GENETIC IMPROVEMENT OF SWITCHGRASS CELL WALL CONTENT, LEAF ANGLE AND FLOWERING TIME

Bin Xu

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Horticulture

Bingyu Zhao
Richard Veilleux
Boris Vinatzer
Andy Pereira
Chuansheng Mei

June 10, 2011
Blacksburg, Virginia

Keywords: *Panicum virgatum* L., Genetic transformation, Lignin, Leaf angle, Flowering time

Genetic improvement of switchgrass cell wall content, leaf angle and flowering time

Bin Xu

ABSTRACT

Switchgrass (*Panicum virgatum* L.) is a candidate bioenergy crop. Somatic embryogenic (SE) calli are used for genetic transformation in switchgrass. A superior switchgrass line, HR8, was developed using recurrent tissue culture selection from cv. Alamo. HR8 SE calli were genetically transformable using *Agrobacterium* at an efficiency of ~12%.

We used HR8 somatic embryogenic calli for genetic improvement of switchgrass. The lignin content of feedstock has been proposed as one key trait impacting biofuel production. 4-Coumarate: Coenzyme A ligase (4CL) is one of the key enzymes involved in the monolignol biosynthetic pathway. Two homologous *4CL* genes, *Pv4CL1* and *Pv4CL2*, were identified in switchgrass. Gene expression patterns and enzymatic activity assays suggested that Pv4CL1 is involved in monolignol biosynthesis. Stable transgenic plants were obtained with *Pv4CL1* downregulated. RNA interference of *Pv4CL1* reduced extractable 4CL activity by 80%, leading to a reduction in lignin content with decreased guaiacyl unit composition. The transgenic plants had uncompromised biomass yield. After dilute acid pretreatment, the low lignin transgenic biomass had significantly increased cellulose hydrolysis (saccharification) efficiency for biofuel production.

Erect leaf is a desirable trait to adjust the overall plant architecture to perceive more solar energy and thereby to increase the plant biomass production in a field population. We overexpressed an Arabidopsis NAC transcriptional factor gene, *LONG VEGETATIVE PHASE ONE* (*AtLOV1*), in switchgrass. Surprisingly, *AtLOV1* induced smaller leaf angle by changing morphologies of epidermal cells in the leaf collar region, affecting lignin content and monolignol composition, and also causing delayed flowering time in switchgrass. Global gene-expression analysis of AtLOV1 transgenic plants demonstrated an array of genes has altered expressions. Potential downstream genes involved in the pleiotropic phenotypic traits of the transgenic plants are discussed.

Acknowledgement

First of all, I am very grateful to my family and my friends for their spiritual and financial supports, and to all my teachers infilling me with knowledge and ways of thinking and behaving! There is no word can tell how much I am indebted for all the things they give to me.

I sincerely appreciate my major advisor, Dr. Bingyu Zhao for his patience, guidance and financial support! I would like to express my sincere appreciation to my committee members, Dr. Richard Veilleux, Dr. Boris Vinatzer, Dr. Andy Pereira, and Dr. Chuansheng Mei for their great supports throughout my graduate study. Specifically, Dr. Zhao, Dr. Veilleux, and Dr. Vinatzer have not only served as academic advisors but also given me many help and genuine advices on my professional career development. I sincerely thank Dr. Roger Harris for recruiting me into the Horticulture program and other staffs in Horticulture Department for various helps and supports!

I specially thank Dr. Zhengxing Shen, Dr. Linkai Huang, Sy Traore, Shuchi Wu, Dr. Qiang Cheng, Xiaoyan Shen, Dr. Changhe Zhou and other lab members for their generous help! I am grateful for my collaborators, Sathitsuksanoh Noppadon, Dr. Maria Balota, Kim Harich, Dr. Xunzhong Zhang, Dr. Greg Welbaum, Dr. Yuhong Tang and Dr. Richard Dixon! Their support and collaboration are essential for the completion of my PhD study.

I will dedicate all my strength and knowledge to the development of agriculture and biological science in the rest of my life!

Table of Contents

Chapter 1. Introduction and Literature Review 1

1. Switchgrass, a potential biofuel crop 1
2. Overview of desirable agronomic traits of switchgrass 3
3. Strategies for manipulating lignin content and compositions 4
4. Strategies for promoting switchgrass vegetative growth 8

4.1 Promoting switchgrass vegetative growth by delaying or aborting flowering 8

4.2 Promoting switchgrass biomass yield by optimizing photosynthetic traits 11

4.2.1 Photosynthetic efficiency 11

4.2.2 Plant architecture 13

Summary 15

References 17

Chapter 2. Selection and characterization of a new switchgrass (*Panicum virgatum* L.) line with high somatic embryogenic capacity for genetic transformation 24

ABSTRACT 25

Introduction 26

Materials and methods 26

Results 30

Discussion 33

Availability 35

Acknowledgements 35

References 43

Chapter 3. Silencing of *4-Coumarate: Coenzyme A ligase* in Switchgrass Leads to Reduced Lignin Content and Improved Fermentable Sugar Yield for Biofuel Production 45

Summary 46

Introduction 47

MATERIALS AND METHODS 49

RESULTS 55

DISCUSSION 59

ACKNOWLEDGEMENTS 63

References 78

Supplementary data 83
SUPPLEMENTAL LITERATURE CITED 87

Chapter 4. Ectopic expression of *AtLOV1* in switchgrass altered plant architecture, lignin content, and flowering time 88
Abstract 89
Introduction 90
MATERIALS AND METHODS 92
RESULTS 97
DISCUSSION 99
SUMMARY 105
ACKNOWLEDGEMENTS 105
References 117
Supplementary Materials and Methods 120
Supplementary Data 120
SUPPLEMENTAL LITERATURE CITED 132

