAPPPPPP--QQQQQLKPPPHCPDLN 178
PvMYB4b IDPVTHRPIADAARNVTISFQPDAPSQQQLSDDAEAPPPPPP--QQQQQLKPPPHCPDLN 178
        ************************************   ****************

PvMYB4a LDLCISPPCHKEEEDQELIKPAAVKREMLQAGHGTLGLCFGCSLGLQKGAAGCTCSSNSH 240
PvMYB4c LDLCISPPCHKEEEDQELIKPAAVKREMLQAGHGTLGLCFGCSLGLQKGAAGCTCSSNSH 240
PvMYB4e LDLCISPPCHKEEEDQELVKPAAVKREMLQAGHGTLGLCFGCSLGLQKGAAGCTCSSNSH 238
PvMYB4d LDLCISPPCHKEEEDQELVKPAAVKREMLQAGHGTLGLCFGCSLGLQKGAAGCTCSSNSH 238
PvMYB4b LDLCISPPCHKEEEDQELVKPAAVKREMLQAGHGTLGLCFGCSLGLQKGAAGCTCGSNSH 238
        ******************:***********************************.****

PvMYB4a FLGLRVGMLLDFRGLEMK 258
PvMYB4c FLGLRVGMLLDFRGLEMK 258
PvMYB4e FLGLRVGMLLDFRGLEMK 256
PvMYB4d FLGLRVGMLLDFRGLEMK 256
PvMYB4b FLGLRVGMLLDFRGLEMK 256
        ******************
```

B

Gene name	Accession #	Different ESTs (variants)
PvMYB4a	JF299185	>7L-8-17 >7L-11 >7L-16-19 >7L-37-26
		>7L-6 >7S-31 >7S-39 >MYB_6497_S3
PvMYB4b	JF299186	>7L-14
PvMYB4c	JF299187	>7L-20-22
PvMYB4d	JF299188	>7L-25 >7S-41 >7S-46 >7L-21-2 >MYB_6497_S11 >MYB_6497_S8
PvMYB4e	JF299189	>7L-32

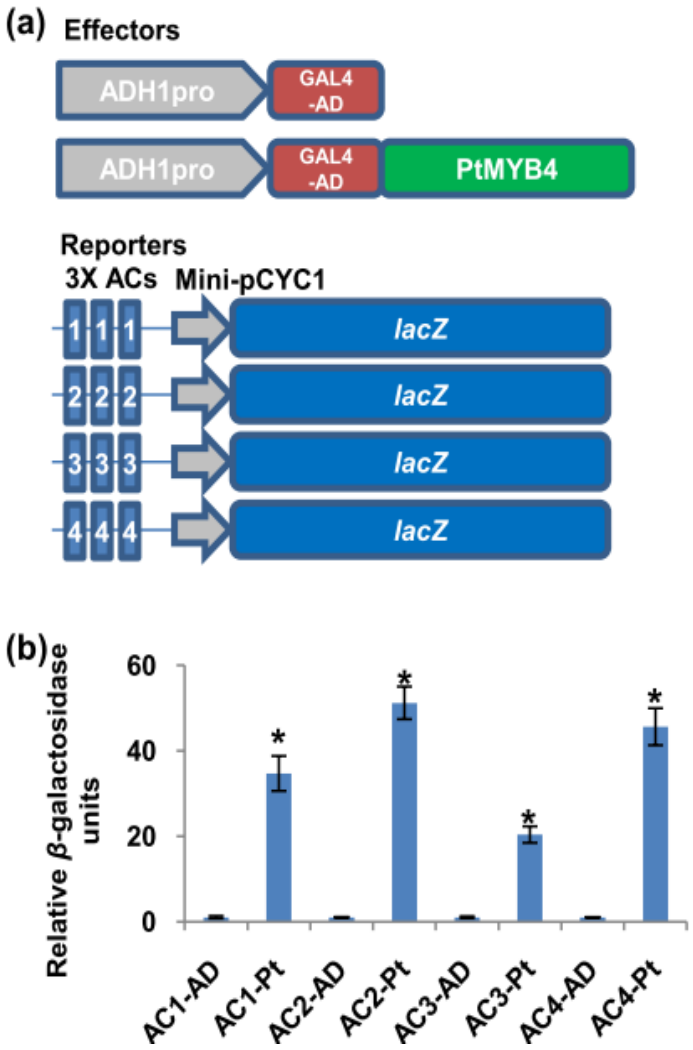


Figure 3-12 PtMYB4 binds to AC elements in yeast system.

(a) Effector and reporter plasmids for testing the binding of PtMYB4 to AC elements in yeast. (b) PtMYB4 can bind to AC elements in yeast to activate the lacZ reporter gene. AC (1,2,3,4) -AD and AC (1,2,3,4)-Pt refer to assays with reporters containing three consecutive copies of the AC I, II, III, or IV elements fused to the pCYC1 minimal promoter and lacZ (as shown in panel B with either GAL4AD or GAL4AD-PtMYB4 as effector). Data are means ±SE (n=4). Asterisks on top of the bars indicate values that were determined by the Student's t test to be significantly different from their equivalent control ($P < 0.05$).

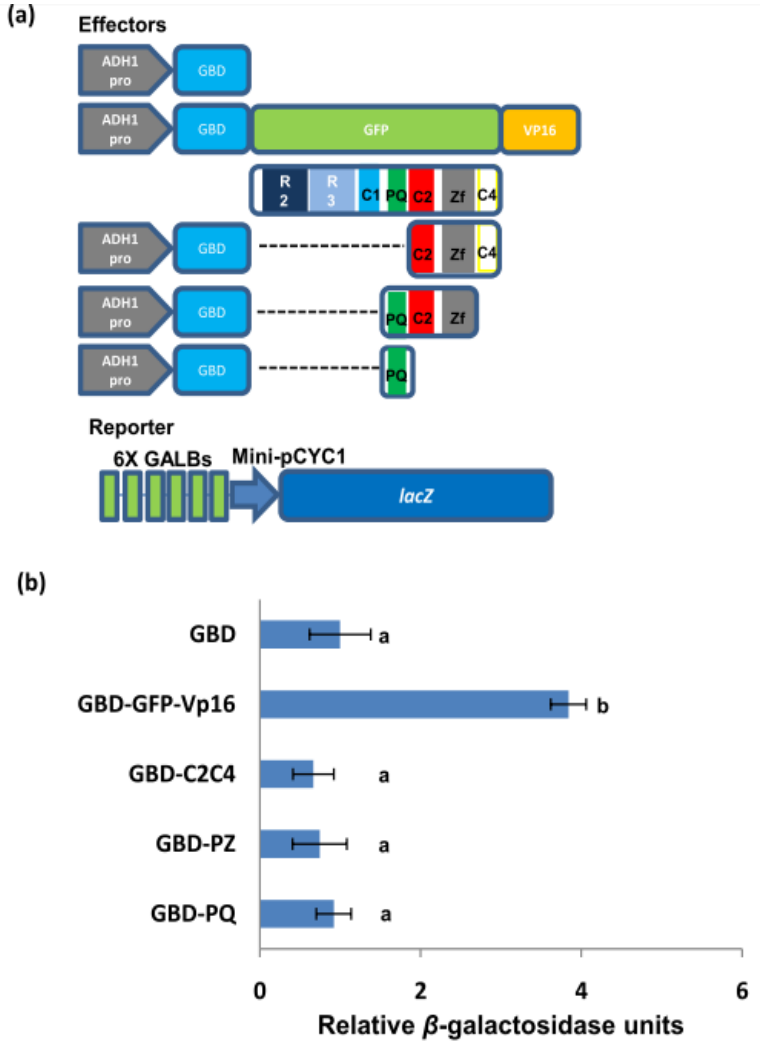


Figure 3-13 The PQ-rich motif has no transcriptional activation activity.

(a) Effectors and reporter constructs for testing the PQ-rich motif in yeast. GBD, GAL4 DNA binding domain; VP16, activation motif of the VP16 protein; GALBs, GAL4 protein binding sites. (b) β-Galactosidase assays showing activation of lacZ expression from the reporter 6GALBs-LacZ by effectors with different C-terminal motifs. Data were normalized to the vector control pGBT-9 (GBD), and are means ±SE (n=4). Different letters on the bars of figures indicate significant differences at $p \leq 0.05$ level as tested by Duncan grouping with SAS software (SAS Institute Inc., Cary, NC, USA).

Figure 3-14 Overexpression of PvMYB4 alters phenylpropanoid metabolism in transgenic tobacco plants.

(a, b) Quantification of the peaks in of HPLC chromatograms (Figure 3-5a and 3-5b) of methanol/water soluble fractions (a) and ester-linked wall bound phenolics (b) from stems of wild-type and PvMYB4-ox tobacco lines. The major soluble phenolic compound is chlorogenic acid (Rt=14.87 min). Wall bound ester-linked phenolic compounds: vanillin (Rt=19.63); syringyl aldehyde (Rt=22.59); p-coumaric acid (p-CA, Rt=23.78) and ferulic acid (FA, Rt=27.14) (c) Phloroglucinol-HCl staining (left panel) and Mäule staining (right panel) of cross-sections of the leaf petioles of wild-type and PvMYB4-OX tobacco transgenic plants.

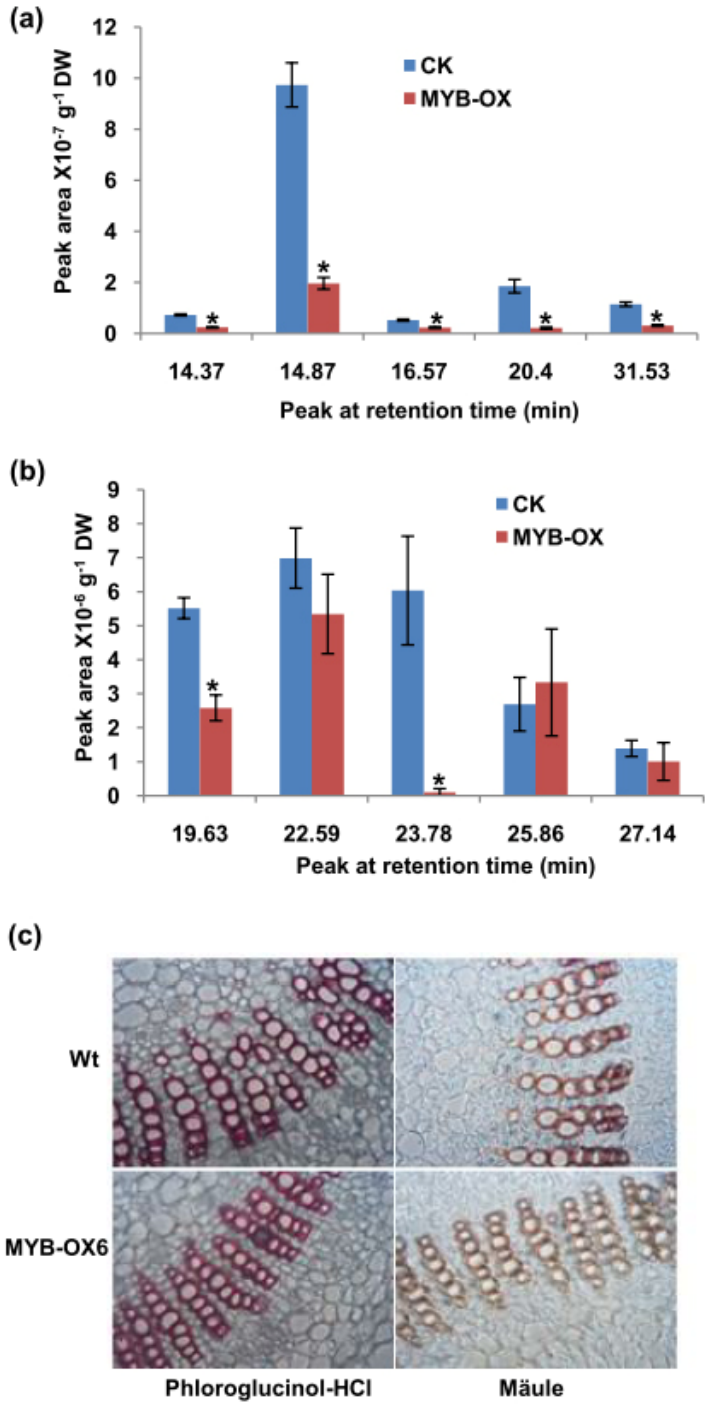
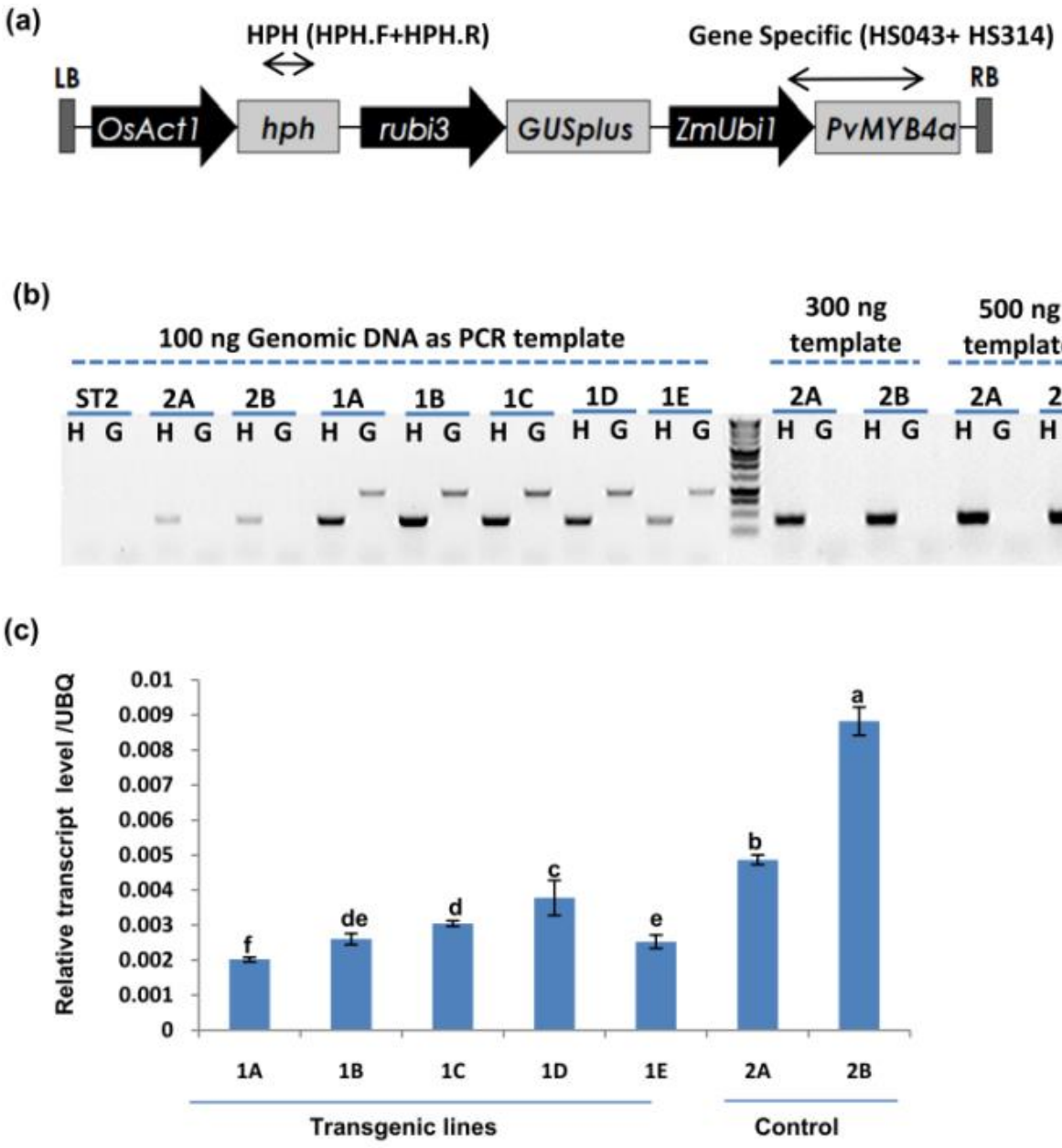

(a)
Peak area X10-7 g-1 DW
12
10
8
6
4
2
0
CK
MYB-OX
*
14.37
14.87
16.57
20.4
31.53
Peak at retention time (min)
(b)
Peak area X10-6 g-1 DW
9
8
7
6
5
4
3
2
1
0
CK
MYB-OX
*
19.63
22.59
23.78
25.86
27.14
Peak at retention time (min)
(c)
Wt
MYB-OX6
Phloroglucinol-HCl
Mäule

Figure 3-15 Genomic DNA PCR and qRT-PCR analysis of PvMYB4-OX transgenic

(a) The pANIC2B-PvMYB4 construct used for switchgrass transformation. OsAct1, promoter of rice Actin1 gene; rubi3, promoter of rubisco3 gene; ZmUbi1, promoter of maize ubiquitin 1 gene. hph, hygromycin resistance gene for transformation selection marker. LB and RB, left and right borders of the T-DNA sequences. The double arrows show the primer pairs used for genomic DNA PCR testing. HPH, hygromycin specific gene primer pairs; Gene specific, primer pairs covering the ZmUBI1 promoter region and PvMYB4 gene specific region. (b) Genomic DNA PCR. 2A and 2B are two control lines that only have the half T-DNA (HPH) part of the construct. (c) qRT-PCR analysis showing the endogenous PvMYB4 gene is down-regulated by the over-expression of PvMYB4 in switchgrass. Primer pairs are HS360 and HS361 (Table 3-1). Data are means ±SE (n=3).


(a)
HPH (HPH.F+HPH.R)
Gene Specific (HS043+ HS314)
LB
RB
OsAct1
hph
rubi3
GUSplus
ZmUbi1
PvMYB4a
(b)
100 ng Genomic DNA as PCR template
300 ng template
500 ng template
ST2
2A
2B
1A
1B
1C
1D
1E
2A
2B
2A
2B
H G
(c)
Relative transcript level /UBQ
0.01
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
f
de
d
c
e
b
a
1A
1B
1C
1D
1E
2A
2B
Transgenic lines
Control

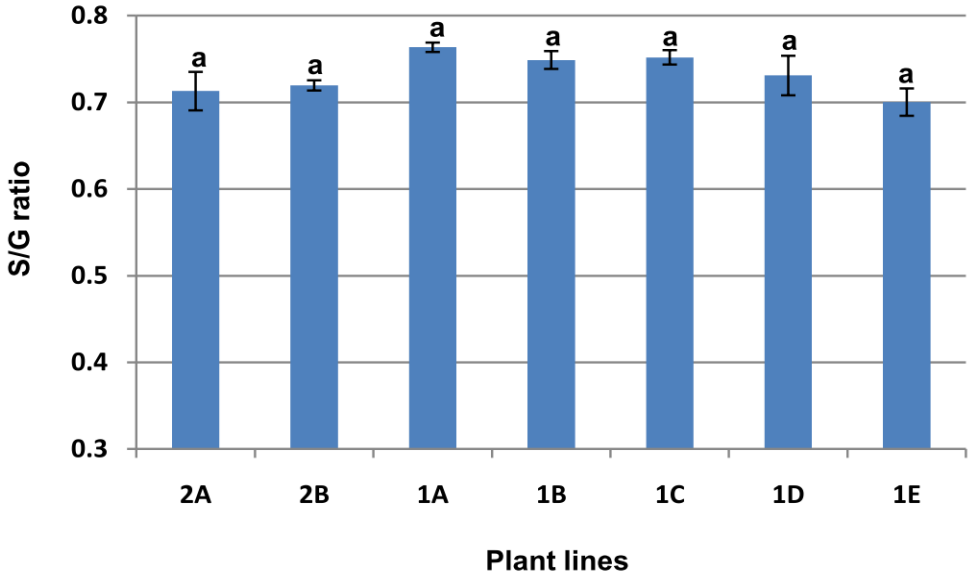


Figure 3-16 The S/G ratio of PvMYB4-OX transgenic switchgrass.

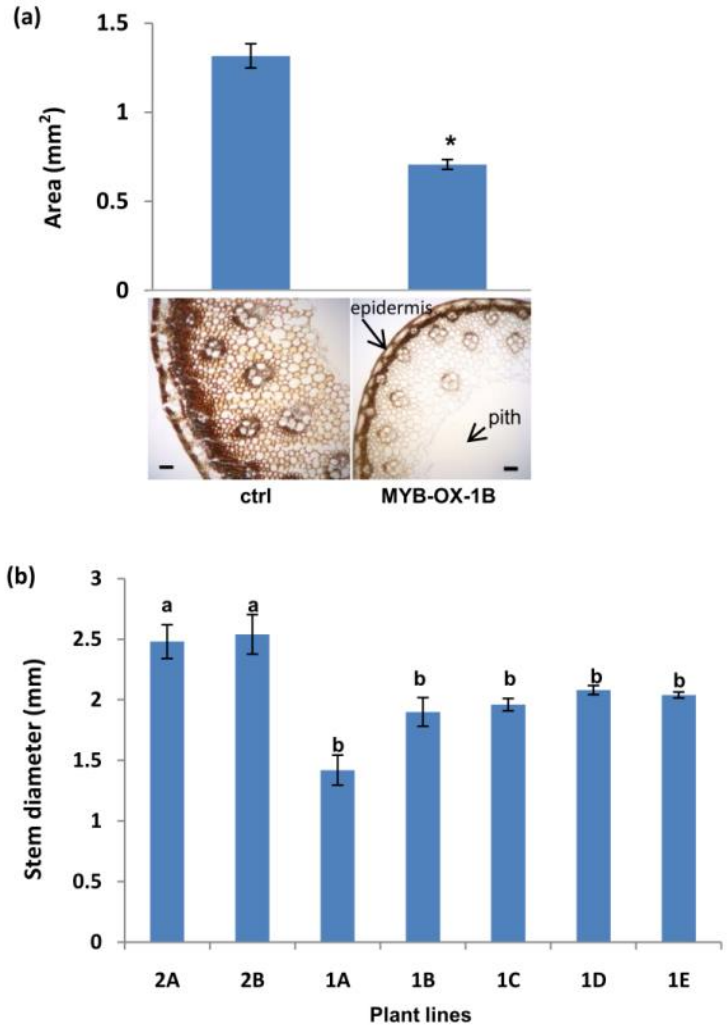


Figure 3-17 PvMYB4-OX transgenic switchgrass have smaller vascular bundles

(a) Vascular bundle areas. The measurement was done with ImageJ (http://rsbweb.nih.gov/ij/) software. Bars show standard errors of the means (n≥8).Only the vascular bundles close to the pith part within the I2 (E4 stage) were measured and compared. The vascular bundles that located in the interfascicular fiber area of the stem were not measured. Lower panel shows Mäule staining of cross-sections (40X) of internode 2 of control and PvMYB4-OX transgenic switchgrass. Bar=100 μm.(b) Tiller diameter. Only the middle parts of I2 internode of E4 stage tillers weremeasured with the venire scale. Data are means ±SE (n≥5).

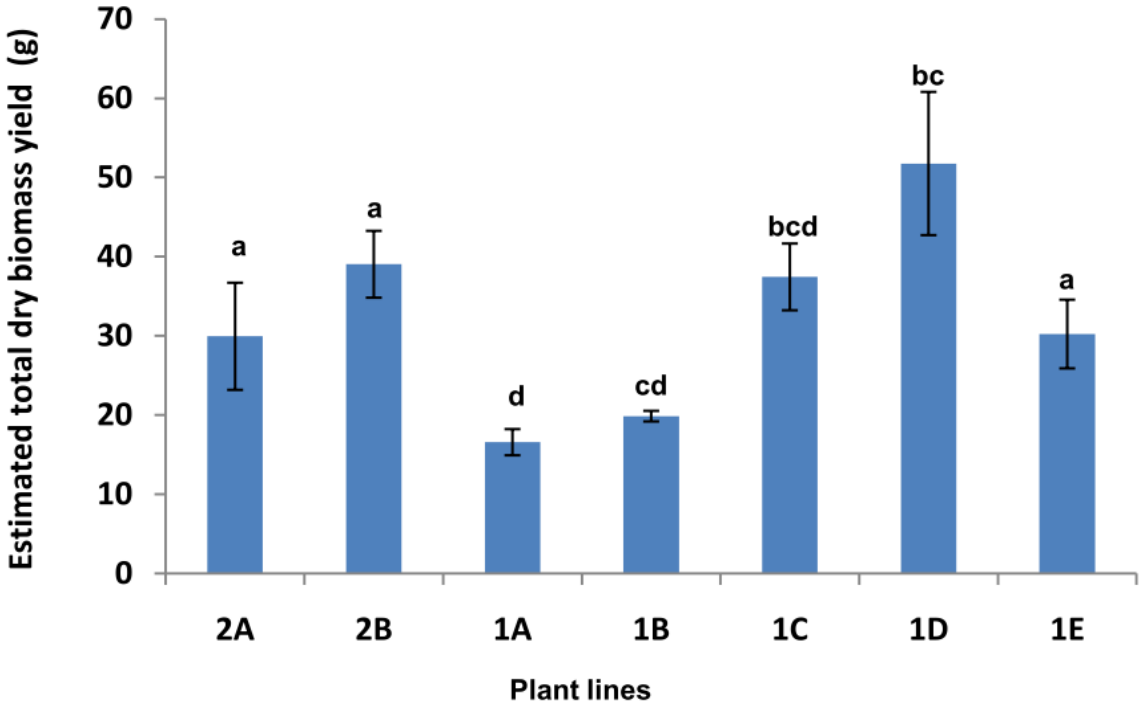


Figure 3-18 Estimation of total dry biomass yield for PvMYB4-OX transgenic

Only the tillers at E4 stage were measured. Total dry biomass was estimated by multiplying the average tiller dry biomass with the total tiller number within the same time period. Data are means ±SE (n≥3).

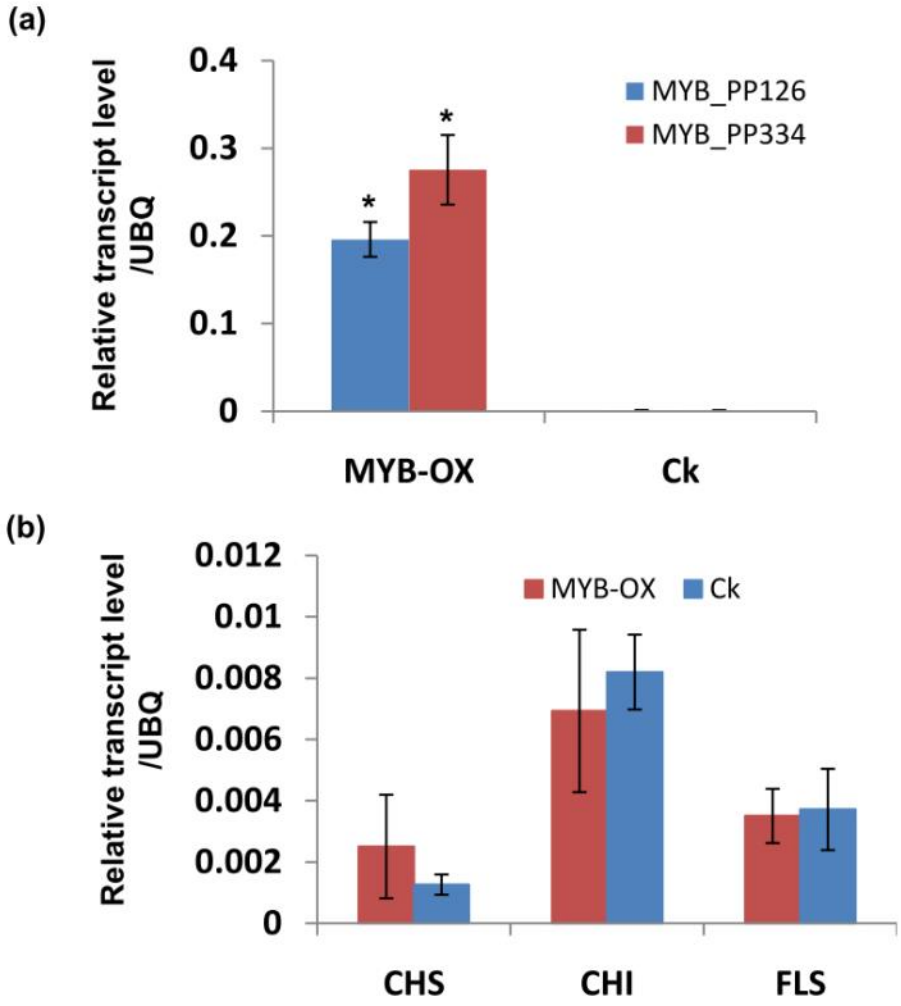


Figure 3-19 qRT-PCR analysis of flavonoid biosynthetic gene transcripts in PvMYB4-OX transgenic tobacco.