List of Tables

Chapter 2

Table 1.Tissue culture media used in this study 36

Table 2. Effects of ABA treatments on seed germination and SE callus formation of two switchgrass lines 37

Chapter 3

Table 1. Kinetic parameters of Pv4CL1 and Pv4CL2 64

Table 2. Growth performance of wild type and T_0 lines 64

Table 3. Growth performance of segregating T_1 plants 65

Table 4. T_0 Transgenic plants have reduced lignin contents 65

Table 5. T_0 Transgenic plants have altered monolignol compositions 66

Table 6. Cell wall compositions of segregating T_1 plants 66

Table S1 Primers used in the study 87

Chapter 4

Table 1. Length of leaf collar region of AtLOV1 transgenic plants compared to wild type plants. 106

Table 2. Lignin content and monolignol composition of independent T_0 AtLOV1 transgenic plants compared to wild type plants. 106

Table 3. AtLOV1 transgenic plants have delayed flowering time compared to wild type plants. 107

Table 4. AtLOV1 transgenic plants have uncompromised whole plant biomass yield but decreased aboveground plant biomass yield. 107

Supplementary Table S1. Primers used for PCR and RT-PCR. 124

Supplementary Table S2. Primers used in qRT-PCR for some microarray data validation 125

Supplementary Table S3. Differentially expressed genes in AtLOV1 transgenic switchgrass. 126

List of Figures

Chapter 2

Fig. 1. Selection of of the highly regenerable line HR8 38

Fig. 2. HR8 has high seed germination, SE callus production percentage, and SE calli of HR8 have high regeneration percentage compared to unselected 'Alamo'. 39

Fig. 3. Tissue culture of switchgrass line HR8 40

Fig. 4. Morphology of switchgrass callus under scanning electron microscope 41

Fig. 5 Agrobacterium-mediated transformation of switchgrass line HR8 with two different plasmid constructs 42

Chapter 3

Fig. 1 Phylogenetic tree of Pv4CL1 and its homologues in some model plants 67

Fig. 2 Transcript abundance of *Pv4CL1* and *Pv4CL2* in different organs 70

Fig. 3 Functional characterization of Pv4CL1 and Pv4CL2 72

Fig. 4 RNAi:Pv4CL1 lines have reduced Pv4CL1 transcript abundances and 4CL enzymatic activity 74

Fig. 5 RNAi:Pv4CL1 lines have altered phenotypes 75

Fig. 6 Lignin histological staining for lignin deposition 76

Fig. 7 RNAi:Pv4CL1 switchgrass biomass has improved saccharification efficiency 77

Fig. S1 Alignment of two switchgrass 4CLs amino acid sequences with their rice orthologs 83

Fig. S2 Key features of the vectors used in this study 85

Fig. S3 Mass balance of Dilute Acid Pretreatment and Enzymatic Hydrolysis of T1 transgenic (a) and wild type (b) biomass 86

Chapter 4

Fig. 1 Overexpression of *AtLOV1* caused erect leaf phenotype in switchgrass. 109

Fig. 2 Plant morphology at the leaf collar region in AtLOV1 transgenic and wild type (WT) control plants. 110

Fig. 3 Microarray analyses on differentially expressed genes in AtLOV1 transgenic plants. 115

Fig. 4 Ectopic overexpression of AtLOV1 in switchgrass induced distinct phenotypes by regulating gene expression in different signaling pathways. 116

Supplementary Fig. S1 *AtLOV1* has two transcripts due to alternative splicing. 120

Supplementary Fig. S2 Phloroglucinol staining of lignin in AtLOV1 transgenic and wild type plants. 121

Supplementary Fig. S3 Cellular localization of AtLOV1.1 and AtLOV1.2. Bars represent 50 μm. 122

Chapter 1

Introduction and Literature Review

1. Switchgrass, a potential biofuel crop

Switchgrass (*Panicum virgatum* L.) is a native warm-season C_4 perennial grass, adapted to various marginal lands of North America, which can be established in the field through seeds or vegetative propagation (McLaughlin & Adams Kszos, 2005). Switchgrass has been used for animal feedstock (Anderson *et al.*, 1988) and for restoring native ecosystems (Casler *et al.*, 2007). Recently, it has been extensively evaluated as a promising lignocellulosic bioenergy crop (McLaughlin & Adams Kszos, 2005). In a large scale study, the average net energy yield of switchgrass (60 Gj/ha/year) was estimated as 540% more than the nonrenewable energy consumed (Schmer et al., 2008). Genetic improvements and optimized agronomic practices may further advance switchgrass biomass yield and feedstock quality for bioenergy production.