(a) qRT-PCR analysis showing the expression level of PvMYB4 in transgenic tobacco with two primer pairs. PP_126, primer pairs HS126 + HS127; PP_334, primer pairs HS334 +HS335. (b) qRT-PCR analysis showing the expression level of flavonoid biosynthetic genes in transgenic tobacco. Data are means ±SE (n=3).

Table 3-1 Sequences and references for the gene-specific primers used in this work

Primer Code	Primer name	Sequence	Reference/Note
		PvMYB4 cloning	
HS010	MYB1_6497.F	caccATGGGGCGGTCGCCGTGCTG	
HS011	MYB1_6497.R	TCAAACAAAAAAAAAACAGCCCAAC	
		Genomic DNA PCR	
HS043	Zmubi.pro.F2	TCTTTTGTCGATGCTCACC	Figure 3-15b
HS314	HS314.MYB1.R1	TCACTTCATCTCGAGGCCTC	Gene specific
	HPH.F	AAGGAATCGGTCAATACACTACATGG	
	HPH.R	AAGACCAATGCGGAGCATATACG	HygromycinB specific
		In Situ Hybridization	
	PvMYB_T7_F	GCGTAATACGACTCACTATAGGGTGCTGCGAGAAGGCGCAC	
	PvMYB_T7_R	GCGTAATACGACTCACTATAGGGTCATCTCGAGGCCTCTGAAG	
	PvMYB_F	TGCTGCGAGAAGGCGCAC	
	PvMYB_R	TCATCTCGAGGCCTCTGAAG	
		PvMYB4 qRT-PCR in switchgrass	
HS543	PvMYB4.F	TCGGCATGCTCCTCGACTTC	Figure 3-6b

Table 3-1 continued

Primer Code	Primer name	Sequence	Reference/Note
HS544	AcV5.R	ACCAGCCGCTCGCATCTTTC	
HS360	PvMYB1_798F	AGGCCTCGAGATGAAGTGAAAC	Figure 3-15c
HS361	PvMYB1_859R	AGCCCAACAAACAAACGAAATT	
HS181	PAL_F	CATATAGTGTGCGTGCGTGTGT	Figure 3-9b
HS182	PAL_R	CTGGCCCGCCAATCG	
HS090	C4H1_1534F	GGGCAGTTCAGCAACCAGAT	
HS091	C4H1_1611R	CGCGTTTCCGGGACTCTAG	
HS094	C3H1_739F	TTGAGATGGTTGTGTCCGCTTA	
HS095	C3H1_803R	AGGCGGTCCCTTCTCTCATT	
	PvCOMT_F461	CAACCGCGTGTTCAACGA	
	PvCOMT_R534	CGGTGTAGAACTCGAGCAGCTT	
HS108	F5H_1720F	CTCTTCTATTTGTGCGTGTAACTGTGT	
HS109	F5H_1792R	CAGCCCTATAGCATCGACATGA	
	4CL1_1179_F	CGAGCAGATCATGAAAGGTTACC	
	4CL1_1251_R	CAGCCAGCCGTCCTTGTC	

Table 3-1 continued

Primer Code	Primer name	Sequence	Reference/Note
HS100	CCOMT_966F	CCGTCTTTCTTTTTTTGGCTCTT	
HS101	CCOMT_1029R	GCATGAAAATGATGACAGTTTCCA	
	PvCCR1. 112_F	GCGTCGTGGCTCGTCAA	
	PvCCR1. 187_R	TCGGGTCATCTGGGTTCCT	
	PvCAD_F116	TCACATCAAGCATCCACCATCT	
	PvCAD_R184	GTTCTCGTGTCCGAGGTGTGT	
	HCT_973_F	GCAGAAGGAGCAGCAGTCATC	
	HCT_1035_R	CGAGCGGCAATAGTCGTTGT	
		qRT-PCR primers for PvMYB4-OX in transgenic tobacco	
HS126	MYB_7X_QRTF1	CTCAATCTCGACCTCTGCATCA	Figure 3-19a
HS127	MYB_7X_QRTR1	ACGAGCTCCTGGTCCTCTTCT	
HS334	MYBqRTR1	ATGAGCTCCTGGTCCTCTTCT	Figure 3-19a

Table 3-1 continued

Primer Code	Primer name	Sequence	Reference/Note
HS335	MYB.R3	GAGCTCCTGGTCCTCTTC	
	PAL.F	GACAAAGTGTTCACAGCAATG	Figure 3-9a
	PAL.R	TAACAGATWGGAAGAGGAGCA	
	C4H.F	TCAACACAATGGTGGAATGC	
	C4H.R	ACTTTGGGACGTTTGGTTCA	
	4CL.F	CTTCTCAACCATCCCAACATT	
	4CL.R	CTAACAACAAAAGCCACTGGA	
	HCT.F	GGCTGCCAATCCATGATGCT	
	HCT.R	GCAACAGATTGACTGCCATCA	
	C3H.F	TGGCTGAGGTGATCAAGAAC	
	C3H.R	TATGGGAGGTTGGGGAAGTC	
	CCoAOMT.F	ACACCCTATGGAATGGATCA	
	CCoAOMT.R	CCTTGTTGAGTTCCAATACGA	
	F5H.F	GAAACTCTACGACTTCACCC	
	F5H.R	TGACTTTGCCGGAATATGGT	

Table 3-1 continued

Primer Code	Primer name	Sequence	Reference/Note
	COMT.F	CCTGCAAATGGGAAGGTGAT	
	COMT.R	CAGTCCTTTCTTTGCCTCCT	
	CAD.F	CTCGGGAGAAAGAGCATCAC	
	CAD.R	CCTCTCCATTGCAGTGTTGA	
	CCR.F	ATGTGACGAAGCCAAGGGTAA	
	CCR.R	GTAGGAATTGGAAGGTGACCT	
	qCHS.F	CACCGCTGTCACGTTTCGT	Figure 3-19b
	qCHS.R	AGGGCTTGCCCAACTAAACTATC	
	qCHI.F	TGCAGAGAGTCAGGCCATTG	
	qCHI.R	CGGCGGGAAGGTTTCATT	
	qFLS.F	GCGAGAAGTTGCGGAGAAGA	
	qFLS.R	CATCATTTCATGGGCTTCTAACC	
		oligos used in EMSA	
HS428	AC1.F	CGGTACCAGTCCACCTACCGCCACCTACCGCCACCTACCGCTGTTCTCGA	Figure 3-3a
HS429	AC1.R	TCGAGAACAGCGGTAGGTGGCGGTAGGTGGCGGTAGGTGGACTGGTACCG	

Table 3-1 continued

Primer Code	Primer name	Sequence	Reference/Note
HS480	AC2.F	CGGTACCAGTCCACCAACCGCCACCAACCGCCACCAACCGCTGTTCTCGA	
HS481	AC2.R	TCGAGAACAGCGGTTGGTGGCGGTTGGTGGCGGTTGGTGGACTGGTACCG	
HS484	AC3.F	CCAGTCCACCTAACTCTACCTAACTCTACCTAACGCTGTTC	
HS485	AC3.R	GAACAGCGTTAGGTAGAGTTAGGTAGAGTTAGGTGGACTGG	
HS400	AC4.F	CCAGTCCACCAAACTCTACCAAACTCTACCAAACGCTGTTC	
HS401	AC4.R	GAACAGCGTTTGGTAGAGTTTGGTAGAGTTTGGTGGACTGG	

Chapter 4: Enhanced characteristics of genetically modified switchgrass (*Panicum virgatum* L.) for high biofuel production

(This chapter was published in Biotechnology for Biofuels 6, 71. with the following authors Hui Shen, Charleson Poovaiah, Angela Ziebell, Timothy J Tschaplinski, Sivakumar Pattathil, Erica Gjersing, Nancy L Engle, Rui Katahira, Yunqiao Pu, Robert Sykes, Fang Chen, Arthur J Ragauskas, Jonathan R Mielenz, Michael G Hahn, Mark Davis, C Neal Stewart Jr. and Richard A Dixon)

Charleson R Poovaiah's contribution was generation of transgenic switchgrass plants, designing and carrying out ethanol fermentation studies and writing the Quantitative saccharification, pretreatment and ethanol fermentation part of the methods and results section; reviewing the manuscript

4.1 Abstract

4.1.1 Background

Lignocellulosic biomass is one of the most promising renewable and clean energy resources to reduce greenhouse gas emissions and dependence on fossil fuels. However, the resistance to accessibility of sugars embedded in plant cell walls (so-called recalcitrance) is a major barrier to economically viable cellulosic ethanol production. A recent report from the US National Academy of Sciences indicated that, “absent technological breakthroughs”, it was unlikely that the US would meet the congressionally mandated renewable fuel standard of 35 billion gallons of ethanol-equivalent biofuels plus 1 billion gallons of biodiesel by 2022. We here describe the properties of switchgrass (*Panicum virgatum* L.) biomass that has been genetically engineered to increase the cellulosic ethanol yield by more than 2-fold.

4.1.2 Results

We have increased the cellulosic ethanol yield from switchgrass by 2.6-fold through overexpression of the transcription factor PvMYB4. This strategy reduces carbon deposition into lignin and phenolic fermentation inhibitors while maintaining the availability of potentially fermentable soluble sugars and pectic polysaccharides. Detailed biomass characterization analyses revealed that the levels and nature of phenolic acids embedded in the cell-wall, the lignin content and polymer size, lignin internal linkage levels, linkages between lignin and xylans/pectins, and levels of wall-bound fucose are all altered in PvMYB4-OX lines. Genetically engineered PvMYB4-OX switchgrass therefore provides a novel system for further understanding cell wall recalcitrance.

4.2 Conclusions

Our results have demonstrated that overexpression of PvMYB4, a general transcriptional repressor of the phenylpropanoid/lignin biosynthesis pathway; can lead to very high yield ethanol production through dramatic reduction of recalcitrance. MYB4-OX switchgrass is an excellent model system for understanding recalcitrance, and provides new germplasm for developing switchgrass cultivars as biomass feedstocks for biofuel production.

4.3 Introduction

Bioethanol from cellulosic feedstocks such as corn stover, switchgrass or wood chips, is a promising renewable and clean energy source, with the potential to reduce greenhouse gas emissions by up to 86% compared with gasoline (Wang et al., 2007). However, ethanol production from lignocellulosic materials faces more challenges than from starch-based feedstocks as a result of the chemical and physical barriers that block accessibility to the sugars (so-called recalcitrance) within the biomass. Pretreatment is required to partially deconstruct the biomass and open up surfaces for enzymatic hydrolysis to release 5- and 6-carbon sugars for fermentation. Pretreatment is not only expensive (Mosier et al., 2005), but also produces inhibitors of microbial ethanol fermentation such as 2-furaldehyde (furfural) and 5-hydroxymethylfurfural (HMF) during acidic pre-treatments (Liu et al., 2010).

Switchgrass has attractive features as a dedicated lignocellulosic feedstock for bioenergy production in the United States (McLaughlin et al., 2005, Schmer et al., 2008, Keshwani and Cheng 2009), and recent studies report partial success in overcoming recalcitrance. For example, down-regulation of cinnamyl alcohol dehydrogenase (CAD), the last enzyme of lignin precursor formation, increases saccharification efficiency up to 23% without acid pretreatment (Saathoff et al., 2011, Fu et al., 2011a). Likewise, down-regulation of caffeic acid 3-O-methyltransferase (COMT), a key enzyme for biosynthesis of the monolignol sinapyl alcohol, increases saccharification efficiency by 29-38% without acid pretreatment (Fu et al., 2011b). However, reduction of sinapyl monolignol production may increase concentrations of fermentation inhibitors (Klinke et al., 2004), and low molecular weight phenolic compounds in COMT down-regulated switchgrass inhibit simultaneous saccharification and fermentation (SSF) by the yeast Saccharomyces cerevisiae unless first removed by hot water pretreatment (Tschaplinski et al.,

2012). Clearly, a better strategy for reducing recalcitrance is required for the development of improved lignocellulosic bioenergy feedstocks.

Overexpression of the switchgrass R2-R3 MYB transcription factor PvMYB4 in switchgrass represses lignin biosynthetic pathway genes and increases saccharification efficiency up to 300% without acid pre-treatment (Shen et al., 2012). Here, we evaluate the bioconversion of such materials to ethanol using yeast-based SSF methods. Metabolite profiling revealed major reductions in levels of phenolic fermentation inhibitors. Furthermore, application of a suite of chemical, immunological, and physical approaches for cell wall characterization revealed that multiple components, including lignin and wall-bound phenolics, pectin lignin and xylan-lignin linkages, and fucosylated xyloglucans and rhamnogalacturonans, could potentially contribute to recalcitrance.

4.4 Materials and methods

4.4.1 Plant materials

Agrobacterium-mediated switchgrass transformation used constructs (Shen et al., 2012) and methods (Burris et al., 2009) described previously. The ST1 and ST2 lines were provided by Dr Zeng-Yu Wang, Noble Foundation. L7, L9 and L10 are transgenic control lines in the ST1 background. L1, L2, L4, L6 and L8 are MYB4-OX lines in the ST1 background. Lines 2A and 2B are vector controls for 1A-E (MYB4-OX) lines in the ST2 gene background.

All plants were grown under greenhouse conditions as described (Shen et al., 2009). Harvested tillers (at R1 stage) were either frozen and milled by a freezer mill (SPEX SamplePrep, Metuchen, NJ) in liquid nitrogen for genomic DNA or RNA isolation, or dried at 40°C for one

week then milled in a Thomas Wiley® Mini-Mill (Thomas Scientific, Swedesboro) through a 0.8 mm screen to 20 mesh for chemical analysis and ethanol fermentation tests. Samples for analysis of lignin content, wall-bound phenolics and solid-state NMR were further milled to 60 mesh size.

4.4.2 Measurement of lignin, phenolic and pectin content

Lignin content and composition of cell wall residues was determined by thioacidolysis followed by GC-MS as described previously (Shen et al., 2009). Soluble phenolics were extracted from 30 mg of freeze-dried tissue powder with 50% (v/v) methanol and assayed by HPLC, which reveals chlorogenic acid derivatives as the majorsoluble phenolics. Total soluble phenolic levels were assayed with Folin-Ciocalteu reagent, and wall-bound phenolics were analyzed as described previously (Shen et al., 2009).

For determination of pectin, plant material was ground in liquid N2, homogenized with 2 volumes of 80% ethanol, and incubated overnight at 4°C. The homogenate was centrifuged at 3,000 rpm for 5 min and the alcohol insoluble cell wall residue (AIR) washed twice with 20 ml of absolute ethanol and dried under N2. One hundred mg of AIR were extracted sequentially with water (20 ml, shaken overnight at room temperature), 0.05 M sodium acetate containing 0.04 M EDTA, pH 4.5 (20 ml, shaken for 4 h at room temperature) and 0.05 M HCl (20ml, incubated at 100°C for 1 h). Two hundred µl of supernatant from the different fractions was further hydrolyzed with 900 µl of H2SO4/sodium tetraborate reagent at 100°C for 5 min. The reaction was stopped on ice and the pectin content was determined by the m-hydroxydiphenyl method (Blumenkrantz and Asboe-Hansen, 1973) with galacturonic acid as standard.

4.4.3 Quantitative saccharification, pretreatment and ethanol fermentation

Quantitative saccharification assays were as described in ASTM E 1758–01 (ASTM 2003) and HPLC method NREL/TP 51–42623. Hot water pretreatment was conducted using the tubular batch method (Yang and Wyman, 2009), except only one sand bath (Omega FSB1, Techne Co., Princeton, NJ) was used to heat the 4□×□0.5 inch pretreatment tubes.

Simultaneous saccharification and fermentation (SSF) with *Saccharomyces cerevisiae* D5A (ATCC 200062) was performed as described in Fu et al. (2011a) with the exception that Accellerase 1500 enzyme (final concentration of 11.5 FPU per gram of cellulose), kindly provided by Genencor International, Inc., was used instead of Spezyme CP and Accellerase BG.

4.4.4 Solvent extraction of switchgrass biomass for solid-state NMR

Sequential extraction was performed as reported previously (Ziebell et al., 2010). Ester-linked wall-bound phenolics were extracted as described previously (Shen et al., 2009). The pellet residue was washed with water until the supernatant was neutral. The solids were then freeze-dried and weighed for solid-state NMR analysis.

4.4.5 Gel permeation chromatography (GPC) of lignin

Ball-milled lignin was isolated from extractives-free switchgrass as described previously (Björkman, 1956). The yields were 1.022% (1C), 1.361% (1D), 2.223% (2A) and 2.286% (2B). GPC: Isolated lignin samples were acetylated and GPC analysis performed using an Agilent HPLC with three polystyrene-divinyl benzene GPC columns (Polymer Laboratories, 300 × 7.5 mm, 10 μm beads) having nominal pore diameters of 104, 103, and 102 Å. The eluent was THF, the flow rate 1.0 ml/min, the sample concentration was ~2 mg/ml and an injection volume of 25

µl was used. The HPLC was attached to a diode array detector measuring absorbance at 260 nm (band width 40 nm). Polystyrene calibration standards were used with molecular weights ranging from 580 Da to 2.95 million Da. Toluene was used as the monomer calibration standard.

4.4.6 Solid, gel and solution-state NMR

Cross-polarization/magic angle spinning (CPMAS) spectra were collected as described previously (Ziebell et al., 2010) with slight modifications: A 7 mm ZiO2 rotor was loaded with approximately 75 mg of dried biomass ground to 60 mesh. CPMAS NMR spectra were collected on a Bruker DSX 200 spectrometer equipped with a 7 mm CPMAS probe and a 4.7 T magnet (1H□=□200.1 MHz and 13C□=□50.32 MHz). A ramped CP pulse with 1H and 13C fields matched at 48 kHz was applied with a contact pulse of 2 ms. An acquisition time of 0.051 s and a recycle delay of 1s were used with 2 k points collected and averaged over 40k scans for each spectrum with MAS□=□7 kHz.

Samples of whole biomass and isolated lignin were prepared for 2D gel state NMR by suspending 20–30 mg of material in 0.5 ml of DMSO-d6 in a 5 mm NMR tube. Samples were then sonicated for 2h (whole biomass) or 30 min- 1 h (isolated lignin).

Gel-state 1H-13C HSQC spectra were collected on a Bruker Avance III 600MHz spectrometer with a 5 mm TCI cyroprobe. HSQC spectra were acquired with a sweep width of 15 ppm, 1024 data points, and an acquisition time of 57 ms in the F2 dimension. For the F1 dimension a sweep width of 166 ppm was used with 256 increments. The recycle delay was set to 1.5 s and 128 scans were collected for each increment for a total experiment time of 14.5 h.

For 2D HSQC NMR spectral analysis, lignin samples were isolated according to modified literature methods (Björkman, 1956, Vanholme et al., 2010, Chang et al., 1975). In brief, 20 mesh switchgrass biomass was Soxhlet-extracted with benzene-ethanol (2:1, v/v) for 24 h to remove extractives. The extracted wall residue was then milled in a porcelain jar (1 l) with ceramic balls using a rotatory ball mill running at 96 rpm under nitrogen for 120 h. The ball milled powder was then suspended in 20 mM sodium acetate, pH 5.0. A mixture of Cellulysin cellulase (EC 3.2.1.4, Calbiochem, http://www.calbiochem.com), Cellobiase (Novozyme 188 from *A. niger*) and xylanase was added and the slurry incubated with shaking at 200 rpm and 37°C for 48 h. The digested cell wall fractions were then extracted twice with dioxane-water (96:4, v/v) under stirring for 24 h. The extract was centrifuged and the supernatant evaporated under reduced pressure, and freeze-dried. The resulting crude lignin-enriched samples were washed with deionized water and purified by liquid-liquid extraction (Björkman, 1956) for NMR characterization.

4.4.7 Glycome profiling

Glycome profiling was carried out by enzyme-linked immunosorbent assays of cell wall extracts using a large collection of plant glycan-directed monoclonal antibodies (http://www.wallmabdb.net) as described previously (Pattathil et al., 2010, 2012) (Table 4-5).

4.4.8 Metabolite profiling

Metabolite profiling of methanol extracts was performed as reported previously (Tschaplinski et al., 2012) with modifications: Ten ml of the extracts were dried under nitrogen. Sorbitol (15 μg) was added as internal standard, and the extracts were silylated for 2 days as described previously (Tschaplinski et al., 2012), and 0.5 μl of the 1-ml reaction volume was analyzed by GC-MS.

4.4.9 Statistical analyses

Metabolite data were averaged by control and PvMYB4-OX lines. Five biological replicates were analyzed for the PvMYB4-OX line and two for the control line, and two technical replicates were averaged for each sample. p-values were determined by Student's t-test (Microsoft Office Excel 2007) and $p < 0.05$ (indicated by asterisks in figures) considered as indicating significant differences. Multiple comparisons were done with SAS software (SAS Institute Inc., Cary, NC). Tukey's honestly significant difference test was used when the null hypothesis was rejected ($p < 0.05$). Means with the same letter, within each variable, are not significantly different at $p < 0.05$.

4.5 Results and discussion

4.5.1 PvMYB4 overexpression in switchgrass

Previously generated PvMYB4-over-expressing (PvMYB4-OX) transgenic switchgrass lines (1A, 1B, 1C, 1D, 1E, 2A and 2B) were in the Alamo ST2 genetic background (Shen et al., 2012), and additional lines were constructed in Alamo ST1 (Figure 4-6). Nine regenerated plants were selected from independent antibiotic resistant calli, and six lines (L1, L2, L4, L6, L8 and L11) were confirmed to be transgene positive by genomic DNA PCR (Figure 4-6b). The PvMYB4 expression level was determined by qRT-PCR analysis (Figure 4-6c). Lines L6 and L8 showed intermediate expression level compared to lines L1, L2, L4 and L11. Overexpression of PvMYB4 repressed endogenous PvMYB4 expression, indicating a negative self-regulatory mechanism (Figure 4-6d). Adult PvMYB4-OX plants showed reduced tiller height and tiller diameter, but increased tiller numbers in both genetic backgrounds under greenhouse conditions (Shen et al., 2012), (Figure 4-6e). Whole tillers (comprised of approximately 48% leaves and

52% stems on a weight basis for both control and transgenic materials) were used in all the following experiments as these represent the material that would be processed in a biorefinery. All materials were harvested at the same developmental stage (R1) according to a recently published protocol designed to facilitate comparisons between transgenic and control switchgrass materials (Hardin et al., 2013).

4.5.2 PvMYB4-OX lines exhibit up to a 2.6-fold increase in ethanol yield

Ethanol yields of control and PvMYB4-OX switchgrass were first assessed by weight loss during yeast-based SSF with or without hot-water pretreatment (Figure 4-1a, b). PvMYB4-OX biomass underwent a faster hydrolysis of cellulose to glucose and faster conversion of the glucose to ethanol and CO2 under both pretreated and non-pretreated conditions than did control material. After 7 days fermentation, the ethanol yield per gram of cellulose or biomass was about 2.6-fold higher for the MYB4-OX lines than the control lines under non-pretreated conditions (Figure 4-1c, d). After hot water pretreatment, the ethanol yield increased significantly in both control and MYB4-OX lines. However, untreated MYB4-OX transgenic biomass had a similar ethanol yield to pretreated control biomass (Figure 4-1c, d). HPLC (High-performance liquid chromatography) analysis indicated that only 0.077 to 0.175 mg glucose per gram of dry biomass was left in the fermentation medium, and no furfural or HMF were detected. The SSF ethanol yield without pretreatment showed a strong positive correlation ($R^2 > 0.8$) with the expression level of PvMYB4 (Figure 4-1e, f). PvMYB4-OX switchgrass produces approximately 1.8-fold more ethanol than COMT-RNAi switchgrass (9) under the same conditions (Figure 4-1g, h).

4.5.3 PvMYB4-OX switchgrass has reduced levels of phenolic fermentation inhibitors

To check for the fermentation inhibitors reported in COMT-RNAi switchgrass (Tschaplinski et al., 2012), we performed metabolite profiling of methanol extractives using GC-MS (Table 4-1). Over 160 peaks were examined. Levels of phenolic fermentation inhibitors such as aromatic aldehydes (p-hydroxybenzaldehyde and coniferaldehyde) and organic acids (p-coumaric (p-CA), ferulic (FA) and sinapic acids) were significantly reduced in the PvMYB4-OX material (Table 4-1). Levels of various lignans were either reduced or increased in PvMYB4-OX lines (Table 4-1). The potential of lignans as fermentation inhibitors is unclear and needs further evaluation.

The content of soluble phenolics extracted by 50% methanol from the whole biomass of PvMYB4-OX lines was reduced by about 10-20% compared to controls (Figure 4-7a). Levels of the monolignols coniferyl alcohol, sinapyl alcohol and its glucoside syringin, and 5-hydroxyconiferyl alcohol were all reduced. Levels of feruloylquinic acid esters declined, whereas levels of caffeoylquinic acid esters were unchanged. Levels of the soluble sugars glucose, fructose, galactose and raffinose were increased in the methanol extractives of MYB4-OX lines, by from 1.6- to 3.5-fold. These increases in monosaccharides, with sucrose unchanged, suggest active production of raffinose (galactose addition to sucrose via galactinol), a storage carbohydrate which accumulated. More uronic acids (2.5-fold) and amino acids (glutamine, tyrosine, alanine, γ-aminobutyric acid) were also found in the MYB4-OX methanol extractives (Table 4-1). Accumulation of most of the soluble sugars measured, coupled with the decline in monolignols, related upstream precursors, downstream lignans, and reduced lignin content with the overexpression of PvMYB4 suggests altered partitioning of carbon away from the lignin pathway (secondary metabolism), consequently benefitting primary metabolism.

4.5.4 Changes in cell wall components in PvMYB4-OX switchgrass

PvMYB4-OX switchgrass transgenic lines have thinner stems with smaller vascular bundles (Shen et al., 2012), although there were no obvious differences in stem structure. The cell walls appeared to be thicker in the control lines based on staining of stem sections (Shen et al., 2012). We measured the wall thickness of the parenchyma cells in the mature stem sections (E4I1 internode); the value for control plants was 4.21 ± 0.52 μm, compared with 1.85 ± 0.50 μm for the PvMYB4-OX transgenics (Student t-test E-value p = 6.0E-20).

After removal of the methanol extractives, the wall-bound (ester- and ether-linked) phenolics in the cell wall residues were released by successive hydrolysis in 2 M NaOH at 37°C for 5 h, and 4 M NaOH with autoclaving for 2 h, respectively, and were then measured by HPLC. Levels of total wall-bound p-CA and FA, and ester-bound and ether-bound p-CA, were reduced in all PvMYB4-OX lines except for L6. There was also a slight reduction in ether-bound FA content in PvMYB4-OX lines in the ST2 background (Figure 4-2a). No changes were observed for ester-bound FA. Thus, both ester-bound and ether-bound p-CA/FA ratios were significantly reduced in lines highly over-expressing MYB4 (Figure 4-7b).

Total lignin thioacidolysis yields were reduced by about 50% in L1 and about 20% in L6 ST1 lines (Figure 4-2b). The SSF ethanol yield, without pretreatment, showed a strong negative correlation with total lignin content (R^2 =0.77) (Figure 4-2c), total wall-bound p-CA (R^2 =0.85), ester-bound p-CA (R^2 =0.86), ether-bound p-CA (R^2 =0.75) (Figure 4-2d) and ether-bound FA (R^2 =0.81) (Figure 4-2e), and a weak negative correlation with ester-bound p-CA/FA ratio (R^2 =0.52) as well as ether-bound p-CA/FA ratio (R^2 =0.60) (Figure 4-2f).

Biomass was extracted with methanol followed by 0.1 M NaOH at 4°C, or with methanol and chloroform followed by 2.0 M NaOH at 37°C to release the ester-linked wall-bound phenolics. Both methanol and alkaline hydrolysis removed more mass from PvMYB4-OX lines than from controls (Figure 4-2a). Extractive-free cell-wall residues were then characterized by solid-state 13C CP/MAS NMR (nuclear magnetic resonance) spectroscopy. Two control (2A and 2B) and five transgenic lines (1A, 1B, 1C, 1Dand 1E) from ST2 were analyzed and gave similar results; however, data are shown for only for 2A and 1C for figure clarity (Figure 4-3b-d). The lignin and aromatic region was assigned to the 110–165 ppm region, which reflects free monomers or wall attached lignin and hydroxycinnamate conjugates; this region was remarkably reduced in the PvMYB4-OX lines (Figure 4-3b-d), in agreement with the chemical analysis. Compared to the methanol extraction, which removes non-wall-attached phenolics and other free compounds (Figure 4-3b), the following dilute base extraction (Figure 4-3c) removed most (if not all) of the sugar acetylation (largely from carboxyl groups of hemicelluloses) as shown by the almost complete disappearance of the hemicellulose acetyl carbonyl signal (168–177 ppm). This then revealed a clear difference in the transgenic versus control line in the region of 162–170 ppm (peak 165 ppm) (Figure 4-3c). The 162–170 ppm region can be seen as a slight shoulder in the control after methanol extraction (Figure 4-3b), and corresponds to Cγ□=□O side chain or the C4 of the phenyl ring of wall-bound hydroxycinnamates. The reduced signal in this region in PvMYB4-OX lines indicates a decrease of ester-bound hydroxycinnamates.