Switchgrass can be classified into two ecological subtypes (upland or lowland ecotypes) according to their natural geographic distributions (Moser & Vogel, 1995). The rich genetic variation was verified by morphological markers, chloroplast DNA polymorphisms, and other molecular markers (Missaoui *et al.*, 2006; Narasimhamoorthy *et al.*, 2008; Zalapa *et al.*, 2011). Lowland switchgrass ecotypes are tall, thick-stemmed, and better adapted to wet conditions, and are predominately tetraploid (2n=4x=36); upland ecotypes are short, rhizomatous, thin-stemmed, and adapted to drier conditions, and are typically octaploid (2n=8x=72) with a small proportion being tetraploid and hexaploid (2n=6x=54) (Moser & Vogel, 1995; Hopkins *et al.*, 1996; Lu *et al.*, 1998). Currently, the average switchgrass biomass yields range from 4-10 tons/acre/year according to field conditions and cultivars (Lemus *et al.*, 2002), and there can be a three to five fold difference between cultivars (Lemus *et al.*, 2002). Because switchgrass is gametophytically self-incompatible, all switchgrass cultivars are synthetic cultivars by recurrent selection (Vogel & Mitchell, 2008). Despite the lack of pure inbred lines for hybrid breeding in switchgrass, hybrid vigor was reported recently, where 30-38% high-

parent heterosis in biomass yield was observed in F_1 hybrids derived from crosses between 'Kanlow' x 'Alamo' and 'Alamo' x 'Kanlow', suggesting the significant impact of genetic improvement on switchgrass cultivar development (Vogel & Mitchell, 2008). By "stacking" certain genetic components in switchgrass either through traditional breeding or biotechnological methods, biomass yield could greatly increased.

Switchgrass molecular breeding programs have been focused on (1) exploring the genetic diversity of switchgrass germplasm; (2) identification of desirable agronomic traits and associated genetic loci by genetic mapping and introgression of the genetic loci into certain cultivars through crossing and marker-aided selection, in which breaking the crossing barrier between ecotypes and different *Panicum* species is important; (3) understanding the heterotic extent between and within switchgrass ecotypic populations, and defining desirable agronomic traits in these populations; (4) functional gene cloning and manipulation in switchgrass based on orthologues or genetic loci with high synteny that have been characterized in closely-related species (e.g. maize, sorghum and rice); (5) direct cloning of genes-of-interests that control important biofuel-related traits in switchgrass, and transforming these genetic components into switchgrass through genetic engineering.

Translating the knowledge gained from the study of model plant species, such as *Arabidopsis* (*Arabidopsis thaliana*), rice, *Medicago truncatula,* and *Brachypodium distachyon*, into crop species has contributed to managing important agronomic problems in major food crops (Zhang *et al.*, 2004; Lawrence & Walbot, 2007). For example, an *Arabidospsis* gene, *Sodium Proton Exchanger 1* (*AtNHX1*), was identified as a key regulator of salt tolerance in *Arabidopsis*. When *AtNHX1* was overexpressed in *Brassica napus*, tomato, and rice, the transgenic plants were significantly more tolerant to salt stress too (Zhang, *et al*. 2004).

Reverse genomics resources and tools are being developed for switchgrass. To date, approximately 720,590 Expressed Sequence Tags (EST), and 195,557 Genome Survey Sequences (GSS) were deposited in GeneBank (NCBI, surveyed in May 18, 2011). Switchgrass full genome sequencing is in progress by joint research groups, including Clemson University, DOE national labs, University of California, and other

institutions (Sharma *et al.*, 2011). A switchgrass oligoarray was recently developed by a research group at Noble Foundation (unpublished). A highly efficient *Agrobacterium*-mediated switchgrass transformation protocol was recently reported in which 50% transformation efficiency was achieved (Li & Qu, 2011), and a highly somatic embryogenic and regenerable (HR) line was also developed suitable for genetic transformation (Xu, *et al.*, 2011a). Transposon-based and chemical induced mutant populations were also under development (unpublished result in our lab), from which loss/gain of function mutations can be identified through TAIL-PCR, Targeting Induced Local Lesions in Genomes (TILLING) (Colbert *et al.*, 2001), or through Next-Generation sequencing for mutation detection.

In this review, we focus on a few important agronomic traits of switchgrass and potential genes or pathways controlling these traits. Strategies for understanding and engineering useful genetic components are also discussed to further improve switchgrass biomass quality and biomass yield.

2. Overview of desirable agronomic traits of switchgrass

In general, four important goals of switchgrass breeding are: (1) improving the bioenergy feedstock quality; (2) increasing the biomass yield under various field and geographic conditions; (3) developing value-added switchgrass biomass feedstock; and (4) decreasing input of switchgrass field production.

Currently, two major biomass processing technologies are being intensively studied to convert lignocellulosic feedstock to bioenergy: thermal conversion and biochemical conversion. Desirable feedstock qualities are largely dependent on the nature of processing technologies (Carroll & Somerville, 2009). For example, in thermal conversion (e.g. direct combustion or pyrolysis), it is more desirable to have feedstock with lower amounts of mineral residues and higher energy content, which often correlate with a high lignin content of biomass (Boateng *et al.*, 2008). On the contrary, in biochemical conversion for biofuel production, feedstock with lower lignin content has higher saccharification efficiency through enzyme hydrolysis and higher enzymatic fermentation efficiency (Fu *et al.*, 2011a). Other cell wall compositions, such as

hemicellulose (Lee *et al.*, 2009) and pectin (Lionetti *et al.*, 2010) also have negative impacts on bioenergy production using biochemical conversion technologies.

Switchgrass breeding programs have aimed to double its biomass yield in the near future (Schubert, 2006). Improving the biomass yield of switchgrass under various field or geographic conditions can be achieved by promoting vegetative growth, increasing the photosynthetic sink-source ratio, increasing resistance to biotic/abiotic stress, and improving water and nutrient use efficiency (WUE and NUE), etc. Improving certain biological traits of switchgrass, such as NUE, can also decrease production input. Producing high value additives, such as plastics, enzymes and secondary metabolic chemicals, can further increase the economics of growing switchgrass (Somleva *et al.*, 2008). For example, ODE-USDA awarded Agrivida Inc. (Medford, MA) a grant for producing switchgrass with cell-wall degrading enzymes which remain inactive during plant growth but become activated after harvest (http://www.energy.gov/8283.htm). Selection and use of plant-growth promoting microbes may also improve grass growth and resistance to stress (Compant *et al.*, 2005). Candidate genetic components and pathways potentially useful for switchgrass improvement are discussed in the following section with an emphasis on lignin, flowering period and photosynthesis traits.