An obvious signal reduction in the aromatic regions of solid-state NMR spectra of methanol extracted and dilute base extracted residues was observed for MYB4-OX lines (Figure 4-8). The region 146–153 ppm is assigned to the C3 of both mono and di-methoxylated aromatic rings, and also to the C4 from mono-methoxylated aromatics and C5 from di-methoxylated aromatics. The

region at 125–135 ppm can be largely attributed to aromatic carbons which have a carbon attached, and the alpha and beta carbons on the propenyl side chains of the aromatic ring. Decreased signals in these two regions are possibly due to de-esterification of lignin or cinnamaldehydes in the PvMYB4-OX lines.

There were small changes in total sugar content of PvMYB4-OX whole biomass (Table 4-2). The major monosaccharides released by acid hydrolysis were glucose, xylose and arabinose, which represent, respectively, about 60%, 32% and 4-5% of the total sugars of the whole biomass. There were no significant differences in total sugar content of cell-wall residues from PvMYB4-OX and control lines after removal of soluble sugars and starch from whole biomass (Table 4-3). About 7–9 mg of total pectin was extracted per gram of alcohol insoluble cell wall residue. Only 25% of this was extractable by water and sodium acetate/EDTA solution, and about 85% (wall-bound pectin) was released by 0.1M HCl at 100°C for 1 h (Table 4-4). More pectin was released from MYB4-OX lines than from controls (Table 4-4). Thus, down-regulation of lignin content in PvMYB4-OX lines leads to increased soluble and wall-bound pectins in the cell walls (Table 4-4).

4.5.5 Overexpression of PvMYB4 reduces lignin size and internal linkages

Lignin molecular weight reduction is linked to reduced recalcitrance of low lignin alfalfa (Medicago sativa) (Ziebell et al., 2010). Isolated lignins were prepared from two control and two MYB4-OX lines, and their molecular weights measured by gel permeation chromatography (GPC) (Figure 4-9). The average molecular weights of the isolated lignins were lower in the PvMYB4-OX lines (1C and 1D), 4,400-4,900 Da as compared to 5,300-5,500 Da in control lines

(2A and 2B). These changes are much smaller than reported in low lignin alfalfa (Ziebell et al., 2010).

To check lignin inter-unit linkages, isolated lignins were analyzed by 2D heteronuclear single quantum coherence (HSQC) NMR based on two-dimensional chemical shifts of protons and carbon linkages. Spectra were collected on controls 2A and 2B and transgenic lines 1C and 1D but data are only presented for 2A and 1C for simplicity and consistency with other figures. The aromatic regions of the 13C-1H HSQC spectra showed no significant differences in aromatic C-H correlations between PvMYB4-OX and control, indicating that the basic monolignol constituents of the lignins are the same (Figure 4-10). However, in the aliphatic regions of the spectra, the C-H correlations of the lignin side chains in β-β linkages (resinols) were decreased in PvMYB4-OX lines, whereas the other two major linkages, β-O-4 and β-5 (phenylcoumaran) were relatively unchanged (Figure 4-4a). Gel-state 2D HSQC NMR spectroscopy also revealed that PvMYB4-OX lines have about a five-fold higher level of fucose residues in the cell walls (Figure 4-4b).

4.5.6 Reduced association of xylan and pectins with lignin in PvMYB4-OX switchgrass

We conducted glycome profiling analyses (Pattathil et al., 2012) of sequential extractions of PvMYB4-OX and control cell wall residues to assess the strengths of association of various polysaccharide polymers within the cell walls (Figure 4-11). Oxalate and carbonate remove mainly pectins and a small portion of hemicellulose from the walls, and the following 1M and 4M KOH treatments extract most of the tightly bound hemicelluloses and pectin. Chlorite removes a significant portion of the lignin, and the post chlorite 4 M KOH extraction releases additional hemicellulose and pectin components. The extracts were then screened by enzyme-

linked immunosorbent assay using a comprehensive set of plant glycan-directed monoclonal antibodies (mAbs) (Table 4-5) that recognize diverse epitopes on most of the major plant cell wall polysaccharides ((Pattathil et al., 2010, 2012). The glycome profiles (represented as heatmaps) were largely similar for control and PvMYB4-OX lines, the most notable differences being in the chlorite and post-chlorite 4M KOH extracts (Figure 4-11). Increased binding intensities in the 4M KOH PC extracts of MYB4-OX lines were for mAbs that recognize pectic arabinogalactan (RG-I/AG) epitopes (Figure 4-11, white boxes and Figure 4-5a). This suggests that lignin in the wild-type secondary cell wall blocks accessibility to such pectic polysaccharides, which are traditionally considered as major components of the primary cell wall and middle lamella. Alternatively, more RG-I/AG polysaccharides are present in the walls of PvMYB4-OX lines. Reduced signals for the pectic RG-Ic, RG-I and HG backbone groups of antibodies were observed in the chlorite extracts from MYB4-OX lines (Figure 4-11, blue boxes and Figure 4-5b). These data reveal associations of pectic polysaccharides with lignin in switchgrass, and the reduced lignin level in PvMYB4-OX lines potentially reduces such associations.

Less xylan epitopes were released during the chlorite extraction in the MYB4-OX lines (Figure 4-11, yellow boxes and Figure 4-5c), suggesting less xylan cross-linking/association with lignin. The chlorite treatment is unlikely to affect other wall components, and hence the release of the carbohydrates by this treatment will arise due to the destruction of the lignin component that ties these epitopes into the wall. Note that only a subfraction of these epitopes is released in the chlorite step; other subfractions of these polymers are not tied to lignin and are released in other extracts.

The chlorite extract of the PvMYB4-OX lines also showed increased binding to the mAb BG-1, which is specific for hemicellulosic β-1,3-1,4 glucan (Meikle et al., 1994) (Figure 4-11, yellow boxes and Figure 4-5c). An increase of fucosylated xyloglucan signal was also revealed by the binding of antibodies such as CCRC-M1, CCRC-M102, and CCRC-M106 in both the 4 M KOH and 4 M KOH PC extractives of the MYB4-OX lines (Figure 4-11, green boxes; Figure 4-5d). The fucose in xyloglucan is (as far as is known to date) located in the terminal position on the side-chains (Hoffman et al., 2005) and the antibodies are specific for the fucose in that position (Puhlmann et al., 1994) thus, it is likely that the antibodies are indeed detecting an increase in fucose level rather than an unmasking of the epitope. This is consistent with the increased wall-associated fucose observed by NMR analysis being due to increased fucosylated xyloglucans in PvMYB4-OX cell walls.

4.5.7 An improved system for high bioethanol production

Overexpression of PvMYB4 reduces the lignin content of switchgrass by 60-70% and increases sugar release efficiency approximately 3-fold without acid pretreatment (Shen et al., 2012). This translates into a 2.6-fold increase in ethanol yield using yeast-based SSF without pretreatment. PvMYB4-OX switchgrass produces approximately 1.8-fold more ethanol than COMT-RNAi switchgrass (Fu et al., 2011b) under the same fermentation conditions. The COMT-RNAi transgenic lines require only 25-30% the level of cellulase for equivalent ethanol fermentation compared to control switchgrass, with an estimated cost reduction for biomass processing of 21-25% for enzyme alone after excluding biomass and capital charges (Fu et al., 2011b). Based on the same calculations, PvMYB4-OX lines could save up to 45% of enzyme costs alone. Without a consolidated bioprocessing fermentation method, the minimum ethanol selling price (MESP) from switchgrass feedstock is $1.42-2.91 /gallon (Humbird et al., 2011). The estimated enzyme

cost savings from use of PvMYB4-OX transgenic switchgrass will give 0.78−1.60/gallon MESP which essentially meets the US Department of Energy′s 1.07/gallon target for 2012.

Fermentation inhibition by low molecular-weight compounds is a critical concern when processing lignin down-regulated biomass (Clark and Mackie, 1984). Increased levels of phenolic aldehydes and acids contribute to inhibition of microbial growth during fermentation of COMT-RNAi switchgrass (Tschaplinski et al., 2012). In contrast, levels of monolignols, phenolic aldehydes, and phenolic acids are all reduced in PvMYB4-OX switchgrass lines, consistent with the improved yeast-based SSF fermentation results.

4.5.8 New insights into recalcitrance of lignocellulosic feedstocks

Multiple factors may contribute to the recalcitrance of lignocellulosic feedstocks towards chemical treatments and/or enzymes, many of which are related to the presence of lignin in cell walls (Himmel et al., 2007). SSF ethanol yields negatively correlate with total lignin content, wall-bound p-CA (both ester-linked and ether-linked), ether-linked FA, and ester-linked p-CA/FA ratio in switchgrass. A decreased ester-linked p-CA/FA ratio has been associated with increased forage digestibility in barley (Du et al., 2009) and increased sugar release efficiency in switchgrass (Shen et al., 2009). FA serves as a bridge between lignin and hemicellulose (Hatfield and Chaptman, 2009) and ferulate-arabinoxylan esters can form ether linkages with lignin polymers (Ralph, 2010). The reduced level of ether-bound FA in MYB4-OX switchgrass suggests a looser wall association between lignin and arabinoxylans, as confirmed by extractability studies and glycome profiling. Reduced lignification or ferulate-lignin cross-linking are also important for improved fiber fermentability in maize suspension cells (Grabber

et al., 2009). Overall, our data suggest that reduced lignin content, polymer size and changes in inter-unit linkages all likely contribute to the reduced recalcitrance of PvMYB4-OX lines.

Fewer pectic epitopes (RG-Ic, RG-I backbone and HG backbone-1 groups) are released from PvMYB4-OX wall residues during chlorite extraction. This suggests that specific sub-populations of these pectic polysaccharides may directly link/associate with lignin. Older literature suggests that pectic arabinogalactans can be removed concurrently with lignin during the delignification of lupin by chemical treatments (Monro et al.,1972, Selvendran et al., 1975). A study in alfalfa suggested that the deposition and distribution of pectin corresponded to the deposition patterns of lignin in the middle lamella (Wi et al., 2005), where much of the pectin in the cell wall is located and lignification is initiated (Donaldson, 2001). A recent study also suggests the presence of critical associations between lignin and pectins in Populus biomass, where hydrothermal pretreatment disrupts lignin-polysaccharide interactions together with a loss of pectins and arabinogalactans (DeMartini et al., 2011). Although a pectin-hemicelluose-cellulose network has been widely accepted, direct lignin-pectin linkages/interactions should be further investigated in view of their potential contribution to recalcitrance.

Lignin and wall-bound phenolics are not the only factors impacting recalcitrance in switchgrass. Glycome profiling and NMR revealed increased levels of wall-associated fucose, possibly in fucosylated xyloglucans, in PvMYB4-OX lines. Fucosylated cell wall components in plants include glycolipids, O- and N-glycoproteins and polysaccharides such as xyloglucans and rhamnogalacturonans (RG). The glycolipids will be removed by methanol extraction and thus do not contribute to the fucose measured in the present study. Cell wall glycoproteins can form ether and aryl linkages through tyrosine, lysine and sulfur-containing amino acids with hydroxycinnamic acids esterified to polysaccharides in the cell wall. The fucosyl residues in RG-

II and xyloglucan are important for the strength of load-bearing elements in cell walls (Ryden et al., 2003, Campbell and Braam, 1999). Fucosylated xyloglucans are thought to have interconnections with the cellulosic matrix (Levy et al., 1997), and in vitro binding assays and computer modeling suggest that the fucosyl groups of xyloglucan may stabilize a xyloglucan conformation and help the polysaccharide to bind more tightly to cellulose in the wall matrix (Levy et al., 1991, Medford et al., 1991). Fucosylated oligosaccharides derived from xyloglucans may also act as signal molecules in plant-pathogen interactions or plant growth regulation (Fry et al., 1993, Vargas-Rechia et al., 1998). The increased fucose content of RG-II and xyloglucan in PvMYB4-OX lines might compensate for the mechanical weakness caused by the reduced lignin levels in the cell walls, explaining why PvMYB4-OX tillers do not show severe lodging when grown in the greenhouse.

4.6 Conclusions

The concept of increased saccharification efficiency and ethanol yield through down-regulation of single lignin biosynthetic genes has been proven successful, while also creating problems, including the accumulation of upstream phenolic metabolites that are fermentation inhibitors. Our results demonstrate that an alternative approach, the overexpression of a general transcriptional repressor of the phenylpropanoid/lignin biosynthesis pathway, can reduce carbon flux into the lignin biosynthetic pathway and produce a bioenergy crop with reduced cell wall recalcitrance, slightly increased polysaccharide content and reduced levels of phenolic fermentation inhibitors. The very large improvement in ethanol yield, proportional to the dramatic reduction of recalcitrance, makes MYB4-OX switchgrass an excellent model system for understanding the chemical basis of recalcitrance, and for the development of economically viable lignocellulosic feedstocks for biofuel production. It is important to note that selection of

specific transgenic events for incorporation in breeding programs is based on multiple considerations. Important in the lignin modification field is the trade-off between reduced recalcitrance and biomass yield. In this respect, line L6 (intermediate high overexpression of PvMYB4) grows much better than more highly overexpressing lines. Although we see a strong correlation between wall-bound phenolic levels and recalcitrance (determined as final ethanol yield) based on our whole population of transgenics, there is no change in wall-bound phenolic levels in line L6, although this line does show improved ethanol yields.

4.7 Acknowledgements

We thank Lisa Jackson and David Huhman for GC-MS analysis of lignin monomers, Tui Ray for assistance with qRT-PCR analysis, Choo Hamilton and Miguel Rodriguez for assistance with ethanol fermentations and HPLC, Jeffrey Miller for assistance with glycome profiling, Dr. Stephen Webb for assistance with statistical analysis, and Professor Rick Nelson and Dr. Yuhong Tang for critical reading of the manuscript. This work was supported by the BioEnergy Science Center, a US Department of Energy Bioenergy Research Center, through the Office of Biological and Environmental Research in the DOE Office of Science. This manuscript has been co-authored by a contractor of the U.S. Government under contract DE-AC05-00OR22725. The CCRC series of plant glycan-directed monoclonal antibodies used in this project were generated with the support of the NSF Plant Genome Program (DCB-041683 and IOS-0923992).

4.8 References

Björkman A. 1956. Studied on finely divided wood. Part I. Extractions of lignin with neutral solvents. *Svensk papperstidn* 59: 477–485.

Blumenkrantz N, Asboe-Hansen G. 1973. New method for quantitative determination of uronic acids. *Analytical Biochemistry* 54: 484–489.

Burris JN, Mann DGJ, Joyce BL, Stewart CN. 2009. An improved tissue culture system for embryogenic callus production and plant regeneration in switchgrass (*Panicum virgatum* L.). *BioEnergy Resource* 2: 267–274.

Campbell P, Braam J. 1999. Xyloglucan endotransglycosylases: diversity of genes, enzymes and potential wall-modifying functions. *Trends in Plant Science* 4: 361–366.

Chang H, Cowling EB, Brown W, Adler E, Miksche G. 1975. Comparative studies on cellulolytic enzyme lignin and milled wood lignin of sweetgum and spruce. *Holzforschung* 29: 153–159.

Clark TA, Mackie KL. 1984. Fermentation inhibitors in wood hydrolysates derived from the softwood Pinus radiata. *Journal of Chemical Technology and Biotechnology* 34: 101–110.

DeMartini JD, Pattathil S, Avci U, Szekalski K, Mazumder K, Hahn MG, Wyman CE. 2011. Application of monoclonal antibodies to investigate plant cell wall deconstruction for biofuels production. *Energy & Environmental Science* 4: 4332–4339.

Dence SYL CW, Timell TE. 1992. *Methods in Lignin Chemistry*. Berlin: Berlin: Springer.

Donaldson LA. 2001. Lignification and lignin topochemistry–an ultrastructural view. *Phytochemistry* 57:859–873.

Du L, Yu P, Rossnagel BG, Christensen DA, McKinnon JJ. 2009. Physicochemical characteristics, hydroxycinnamic acids (ferulic acid, p-coumaric acid) and their ratio, and in situ biodegradability: comparison of genotypic differences among six barley varieties. *Journal of Agricultural Food Chemistry* 57: 4777–4783.

Fry SC, Aldington S, Hetherington PR, Aitken J.1993. Oligosaccharides as signals and substrates in the plant cell wall. *Plant Physiology* 103: 1–5.

Fu C, Xiao X, Xi Y, Ge Y, Chen F, Bouton J, Dixon RA, Wang ZY. 2011a. Downregulation of cinnamyl alcohol dehydrogenase (CAD) leads to improved saccharification efficiency in switchgrass. *BioEnergy Resource* 4: 153–164.

Fu C, Mielenz JR, Xiao X, Ge Y, Hamilton CY, Chen F, Bouton J, Foston M, Dixon RA, Wang Z-Y. 2011b. Genetic manipulation of lignin biosynthesis in switchgrass significantly reduces recalcitrance and improves biomass ethanol production. *Proceedings of the National Academy of Sciences USA* 108: 3803–3808.

Grabber JH, Mertens DR, Kim H, Funk C, Liu F, Ralph J. 2009. Cell wall fermentation kinetics are impacted more by lignin content and ferulate cross-linking than by lignin composition. *Journal of the Science of Food and Agriculture* 89: 122–129.

Hardin CF, Fu C, Hisano H, Xiao X, Shen H, Stewart CN Jr, Parrott W, Dixon RA, Wang Z-Y. 2013. Standardization of switchgrass sample collection for cell wall and biomass trait analysis. *BioEnergy Resource* 1–8.

Hatfield RD, Chaptman AK. 2009. Comparing corn types for differences in cell wall characteristics and p-coumaroylation of lignin. *Journal of Agricultural and Food Chemistry* 57: 4243–4249.

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315: 804–807.

Hoffman M, Jia ZH, Peña MJ, Cash M, Harper A, Blackburn AR, Darvill A, York WS. 2005. Structural analysis of xyloglucans in the primary cell walls of plants in the subclass Asteridae. *Carbohydrate Research* 340: 1826–1840.

Humbird D, Davis R, Kinchin C, Tao L, Hsu D, Aden A, Schoen P, Lukas J, Olthof B, Worley M, Sexton D, Dudgeon D. 2011. Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol. *NREL Technical Report*. NREL/TP-5100–47767. http://www.nrel.gov/docs/fy11osti/47764.pdf

Keshwani DR, Cheng JJ. 2009. Switchgrass for bioethanol and other value-added applications: A review. *Bioresource Technology* 100: 1515–1523.

Klinke HB, Thomsen A, Ahring BK. 2004. Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Applied Microbiology and Biotechnology* 66: 10–26.

Levy S, Maclachlan G, Staehelin LA.1997. Xyloglucan sidechains modulate binding to cellulose during in vitro binding assays as predicted by conformational dynamics simulations. *Plant Journal* 11: 373–386.

Levy S, York WS, Stuike□Prill R, Meyer B, Staehelin LA. 1991. Simulations of the static and dynamic molecular conformations of xyloglucan. The role of the fucosylated sidechain in surface□specific sidechain folding. *Plant Journal* 1: 195–215.

Liu ZL, Blaschek HP. 2010. Biomass conversion inhibitors and in situ detoxification. In *Biomass to Biofuels: Strategies for Global Industries*. Edited by: Vertès AA, Qureshi N, Blaschek HP, Yukawa H. Oxford, UK: Blackwell Publishing Ltd. 10.1002/9780470750025.ch12

McLaughlin SB, Adams Kszos L. 2005. Development of switchgrass (*Panicum virgatum*) as a bioenergy feedstock in the United States. *Biomass and Bioenergy* 28: 515–535.

Meikle PJ, Hoogenraad NJ, Bonig I, Clarke AE, Stone BA. 1994. A (1□→□3,1□→□4)-β-glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1□→□3,1□→□4)-β-glucans. *Plant Journal* 5: 1–9.

Medford JI, Elmer JS, Klee HJ. 1991. Molecular cloning and characterization of genes expressed in shoot apical meristems. *Plant Cell* 3: 359–370.

Monro J, Bailey R, Penny D. 1972. Polysaccharide composition in relation to extensibility and possible peptide linked arabino-galactan of lupin hypocotyl cell walls. *Phytochemistry* 11: 1597–1602.

Mosier N, Wyman C, Dale B, Elander R, Lee Y, Holtzapple M, Ladisch M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology* 96: 673–686.

Pattathil S, Avci U, Miller JS, Hahn MG: Immunological approaches to plant cell wall and biomass characterization: glycome profiling. Biomass Conversion: *Methods and Protocols Methods in Molecular Biology* 908: 61–72.

Pattathil S, Avci U, Baldwin D, Swennes AG, McGill JA, Popper Z, Bootten T, Albert A, Davis RH, Chennareddy C. 2010. A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiology* 153: 514–525.

Puhlmann J, Bucheli E, Swain MJ, Dunning N, Albersheim P, Darvill AG, Hahn MG. 1994. Generation of monoclonal antibodies against plant cell wall polysaccharides. I. Characterization of a monoclonal antibody to a terminal α-(1□→□2)-linked fucosyl-containing epitope. *Plant Physiology* 104: 699–710.

Ralph J. 2010. Hydroxycinnamates in lignification. *Phytochemistry Reviews* 9: 65–83.

Ryden P, Sugimoto-Shirasu K, Smith AC, Findlay K, Reiter WD, McCann MC. 2003. Tensile properties of Arabidopsis cell walls depend on both a xyloglucan cross-linked microfibrillar network and rhamnogalacturonan II-borate complexes. *Plant Physiology* 132: 1033–1040.

Saathoff AJ, Sarath G, Chow EK, Dien BS, Tobias CM. 2011. Downregulation of cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. *PLoS One* 6: e16416.

Schmer MR, Vogel KP, Mitchell RB, Perrin RK. 2008. Net energy of cellulosic ethanol from switchgrass. *Proceedings of the National Academy of Sciences USA* 105: 464–469.

Selvendran RR, Davies A, Tidder E. 1975. Cell wall glycoproteins and polysaccharides of mature runner beans. *Phytochemistry* 14: 2169–2174.

Shen H, Fu C, Xiao X, Ray T, Tang Y, Wang Z, Chen F. 2009. Developmental control of lignification in stems of lowland switchgrass variety Alamo and the effects on saccharification efficiency. *BioEnergy Resource* 2: 233–245.

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

Tschaplinski TJ, Standaert RF, Engle NL, Martin MZ, Sangha AK, Parks JM, Smith JC, Samuel R, Jiang N, Pu Y. 2012. Down-regulation of the caffeic acid O-methyltransferase gene in switchgrass reveals a novel monolignol analog. *Biotechnology for Biofuels* 5: 71.

Vanholme R, Ralph J, Akiyama T, Lu F, Pazo JR, Kim H, Christensen JH, Van Reusel B, Storme V, De Rycke R. 2010. Engineering traditional monolignols out of lignin by concomitant up□regulation of F5H1 and down□regulation of COMT in Arabidopsis. *Plant Journal* 64: 885–897.

Vargas-Rechia C, Reicher F, Sierakowski MR, Heyraud A, Driguez H, Liénart Y. 1998. Xyloglucan octasaccharide XXLGol derived from the seeds of Hymenaea courbaril acts as a signaling molecule. *Plant Physiology* 116: 1013–1021.

Wang M, Wu M, Huo H. 2007. Life-cycle energy and greenhouse gas emission impacts of different corn ethanol plant types. *Environmental Research Letters* 2: 024001.

Wi S, Singh A, Lee K, Kim Y. 2005. The pattern of distribution of pectin, peroxidase and lignin in the middle lamella of secondary xylem fibres in alfalfa (*Medicago sativa*). *Annals of Botany* 95: 863–868.

Yang B, Wyman CE. 2009. Dilute acid and autohydrolysis pretreatment. Biofuels *Methods and Protocols Series, Methods in Molecular Biology: Biofuels* 581: 103–114.

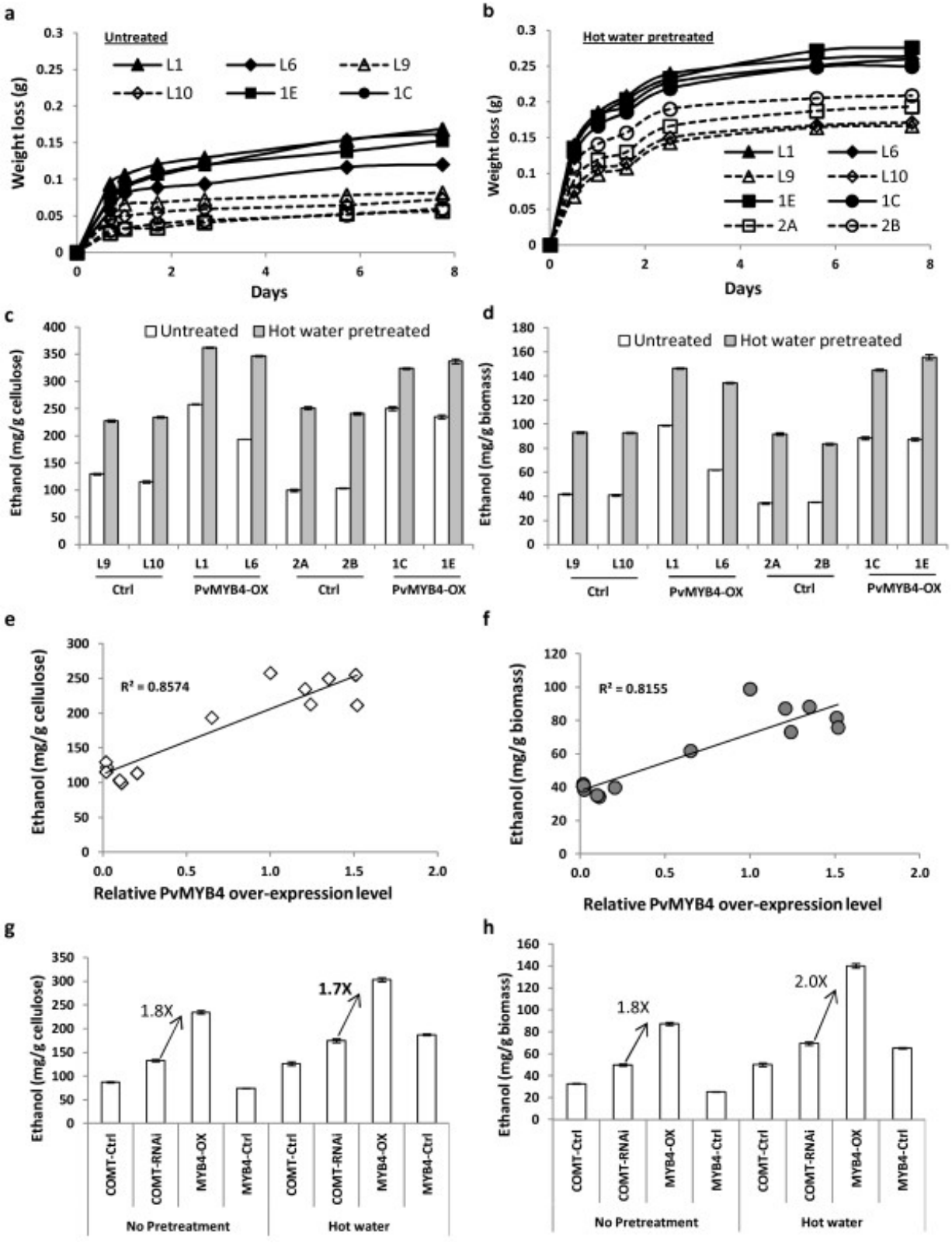
Ziebell A, Gracom K, Katahira R, Chen F, Pu Y, Ragauskas A, Dixon RA, Davis M. 2010. Increase in 4-coumaryl alcohol units during lignification in alfalfa (*Medicago sativa*) alters the extractability and molecular weight of lignin. *Journal of Biological Chemistry* 285: 38961–38968.

4.9 Appendix

4.9.1 Figures and tables

Figure 4-1 Bioconversion of PvMYB4-OX transgenic switchgrass biomass to ethanol with or without hot water pretreatment using Saccharomyces cerevisiae D5A.