3. Strategies for manipulating lignin content and compositions

About 80% of the dry plant biomass is comprised of plant cell walls which store most of the biomass energy (Vogel & Jung, 2001). Cellulose, hemicellulose, and pectin are the polysaccharide components of plant cell walls, among which cellulose is the primary component for biofuel (ethanol) production through fermentation (Carroll & Somerville, 2009). Cell walls, especially secondary cell walls, are strengthened by lignin, a phenolic polymer derived from hydroxycinnamy alcohols via combinatorial radical coupling reactions (Boudet, 2007). Lignin deposition reinforces pant cell walls to enable water transport, provides mechanical support and a barrier to pathogens, and is involved in abiotic tolerance (Halpin, 2004; Boudet, 2007). However, high lignin content is not desirable for bioconversion of lignocellulosic feedstock to biofuel because it prevents the access of hydrolytic enzymes to the polysaccharides, absorbs hydrolytic enzymes, and inhibits the activities of the hydrolytic and fermentable enzymes used in the biofuel

conversion process (Halpin, 2004; Endo *et al.*, 2008; Abramson *et al.*, 2009). Studies using different alfafa transgenic lines with variable reduced lignin content proved the negative correlation between lignin content and fermentable sugar release efficiency (Chen & Dixon, 2007). Therefore, there is a strong interest in developing low-lignin content switchgrass cultivars for biofuel production.

The grass lignin polymer is usually composed of three monolignols [hydroxyphenyl (H), guaiacyl (G), and syringyl (S)] (Hatfield *et al.*, 1999). Monolignols are derived from the amino acid phenylalanine through the monolignol biosynthesis pathway. The pathway has about ten key enzymes catalyzing the reaction steps and the pathway is evolutionarily conserved across angiosperms (Smita & Nath, 2008). Gene families encoding these key enzymes went through a rapid expansion after the divergence of monocots and dicots (Xu *et al.*, 2009). By BLASTing against the switchgrass EST database and phylogenetic analysis, we can find switchgrass orthologs of all monolignol biosynthesis genes in model plants. Through gene-expression patterns, *in vitro* enzymatic assays and the generation of stable RNAi transgenic plants, a few switchgrass genes [*4 Coumarate：Coenzyme A Ligase* (*4CL*), *Cinnamyl Alcohol Dehydrogenase* (*CAD*), *Catechol-O-methyltransferase* (*COMT*)] in the monolignol biosynthesis pathway were identified. RNAi: *PvCOMT* and RNAi: *PvCAD2* transgenic plants have significantly less lignin content than wild type plants (Fu *et al.*, 2011a; Fu *et al.*, 2011b; Saathoff *et al.*, 2011a; Saathoff *et al.*, 2011b). We have identified two *4CL* genes in switchgrass; phylogenetic and gene-expression pattern analyses suggested that *Pv4CL1*, but not *Pv4CL2*, is involved in monolignol biosynthesis. RNAi downregulating the expression of *Pv4CL1* also generated low lignin-content switchgrass plants (Xu, *et al.*, 2011b). Despite of low-lignin contents, RNAi: *PvCOMT* (Fu *et al.*, 2011a) and our RNAi: *Pv4CL1* transgenic plants grown under greenhouse conditions have normal plant growth and similar biomass yield to that of wild type plants. Field tests on these RNAi transgenic plants are required to support the claim that the low lignin content and normal biomass yield of transgenic lines are independent of environmental cues and stress. Nonetheless, current studies suggest that it is feasible to generate low-lignin switchgrass by generating transgenic plants with monolignol biosynthesis gene(s) silenced or by selecting null mutants of monoligniol biosynthesis gene(s).

The lignin polymer can be composed of different ratios of S, G and H, and can tolerate incorporation of other phenolic components. For example, in interfascicular fibers of *Arabidopsis* stems, lignin has a high proportion of G lignin; but in vascular bundles of *Arabidopsis* stems, lignin is primarily composed of S (Chapple *et al.*, 1992). It is also shown that lignin can comprise about 90% of the benzodioxane units in transgenic *Arabidopsis* with upregulated *Ferulate 5-Hydroxylase* (*F5H*) and downregulated *COMT* (Vanholme *et al.*, 2010; Weng *et al.*, 2010). Notably, transgenic *Arabidopsis* has a dwarf stature but still produces viable seeds (Vanholme *et al.*, 2010; Weng *et al.*, 2010), echoing the flexibility of lignin polymers. Lignin polymers with different compositions often have different strengths of chemical bonds, which impact lignin digestion and degradation. The frequency of resistant bonds (or condensed bonds) in lignin can be detected by the monolignol yield in thioacidolysis such that a higher frequency of resistant bonds results in a lower thioacidolysis yield (Berthet *et al.*, 2011). S-enriched lignin is thought to have fewer crosslinked bonds than G-enriched lignin, and thereby an increased S/G ratio could lead to easier lignin digestion and degradation (Abramson *et al.*, 2009). For example, transgenic *Arabidopsis* with S-enriched lignin has higher enzymatic hydrolysis efficiency than wild type plants after hot-water pretreatment (Li, X *et al.*, 2010). Although the correlation between S/G ratio and enzymatic hydrolysis efficiency is not universally recognized (Chen & Dixon, 2007), the hypothesis and current experimental results suggest engineering lignin composition may decrease the strength of lignin bonds to facilitate the enzymatic hydrolysis of plant cell walls.