(a, b) Time courses of fermentation of whole plant material without (a) and with (b) hot water pre-treatment in fermentation broths measured by weight loss. (c, d) Final ethanol yields calculated as mg/g cellulose (c) or mg/g biomass (d) from hot water pre-treated and non-treated biomass, with comparison of different control (Ctrl) and PvMYB4-OX lines. (e, f) Correlation between the PvMYB4 transcript level and SSF ethanol yield without pretreatment. Data from lines L1, L6, L7, L8, L9 and L10 (Alamo ST1 background), and 1A, 1B, 1C, 1D, 1E, 2A and 2B) Alamo ST2 genetic background) are used. (g, h) SSF ethanol yield comparison of PvMYB4-OX and PvCOMT-RNAi (9) transgenic lines. All data are means□±□SE (n□=□3).


a
Untreated
L1
L6
L9
L10
1E
1C
Weight loss (g)
Days
b
Hot water pretreated
L1
L6
L9
L10
1E
1C
2A
2B
Weight loss (g)
Days
c
Untreated
Hot water pretreated
Ethanol (mg/g cellulose)
L9
L10
L1
L6
2A
2B
1C
1E
Ctrl
PvMYB4-OX
Ctrl
PvMYB4-OX
d
Untreated
Hot water pretreated
Ethanol (mg/g biomass)
L9
L10
L1
L6
2A
2B
1C
1E
Ctrl
PvMYB4-OX
Ctrl
PvMYB4-OX
e
R² = 0.8574
Ethanol (mg/g cellulose)
Relative PvMYB4 over-expression level
f
R² = 0.8155
Ethanol (mg/g biomass)
Relative PvMYB4 over-expression level
g
1.8X
1.7X
Ethanol (mg/g cellulose)
COMT-Ctrl
COMT-RNAi
MYB4-OX
MYB4-Ctrl
COMT-Ctrl
COMT-RNAi
MYB4-OX
MYB4-Ctrl
No Pretreatment
Hot water
h
1.8X
2.0X
Ethanol (mg/g biomass)
COMT-Ctrl
COMT-RNAi
MYB4-OX
MYB4-Ctrl
COMT-Ctrl
COMT-RNAi
MYB4-OX
MYB4-Ctrl
No Pretreatment
Hot water

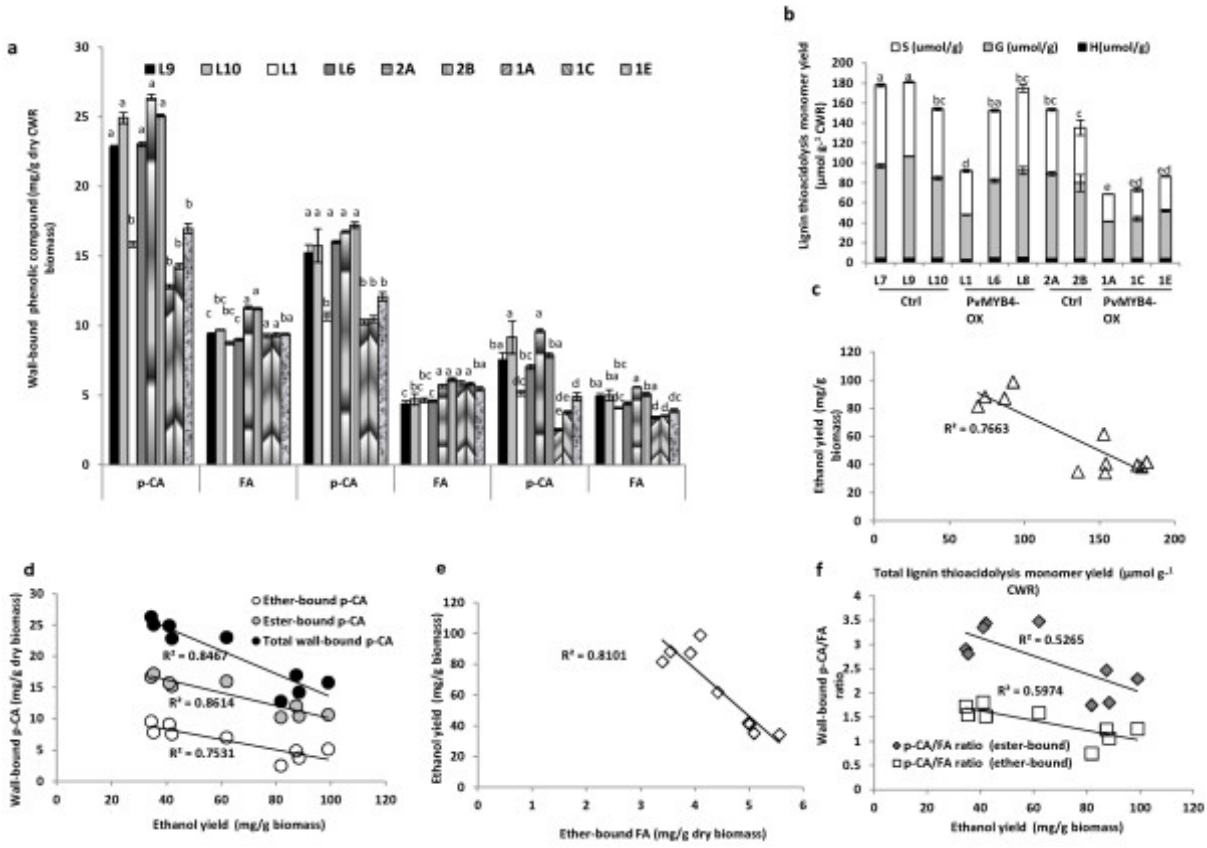
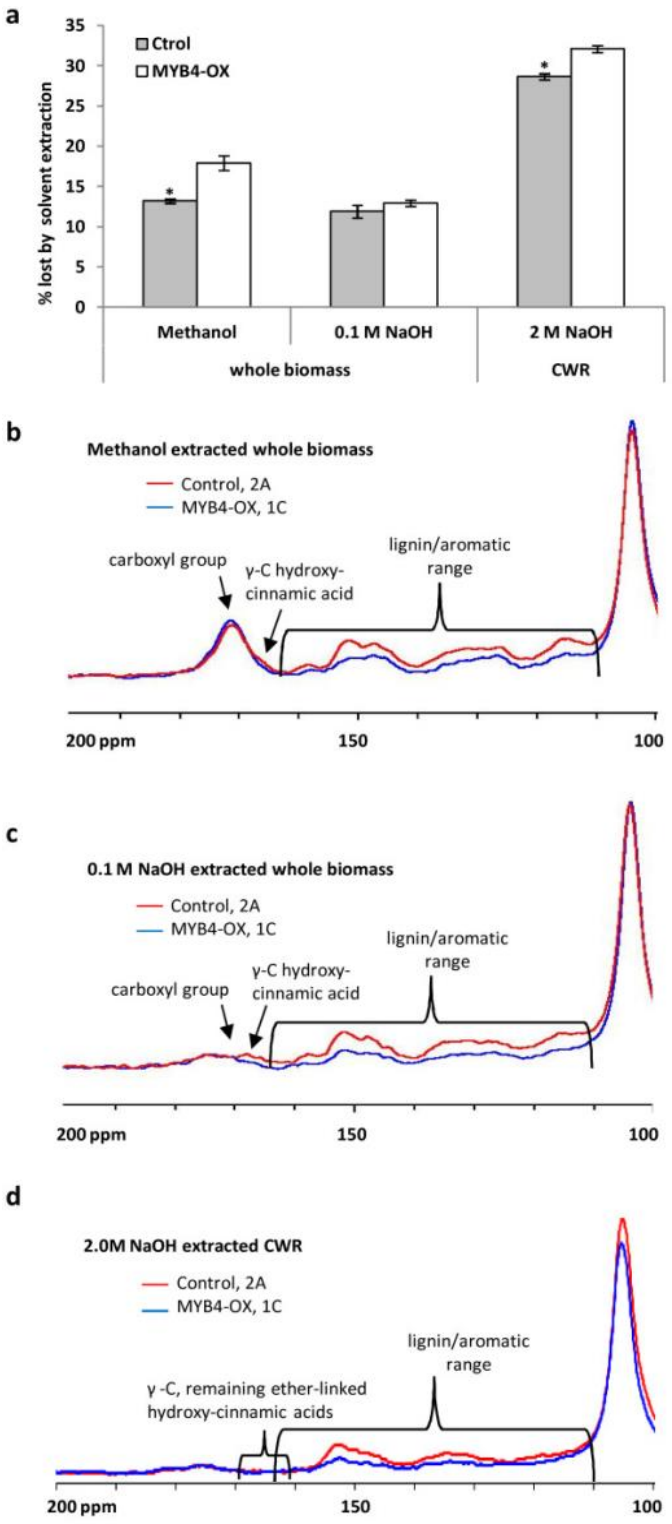


Figure 4-2 Correlations between lignin content, wall-bound phenolics and SSF ethanol yield.

(**a**) The content of ester-linked, ether-linked and total wall-bound **p**-coumaric and ferulic acids in the cell wall residues (CWR) of switchgrass whole tillers. (**b**) Lignin composition of control and PvMYB4-OX switchgrass whole tillers determined by thioacidolysis. Ctrl: control lines; S, syringyl unit; G, guaiacyl unit; H, **p**-hydroxyphenyl unit. (**c**-**f**) Correlations with SSF ethanol yield without pretreatment. (**c**) lignin content. (**d**) wall-bound **p**-coumaric acid. (**e**) ether-lined ferulic acid. (**f**) wall-bound p-CA/FA ratio. All data are means□±□SE (n□=□3). The letters indicates significant differences of total lignin content at the **p**□<□0.05 level. Mean comparisons, based on mean separation test results, cannot be compared across variables in Figure 4-2a

Figure 4-3 Cell wall extractabilities and characteristics determined by solid-state 13 C CP/MAS NMR spectroscopy.

(a) Cell wall integrity of whole biomass or cell wall residues (CWR) determined by the percentage of mass lost under different extraction conditions. 0.1M NaOH, 0.1 M NaOH extraction overnight at 5°C after 91-92% methanol extraction overnight at 5°C. 2M NaOH: extraction of ester-linked wall-bound phenolics from CWR. *Asterisks indicate values that were determined by the Student t-test to be significantly different from their equivalent control ($p < 0.05$). All data are means ± SE ($n = 3$). (b-d) Solid-state 13C CP/MAS NMR spectra of CWRs of control (2A) and PvMYB4-OX (1C) biomass that had been extracted with methanol (b), 0.1M NaOH (c) or 2M NaOH (d). Whole biomass, non-extracted biomass from the whole tillers. CWR, cell wall residues of the whole biomass extracted by methanol: chloroform, methanol, methanol: H_2O and H_2O.


a
35
30
25
20
15
10
5
0
% lost by solvent extraction
Ctrol
MYB4-OX
*
*
Methanol
0.1 M NaOH
2 M NaOH
whole biomass
CWR
b
Methanol extracted whole biomass
Control, 2A
MYB4-OX, 1C
carboxyl group
γ-C hydroxy-cinnamic acid
lignin/aromatic range
200 ppm
150
100
c
0.1 M NaOH extracted whole biomass
Control, 2A
MYB4-OX, 1C
lignin/aromatic range
γ-C hydroxy-cinnamic acid
carboxyl group
200 ppm
150
100
d
2.0M NaOH extracted CWR
Control, 2A
MYB4-OX, 1C
lignin/aromatic range
γ -C, remaining ether-linked hydroxy-cinnamic acids
200 ppm
150
100

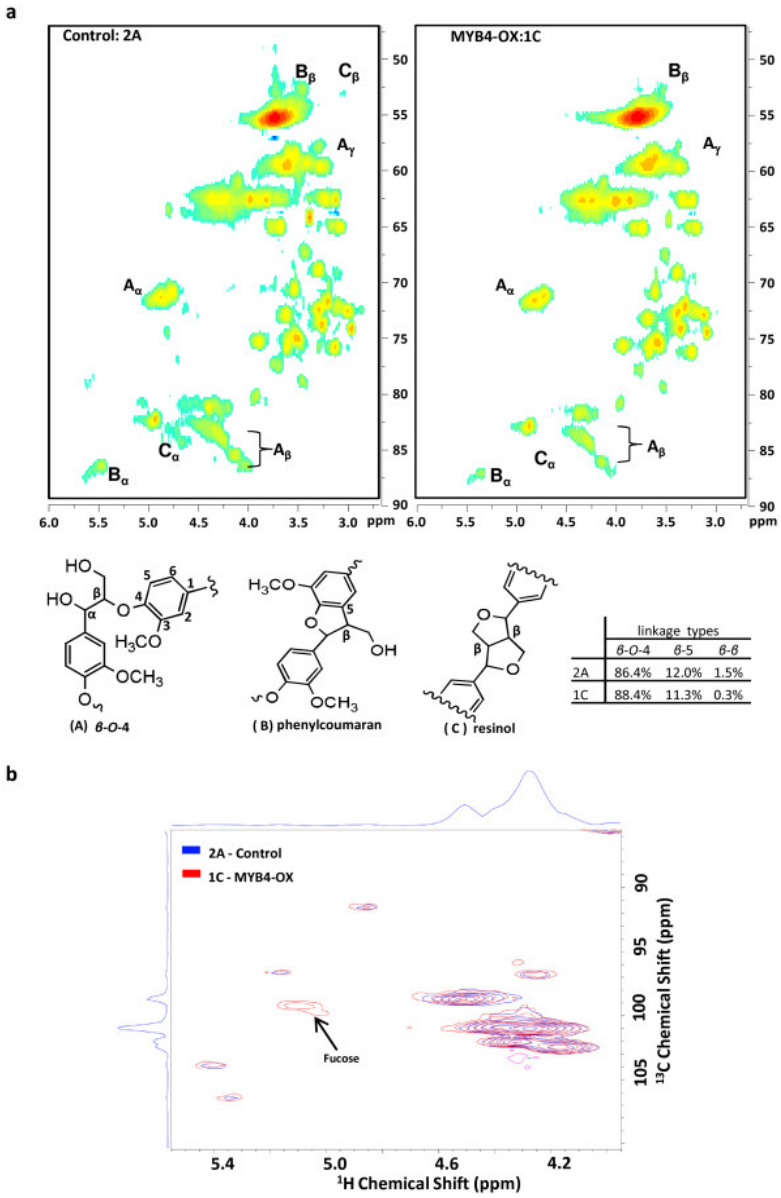


Figure 4-4 Partial HSQC spectra showing the side-chain C-H correlations from the three main linkages (β-O-4-, β -5-, and β-β) in lignins and presence of wall-associated fucose.

Partial HSQC spectra of isolated lignins. The carbon positions in aromatic ring and side-chain are labeled on the molecular structures of the A, B and C types of linkages. The insert table shows the relative content of different linkage types. (b) Partial 2D HSQC NMR analysis of gel-state CWR showing increased wall-associated fucose in PvMYB-OX. Red, control 2A. Blue, PvMYB4-OX

Figure 4-5 Glycome profiling heatmaps of specific cell wall extracts showing areas thatwith indication of antibodies.

The white, blue, yellow and green colored boxes highlighted in Additional file 1 are enlarged in (a-d) to show differences in antibody binding signals. (e) Heatmap scale. The black, blue, red and yellow colors indicate the binding intensities of groups of plant glycan-directed monoclonal antibodies (with black color depicting no binding and bright yellow depicting strongest binding) that are selective for the different cell wall polysaccharides as labeled on the right-hand side of the figure.

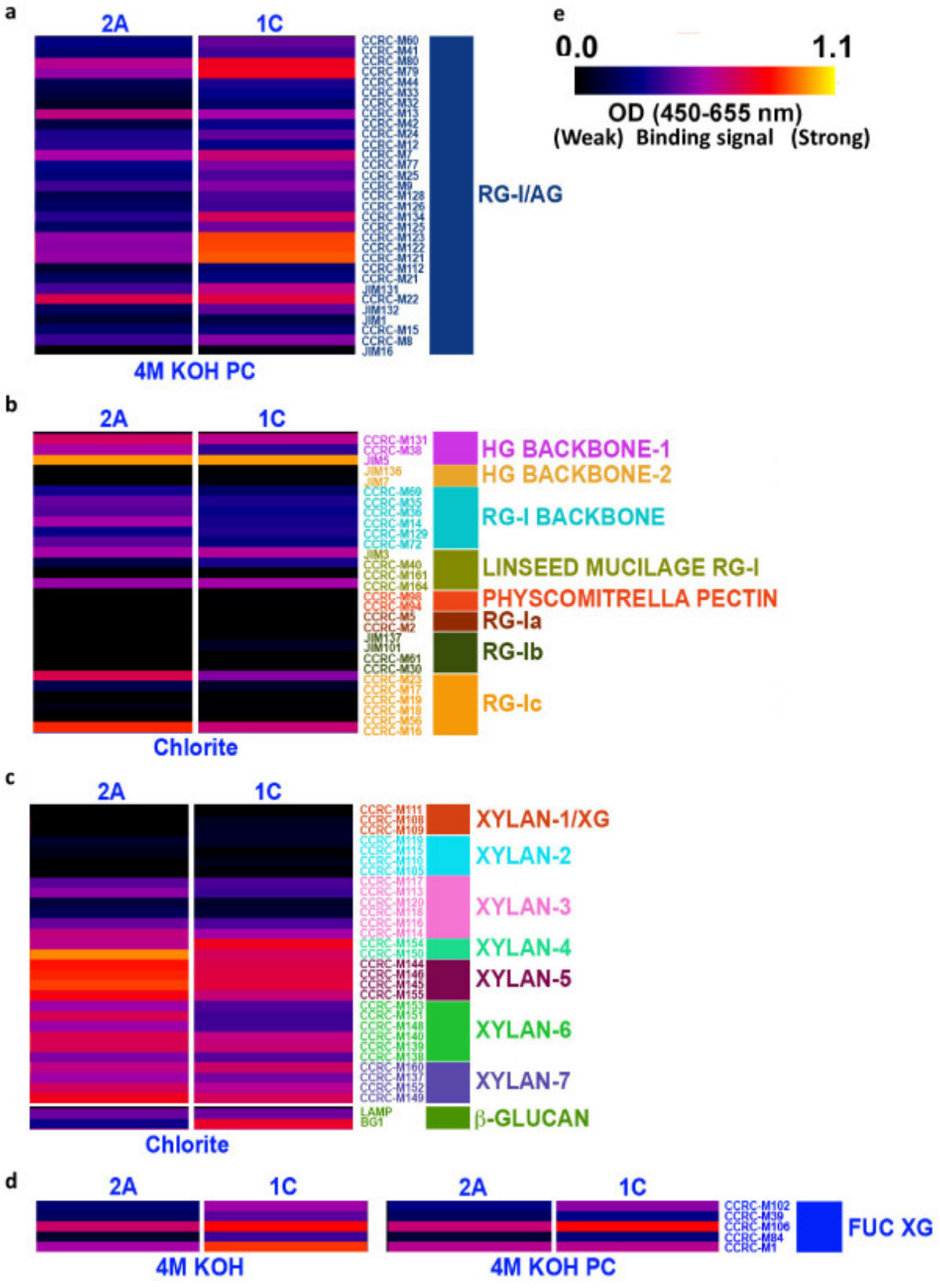
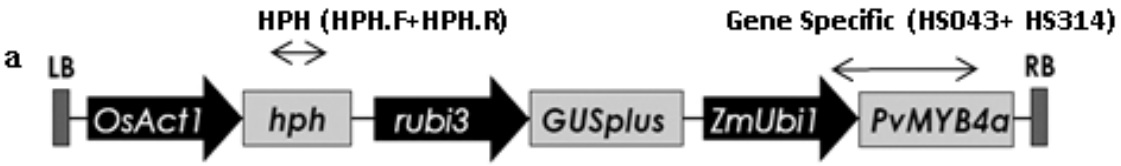
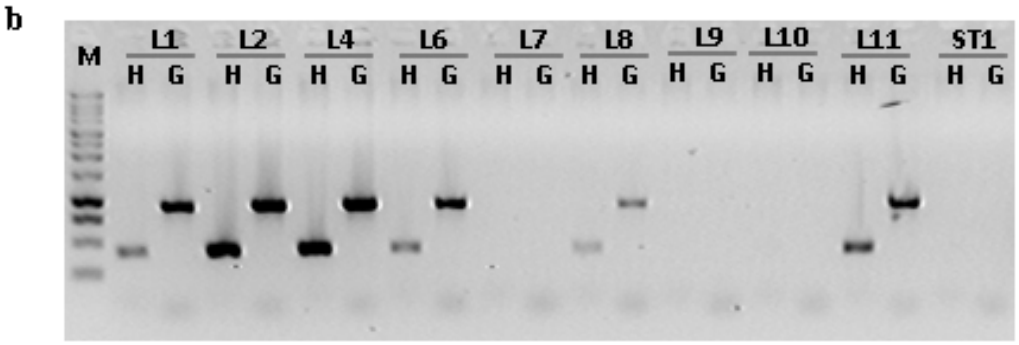
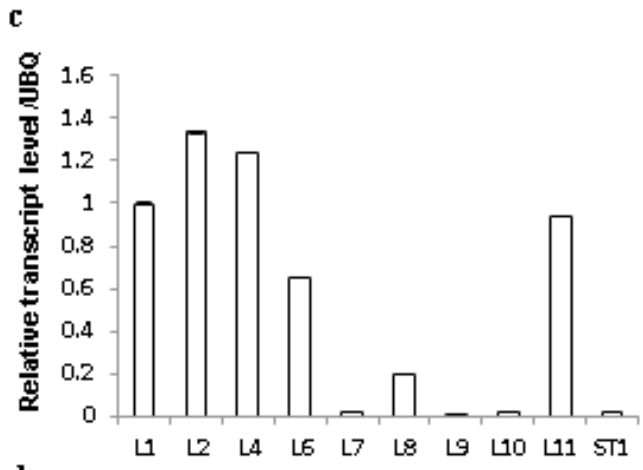

a
2A
1C
CCRC-M60
CCRC-M61
CCRC-M60
CCRC-M79
CCRC-M44
CCRC-M33
CCRC-M32
CCRC-M13
CCRC-M42
CCRC-M24
CCRC-M12
CCRC-M7
CCRC-M77
CCRC-M25
CCRC-M9
CCRC-M128
CCRC-M126
CCRC-M134
CCRC-M125
CCRC-M123
CCRC-M122
CCRC-M121
CCRC-M112
CCRC-M21
JIM131
CCRC-M22
JIM132
JIM1
CCRC-M15
CCRC-M8
JIM16
RG-I/AG
4M KOH PC
e
0.0
1.1
OD (450-655 nm)
(Weak) Binding signal (Strong)
b
2A
1C
CCRC-M131
CCRC-M38
JIM5
JIM136
JIM7
CCRC-M69
CCRC-M35
CCRC-M36
CCRC-M14
CCRC-M129
CCRC-M72
JIM5
CCRC-M40
CCRC-M161
CCRC-M164
CCRC-M66
CCRC-M54
CCRC-M5
CCRC-M2
JIM137
JIM101
CCRC-M61
CCRC-M30
CCRC-M23
CCRC-M17
CCRC-M19
CCRC-M18
CCRC-M56
CCRC-M16
HG BACKBONE-1
HG BACKBONE-2
RG-I BACKBONE
LINSEED MUCILAGE RG-I
PHYSCOMITRELLA PECTIN
RG-Ia
RG-Ib
RG-Ic
Chlorite
c
2A
1C
CCRC-M111
CCRC-M108
CCRC-M109
CCRC-M119
CCRC-M115
CCRC-M110
CCRC-M105
CCRC-M117
CCRC-M113
CCRC-M120
CCRC-M118
CCRC-M116
CCRC-M114
CCRC-M154
CCRC-M150
CCRC-M144
CCRC-M146
CCRC-M145
CCRC-M155
CCRC-M153
CCRC-M151
CCRC-M148
CCRC-M140
CCRC-M139
CCRC-M138
CCRC-M160
CCRC-M137
CCRC-M152
CCRC-M149
LAMP
BG1
XYLAN-1/XG
XYLAN-2
XYLAN-3
XYLAN-4
XYLAN-5
XYLAN-6
XYLAN-7
β-GLUCAN
Chlorite
d
2A
1C
2A
1C
CCRC-M102
CCRC-M38
CCRC-M106
CCRC-M84
CCRC-M1
FUC XG
4M KOH
4M KOH PC

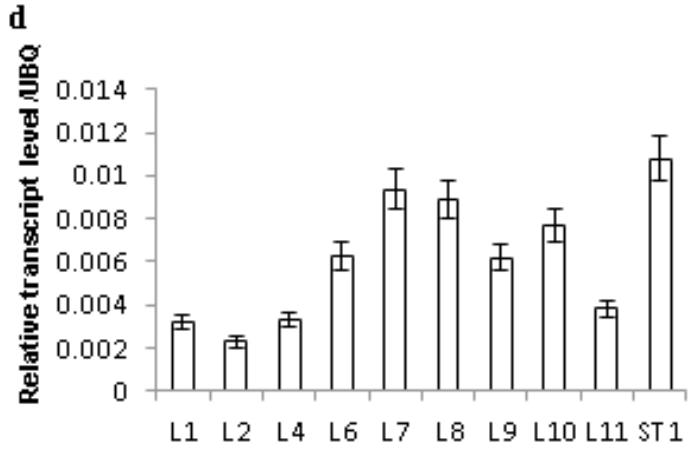
Figure 4-6 Genomic DNA PCR and qRT-PCR for PvMYB4-OX lines generated from Alamo ST1 background.

(a) The *pANIC2B-PvMYB4* construct used for switchgrass transformation. *OsAct1*, promoter of rice *Actin1* gene; *rubi3*, promoter of *rubisco*3 gene; *ZmUbi1*, promoter of maize *ubiquitin 1* gene. *hph*, hygromycin resistance gene for transformation selection marker. LB and RB, left and right borders of the T-DNA sequences. The double arrows show the primer pairs used for genomic DNA PCR testing. HPH, hygromycin specific gene primer pairs; Gene specific, primer pairs covering the ZmUBI1 promoter region and PvMYB4 gene specific region. (b) Genomic DNA PCR. L7, L9, L10 and ST1 are transgenic and non-transgenic control lines. (c,d) qRT-PCR analysis showing (c) the *PvMYB4* transcripts in control and PvMYB4-OX transgenic switchgrass and (d) endogenous *PvMYB4* gene is down-regulated by the over-expression of *PvMYB4* in switchgrass. Primer pairs used as been reported before. Data are means ±SE (n=3). The L1 (high overexpression), L6 (intermediate overexpression), L9 and L19 (transgenic controls) are chosen for further analysis. (e) Visible phenotypes of control (Line 2A) and PvMYB4-OX switchgrass (Line 1C).


a
HPH (HPH.F+HPH.R)
Gene Specific (HS043+ HS314)
LB
RB
OsAct1
hph
rubi3
GUSplus
ZmUbi1
PvMYB4a


b
M
L1
L2
L4
L6
L7
L8
L9
L10
L11
ST1
H G

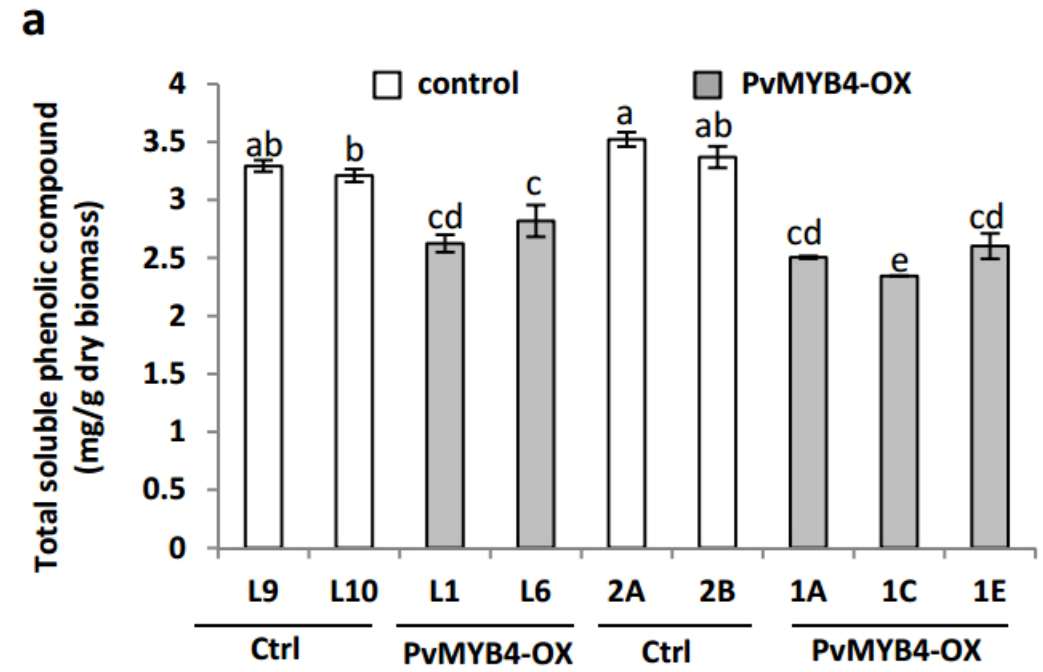

c
Relative transcript level /UBQ
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
L1 L2 L4 L6 L7 L8 L9 L10 L11 ST1

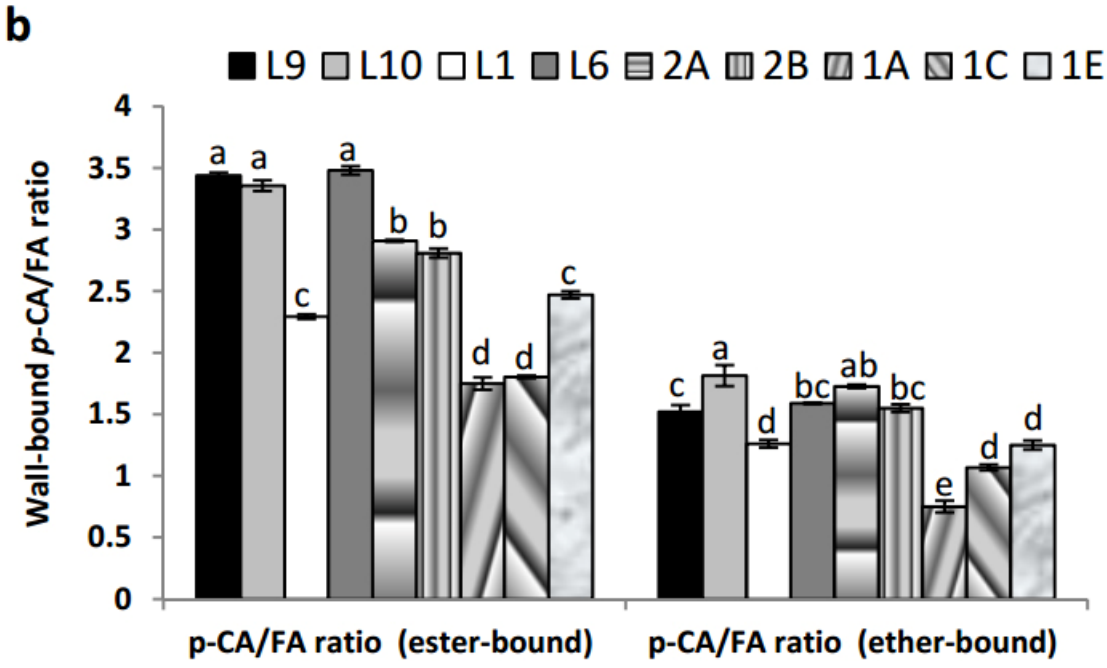

d
Relative transcript level /UBQ
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
L1 L2 L4 L6 L7 L8 L9 L10 L11 ST1


e
Ctrl
PvMYB4-OX

Figure 4-7 PvMYB4-OX transgenic switchgrass lines have reduced levels of soluble phenolic compounds and reduced ester- and ether-linked p-coumaric/ferulic acid ratios.

(a) Levels of total soluble phenolic compounds as determined by reaction with Folin Ciocalteau reagent. (b) Levels of ester- and ether-linked wall-bound p-coumaric/ferulic acid (p-CA/FA) ratios. All data are means ±SE (n=3). Different letters on the bars of figures indicate significant differences at the $p<0.05$ level. Mean comparisons, based on mean separation test results, cannot be compared across variables in panel b. Lines L9, L10, L1 and L6 are in the Alamo ST1 genetic background, lines 2A, 2B, 1A, 1C and 1E in the Alamo ST2 genetic background.


a
control
PvMYB4-OX
Total soluble phenolic compound (mg/g dry biomass)
4
3.5
3
2.5
2
1.5
1
0.5
0
ab
b
cd
c
a
ab
cd
e
cd
L9
L10
L1
L6
2A
2B
1A
1C
1E
Ctrl
PvMYB4-OX
Ctrl
PvMYB4-OX


b
L9
L10
L1
L6
2A
2B
1A
1C
1E
Wall-bound p-CA/FA ratio
4
3.5
3
2.5
2
1.5
1
0.5
0
a
a
c
a
b
b
d
d
c
c
a
d
bc
ab
bc
e
d
d
p-CA/FA ratio (ester-bound)
p-CA/FA ratio (ether-bound)

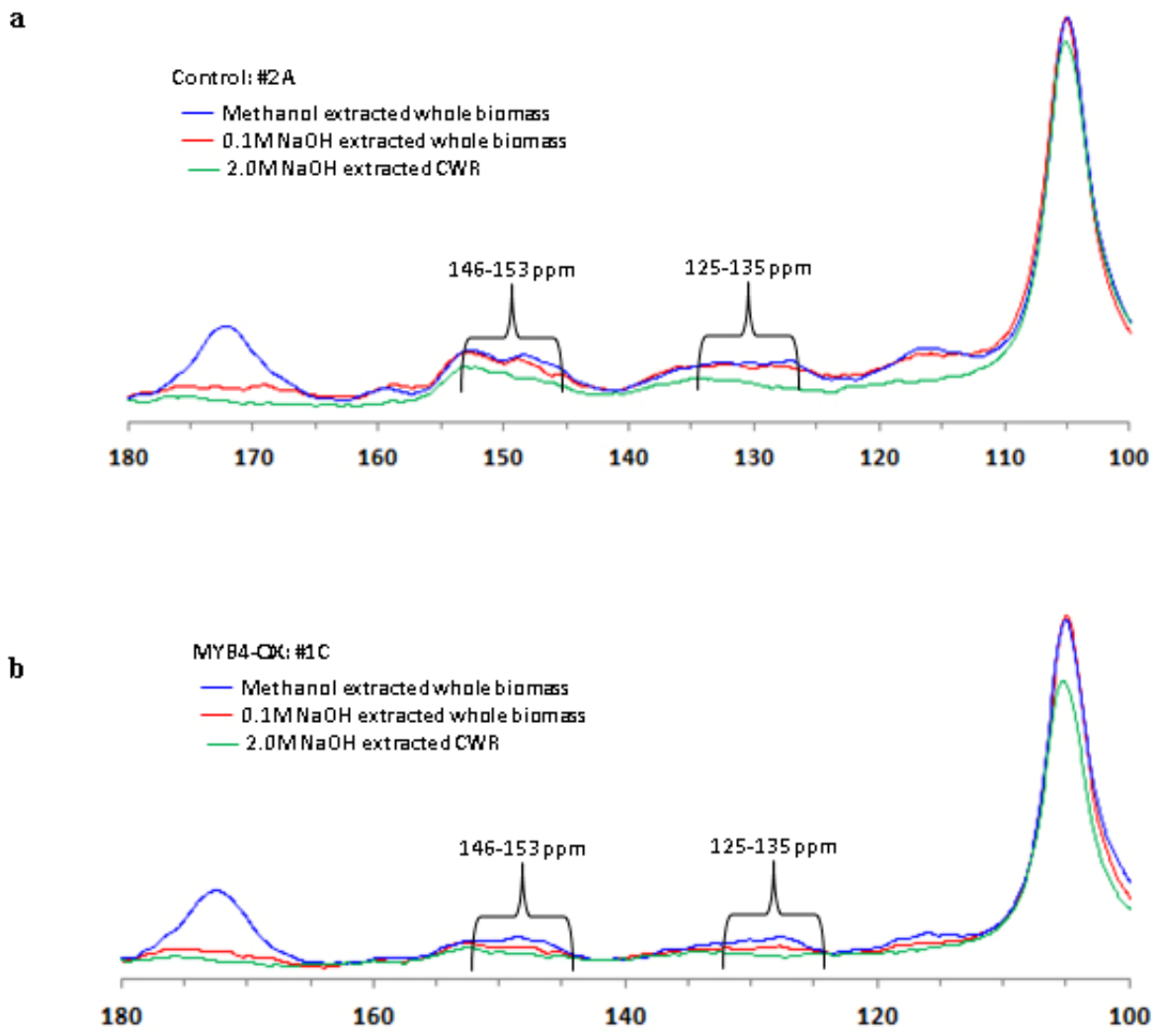


Figure 4-8 Comparisons of methanol (blue), 0.1 M NaOH (red) and 2.0 M NaOH (green) extracted cell wall residues determined by solid-state 13C CP/MAS NMR spectroscopy.

Panels(a) and (b) show the NMR spectrum for control line 2A and MYB4-OX line 1C.

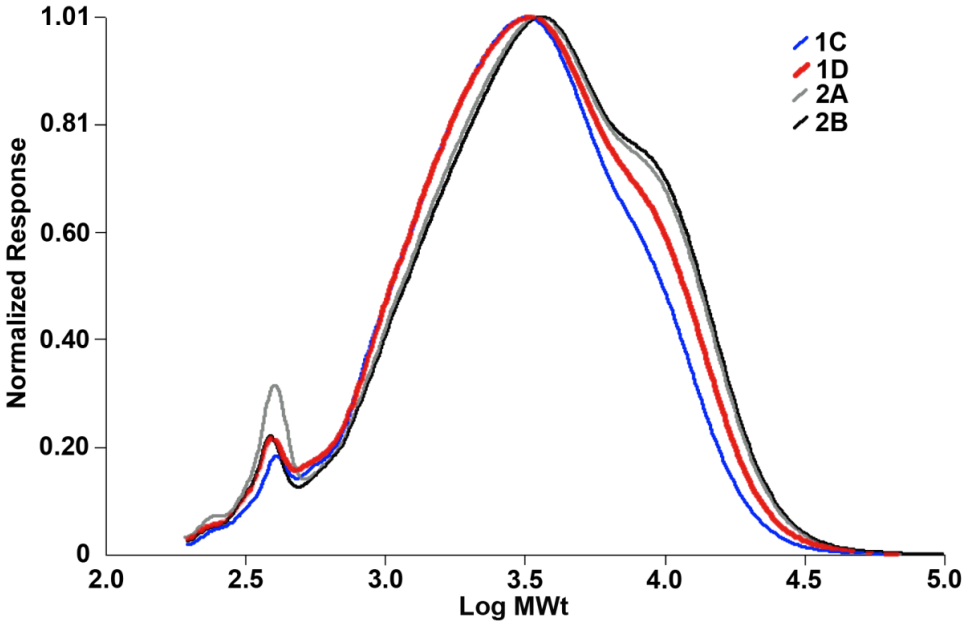
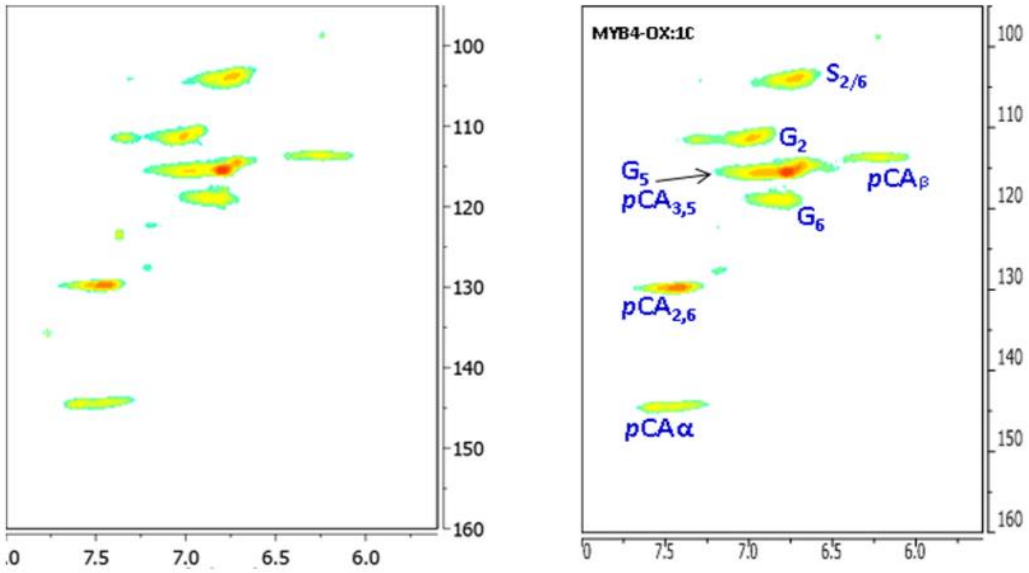


Figure 4-9 Gel permeation chromatography (GPC) analysis of isolated lignin s from PvMYB-OX and control switchgrass.

Control lines: 2A and 2B. MYb4-OX lines: 1C and 1D.



Aromatic regions of a 13C-1H HSQC spectrum switchgrass lignin

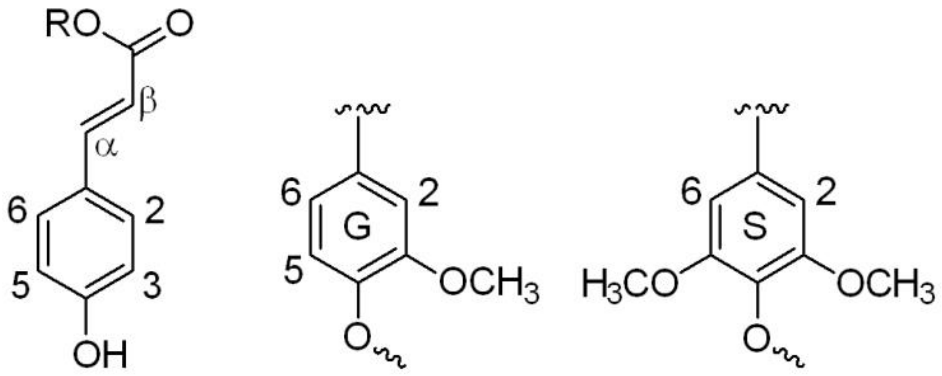


Figure 4-10 Partial 13C-1H HSQC spectra of milled lignins showing the C–H correlations from aromatic rings of the structural units.

*p*CA, p-coumaryl unit. G, coniferyl lignin unit. S, sinapyl lignin unit.

Figure 4-11 Glycome profiling of extracts from alcohol insoluble cell wall residues from switchgrass

Sequential extracts were prepared from the cell wall materials (AIR) prepared from control and PvMYB4-OX transgenic lines using a set of reagents with increasing harshness (as labeled in the bottom panel of the figure). The bars on the top panels of the figure indicates the amount of carbohydrates released by sequential extraction of the wall residues.

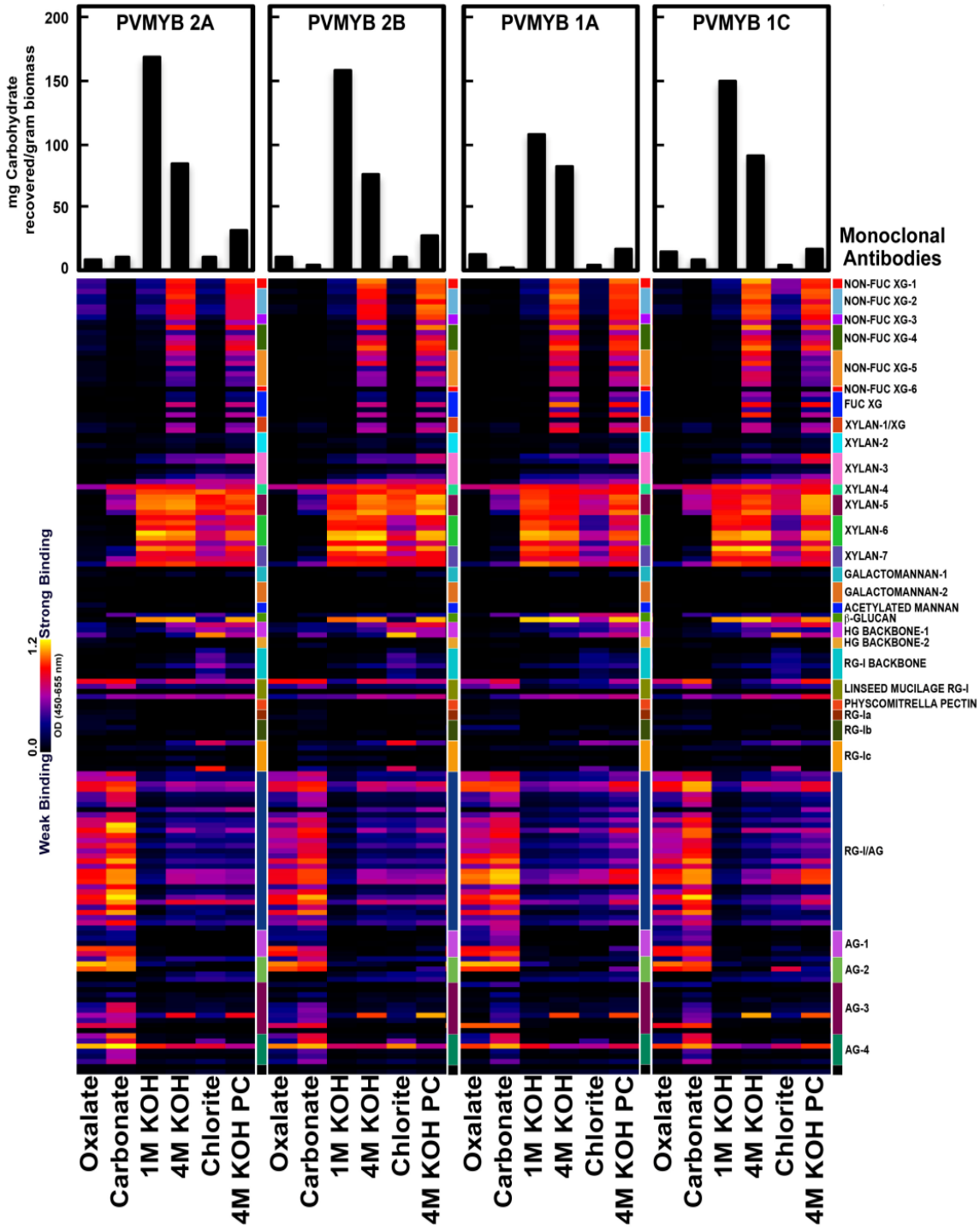

PVMYB 2A
PVMYB 2B
PVMYB 1A
PVMYB 1C
mg Carbohydrate recovered/gram biomass
200
150
100
50
0
Monoclonal Antibodies
NON-FUC XG-1
NON-FUC XG-2
NON-FUC XG-3
NON-FUC XG-4
NON-FUC XG-5
NON-FUC XG-6
FUC XG
XYLAN-1/XG
XYLAN-2
XYLAN-3
XYLAN-4
XYLAN-5
XYLAN-6
XYLAN-7
GALACTOMANNAN-1
GALACTOMANNAN-2
ACETYLATED MANNAN
β-GLUCAN
HG BACKBONE-1
HG BACKBONE-2
RG-I BACKBONE
LINSEED MUCILAGE RG-I
PHYSCOMITRELLA PECTIN
RG-Ia
RG-Ib
RG-Ic
RG-I/AG
AG-1
AG-2
AG-3
AG-4
Strong Binding
1.2
OD (450-655 nm)
0.0
Weak Binding
Oxalate
Carbonate
1M KOH
4M KOH
Chlorite
4M KOH PC

Table 4-1 Metabolite concentrations (ng/ml) of methanol extracts as determined by GC-MS

Metabolite	MYB4-OX	Control	MYB4OX/ Ctrl	*P*-value	Classification
16.07 375 583 513 411 427 204	6	1	8.66	0.000	glucoside
16.23 488 327 265 syringyl lignan	4	1	6.33	0.018	lignan
15.97 583 375 285	70	12	5.98	0.001	N/A
11.01 450 217 sugar	45	11	4.29	0.030	modified sugar
9.96 281 383 354	289	71	4.07	0.022	phenolic
10.27 328 343 284 254	999	245	4.08	0.005	N/A
raffinose	933	267	3.49	0.022	sugar - trisaccharide
galactose	1193	345	3.46	0.056	sugar - monosaccharide
19.47 496 481 209 lignan	36	11	3.31	0.050	lignan
α-tocopherol	18	6	3.20	0.060	vitamin
γ-tocopherol	11	4	2.73	0.018	vitamin
13.93 375 292 305 275 uronic acid	176	69	2.54	0.002	sugar acid conjugate
dehydroabietic acid	76	32	2.36	0.048	resin acid
19.09 483 498 lignan	11	5	2.27	0.047	lignan
Tryptophan	264	120	2.21	0.036	amino acid
16.11 368 600 585 353 255	11	5	2.18	0.001	N/A
Fructose	8160	3897	2.09	0.083	sugar - monosaccharide
Bornesitol	2914	1403	2.08	0.000	cyclitol
Glutamine	605	315	1.92	0.081	amino acid
Tyrosine	610	320	1.91	0.058	amino acid
α-linolenic acid	1132	674	1.68	0.060	fatty acid
Alanine	3061	1835	1.67	0.002	amino acid
Glucose	6307	3799	1.66	0.040	sugar - monosaccharide
16.85 caffeic acid conjugate	17	11	1.59	0.031	phenylpropanoid
dodecanoic acid	104	68	1.53	0.090	fatty acid
3-*O*-caffeoylquinic acid	414	297	1.39	0.146	phenylpropanoid

Table 4-1 continued

Metabolite	MYB4-OX	Control	MYB4OX/Ctrl	*P*-value	Classification
5-*O*-caffeoylquinic acid	24	17	1.37	0.073	phenylpropanoid
caffeic acid	42	30	1.37	0.040	phenylpropanoid
	7451	5457	1.37	0.050	amino acid
GABA (γ-aminobutyric acid)					
shikimic acid	6025	4626	1.30	0.423	organic acid
Campesterol	54	43	1.26	0.078	sterol
4-*O*-caffeoylquinic acid	28	23	1.23	0.405	phenylpropanoid
quinic acid	2716	2670	1.02	0.932	organic acid
sinapic acid	4	6	0.78	0.075	phenylpropanoid
5-*O*-feruloylquinic acid	134	185	0.72	0.050	phenylpropanoid
19.14 572 498 483	2	2	0.70	0.001	N/A
5-hydroxyconiferyl alcohol	1	1	0.68	0.001	phenylpropanoid
ferulic acid	26	42	0.61	0.001	phenylpropanoid
sinapyl alcohol	6	10	0.58	0.004	phenylpropanoid
Sucrose	3600	6212	0.58	0.226	sugar - disaccharide
4-*O*-feruloylquinic acid	98	173	0.57	0.013	phenylpropanoid
12.88 553 463 373 283	27	49	0.56	0.001	N/A
coniferaldehyde	1	1	0.53	0.006	phenylpropanoid
p-coumaric acid	102	236	0.43	0.000	phenylpropanoid
3-*O*-feruloylquinic acid	171	413	0.41	0.007	phenylpropanoid
16.32 327 syringyl lignan	227	562	0.41	0.000	lignan
coniferyl alcohol	4	12	0.36	0.000	phenylpropanoid
16.11 327 297 syringyl lignan	23	66	0.35	0.002	lignan
16.76 354 482 439 323 297 lignan	1	3	0.32	0.000	lignan
Syringing	3	10	0.31	0.000	phenylpropanoid
16.82 354 456 203 188	0	1	0.30	0.002	N/A

Table 4-1 continued

Metabolite	MYB4-OX	Control	MYB4OX/ Ctrl	*P*-value	Classification
16.06 297 guaiacyl lignan	644	2184	0.30	0.000	lignan
9.99 275	1	4	0.20	0.000	N/A
p-hydroxybenzaldehyde	4	23	0.18	0.000	phenylpropanoid
guaiacylglycerol	45	287	0.16	0.000	phenylpropanoid
15.84 412 323 297 209 lignan	29	207	0.14	0.000	lignan
15.12 518 shikimic acid conjugate	0	2	0.13	0.001	organic acid
16.62 486 576 546 456 209 lignan	9	72	0.13	0.000	lignan

Table 4-2 Sugar composition (mg/g dry biomass) of whole tiller biomass determined by quantitative saccharification assay

Lines	Glucose	Xylose	Galactose	Arabinose	Total sugars	Glucose (% of control)
Control (ST1)	339.2±7.1	158.0±10.1	20.8±0.5	29.4±0.4	549.3±17.1	100
L1	383.9±1.8*	161.8±1.1	28.2±0.5*	36.1±0.2*	613.1±3.2*	113.2
L6	319.7±12.2	129.0±13.5	16.9±0.7*	30.2±1.7	497.8±27.8	94.3
L8	351.3±1.4	157.8±4.2	16.6±0.1*	29.3±0.7	557.5±6.3	103.6
Control (ST2)	343.1±2.6	190.0±4.2	17.7±0.3	30.2±0.3	583.2±3.9	100
1A	320.2±25.4	142.6±24.7	25.0±1.6*	34.2±4.4	524.9±56.0	93.3
1B	359.0±1.4*	178.5±1.7*	24.7±0.1*	35.3±0.1*	601.1±3.0*	104.6
1C	353.6±3.3	169.1±0.9*	25.1±0.5*	34.5±0.2*	585.5±4.5	103.1
1D	344.1±3.3	179.5±2.4*	24.0±0.5*	34.1±0.2*	584.5±6.3	100.3
1E	371.8±3.7*	190.8±2.2	21.7±0.1*	34.0±0.4*	620.8±6.0*	108.4

Control for ST1 lines is the average of control lines L9 and L10. Control for ST2 lines is the average of control lines 2A and 2B. All data are means ±SE (n=6 for control lines and n=3 for transgenic lines). Asterisks indicate values that were determined by the Student t-test to be significantly different from their equivalent control ($p<0.05$).

Table 4-3 Levels of total sugars (mg/mg dry cell wall residues) from control and MYB4-OX switchgrass after removal of soluble sugars and starch

Lines	Glucose	Xylose	Galactose	Arabinose	Total sugars	Glucose (% of control)
Control (ST1)	468.2±8.6	254.0±2.2	12.8±0.6	24.1±0.7	759.0±9.6	100
L1	474.0±7.9	261.4±0.5	14.7±0.1	29.7±0.2	779.8±8.8	101.2
L6	477.1±2.1	261.2±1.3.	14.5±0.6	28.7±1.9	781.5±3.3	101.9
L8	472.3±5.1	259.2±0.2	13.0±0.7	25.8±2.3	770.3±2.3	100.9
Control (ST2)	426.1±5.7	270.5±3.2	12.1±0.3	25.0±0.9	733.7±8.0	100
1A	438.1±1.8	257.9±0.2	16.2±0.3	33.2±1.3	745.3±0.4	102.8
1B	443.9±1.8*	270.1±0.3	12.6±0.1	30.9±0.5	757.5±0.9	104.2
1C	448.2±0.0*	264.5±2.2	13.5±0.3	31.8±0.5	758.0±1.4	105.2
1D	444.8±7.2	274.5±4.3	13.1±1.1	31.3±3.9	763.7±6.5	104.4
1E	423.8±3.5	271.5±4.1	13.3±0.1	33.5±0.4	742.0±7.3	99.5

No significant differences among lines were detected. Control for ST1 lines is the average of control lines L9 and L10. Control for ST2 lines is the average of control lines 2A and 2B. All data are means ±SE (n=4 for control lines and n=2 for transgenic lines). Asterisks indicate values that were determined by the Student t-test to be significantly different from their equivalent control ($p<0.05$).

Table 4-4 Levels of pectin extracted from control and MYB4-OX switchgrass

Lines	Water extracted	Na-AC/EDTA extracted	HCl extracted	Total pectin
2A (control)	0.73±0.05	0.33±0.14	5.93±0.08	6.98±0.18
1C	0.93±0.04 *	0.52±0.13 *	6.97±0.17 *	8.42±0.22 *
1D	0.77±0.07	0.50±0.13	7.10±0.27 *	8.37±0.34 *
L10 (control)	0.96±0.04	0.40±0.01	4.82±0.12	6.14±0.10
L1	1.04±0.06	0.49±0.23	7.13±0.08 *	8.66±0.32 *

All data are means ±SE (n=3). Asterisks indicate values that were determined by the Student t-test to be significantly different from their equivalent control ($p<0.05$).

Table 4-5 Plant cell wall glycan-directed monoclonal antibodies (mAbs) used for glycome profiling analyses.