Some transcriptional factors directly regulating monolignol biosynthetic genes have been identified (Zhou *et al.*, 2009; Zhao, Q *et al.*, 2010; Zhao, Q. *et al.*, 2010; Ambavaram *et al.*, 2011). In *Arabidopsis*, MYB Domain Protein 58 (MYB58) directly regulates expression of genes involved in monolignol biosynthesis except *F5H*; and the expression of *MYB58* is regulated by "master regulator genes" for secondary cell wall formation--- NST1/NST2/MYB46/VND6/VND7 (NST stands for NAC Secondary Wall Thickening Promoting Factor, VND for Vascular-related NAC-domain Protein) (Zhou et al., 2009). Interestingly, F5H is directly regulated by NST1 and Secondary Wall-associated NAC Domain Protein 1 (SND1) (Zhao, Q. et al., 2010). On the other hand, since secondary cell wall structure is hallmarked by not only lignification but also higher

cellulose and hemicellulose contents, other transcriptional factors downstream of NST1/SND1/VND6/VND7 should be involved in the activation of cellulose and hemicellulose biosynthesis. Recently, an *Arabidopsis* gene *SHN*, when overexpressed in rice caused increased cellulose but decreased lignin content by directly binding to promoters of rice MYB58/63, NST1/2/SND1, VND4/5/6, and MYB20/43 to downregulate genes in monolignol biosynthesis and to upregulate genes in cellulose biosynthesis (Ambavaram *et al.*, 2011).

Most monolignol biosynthetic genes (except *F5H*) have AC elements (ACCT/AAC/AC) in their promoter regions (Raes *et al.*, 2003). The AC cis-element can be bound by some MYB proteins, such as transactivators MYB58/63/85 and transrepressors MYB4/32 (Goicoechea *et al.*, 2005; Zhou *et al.*, 2009; Zhao & Dixon, 2011). Many MYB transcription factors are regulated by environmental cues and plant hormones, which at least partially explains why cell wall lignification is largely influenced by plant growth conditions (see review by Zhao & Dixon, 2011). Understanding the functions of these transcriptional factors may assist to engineer low lignin content switchgrass independent of field conditions.

Monolignols are synthesized inside cytoplasm, and then transported across the cell membrane to the cell wall where they oxidize and polymerize into lignin polymer (Miao & Liu, 2010). Monolignol transportation is mediated by ATP-dependent ATP-binding cassette-like transporters. The transporters in the plasma membrane preferentially transport monolignol aglycones, whereas transporters in the vacuolar membrane prefer gluco-conjugated monolignols for vacuole storage (Miao & Liu, 2010). Genes encoding proteins for these transporters are not identified and it is unclear whether different transporters have preferences for different monolignols. Nonetheless, it is possible to reduce lignin content or alter lignin composition by engineering these transporters in the future.

Laccases and guaiacol peroxidases (class III peroxidases) are proposed to oxidize monolignols to form lignin polymers. In *Arabidopsis*, 73 peroxidases and 17 laccase-like genes were identified (Berthet *et al.*, 2011). The high number of guaiacol peroxidase genes potentially involved in the oxidization of lignin polymer makes it difficult to assign

the specific function to each gene (Mathé *et al.*, 2010). Certain laccase genes are expressed exclusively in lignifying cells (Boerjan *et al.*, 2003). Recently, a study showed that *Laccase 4* (*LAC4*) and *LAC17* are involved in the lignification of stems because *lac17* single mutant has a reduced G lignin deposition and *lac4/17* double mutant has ~40% less overall lignin content (Berthet *et al.*, 2011). Another laccase gene, *LAC15*, is specifically involved in oxidative polymerization of flavonoids and monolignols in *Arabidopsis* seed coats (Liang *et al.*, 2006). Notably, MYB58 and MYB63 can directly transactivate the expression of *LAC4* gene (Zhou *et al.*, 2009). Interestingly, the double mutant *lac4/17* has semi-dwarfed peduncles under long-day conditions but retains normal plant size under continuous light; whereas lignin content is consistently reduced under both conditions (Berthet *et al.*, 2011). This result further suggests that many genes involved in lignin-synthesis are affected by environmental cues but reducing lignin content is not necessarily associated with reduced biomass yield.

4. Strategies for promoting switchgrass vegetative growth

4.1 Promoting switchgrass vegetative growth by delaying or aborting flowering

Biomass accumulation ends when switchgrass begins flowering, and the reproduction process (flowering and seed setting) is energy consuming. Therefore, delayed or aborted flowering can extend and promote vegetative growth. For example, lowland cultivars flower late in high latitude areas and produce higher biomass yields than upland cultivars (Lemus *et al.*, 2002). However, growing lowland cultivars in high latitudes is challenging because of the winter hardiness and drought stress of the high latitude in the long run (Lemus *et al.*, 2002). In addition to genotype effect, environmental factors also have a significant impact on switchgrass flowering time, which have not been well studied yet. Under the same growth conditions, different switchgrass cultivars normally have very similar final leaf numbers; tillers that emerged in the spring season produced a final leaf number ranging from nine to 11, and tillers that emerged in and after the summer season gave a final leaf number of less than seven in Texas, USA (Van Esbroeck *et al.*, 1997). In the winter season, we observed switchgrass cultivars (both upland and lowland) flower even when they only have 2-3 leaves in the

greenhouse (temperature set at 22-28 ℃, under natural light) at Blacksburg, Virginia, USA. These observations suggest that switchgrass flowering time can be regulated through photoperiod and autonomous pathways.