Glycan groups	mAb names
Non-Fucosylated Xyloglucan-1	CCRC-M95 CCRC-M101
Non-Fucosylated Xyloglucan-2	CCRC-M104 CCRC-M89 CCRC-M93 CCRC-M87 CCRC-M88
Non-Fucosylated Xyloglucan-3	CCRC-M100 CCRC-M103
Non-Fucosylated Xyloglucan-4	CCRC-M58 CCRC-M86 CCRC-M55 CCRC-M52 CCRC-M99
Non-Fucosylated Xyloglucan-5	CCRC-M54 CCRC-M48 CCRC-M49 CCRC-M96 CCRC-M50 CCRC-M51 CCRC-M53
Non-Fucosylated Xyloglucan-6	CCRC-M57
Fucosylated Xyloglucan	CCRC-M102 CCRC-M39 CCRC-M106 CCRC-M84 CCRC-M1
Xylan-1/XG	CCRC-M111

Table 4-5 continued

Glycan groups	mAb names
	CCRC-M108 CCRC-M109
Xylan-2	CCRC-M119 CCRC-M115 CCRC-M110 CCRC-M105
Xylan-3	CCRC-M117 CCRC-M113 CCRC-M120 CCRC-M118 CCRC-M116 CCRC-M114
Xylan-4	CCRC-M154 CCRC-M150
Xylan-5	CCRC-M144 CCRC-M146 CCRC-M145 CCRC-M155
Xylan-6	CCRC-M153 CCRC-M151 CCRC-M148 CCRC-M140 CCRC-M139 CCRC-M138
Xylan-7	CCRC-M160 CCRC-M137 CCRC-M152 CCRC-M149
Galactomannan-1	CCRC-M75 CCRC-M70

Table 4-5 continued

Glycan groups	mAb names
	CCRC-M74
Galactomannan-2	CCRC-M166 CCRC-M168 CCRC-M174 CCRC-M175
Acetylated Mannan	CCRC-M169 CCRC-M170
β-Glucan	LAMP BG1
HG Backbone-1	CCRC-M131 CCRC-M38 JIM5
HG Backbone-2	JIM136 JIM7
RG-I Backbone	CCRC-M69 CCRC-M35 CCRC-M36 CCRC-M14 CCRC-M129 CCRC-M72
Linseed Mucilage RG-I	JIM3 CCRC-M40 CCRC-M161 CCRC-M164
Physcomitrella Pectin	CCRC-M98 CCRC-M94
RG-Ia	CCRC-M5 CCRC-M2
RG-Ib	JIM137

Table 4-5 continued

Glycan groups	mAb names
RG-Ib	JIM101
	CCRC-M61
	CCRC-M30
RG-Ic	CCRC-M23
	CCRC-M17
	CCRC-M19
	CCRC-M18
	CCRC-M56
	CCRC-M16
RG-I/Arabinogalactan	CCRC-M60
	CCRC-M41
	CCRC-M80
	CCRC-M79
	CCRC-M44
	CCRC-M33
	CCRC-M32
	CCRC-M13
	CCRC-M42
	CCRC-M24
	CCRC-M12
	CCRC-M7
	CCRC-M77
	CCRC-M25
	CCRC-M9
	CCRC-M128
	CCRC-M126
	CCRC-M134
	CCRC-M125
	CCRC-M123
	CCRC-M122
	CCRC-M121
	CCRC-M112
	CCRC-M21
	JIM131
	CCRC-M22
	JIM132

Table 4-5 continued

Glycan groups	mAb names
RG-I/Arabinogalactan	JIM1
	CCRC-M15
	CCRC-M8
	JIM16
Arabinogalactan-1	JIM93
	JIM94
	JIM11
	MAC204
	JIM20
Arabinogalactan-2	JIM14
	JIM19
	JIM12
	CCRC-M133
	CCRC-M107
Arabinogalactan-3	JIM4
	CCRC-M31
	JIM17
	CCRC-M26
	JIM15
	JIM8
	CCRC-M85
	CCRC-M81
	MAC266
	PN 16.4B4
Arabinogalactan-4	MAC207
	JIM133
	JIM13
	CCRC-M92
	CCRC-M91
	CCRC-M78
Unidentified	MAC265
	CCRC-M97

The groupings of antibodies are based on a hierarchical clustering of ELISA data generated from a screen of all mAbs against a panel of plant polysaccharide preparations1,2 that groups the mAbs according to the predominant polysaccharides that they recognize. The majority of listings link to the WallMabDB plant cell wall monoclonal antibody database (http://www.wallmabdb.net) that provides detailed descriptions of each mAb, including immunogen, antibody isotype, epitope structure (to the extent known), supplier information, and related literature citations.

Chapter 5: Overexpression of switchgrass sucrose synthase (*PvSUS1*) increases plant biomass in transgenic switchgrass (*Panicum virgatum* L.)

(This chapter has been submitted to Plant Cell Reports with the following authors Charleson R. Poovaiah, Mitra Mazarei, Stephen R Decker, Geoffrey B Turner, Robert W Sykes, Mark F Davis and C. Neal Stewart Jr.)

Charleson R Poovaiah designed and conducted the experiments except lignin and sugar release assays and wrote the manuscript

5.1 Abstract

Sucrose synthase converts sucrose and uridine di-phosphate (UDP) into UDP-glucose and fructose. UDP-glucose is used by the cellulose synthase biosynthesis machinery to produce cellulose for cell walls. For lignocellulosic feedstocks such as switchgrass (*Panicum virgatum* L.), the manipulation of cell walls is needed to reduce recalcitrance of conversion of biomass into biofuels. Four sucrose synthase (SUS) genes were identified in switchgrass by bioinformatic analysis of genomics data. When the sequence of the derived amino acids were compared with other plant SUS genes, the cluster analysis showed that the SUS isoforms were clustered into the three SUS groups, and each gene is clustered with its homolog from other monocot species. Each switchgrass SUS (*PvSUS*) gene contains 14 or 15 introns, similar to the structure of homologs in other species. *PvSUS1* was expressed ubiquitously in the tissues tested. *PvSUS2* and *PvSUS6* were highly expressed in internodes and root respectively. *PvSUS4* was expressed in low levels in all the tissues tested. cDNA was isolated and characterized. Transient expression analysis of *PvSUS1*for subcellular localization using a fluorescent protein marker showed that the *Pv*SUS1 protein was localized to the plant plasma membrane. Transgenic switchgrass plants overexpressing *PvSUS1* had increased in biomass and cellulose content. For switchgrass and other bioenergy feedstocks, the overexpression of *SUS1* genes might be a feasible strategy to increase both plant biomass and cellulose content to increase biofuel production per land area cultivated.

Keywords Sucrose synthase, Switchgrass, Gene structure, Subcellular localization, Overexpression, Biomass

5.2 Introduction

In higher plants photosynthates are transported from source tissues to sink tissues, mostly in the form of sucrose. Sucrose in cleaved by the enzymes invertase and sucrose synthase (SUS) in sink tissues to yield glucose and fructose which is used in various metabolic processes. Sucrose synthase (SUS), a glycosyltransferase, catalyzes the reversible conversion of sucrose to uridine di-phosphate (UDP) to UDP-glucose and fructose (Geigenbergere and Stitt 1993). The role of SUS in higher plants in phloem loading and unloading and in carbon partitioning for storage and metabolism in the plant cell has been well-characterized in many species (Haigler, et al. 2001; Ruan, et al. 2008). The breakdown of sucrose by SUS is correlated with the sink strength of starch storage organs (Fu and Park 1995; Zrenner, et al. 1995). Activity of sucrose synthase correlates with sink strength of organs storing carbohydrates such as potato tubers (Zrenner, et al. 1995) and maize kernels (Chourey, et al. 1998).

Sucrose synthase can be accumulated in the cytosol or associated with the plasma membrane where it is linked to the synthesis of cell wall polysaccharides callose and cellulose (Amor, et al. 1995). Sucrose synthase has been found to accumulate to high levels along the plasma membrane and in microtubules beneath the secondary cell wall thickenings in tracheary elements (Salnikov, et al. 2001). *ZmSUS1*, one of the membrane-associated forms of SUS, from maize, is primarily expressed during stem elongation (Duncan, et al. 2006). Sucrose synthase has been identified as an integral part of the cellulose synthase machinery and the catalytic unit that is part of the plasma membrane associated rosette structure involved in cellulose synthesis (Fujii, et al. 2010). All four identified SUS genes in barley identified to date exist in both soluble and membrane associated forms, although their exact functions have not been elucidated (Barrero-Sicilia, et al. 2011).

Overexpression of *SUS* genes has been established to increase cellulose content in *Acetobacter xylinum* (Nakai, et al. 1999) and poplar (Coleman, et al. 2009). When overexpressed in tobacco (Coleman, et al. 2006) and cotton (Jiang, et al. 2012), the transgenic plants had increased biomass. Conversely, suppression of SUS decreased cellulose content in carrots (Tang and Sturm 1999). However, the membrane association of these particular gene isoforms has not been established.

In our study, we identified four SUS genes in the lignocellulosic biofuel feedstock switchgrass (*Panicum virgatum* L.) and compared their sequences to previously identified SUS genes. Subcellular localization of PvSUS1 was studied using a translational fusion construct to a fluorescent protein gene. Stable transgenic switchgrass plants were made whereby overexpression of *PvSUS1* yielded plants with increased biomass and cellulose content. Our findings indicate PvSUS1 is membrane associated and could play a role in cellulose synthesis, which renders it a good candidate for improving biomass and biofuel yield for sustainable energy production.

5.3 Material and methods

5.3.1 Bioinformatic and cluster analyses

Rice SUS sequences were used for BLAST searches against the switchgrass sequence database at Phytozome (www.phytozome.net) to identify putative switchgrass members of the SUS gene family. Exon-intron structure was predicted by comparing genomic sequences with mRNA and cDNA sequences. To identify conserved domains in SUS proteins, a search was done with InterPro program using the Pfam database (Bateman, et al. 2002).

To enable cluster analysis, amino acid sequences of SUS genes were obtained from Genbank. Protein accession numbers of SUS genes from *Eucalyptus*, poplar, sugarcane, potato, beet, cotton, pea, *Arabidopsis*, *Phaseolus*, *Vigna*, lolium, bamboo, maize, rice, barley, tomato and *Physcomitrella* were used: EgSUS1 (ABB53601), EgSUS3 (ABB53602), PtrSUS1 (AAR03498), PtrSUS2 (ABW76422), SoSUS2 (AAM68126), StSUS (AAO67719), StSUS2 (AAO34668), BvSUS (AAK65960), GhSUS1 (ACV72640), GaSUS3 (AAD28641), Ps-NeSUS (AAC28107), MsSUS (AAC17867), BvSUS2 (AAR19769), AtSUS1-6 (NP_197583, NP_199730, NP_192137, NP_566865, NP_198534, NP_177480), PvSUS (ACP17902), VaSUS (BAH56282), SbSUS2 (ACM69042), CcSUS2 (ABI17891), MeSUS (ABD96570), PtrSUS2 (ABW76422), LpSUS (BAE79815), BoSUS1-4 (AAV64256, AAL50571, AAL50570, AAL50572), ZmSUS1 (NP_001105323), ZmSUS2 (NP_001105194), ZmSUS-SH1 (NP_001266691), OsSUS1-7 (P31924, P30298, Q43009, Q10LP5, H6TFZ4, Q6K973, Q7XNX6), HvSUS1-4 (P31922, P31923, CAZ65725, AEH16642), Le-FSUS (AAA34196), PpSUS (XP_001781447). The amino acid sequences were aligned using the online version of MAFFT v 7.0 (Katoh and Standley 2013). Model testing for producing the best-fit model was done using ProtTest v 2.4 (Abascal, et al. 2005). The maximum likelihood tree was constructed by PhyML (Guindon, et al. 2010) using the model identified by ProtTest and bootstrapping with 100 trees. The resulting tree was visualized in FigTree v.1.4.0 (http://tree.bio.ed.ac.uk/software/figtree) and rooted using *Physcomitrella patens* as the outgroup.

5.3.2 Cloning of *PvSUS1*

RNA was extracted from E2- E3 stage stem and leaf of switchgrass cultivar Alamo using TRIREAGENT (Molecular Research Center, Inc. Cincinnati, OH) following manufacturers'

guidelines. Complementary DNA was synthesized using random hexamers and oligo dT using iScript™ cDNA synthesis kit (Bio-Rad Laboratories, Inc. Hercules CA) after DNAse treatment. PvSUS1 was amplified from the cDNA using the primers CP33 and CP34 (Table 5-1). *PvSUS1* was cloned into pCR8-GW-TOPO vector (Life Technologies, Carlsbad, CA) for sequence verification. *PvSUS1* was subsequently cloned into pANIC-10A vector (Mann, et al. 2012) through GATEWAY recombination reaction for overexpression in switchgrass.

5.3.3 Subcellular localization experiments

To construct plasmids for subcellular localization *PvSUS1* was amplified with restriction sites *HindIII* and *BamHI* using the primers CP210 and CP211 (Table 5-1) and then subcloned into the vector pSAT6-RFP-C1 (Tzfira, et al. 2005) at *HindIII* and *BamHI* sites to generate the plasmid pSAT6-RFP-C1-*PvSUS1*. The procedure for transient expression in living onion epidermal cells using particle bombardment was done as described previously (Varagona, et al. 1992) with slight modifications. Briefly, 10 µg of pSAT6-RFP-C1, subcellular localization control plasmid DNA and 10 µg pSAT6-RFP-C1-*PvSUS1* were coated on 0.6 um gold particles, bombarded into onion epidermal cells and incubated for 48 hours before observation under the epifluorescence microscope using Texas Red filters (Em-D560/40, BS-595DCLP, EM-D630/60). The cells were plasmolyzed with one or two drops of 40% sucrose solution and visualized after 30 minutes.

5.3.4 Switchgrass transformation and plant growth

Switchgrass callus generated from immature inflorescence was transformed with a binary vector overexpressing *PvSUS1* as described in Xi et al., (2009). Embryogenic calli induced from immature inflorescences of switchgrass cultivar Alamo genotype ST1 (Shen, et al. 2013), was co-cultivated with *Agrobacterium tumefaciens* containing the binary vector pANIC10A-*PvSUS1*

(Mann, et al. 2012). Transgenic callus was selected on MS medium (Murashige and Skoog 1962) supplemented with 60 mg/L hygromycin (EMD Millipore, Billerica, MA). The transgenic hygromycin resistant callus was regenerated on MS medium supplemented with 0.9 µM kinetin. Regenerated shoots were transferred to a hormone free medium for root development. The regenerated plants were PCR-screened using the primers CP43 and CP44 (Table 5-1) to assay for the presence of the insert. The plants were transferred to 12 L pots in the greenhouse after acclimatization. For plant biomass experiments, plants were grown from a single tiller in 4x4x6 pots in triplicate at the same growth stage. The plants were grown from 4 months in the greenhouse before all above ground biomass was harvested. Plant height and tiller number were taken before harvesting. Plant biomass was dried in the oven for 3 days at 40 ^{o}C and weighed. The dry biomass was milled to 20 mesh using a Wiley mill for lignin and sugar release assays.

5.3.5 Quantification of transgene expression levels by qRT-PCR

RNA was extracted from leaves at E4 stage from transgenic switchgrass plants overexpressing *PvSUS1* and control switchgrass plants using TRIREAGENT (Molecular Research Center Inc., Cincinnati, Ohio) according to manufacturer's instructions. Following extraction, 5 µg of total RNA was treated with 1 unit of DNase I (Promega Corporation, Madison, Wisc.) in a 10 µl reaction volume. The reaction was incubated at 37 ^{o}C for 30 min and then incubated at 65 ^{o}C with 1 µl stop solution for 1 min to inactivate the DNase.

For expression analysis of *PvSUS* genes, RNA was extracted from leaf blade, leaf sheath, internode, node and panicle of R0 stage switchgrass plants. One microgram of RNA was used for synthesizing cDNA after DNase treatment. Complementary DNA was synthesized using a High

Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, Calif.). Quantitative PCR was performed on the ABI HT7900 instrument (Applied Biosystems) with the Power SYBR I kit for PCR (Applied Biosystems) according to manufacturer's instructions. Each reaction was performed with 1 µl of 1:100 dilutions of the first cDNA strands in a total volume of 10 µl. The reactions were incubated at 95 ^{o}C for 10 min to activate the hot start recombinant polymerase, followed by 40 cycles of 10 s at 95C, 30 s at 56 ^{o}C, and 30 s at 72 ^{o}C. The specificity of the PCR amplification was checked with a heat dissociation protocol (from 60 °C to 95 °C) following the final cycle of the PCR. The results obtained for the different samples analyzed were standardized to the constitutive *PvActC* gene expression level (Mann, et al. 2011). The primers used for each gene is listed in Table 1. Data was analyzed with one-way ANOVA, and letter groupings were obtained using Fisher's least significant difference method.

5.3.6 Lignin and sugar release assays

Total lignin content was determined by high throughput pyrolysis molecular beam mass spectrometry. Briefly ~4 µg of dry biomass was loaded in 80 µl stainless steel cups and pyrolyzed at 500˚C in a quartz reactor using a Frontier py2020 autosampler. The resulting pyrolysis vapors are analyzed using a custom Extrel single quadrapole molecular beam mass spectrometer (Sykes, et al. 2009).

Sugar release efficiency was determined by high throughput screening as described by Selig, et al. (2010). Briefly, dry, milled biomass (-20 mesh) was dispensed in triplicate into custom 96 well Hastelloy reactor plates. Biomass-only controls, sugar standards, enzyme-only and blank wells were included in each plate. The total capacity of each well in the plate was 417 µl.

Subsequently, 300 µl water was added, and the plates were sealed with PTFE film tape containing a silicone-based adhesive (3M Corporation, St. Paul, Minn.), stacked and pretreated in a steam chamber to 180C for 40 min. The plates were rapidly cooled with water, centrifuged and the tape was removed in order to add 40 µl enzyme in citrate buffer (1.0 M, pH 5.0) to each well. Enzyme cocktail consisting of Novozymes Cellic CTec2 cellulase (Novozymes North America, Franklinton N.C.) and Novo188 β-glucosidase (Novozymes) were loaded at 70 mg protein/g initial biomass, and 2.5 mg protein /g initial biomass respectively. The plates were resealed and incubated for 3 days at 54ºC. The plates were analyzed for glucose and xylose release using enzyme linked sugar assay kits (Megazyme International).

5.3.7 Histology

Tissues from the 2nd internode of R1 stage plants were fixed and embedded with JB4 embedding kit (Polysciences, Inc. Warrington, PA) and sectioned to 7 µm with a microtome using glass knives. The tissues were stained with one drop of calcofluor white (comprising 1 g/l calcofluor White M2R and 0.5 g/l Evans blue; Sigma-Aldrich, St. Louis, MO, USA), and one drop of 10% potassium hydroxide was applied on top of the slides, a cover slip was applied. The slide was then left to stand for 1 – 2 min before being examined by fluorescence microscopy using blue light excitation and DAPI filters (Ex-D360/40, BS-400DCLP, and Em-D460).

5.4 Results

5.4.1 Phylogenetic analysis of SUS proteins

In order to identify SUS homologs in switchgrass, deduced amino acid sequences of rice and maize SUS genes (Duncan, et al. 2006; Hirose, et al. 2008) were used to search the Phytozome

switchgrass database for cDNA sequences using BLAST. Four switchgrass SUS sequences *PvSUS1* (Pavirv00026879m), *PvSUS2* (Pavirv00046631m), *PvSUS4* (Pavirv00031455m) and *PvSUS6* (Pavirv00054750m) were identified. The identified model LG+I+G for the tree, with a gamma distribution of 0.959 and proportion of invariable sites of 0.137, was selected using Akaike information criterion. SUS sequences from monocots and dicots fell into distinct groups as described by Komatsu, et al. (2002) and Barrero-Sicilia, et al. (2011). Cluster and sequence analyses indicated the presence four different SUSs in switchgrass (Figures 5-1, 5- 7). Monocot group 1 was comprised of *PvSUS1* and *PvSUS2* along with orthologs from maize *ZmSUS1* and *ZmSUS*-SH1, rice *OsSUS1*, *OsSUS2* and *OsSUS3* and barley *HvSUS1* and *HvSUS2*. SUS group 2 contained *PvSUS4* along with *ZmSUS4* from maize, *OsSUS4* from rice and *HvSUS3* from barley. SUS group 3 was comprised of *PvSUS6* and orthologous sequences from rice *OsSUS5*, *OsSUS6* and *OsSUS7* and from barley *HvSUS4*. Each of the SUS groups contained at least one member from switchgrass. Switchgrass sequence will be deposited into Genbank.

5.4.2 Characterization of switchgrass sucrose synthases

PvSUS1 clustered with *ZmSUS1*, *OsSUS1* and *OsSUS3* (98.4%, 95.6% and 90.5% identity respectively). *PvSUS2* appeared to be closely related to *ZmSUS-SH1* and *OsSUS2* (87.5% and 84.7% respectively). A high percentage of amino acid identity was found between *PvSUS4, ZmSUS2* and *OsSUS4* (94.6% and 93.7% respectively) while identity percentage of *PvSUS6* with *OsSUS7*, *OsSUS5* and *OsSUS6* was 89.3%, 89.1% and 80.8%. The identity of *PvSUS1* with *PvSUS2, PvSUS4* and *PvSUS6* was 71.9%, 70.5% and 53.8% respectively (Table 5-1). The molecular mass of *PvSUS1*, *PvSUS2*, and *PvSUS4* were deduced to be approximately 92 kDa

whereas *PvSUS6* had a higher molecular mass of 97 kDa. Amino acid sequence analysis revealed all 4 switchgrass SUSs share the SUS domain and glycosyltransferase domain (Table 5-2).

The predicted intron/exon structures of the SUS genes were analyzed between the 5` untranslated region (UTR) and 3` UTR including the noncoding exon. Comparison between the genomic sequence and the cDNA sequence revealed that all the genes have 14 or 15 coding exons. The exon-intron structure is presented as a schematic representation in Figure 5-2. Almost all the introns are similarly distanced, although they are not of the same size. *PvSUS1* and *PvSUS2* have large introns in the 5` non-coding region and *PvSUS1* has a 1.34 kb intron between $4^{th}$ and $5^{th}$ exons. Non-coding exons are absent in both *PvSUS4* and *PvSUS6*. In addition, in *PvSUS6* the 3` non-coding region is split by an intron.

5.4.3 Expression analysis of *PvSUS* genes

The relative expression profile of the four members of *PvSUS* gene family were investigated in R0 stage switchgrass. Specific primers were designed to determine the expression profile of each member of the family in different organs of the plant. The relative transcription levels of the four *PvSUS* genes were standardized to *PvActinC* (ActC) expression level (Mann, et al. 2011). The detection of all four members of the *PvSUS* gene family indicated that all are transcriptionally active. In the set of organs tested, *PvSUS1* seems to be expressed ubiquitously with higher expression levels in internodes, nodes and roots. Transcript level of *PvSUS2* was high in internodes (Figure 5-3A and B). The expression of *PvSUS4* is relatively high in roots compared to other tissues tested; expression is lower level in all tissues relative to the other *PvSUS* genes. *PvSUS6* was especially highly expressed in roots (Fig 5-3C and D).

5.4.4 Sub-cellular localization

We performed subcellular co-localization assay in onion epidermal cells to determine subcellular distribution of *PvSUS1* and its association with the plasma membrane. The open reading frame of *PvSUS1* was translationally fused to the DsRed fluorescent protein (RFP), was cloned to be under the control of CaMv35S promoter (pSAT6-RFP-C1-*PvSUS1*), and transiently expressed in onion epidermal cells. The plasmid pSAT6-RFP-C1 containing 35S-*RFP* was used as control for subcellular localization. RFP fluorescence was seen throughout the cytoplasm in cells with control plasmid (Figure 5-4B) while in RFP-tagged *PvSUS1* fluorescence was seen predominantly along the plasma membrane with some expression in the cytoplasm (Figure 5-4H). When epidermal cells were plasmolyzed with 40% sucrose solution, fluorescence was observed throughout the cytoplasm in cells with control plasmid (Figure 5-4E) and fluorescence was seen predominantly along the plasma membrane in cells with RFP tagged *PvSUS1* (Figure 5-4K) along with low levels of fluorescence in the cytoplasm.

5.4.5 *PvSUS1* overexpression in switchgrass

To generate *PvSUS1* overexpression (PvSUS1-OE) lines, full length *PvSUS1* open reading frame was cloned in to pANIC-10A binary vector by Gateway recombination under the control of the maize ubiquitin 1 promoter, and fused in-frame at the C-terminal with AcV5 epitope tag (Monsma and Blissard 1995; Mann et al. 2012). Switchgrass lines overexpressing *PvSUS1* were generated by *Agrobacterium*-mediated transformation (Xi, et al. 2009). Five independent transgenic lines were regenerated from hygromycin-resistant calli. Genomic PCR confirmed that all plants were transgenic (data not shown). Quantitative PCR (qPCR) demonstrated that the *PvSus1* was overexpressed between 2- and 8-fold increase in the transgenic plants (Figure 5-5B).

The level of expression of the endogenous *PvSUS1* was not affected and was similar to expression levels in non-transgenic control (Figure 5-5B).

We examined plant growth parameters height and tiller number in the *PvSUS1* transgenic plants. Plant height was significantly increased in line 2 (37%) compared to the control, whereas it decreased significantly in lines 1 and 5 (21% and 12% respectively) (Figure 5-5C). There was no significant difference in the other two lines compared to control plants. Line 5 had a significant increase in number of tillers (79%) compared to control, whereas other lines were not significantly different from control (Figure 5-5D). Differences on dry weight basis were significant in lines 2, 4 and 5 compared to control (Figure 5-5E). Percentage of increase in dry weight in lines 2, 4 and 5 was 13.6%, 11.5% and 10.6% respectively.

Various cell wall analyses provide insight into how overexpression of *PvSUS1* could be used to alter cell walls for potentially improved biofuels. Lignin content increased in switchgrass plants overexpressing *PvSUS1* compared to control plants (Figure 5-9A). The S:G monomer lignin ratio also increased in the transgenic plants compared to control (Figure 5-9B). All transgenic plants had lower glucose release (Figure 5-9C) efficiency compared to control plants. Xylose release efficiency was not significantly different from control plants (Figure 5-9D).

Staining with calcoflor shows increased fluorescence in *PvSUS1* plants from an increase in deposition of cellulose in cell walls. Additionally, the cell walls appeared to be thicker in transgenic relative to controls (Figure 5-6). We quantified the transcript level of fructokinase in the transgenic plants to determine if fructose produced by the breakdown of sucrose was recycled into the hexose-phosphate pool. All lines except line 4 had decreased levels of fructokinase (Figure 5-8).

5.5 Discussion

In this study, we identified switchgrass SUS genes *PvSUS1*, *PvSUS2*, *PvSUS4* and *PvSUS6* that have homologues with other species. By comparison, barley has four SUS genes with *HvSUS1* and *HvSUS2* in group 1, *HvSUS3* in group 2 and *HvSUS4* in group 3 (Barrero-Sicilia, et al. 2011) corresponding to SUS1 (Group 1), SUSA (Group 2) and NG (Group 3) groups (Jiang, et al. 2012; Komatsu, et al. 2002). Although the monocots rice and *Brachypodium* have at least six SUS isoforms, only four SUS isoforms were identified in switchgrass during our search in the Phytozome switchgrass database. Given that lowland switchgrass is a allotetraploid, there are likely additional SUS genes in the species that were not described here.

Expression analysis shows that *PvSUS1* is ubiquitously expressed in all the tissues. *PvSUS1* falls in the same clade as *OsSUS1*, *OsSUS3*, *ZmSUS1* and *HvSUS2*. *PvSUS1* expression is similar to homologous gene *OsSUS1* in rice (Hirose, et al. 2008) where they were highly expressed in internodes. *PvSUS2*, which is similar to *OsSUS2*, *HvSUS2* and *ZmSUS-SH1* is highly expressed in internodes. While *HvSUS2* and *ZmSUS-SH1* are highly expressed in embryos, the expression of *PvSUS2* is similar to that of *OsSUS2* (Hirose, et al. 2008). *PvSUS4* transcript levels were very low in all tissues except roots and were not similar to other homologous genes in rice *OsSUS4* and barley *HvSUS3*. *HvSUS3* was mainly expressed in dry kernels and was induced by dehydration in leaves (Barrero-Sicilia, et al. 2011). Expression of *PvSUS6* is similar to *HvSUS4*, *OsSUS5* and *OsSUS6* in leaf blade and internodes, but *PvSUS6* expression is highest in roots compared to other tissues, in contrast to its homologs.

Membrane associated forms of SUS have been implicated in the synthesis of both callose and cellulose (Amor et al. 1995) and these functions have been identified in maize and barley

(Carlson and Chourey 1996; Barrero-Sicilia et al. 2011). The transmembrane prediction program TMPred predicted the presence of two transmembrane domains in *PvSUS1* which corroborates our subcellular co-localization results. In our plasmolysis experiments after 40% sucrose treatment, the RFP fluorescent signal appeared to be co-localized with the retracted plasma membrane indicating an association with the plasma membrane. SUS has been identified to be a component of the cellulose synthase biosynthesis machinery in Asuki bean (Fujii, et al. 2010) and is expressed along with cellulose synthase during secondary cell wall formation in poplar (Coleman, et al. 2009). Elevation of cellulose levels in plants overexpressing *PvSUS1* along with its association with the plasma membrane supports its purported role in cellulose production and carbon partitioning.

Switchgrass plants overexpressing *PvSUS1* had increased plant biomass. When overexpressed, SUS has been demonstrated to increase plant height and biomass. In cotton, overexpression of SUS increased plant biomass and increased cotton fiber quality and yield (Jiang, et al. 2012) and in tobacco and poplar, overexpression of SUS increased plant height and soluble sugar content (Coleman, et al. 2006; Konishi, et al. 2004). The role of SUS in the control of sink strength is consistent with the increase of plant height and biomass when overexpressed and is corroborated by the reduced plant size in carrots by the suppression of SUS (Tang and Sturm 1999).

Lignin content was increased in the transgenic plants. In poplar and tobacco plants overexpressing SUS, no significant increase in lignin content was observed though there was an increase in biomass (Coleman, et al. 2010; Coleman, et al. 2009). Increases in lignin content seemed to be associated with a decrease in glucose- and xylose-release efficiency.

In the presence of UDP, a plasma membrane associated form of SUS breaks down sucrose to UDP-glucose and fructose. UDP-glucose is a substrate for cellulose synthase to produce cellulose and callose for cell wall synthesis (Amor, et al. 1995). While UDP-glucose is available immediately for cellulose synthesis, fructose is recycled into the hexose-phosphate pool (Konishi, et al. 2004). The first step in this direction is the phosphorylation of fructose to fructose-6-phosphate, catalyzed by the enzyme fructokinase. Accumulation of fructose causes inhibition of SUS activity by feedback inhibition (Granot, et al. 2013). Transcript levels of fructokinase in *PvSUS1*-overexpression transgenic switchgrass were decreased in all lines except line 4 (Figure S3) and was not correlated with the increase in biomass. Anti-sense suppression of SUS did not change levels of fructokinase and glucokinase in carrots (Tang and Sturm 1999). Therefore it is possible that another mechanism might also be present to recycle fructose.

In conclusion, we have identified four isoforms of SUS in switchgrass. Expression analyses indicate that the four isoforms may have specific roles in different tissues. PvSUS1 is most likely to be associated with the plasma membrane. Overexpression of *PvSUS1* increased biomass yield confirming earlier reports of biomass increase in other species. These results have significant implications in improving biomass yield in bioenergy crops.

5.6 Acknowledgements

We thank Joshua Grant for assistance with histology, Jason N. Burris, and Ellen Reeves for assistance with switchgrass tissue culture and transformation. This work was supported by the BioEnergy Science Center, a US Department of Energy Bioenergy Research Center, through the Office of Biological and Environmental Research in the DOE Office of Science.

5.7 References

Abascal F, Zardoya R, Posada D. 2005. ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* 21: 2104-2105.

Amor Y, Haigler CH, Johnson S, Wainscott M, Delmer DP. 1995. A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. *Proceedings of the National Academy of Sciences, USA* 92: 9353-9357.

Barrero-Sicilia C, Hernando-Amado S, Gonzalez-Melendi P, Carbonero P. 2011. Structure, expression profile and subcellular localisation of four different sucrose synthase genes from barley. *Planta* 234: 391-403.

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer ELL. 2002. The Pfam protein families database. *Nucleic Acids Research* 30: 276-280.

Baud S, Vaultier MN, Rochat C. 2004. Structure and expression profile of the sucrose synthase multigene family in *Arabidopsis*. *Journal of Experimental Botany* 55: 397-409.

Cho JI, Kim HB, Kim CY, Hahn TR, Jeon JS. 2011. Identification and characterization of the duplicate rice sucrose synthase genes OsSUS5 and OsSUS7 which are associated with the plasma membrane. *Molecules and Cells* 31: 553-561.

Chourey PS, Taliercio EW, Carlson SJ, Ruan YL. 1998. Genetic evidence that the two isozymes of sucrose synthase present in developing maize endosperm are critical, one for cell wall integrity and the other for starch biosynthesis. *Molecular and General Genetics* 259: 88-96.

Coleman H, Beamish L, Reid A, Park J-Y, Mansfield S. 2010. Altered sucrose metabolism impacts plant biomass production and flower development. *Transgenic Research* 19: 269-283.

Coleman HD, Ellis DD, Gilbert M, Mansfield SD. 2006. Up-regulation of sucrose synthase and UDP-glucose pyrophosphorylase impacts plant growth and metabolism. *Plant Biotechnology Journal* 4: 87-101.

Coleman HD, Yan J, Mansfield SD. 2009. Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proceedings of the National Academy of Sciences, USA* 106: 13118-13123.

Duncan KA, Hardin SC, Huber SC. 2006. The three maize sucrose synthase isoforms differ in distribution, localization, and phosphorylation. *Plant & Cell Physiology* 47: 959-971.

Fu HY, Park WD. 1995. Sink-associated and vascular-associated sucrose synthase functions are encoded by different gene classes in potato. *Plant Cell* 7: 1369-1385.

Fujii S, Hayashi T, Mizuno K. 2010. Sucrose synthase is an integral component of the cellulose synthesis machinery. *Plant & Cell Physiology* 51: 294-301.

Geigenberger P, Stitt M. 1993. Sucrose synthase catalyses a readily reversible reaction in vivo in developing potato tubers and other plant tissues. *Planta* 189: 329-339.

Granot D, David-Schwartz R, Kelly G. 2013. Hexose kinases and their role in sugar-sensing and plant development. *Frontiers in plant science* 4.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307-321.

Haigler CH, Ivanova-Datcheva M, Hogan PS, Salnikov VV, Hwang S, Martin K, Delmer DP. 2001. Carbon partitioning to cellulose synthesis. *Plant Molecular Biology* 47: 29-51.

Hirose T, Scofield GN, Terao T. 2008. An expression analysis profile for the entire sucrose synthase gene family in rice. *Plant Science* 174: 534-543.

Huang JW, Chen JT, Yu WP, Shyur LF, Wang AY, Sung HY, Lee PD, Su JC. 1996. Complete structures of three rice sucrose synthase isogenes and differential regulation of their expressions. *Bioscience Biotechnology and Biochemistry* 60: 233-239.

Jiang Y, Guo W, Zhu H, Ruan YL, Zhang T. 2012. Overexpression of GhSusA1 increases plant biomass and improves cotton fiber yield and quality. *Plant Biotechnology Journal* 10: 301-312.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* 30: 772-780.

Komatsu A, Moriguchi T, Koyama K, Omura M, Akihama T. 2002. Analysis of sucrose synthase genes in citrus suggests different roles and phylogenetic relationships. *Journal of Experimental Botany* 53: 61-71.

Konishi T, Ohmiya Y, Hayashi T. 2004. Evidence that sucrose loaded into the phloem of a poplar leaf is used directly by sucrose synthase associated with various beta-glucan synthases in the stem. *Plant Physiology* 134: 1146-1152.

Mann DGJ, King Z, Liu W, Joyce B, Percifield R, Hawkins J, LaFayette PR, Artelt B, Burris J, Mazarei M, Bennetzen J, Parrott W, Stewart Jr CN. 2011. Switchgrass (*Panicum virgatum* L.) polyubiquitin gene (PvUbi1 and PvUbi2) promoters for use in plant transformation. *BMC Biotechnology* 11: 74.

Mann DGJ, LaFayette PR, Abercrombie LL, King ZR, Mazarei M, Halter MC, Poovaiah CR, Baxter H, Shen H, Dixon RA, Parrott WA, Neal Stewart Jr C. 2012. Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (Panicum virgatum L.) and other monocot species. *Plant Biotechnology Journal* 10: 226-236.

Murashige T, Skoog F. 1962. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum* 15: 473-497.

Nakai T, Tonouchi N, Konishi T, Kojima Y, Tsuchida T, Yoshinaga F, Sakai F, Hayashi T. 1999. Enhancement of cellulose production by expression of sucrose synthase in *Acetobacter xylinum*. *Proceedings of the National Academy of Sciences, USA* 96: 14-18.

Ruan YL, Llewellyn DJ, Liu Q, Xu SM, Wu LM, Wang L, Furbank RT. 2008. Expression of sucrose synthase in the developing endosperm is essential for early seed development in cotton. *Functional Plant Biology* 35: 382-393.

Salnikov VV, Grimson MJ, Delmar DP, Haigler CH. 2001. Sucrose synthase localizes to cellulose synthesis sites in tracheary elements. *Phytochemistry* 57: 823-833.

Selig MJ, Tucker MP, Sykes RW, Reichel KL, Brunecky R, Himmel ME, Davis MF, Decker SR. 2010. Lignocellulose recalcitrance screening by integrated high-throughput hydrothermal pretreatment and enzymatic saccharification. *Industrial Biotechnology* 6: 104-111.

Shen H, Poovaiah CR, Ziebell A, Tschaplinski TJ, Pattathil S, Gjersing E, Engle NL, Katahira R, Pu Y, Sykes R, Chen F, Ragauskas AJ, Mielenz JR, Hahn MG, Davis M, Stewart CN, Jr., Dixon RA. 2013. Enhanced characteristics of genetically modified switchgrass (*Panicum virgatum* L.) for high biofuel production. *Biotechnolology for Biofuels* 6: 71.

Sykes R, Yung M, Novaes E, Kirst M, Peter G, Davis M 2009. High-throughput screening of plant cell-wall composition using pyrolysis molecular beam mass spectroscopy. *Biofuels*: Springer, 169-183.

Tang GQ, Sturm A. 1999. Antisense repression of sucrose synthase in carrot (*Daucus carota* L.) affects growth rather than sucrose partitioning. *Plant Molecular Biology* 41: 465-479.

Tzfira T, Tian GW, Lacroix B, Vyas S, Li J, Leitner-Dagan Y, Krichevsky A, Taylor T, Vainstein A, Citovsky V. 2005. pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. *Plant Molecular Biology* 57: 503-516.

Varagona MJ, Schmidt RJ, Raikhel NV. 1992. Nuclear-localization signal(s) required for nuclear targeting of the maize regulatory protein opaque-2. *Plant Cell* 4: 1213-1227.

Xi Y, Fu C, Ge Y, Nandakumar R, Hisano H, Bouton J, Wang Z-Y. 2009. *Agrobacterium*-mediated transformation of switchgrass and inheritance of the transgenes. *BioEnergy Research* 2: 275-283.

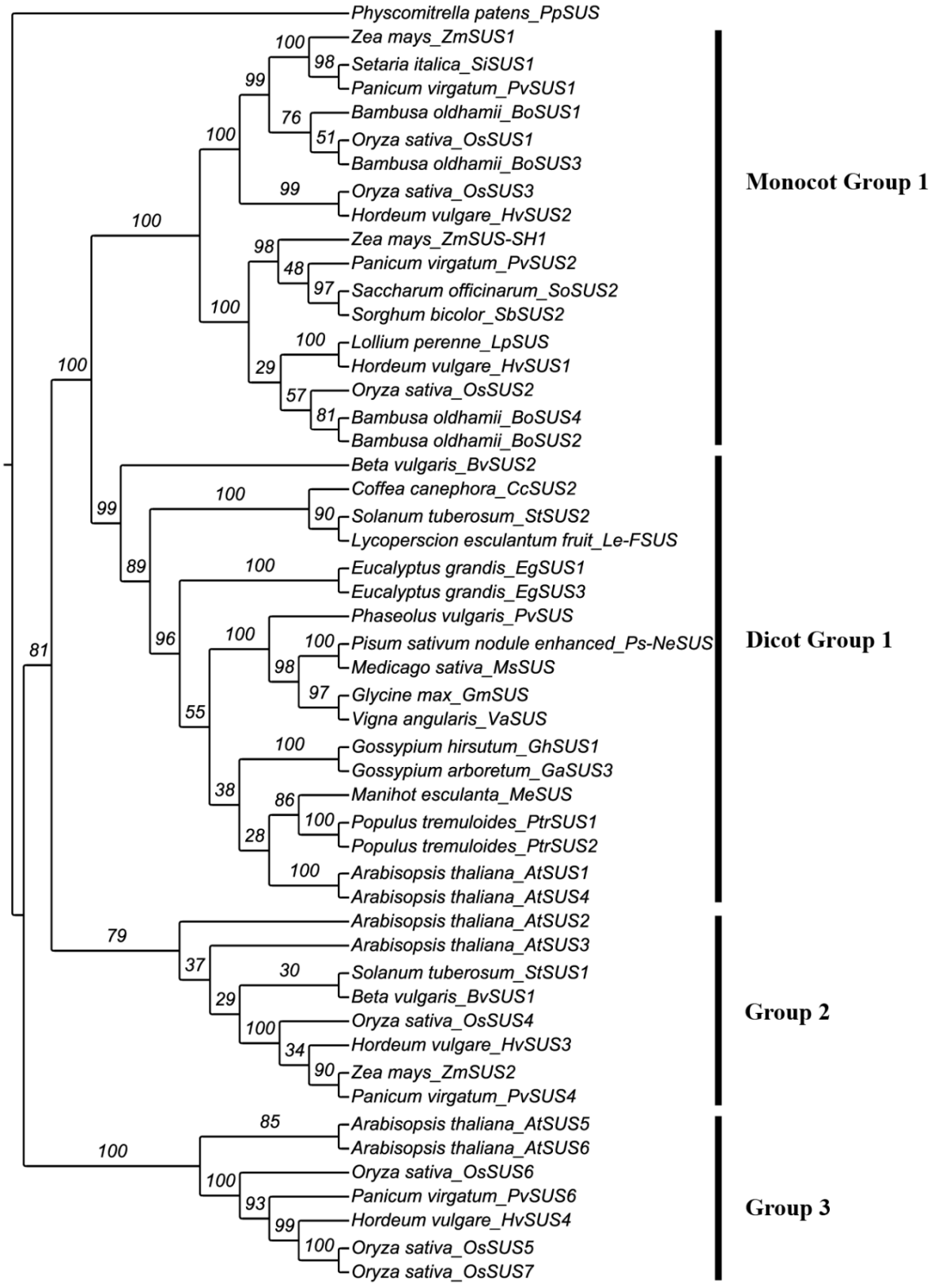
Zrenner R, Salanoubat M, Willmitzer L, Sonnewald U. 1995. Evidence of the crucial role of sucrose synthase for sink strength using transgenic potato plants (*Solanum tuberosum* L). *Plant Journal* 7: 97-107.

5.8 Appendix

5.8.1 Figures and tables

Figure 5-1 Maximum likelihood tree of switchgrass sucrose synthase deduced amino acid sequences and other known sucrose synthase amino acid sequences

Protein accession numbers of SUS genes are from *Eucalyptus*, poplar, sugarcane, potato, beet, cotton, pea, *Arabidopsis*, *Phaseolus*, *Vigna*, *Lolium*, bamboo, maize, rice, barley, tomato and *Physcomitrella* accessions are matched with their node labels: EgSUS1 (ABB53601), EgSUS3 (ABB53602), PtrSUS1 (AAR03498), PtrSUS2 (ABW76422), SoSUS2 (AAM68126), StSUS (AAO67719), StSUS2 (AAO34668), BvSUS (AAK65960), GhSUS1 (ACV72640), GaSUS3 (AAD28641), Ps-NeSUS (AAC28107), MsSUS (AAC17867), BvSUS2 (AAR19769), AtSUS1-6 (NP_197583, NP_199730, NP_192137, NP_566865, NP_198534, NP_177480), PvSUS (ACP17902), VaSUS (BAH56282), SbSUS2 (ACM69042), CcSUS2 (ABI17891), MeSUS (ABD96570), PtrSUS2 (ABW76422), LpSUS (BAE79815), BoSUS1-4 (AAV64256, AAL50571, AAL50570, AAL50572), ZmSUS1 (NP_001105323), ZmSUS2 (NP_001105194), ZmSUS-SH1 (NP_001266691), OsSUS1-7 (P31924, P30298, Q43009, Q10LP5, H6TFZ4, Q6K973, Q7XNX6), HvSUS1-4 (P31922, P31923, CAZ65725, AEH16642), Le-FSUS (AAA34196), PpSUS (XP_001781447). PvSUS1 and PvSUS2 fall in to SUS Group1; PvSUS4 in Group 2 and PvSUS6 in Group3. Bootstrapping values are indicated as node values.


Physcomitrella patens_PpSUS
Zea mays_ZmSUS1
Setaria italica_SiSUS1
Panicum virgatum_PvSUS1
Bambusa oldhamii_BoSUS1
Oryza sativa_OsSUS1
Bambusa oldhamii_BoSUS3
Oryza sativa_OsSUS3
Hordeum vulgare_HvSUS2
Zea mays_ZmSUS-SH1
Panicum virgatum_PvSUS2
Saccharum officinarum_SoSUS2
Sorghum bicolor_SbSUS2
Lollium perenne_LpSUS
Hordeum vulgare_HvSUS1
Oryza sativa_OsSUS2
Bambusa oldhamii_BoSUS4
Bambusa oldhamii_BoSUS2
Beta vulgaris_BvSUS2
Coffea canephora_CcSUS2
Solanum tuberosum_StSUS2
Lycoperscion esculantum fruit_Le-FSUS
Eucalyptus grandis_EgSUS1
Eucalyptus grandis_EgSUS3
Phaseolus vulgaris_PvSUS
Pisum sativum nodule enhanced_Ps-NeSUS
Medicago sativa_MsSUS
Glycine max_GmSUS
Vigna angularis_VaSUS
Gossypium hirsutum_GhSUS1
Gossypium arboretum_GaSUS3
Manihot esculanta_MeSUS
Populus tremuloides_PtrSUS1
Populus tremuloides_PtrSUS2
Arabisopsis thaliana_AtSUS1
Arabisopsis thaliana_AtSUS4
Arabisopsis thaliana_AtSUS2
Arabisopsis thaliana_AtSUS3
Solanum tuberosum_StSUS1
Beta vulgaris_BvSUS1
Oryza sativa_OsSUS4
Hordeum vulgare_HvSUS3
Zea mays_ZmSUS2
Panicum virgatum_PvSUS4
Arabisopsis thaliana_AtSUS5
Arabisopsis thaliana_AtSUS6
Oryza sativa_OsSUS6
Panicum virgatum_PvSUS6
Hordeum vulgare_HvSUS4
Oryza sativa_OsSUS5
Oryza sativa_OsSUS7
Monocot Group 1
Dicot Group 1
Group 2
Group 3

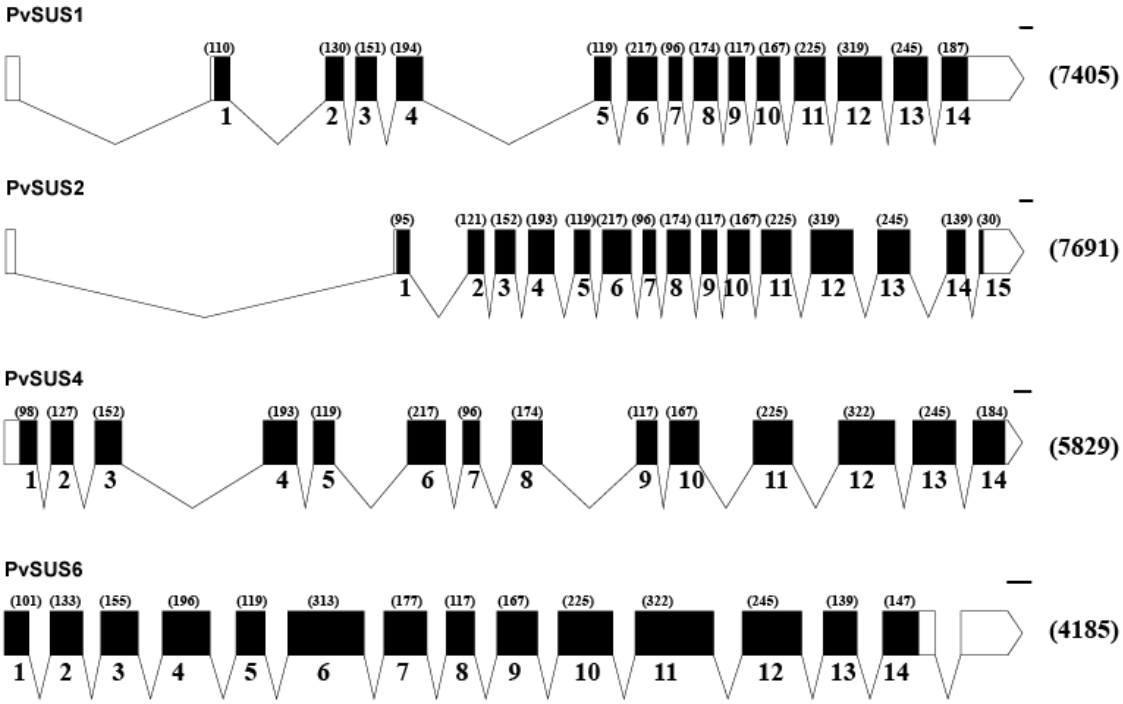
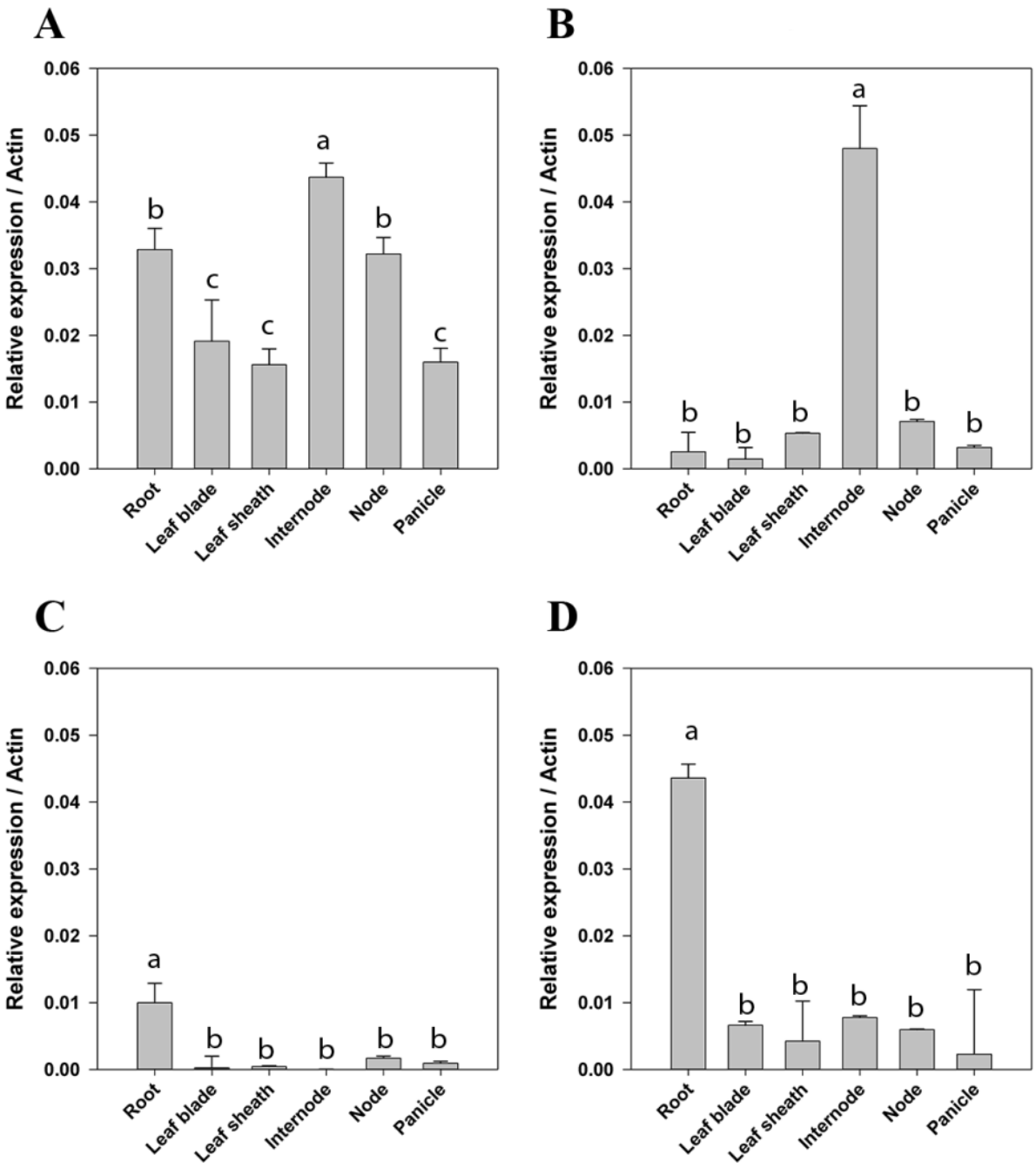


Figure 5-2 Intron-exon structure of switchgrass sucrose synthase genes

Solid black boxes represent coding exons, whereas hollow boxes are non-coding exons. Introns are indicated by lines. Coding exons are numbered below the exons and the size of the coding exons is numbered in parenthesis above the exons. The rightmost numbers represent the number of bases comprising each genomic sequence. Scale bar represents 100 bp.

Figure 5-3 Relative expression profile of the four PvSUS genes in various switchgrass tissues

Relative expression level of A) *PvSUS1* ($p<0.0015$) B) *PvSUS2* ($p<0.0001$) C) *PvSUS4* ($p=0.0023$) and D) *PvSUS6* ($p<0.0001$). Relative transcript levels were measured by qRT-PCR and standardized to the constitutive *PvACTC* gene expression level. The mRNA was extracted from roots, leaf blade, leaf sheath, internodes nodes and panicle at R0 growth stage. All data are means ± SE. Means were analyzed with one-way ANOVA, and letter groupings were obtained using Fisher's least significant difference method. Bars with the same letter do not differ significantly at the 5% level.