Flowering pathways have been well studied in the dicot model *Arabidopsis*. In *Arabidopsis*, flowering time is mainly controlled by interactions between photoperiod, vernalizatin, gibberellic acid (GA)-response, and autonomous pathways (Corbesier & Coupland, 2006). In monocot plant species, flowering pathways are poorly understood, partially because of the redundancy of gene families involved in flowering time. For example, maize (*Zea mays*) has more than 1000 orthologs of *Arabidopsis* flowering genes (Buckler *et al.*, 2009), where numerous small-effect quantitative trait loci (QTLs) contribute to maize flowering architecture (Buckler *et al.*, 2009). In contrast to the out-crossing plant species (e.g. maize), the flowering pathways in selfing plant species (e.g. *Arabidopsis* and rice) are mainly controlled by a set of large-effect genetic components (Buckler *et al.*, 2009), possibly because selfing plants can tolerate large changes in flowering time with seed production while out-crossing species cannot (Maloof, 2010).

Despite the difficulty of directly identifying functional orthologs in grasses by BLAST searching with *Arabidopsis* flowering genes, many known grass flowering genes share common signatures with *Arabidopsis* flowering genes and are involved in similar signaling pathways. For example, *CONSTANS* (*CO*) and *FLOWERING LOCUS T* (*FT*) orthologs are key regulatory genes in the photoperiod flowering pathway in both the long-day plant (LDP) *Arabidopsis* and the short-day plant (SDP) rice (Maloof, 2010). One rice *FT* orthorlog, *Hd3a*, shares key features of *Arabidopsis FT*, and can complement the *Arabidopsis ft* mutant (Izawa *et al.*, 2002). In *Arabidopsis*, stabilized CO protein acts as positive regulator of flowering by activating the transcription of *FT* gene, where FT protein translocates to the apical meristem and induces flowering (Yanovsky & Kay, 2002; Turck *et al.*, 2008). However, in rice, *CO* orthologs [(e.g. *Heading date 1* (*Hd1*)] act as a repressor but not an activator of flowering in the presence of light (Tamaki *et al.*, 2007; Komiya *et al.*, 2009). The photoperiod flowering pathway is also more complex in rice than in *Arabidopsis*. For example, a grass-specific gene, *Early Heading Date 1* (*Ehd1*), promotes SD flowering in *hd1* mutant rice by inducing the expression of rice *FT-like* genes, suggesting an additional grass (or rice)-specific signaling cascade present in

the photoperiodic pathway in rice (Doi *et al.*, 2004). In cereal plants (e.g. wheat and barley), flowering time may be largely controlled by perception of cold (vernalization) and photoperiod through regulations of large-effect genes, such as vernalization genes *VRN1*, *VRN2*, *VRN3* and orthologs of *CO*, (see review by Distelfeld *et al.*, 2009), which also differs from, but still shares common signatures with that in *Arabidopsis*.

Switchgrass and maize are in the same subfamily of the PACCAD clade (Lawrence & Walbot, 2007), and are both out-crossing plant species. Therefore, switchgrass may share some common features in the flowering pathway with maize. Although there are many small-effect QTLs fine-tuning the flowering time in maize, some maize null mutants have dramatic postponed flowering. For example, a homozygous null mutant of maize in *Indeterminate1* (*Id1*) has a late flowering time and prolonged vegetative growth compared to wild type maize (32 leaves in *id1* mutant compared to 13 in wild type at flowering time) (Colasanti *et al.*, 1998). A set of microRNAs, miR156 and miR172, have recently been found important for regulating plant development and flowering independent of environmental cues (autonomous pathway), and are conserved across *Arabidopsis*, maize and rice (Xie *et al.*, 2006; Chuck *et al.*, 2007; Wu *et al.*, 2009). miR156 promotes the juvenile vegetative growth phase at least partially by repressing expression of certain *SPL* genes which are positive regulators of miR172 and other flowering genes; and the expression level of miR156 declines with the development/age of plants (Xie *et al.*, 2006; Wu *et al.*, 2009). miR172 acts downstream of miR156 through *SPL9*, and targets repression of *SMZ* where SMZ directly represses the transcription of *FT* (Mathieu *et al.*, 2009). In contrast to miR156, the expression level of miR172 increases with the development/age of plants (Wu *et al.*, 2009). Overexpression of miR156 significantly promoted the vegetative growth and tiller number in maize (Chuck *et al.*, 2007).

In summary, it is highly possible to isolate homologs of conserved genetic components (such as miR156 and *Id1*) involved in flowering time of switchgrass. Selecting mutants of or directly manipulating these genetic components may lead to a significantly postponed or aborted flowering; an extended vegetative growth, and an increased biomass yield.

4.2 Promoting switchgrass biomass yield by optimizing photosynthetic traits

In theory, the yield of a plant is the product of solar energy that the plant intercepts, utilizes, expends, and stores in harvestable plant biomass (Heaton *et al.*, 2008). The amount of solar energy that a field of plants can intercept depends on the period and length of vegetative growth, the plant architecture and canopy, and planting density in a field. The solar energy utilization of a plant is largely determined by its net photosynthetic efficiency. The C_3 and C_4 photosynthetic pathway and theoretical solar energy conversion efficiency, and their implications on bioenergy grass improvement were recently reviewed (Heaton *et al.*, 2008; Zhu *et al.*, 2008). Here, a few phenotypic traits, potential genes and genetic pathways contributing to these traits are addressed.

4.2.1 Photosynthetic efficiency

C_4 plants have greater photosynthetic efficiency primarily because of the C_4 cycle and CO_2 concentrating mechanism resulting in the avoidance of photorespiration, increased net CO_2 assimilation and higher water use efficiency (Schmitt & Edwards, 1981; Zhu *et al.*, 2008). In most C_4 plants, CO_2 assimilation was processed in two distinct cell types, Kranz mesophyll cells and bundle sheath cells. Firstly, CO_2 reacts with phospoenopyruvate (PEP) by PEP carboxylase (PEPC) and converts into the C_4 acid oxaloacetate (OAA) in Kranz mesophyll cells. Photosynthetic intermediates (e.g. C_4 acids) are diffused or transported between mesophyll cells and bundle sheath cells through plasmodesmata, and decarboxylated to release CO_2 in bundle sheath cells which is used by RubisCO as a substrate (Sowinski *et al.*, 2008). According to the decarboxylation routes, C_4 plants have three subtypes: NADP-Malic Enzyme (NADP-ME) subtype, NAD-Malic Enzyme (NAD-ME) subtype, and PEP-Carboxykinase (PEP-CK) subtype (Edwards *et al.*, 2004; Weber & von Caemmerer, 2010).

Most major C_4 crops, such as maize, sorghum, sugar cane, and miscanthus belong to the NADP-ME subtype, which is more efficient than the other subtypes (Zhu *et al.*, 2008). However, switchgrass is a NAD-ME subtype species. Notably, it is fairly unique that *Panicum* species have all photosynthesis types: C_3, C_3-C_4 intermediate species, and all three C_4 subtypes (Ohsugi & Murata, 1986; Ohsugi & Huber, 1987). In the same *Panicum* genus, the NADP-ME subtype species (e.g. *P. antidotale*) have about 1.5 to 2

times higher PEPC and RubisCO activities than NAD-ME subtype species (e.g. *P. coloratum*), and about 2 times higher PEPC activity but 1.5 time lower RubisCO activity than PEP-CK subtype species (e.g. *P. maximum*) (Ohsugi & Huber, 1987). The huge difference of PEPC and RubisCO activities between species in the *Panicum* genus indicates a spacious room for improving photosynthesis efficiency of switchgrass by improving the activities of these two key enzymes.

The difference of photosynthetic intermediates between NADP-ME and NAD-ME subtypes was nicely reviewed recently (Weber & von Caemmerer, 2010). In NADP-ME subtype species, OAA, once synthesized in the cytosol of mesophyll cells, is directly transported back to the chloroplast and converted into malate in mesophyll cells. Then malate is diffused or transported into chloroplasts of bundle sheath cells, and decarboxylated by NADP-malic enzyme to release CO_2. However, in NAD-ME subtype species, OAA, once synthesized in the cytosol of mesophyll cell, is not transported back to chloroplasts but instead converted into aspartate in the cytosol. Aspartate is then diffused or transported into the mitochondria of bundle sheath cells and converts into malate, which decarboxylates by NAD-malic enzymes in the mitochondria and release CO_2 in bundle sheath cells (Weber & von Caemmerer, 2010). Therefore, the branch point between NADP-ME and NAD-ME subtypes starts from the catalysis of OAA into aspartate or malate, and further diverges in their subcellular transportations possibly because of the presence of selective membrane transporters guarding chloroplasts and the cytosol. Identifying the genetic components behind these differences and engineering the pathway in switchgrass are likely to convert switchgrass into a "synthetic" NADP-ME subtype species with greater photosynthetic efficiency.

Although a finely constructed genomic map is not yet available in switchgrass and other NAD-ME subtype C_4 grasses, high quality genomic sequences of several C_4 grasses (maize, sorghum) and C_3 grasses (rice, *Brachypodium*) are publicly available now (Goff *et al.*, 2002; Yu *et al.*, 2002; Paterson *et al.*, 2009; Schnable *et al.*, 2009; Vogel *et al.*, 2010). Comparative genomic studies have revealed that certain genetic components contribute to the difference of C_3 and C_4 photosynthesis. For example, the comparison between genomes of sorghum and rice showed that "evolution of C_4 photosynthesis in the Sorghum lineage involved redirection of C_3 progenitor genes as well as recruitment and

functional divergence of both ancient and recent gene duplicates" (Paterson *et al.*, 2009). The number of genetic components causing differences between C_4 subtypes should be less than those between C_3 and C_4 plants. Moreover, closely related *Panicum* species comprise a natural pool of photosynthesis types and subtypes. Comparative studies on transcriptomes or genomes between representative *Panicum* species and functional studies on candidate genes will assist in the revelation of the mystery of photosynthesis types and subtypes. The resultant knowledge can be readily used for genetic improvement of switchgrass and other economic plants.

4.2.2 Plant architecture

The amount of light intercepted by a field of plants was largely determined by plant architecture (leaf angle and shapes, plant height and tillering number), planting density and vegetative growth period (Heaton *et al.*, 2008; Wang & Li, 2008). The vegetative growth period can be prolonged by promoting early emergence of tillers of perennial grasses and by delaying flowering as mentioned above. The planting density of a field is dependent on plant architecture. Grass tillering number is important for field establishment. For grass cultivars with less tillering potential, cultivation strategies (e.g. dense planting) can compensate for their disadvantages. Here, we focus on research progress on a few aspects of plant architecture, such as leaf angle, leaf shape and plant height.