A
B
C
D
Relative expression / Actin
0.06
0.05
0.04
0.03
0.02
0.01
0.00
Root
Leaf blade
Leaf sheath
Internode
Node
Panicle
a
b
c

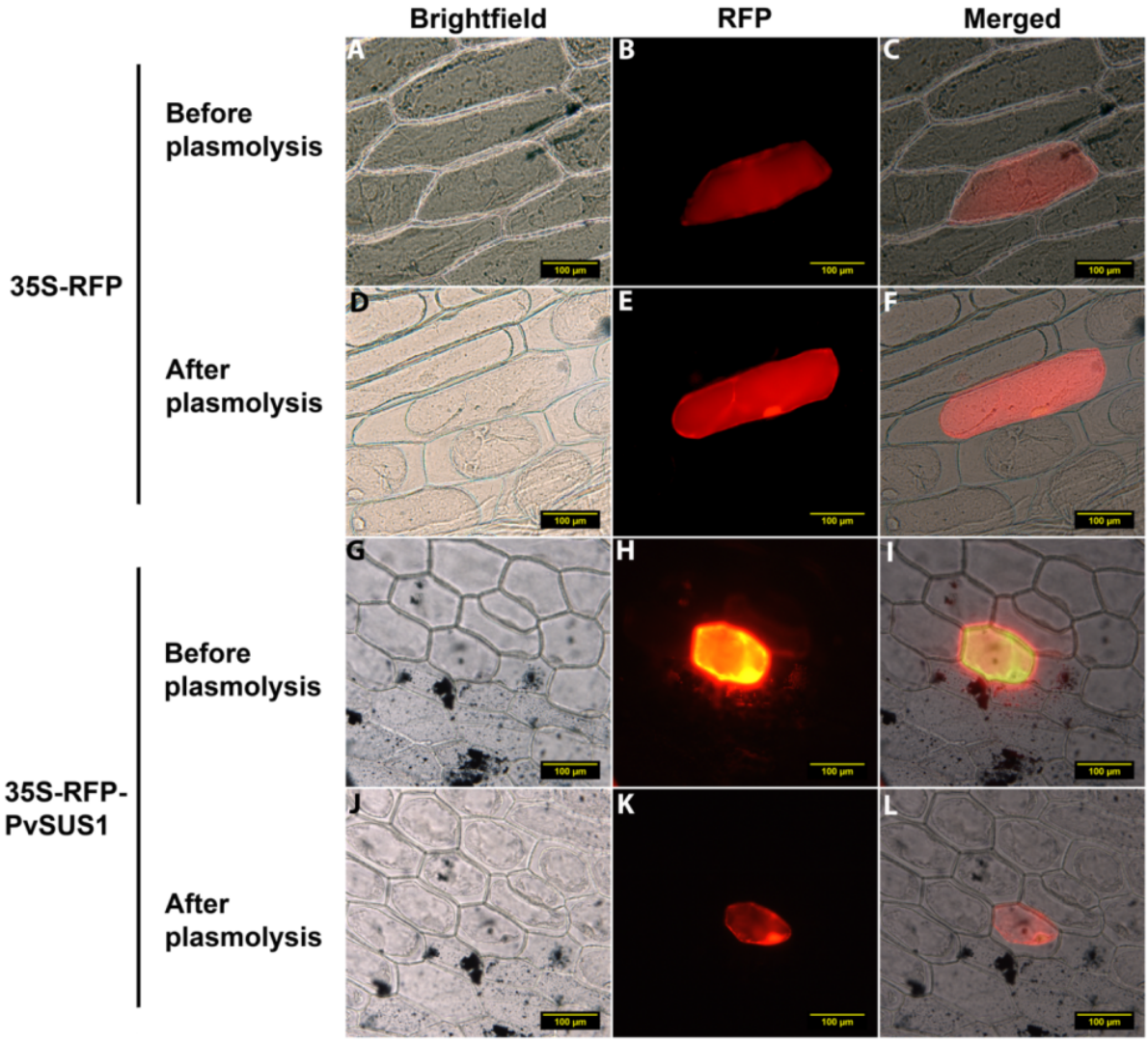
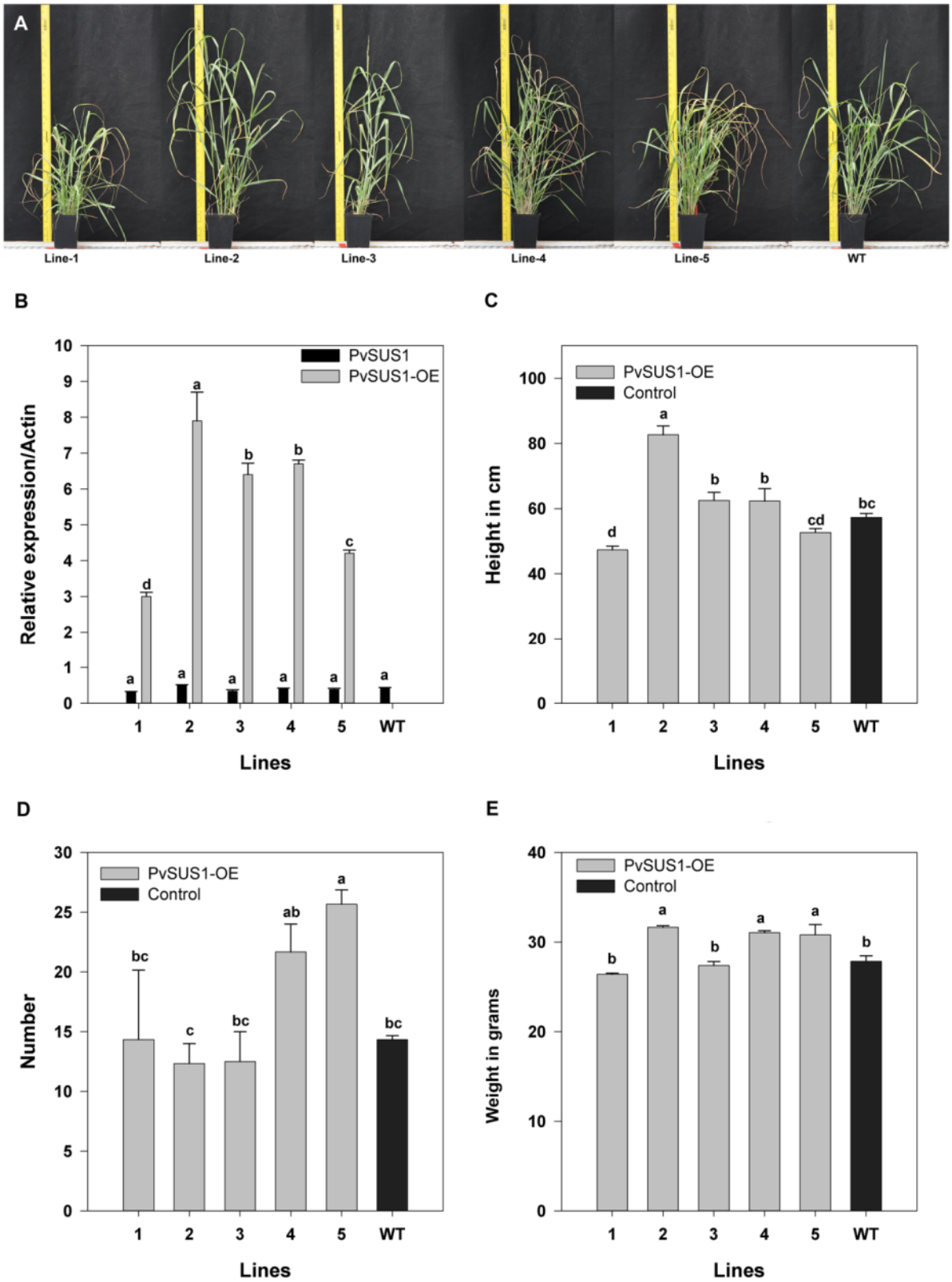


Figure 5-4 Subcellular co-localization of PvSUS1

Control plasmid containing 35S-RFP and PvSUS1-RFP fusion plasmid 35S-RFP-PvSUS1 were coated on 0.6 um gold particles, bombarded into onion epidermal cells and incubated for 48 hours before observation under epifluorescence microscope using Texas REd filters (Em-D560/40, BS-595DCLP, EM-D630/60). The cells were plasmolyzed with 40% sucrose solution and viewed under epi-fluorescence microscope after 30 minutes. Cells bombarded with 35S-RFP control plasmid show fluorescence throughout the cell while cells bombarded with 35S-RFP-PvSUS1 plasmid is localized mainly along the plasma membrane after plasmolysis.

Figure 5-5 Phenotype, relative expression, height, tiller and dry biomass weight of PvSUS1-OE plants

A) Regenerated PvSUS1-OE plants in greenhouse. B) Relative expression level of *PvSUS1* and PvSUS1-OE transcripts in PvSUS1-OE transgenic switchgrass lines. Different letters on the bars of figures indicate significant differences at the $P \leq 0.05$ level as tested by least significant difference (LSD) grouping with SAS software (SAS Institute Inc.) C) Height of PvSUS1-OE plants. Line 2 had significant increase in height compared to control ($p<0.0001$). D) Number of tillers in PvSUS1-OE. Lines 4 and 5 had more tillers compared to control ($p= 0.0379$). E) Dry biomass weight in PvSUS1-OE plants. Lines 2, 4 and 5 had significantly greater biomass yield compared to control ($p= 0.0007$). Data are means ± standard error (n=3).


A
Line-1
Line-2
Line-3
Line-4
Line-5
WT
B
Relative expression/Actin
PvSUS1
PvSUS1-OE
Lines
C
Height in cm
PvSUS1-OE
Control
Lines
D
Number
PvSUS1-OE
Control
Lines
E
Weight in grams
PvSUS1-OE
Control
Lines

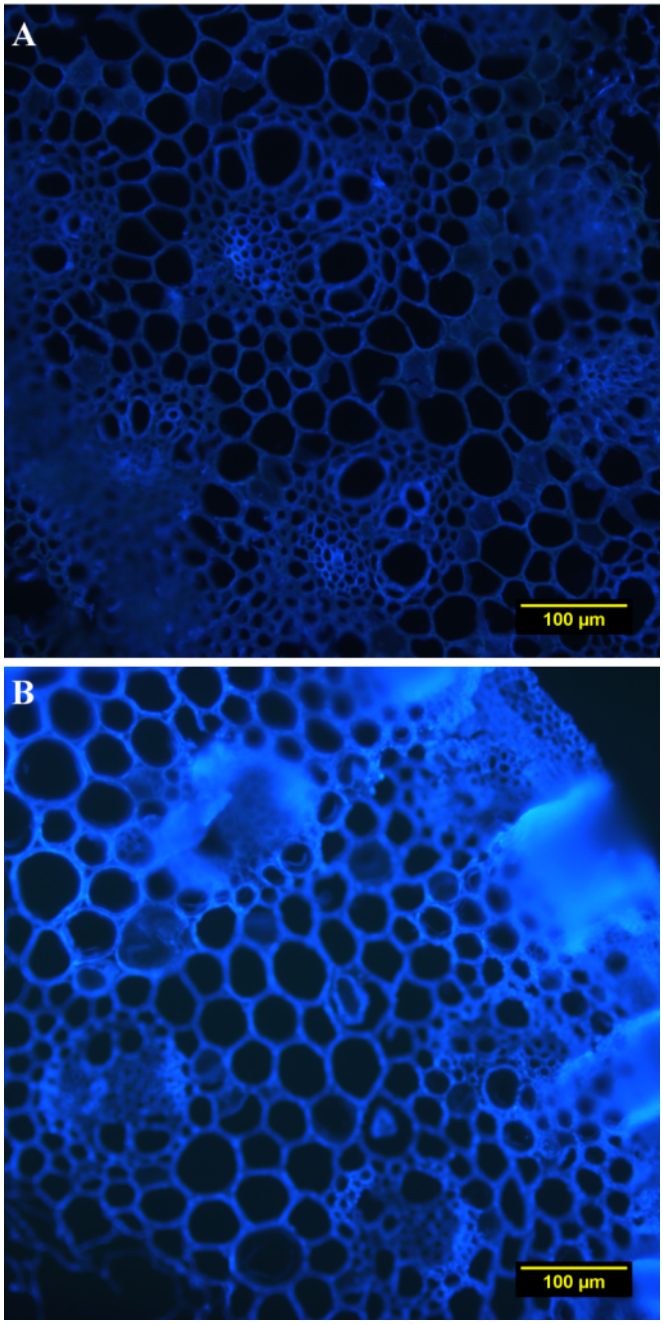
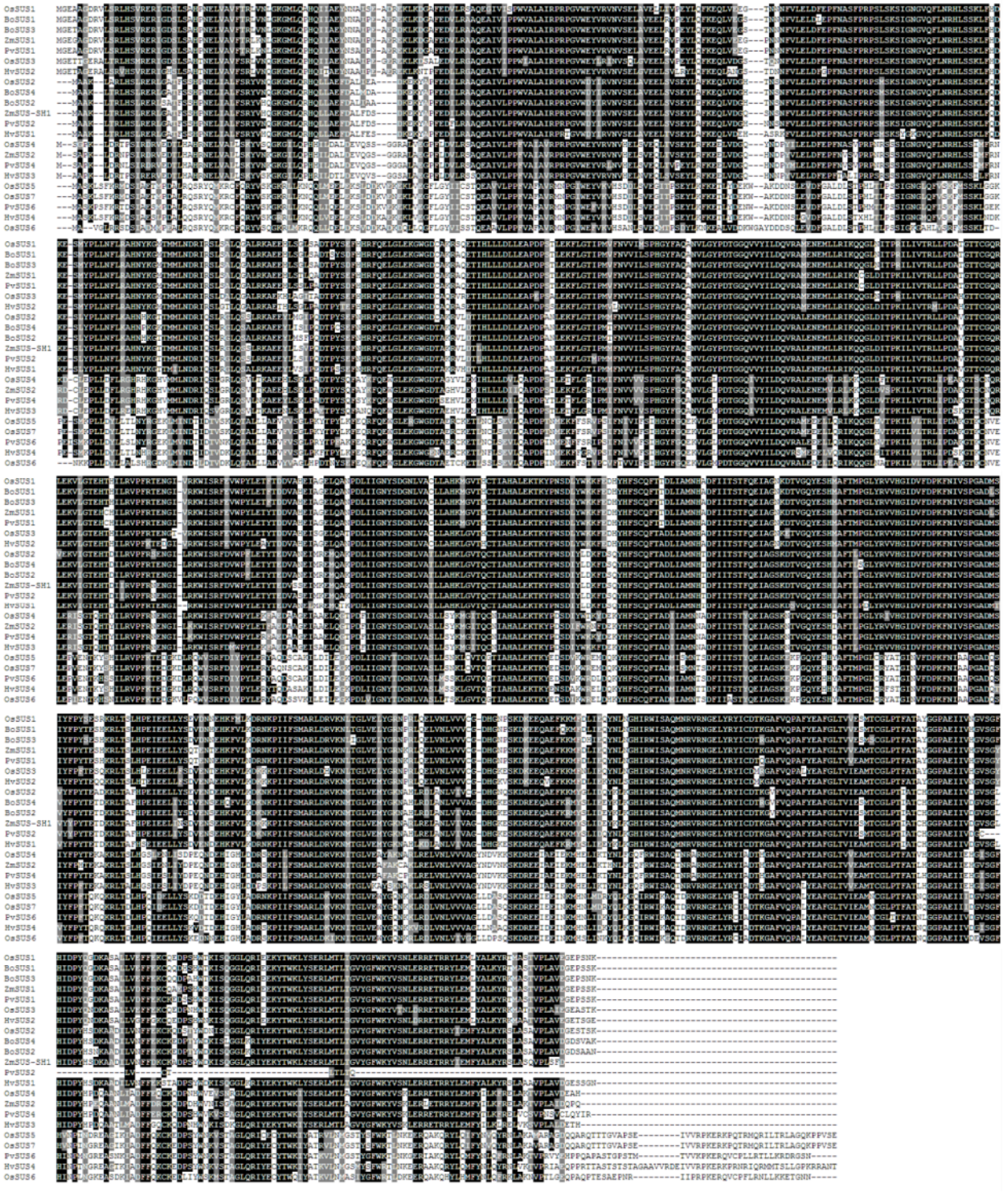


Figure 5-6 Calcofluor staining

Control (A) and PvSUS1-OE switchgrass (B) Sections were taken from 2nd internode of R1 stage plants. Sections were stained with one drop of calcofluor white and one drop of 10% potassium hydroxide and exposed for 15 ms and viewed under DAPI filters (Ex-D360/40, BS-400DCLP, and Em-D460).

Figure 5-7 Multiple sequence alignment of sucrose synthase amino acid sequences from monocots

Amino acid sequences of bamboo, maize, rice barley were taken from Genbank. Switchgrass sequences were taken found searching phytozome switchgrass database using rice amino acid sequences. Sequence alignment was done using online version of MAFFT v7.0. Accession numbers are BoSUS1-4 (AAV64256, AAL50571, AAL50570, AAL50572), ZmSUS1 (NP_001105323), ZmSUS2 (NP_001105194), ZmSUS-SH1 (NP_001266691), OsSUS1-7 (P31924, P30298, Q43009, Q10LP5, H6TFZ4, Q6K973, Q7XNX6), HvSUS1-4 (P31922, P31923, CAZ65725, AEH16642).



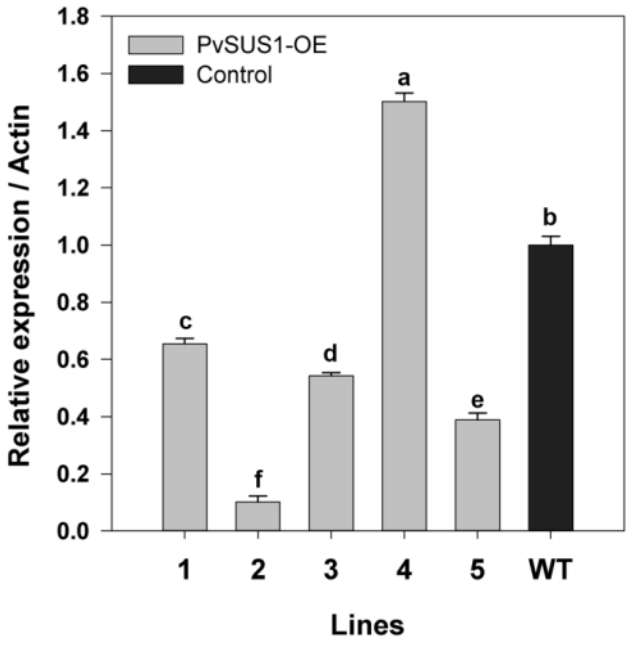


Figure 5-8 Relative expression of fructokinase (*PvFRK1*)

Relative transcript levels of *PvFRK1* were measured by qRT-PCR and standardized to the constitutive *PvACTC* gene expression level. The mRNA was extracted from whole tillers at E4 growth stage. All lines except line 4 had decreased levels of *PvFRK1* transcript ($p<0.0001$). Data are means ± standard error (n=3). Means were analyzed with one-way ANOVA, and letter groupings were obtained using Fisher's least significant difference method. Bars with the same letter do not differ significantly at the 5% level.

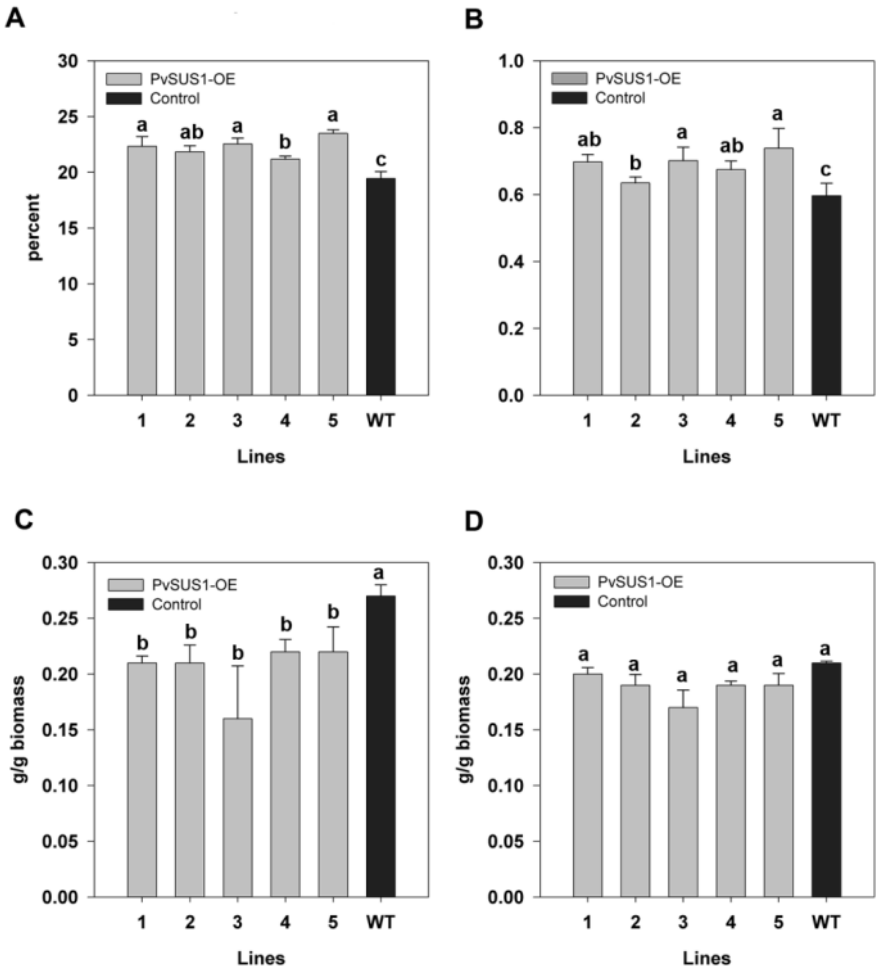


Figure 5-9 Lignin content, S/G ratio and sugar release efficiency of PvSUS1 overexpressing transgenic plants

Lignin content (%) and S/G ratio was determined by MBMS. Lignin content (A) and S/G ratio (B) were statistically significantly different at the P=0.05 level in all transgenic events relative to the control. Glucose release efficiency (C) was significantly different in all transgenic events relative to the control (p=0.031). Xylose release efficiency (D) was not significantly different (p=0.689). All data are means ± standard error (n=3). Means were analyzed with one-way ANOVA, and letter groupings were obtained using Fisher's least significant difference method. Bars with the same letter do not differ significantly at the 5% level.

Table 5-1 Protein identity matrix of switchgrass sucrose synthases and other monocot orthologs

	OsSUS1	ZmSUS1	PvSUS1	OsSUS3	OsSUS2	ZmSUS-SH1	PvSUS2	OsSUS4	ZmSUS2	PvSUS4	OsSUS5	OsSUS7	PvSUS6
ZmSUS1	95.47%	-											
PvSUS1	95.58%	98.49%	-										
OsSUS3	90.47%	90.47%	90.47%	-									
OsSUS2	80.95%	81.06%	81.53%	80.02%	-								
ZmSUS-SH1	79.55%	80.02%	80.02%	79.32%	94.07%	-							
PvSUS2	71.42%	71.77%	71.89%	71.42%	84.66%	87.45%	-						
OsSUS4	70.73%	70.49%	70.96%	69.10%	71.54%	72.12%	64.11%	-					
ZmSUS2	70.38%	70.49%	70.73%	68.52%	71.31%	71.54%	63.76%	94.42%	-				
PvSUS4	70.26%	70.03%	70.49%	68.29%	70.73%	71.19%	63.53%	93.72%	94.65%	-			
OsSUS5	52.14%	52.14%	52.38%	51.56%	52.49%	52.38%	47.50%	54.23%	54.23%	54.12%	-		
OsSUS7	52.38%	52.38%	52.61%	51.68%	52.61%	52.49%	47.38%	54.47%	54.47%	54.35%	99.30%	-	
PvSUS6	53.65%	53.54%	53.77%	52.84%	53.54%	53.65%	48.19%	56.09%	55.86%	55.86%	89.08%	89.31%	-
OsSUS6	51.33%	50.98%	51.21%	51.10%	52.49%	52.61%	47.50%	55.05%	54.58%	54.35%	78.16%	78.39%	80.83%

Amino acid sequences aligned with MAFFT v7.0 were used and identity was calculated as the number of identical positions, including gaps divided by the length of the alignment

Table 5-2 Domains present in switchgrass sucrose synthases, length and molecular mass of protein

Name	No of exons	Length of genomic sequence (bp)	Length of coding sequence (bp)	Length of protein sequence (aa)	Sucrose synthase domain	Glycosyl transferase domain	Molecular mass (kDa)
PvSUS1	14	7405	2451	816	12-559	564-745	92.95
PvSUS2	15	7691	2409	802	6-551	555-733	91.76
PvSUS4	14	5829	2436	811	8-553	558-728	92.63
PvSUS6	14	4185	2556	851	9-560	571-735	97.02

Table 5-3 List of primers

Primer No	Primer name	Primer sequence 5`-> 3`
CP33	PvSUS1-R	TCACTTGCTGGAGGGCTCTCCCTC
CP34	PvSUS1-F	ATGGGGGAAGCTGCCGGCGACCGC
CP43	PCR-F	TCGATGCTCACCCTGTTGTTTGG
CP44	PCR-R	CGATCATAGGCGTCTCGCATATCTC
CP101	PvSUS1-qRT-F1	AATTGTTCGCAAGTGGATCTC
CP102	PvSUS1-qRT-R1	ACGTCATCGGTGTAAGTCTC
CP104	PvSUS2-qRT-F1	TCGTGTTCCATTCAGAAATGAG
CP105	PvSUS2-qRT-R1	GCAACATCCTCAGTGTATGTC
CP181	PvSUS1-F	CCCTCCAGCAAGTGAGG
CP182	PvSUS1-R	CTCCGAATGAGGAAGGAAAC
CP210	PvSUS1-HindIII-F	AAGCTTTAATGGGGGAAGCTGCCGGCG
CP211	PvSUS1-BamHI-R	GGATCCTCACTTGCTGGAGGGCTCTCC
CP277	PvFRK-F1	ATCGTTGTTGCCTGAATGA
CP278	PvFRK-R1	GCACCGCCAAGATATCAAT
CP181	PvSUS1-F	CCCTCCAGCAAGTGAGG
CP182	PvSUS1-R	CTCCGAATGAGGAAGGAAAC
CP295	PvSUS4-qF4	TCCCAACTCTGTTTGCTTAC
CP296	PvSUS4-qR4	AACCTAATCTTGTCGGAACG
CP285	PvSUS6-qF2	AGATTTGTTCCCCATCAAGC
CP286	PvSUS6-qR2	TTTACTGAAGAGAGACCGGG
	PvAct-RT-F1	CAAGATTTGGAGATCCCGTG
	PvAct-RT-R1	AATGCTCCACGGCGAACAG

Chapter 6: Conclusion

Lignocellulosic feedstocks can be converted to biofuels, by breaking down the cellulosic sugars into simple sugars for fermentation into ethanol or conversion to other biofuels. These biofuels can conceivably replace a large fraction of fossil fuels currently used for transportation. However, lignin, a prominent constituent of secondary cell walls, binds the cellulosic sugars with hemicellulose and pectin. Therefore, lignin is an impediment to the breakdown and conversion of cellulosic biomass to biofuel. This resistance to degradation is called recalcitrance. Pretreatment of biomass to remove lignin is the most expensive step in the production of lignocellulosic biofuels. Even though most of the genes in the monolignol biosynthetic pathway have been identified, the role of lignin in plant biology is not fully understood. Rational design of cell walls engineered for lignocellulosic feedstocks cannot be done without full knowledge of role of lignin and its biosynthesis.

Engineering of plant cell walls for biofuel production will likely require the manipulation of genes involved in the monolignol biosynthetic pathway. Transcription factors can act as activators or repressors of the genes in this and other biosynthetic pathways. Overexpression of transcriptional repressors of the monolignol biosynthetic pathway has led to reduced lignin content in a range of dicot plants in the past, which led to using orthologous genes to manipulate lignin production in grasses. *PvMYB4* is a transcriptional repressor of monolignol biosynthetic pathway. It binds to AC elements in the promoters of lignin genes and represses their activity. When overexpressed *PvMYB4* significantly reduced total lignin content. The composition of syringyl and guaiacyl units – the units of lignin - was reduced, though there was no change in their ratio. Indeed, the *PvMYB4* transgenic switchgrass lines (PvMYB4-OX) have up to three-fold increased saccharification efficiency compared with the control lines. In addition, ethanol yield using simultaneous saccharification and fermentation was up to 2.6-fold higher in

transgenic plants compared with nontransgenic controls. The transgenic plants also had faster conversion of the glucose to ethanol and CO_2 under both hot-water pretreated and non-pretreated conditions than did controls. Phenolic inhibitors of fermentation were greatly reduced in PvMYB-OX lines which contributed to the increase in ethanol yield. These results have demonstrated that overexpression of a general transcriptional repressor of the phenylpropanoid/lignin biosynthesis pathway, can reduce cell wall recalcitrance. Large improvement in sugar release efficiency, proportional to reduction in lignin content is vital for the development of economically viable lignocellulosic feedstocks for biofuel production.

Apart for reducing cell wall recalcitrance, increasing biomass is also important in lignocellulosic crops. Sucrose synthase (SUS) catalyzes the reversible breakdown of sucrose in the presence of uridine di-phosphate (UDP) in to UDP-glucose and fructose. UDP-glucose is used by cellulose synthase in the biosynthesis of cellulose for cell wall production. Four *SUS* genes in switchgrass were identified. The deduced amino acid sequences of the SUS genes clustered together with *SUS* genes from other species. At least one member of switchgrass *SUS* gene family is clustered with each of the three SUS groups. Subcellular localization of *PvSUS1* indicated that it is localized to the plasma membrane. Some transgenic switchgrass overexpressing *PvSUS1* transgenic switchgrass plants had increased height and tiller number, although these traits were not correlated with each other. Finally, *PvSUS1* transgenic plants had increased overall dry biomass. This suggests that constitutive overexpression of *PvSUS1* in switchgrass can be used to increase biomass of this lignocellulosic biofuel crop, and perhaps other crops as well.

Vita

Charleson Rajendran Poovaiah was born on July 28, 1975 in Tirunelveli, Tamil Nadu, India to Mr. G. Poovaiah and Retnavathy Poovaiah. He graduated from Rosemary Higher Secondary Matriculation School and started his college education in 1993 at Allahabad Agriculture Institute, Allahabad, India for his Bachelors of Science in Agriculture. He obtained a Masters of Science in Horticulture from the same institute in 2000 focusing on the postharvest preservation of mangoes. He worked in the landscaping industry for 3 years before joining Purdue University in 2003 for Masters of Science in Horticulture, and developed regeneration protocols for native and Scotch spearmint. After graduation, he worked in the laboratory of Dr. Steven H. Strauss at Oregon State University as a Faculty Research Associate from 2005 to 2008. In 2009, he started his Ph.D studies in the Stewart laboratory in Department of Plant Sciences at the University of Tennessee, Knoxville. His research focused on developing strategies for genetically engineered biofuel crop production. He plans to graduate with a Doctor of Philosophy degree in Plants, Soils, and Insects with a concentration of Plant Molecular Genetics in December 2013. He plans to do a postdoctoral training at Syracuse University beginning in October 2013.