Erect leaves (small leaf angle against stem) enhance light interception in densely planted fields (higher leaf area index), and thereby may increase biomass yield (Sakamoto *et al.*, 2005). Decreasing brassinosteroid (BR) content or sensitivity by selecting BR-deficient mutants [e.g. *brassinosteroid-dependent 1 (brd1*), *ebisu dwarf* (*d2*), *dwarf11*, *osdwarf4-1*)], or BR-insensitive mutants [*dwarf 61* (*d61*) and *leaf and tiller angle increased controller* (*oslic*)] can effectively induce erect leaves in rice by altering lamina joint bending (Yamamuro *et al.*, 2000; Hong *et al.*, 2002; Sakamoto *et al.*, 2005; Morinaka *et al.*, 2006; Wang *et al.*, 2008). Specifically, the rice *osdwarf4-1* mutant has erect leaves but no alteration in reproductive development and thereby produced higher grain yields under dense planting conditions without extra fertilizer (Sakamoto *et al.*, 2005). All these BR-related mutants have erect and dark green (higher chlorophyll

content) leaves. However, these mutants are dwarf or semi-dwarf. The dwarf to semi-dwarf stature is important for rice stand and grain yield, as the selection of one semi-dwarf mutant in GA-biosynthesis gene *OsGA20ox2* (*sd1*) successfully led to the development of elite rice cultivars in the "Green Revolution" (Sakamoto *et al.*, 2004); however, dwarf stature is not a desirable trait for bioenergy crops where the above-ground vegetative organs account for the biomass yield. The semi-dwarf to dwarf phenotype in BR-related mutants is caused by failure in the organization and polar elongation of the leaf and stem cells (Yamamuro *et al.*, 2000). On the contrary, GA can positively regulate plant stem elongation (Kende *et al.*, 1998). GA and BR may antagonistically regulate the expression of some downstream genes (Bouquin *et al.*, 2001). Recently, a rice GA-stimulated transcript family gene *OsGSR1* involved in the crosstalk between GA and BR was identified; the study showed that *OsGSR1* is a positive regulator of both GA signaling and BR biosynthesis (Wang *et al.*, 2009). Yet, it is not reported that GA can alter grass leaf angles so far. Therefore, it is possible to engineer grasses with erect leaves but normal or increased plant height by simultaneously manipulating BR and GA-related genes.

Leaf angle and leaf shape are often correlated. In BR mutants, those erect leaves are often short because of the failure of elongation of leaf cells (Yamamuro *et al.*, 2000). In several other cases, rolling (typically upward-curling) leaves cause more erect leaves in rice (Shi *et al.*, 2007; Zhang *et al.*, 2009; Li, L *et al.*, 2010). Rolling leaves may also help preventing water loss by increasing stomata resistance, decreasing leaf temperature, and reducing light interception per leaf but increasing light transmission rates to lower leaves of a plant (O'Toole & Cruz, 1980). Altered expression of a few genes in rice caused rolling leaves, but anatomical reasons for the leaf-curling are different. Rice null mutant of *Shallot-like 1* (*SLL1*, a *KANADI* family gene) has a broader distribution of mesophyll cells in the region where sclerenchymatous cells distribute in wild type rice and bulliform cells on the abaxial side, which thereby induced upward-curling leaves (Zhang *et al.*, 2009). Studies in *Arabidopsis* showed that a group of *YABBY* and *KANADI* family genes regulate abaxial organ identity (Emery *et al.*, 2003; Eshed *et al.*, 2004; Eckardt, 2010); while a group of HD-ZIP III family genes [e.g. *PHABULOSA* (*PHB*) and *PHAVOLUTA* (*PHV*)] promote adaxial organ identity; and it could be the antagonism of

these genes coordinating normal leaf polarity and leaf shape (Emery *et al.*, 2003). Similarly, a maize *KANADI* family gene *Milkweed Pod 1*(*MWP1*), also functions for defining abaxial cell identity (Candela *et al.*, 2008). Recent studies showed that miRNA and genes in other families are also involved in leaf shape formation. For example, overexpression of rice *Argonaute 7* (*OsAGO7*), a gene presumably involved in miRNA metabolism, caused upward-curling leaves in rice (Shi *et al.*, 2007). Overexpression of *Abaxially Curled Leaf 1* (*ACL1*) induced downward rolling (abaxial-curling) leaves by increasing the number and size of bulliform cells in the adaxial side of leaf (Li, L *et al.*, 2010).

In our lab, we obtained erect leaf switchgrass by overexpressing an *Arabidopsis* NAC domain gene *Long Vegetative Phase 1* (*AtLOV1*). Interestingly, the transgenic switchgrass has a typical phenotype of BR-mutants (dark-green and erect leaf) but is not obviously dwarfed (except one transgenic line with an extreme phenotype). Differential gene expression analysis by microarray did not show significant expression changes of identified BR or GA-related genes in the transgenic plants (unpublished). Overexpression of *AtLOV1* in rice induced dark green leaves and a dramatically dwarfed stature, but did not change leaf angle (unpublished). Although the mechanism controlling the phenotype of transgenic switchgrass and rice is unclear, it is possible to alter leaf angle without causing a dramatic negative effect on other vegetative growth traits.

Summary

In short, several important agronomic traits of switchgrass production and potential genes/genetic pathways underlying these traits are reviewed. Translational and functional genomics study will allow us to understand functions of gene(s) and pictures of pathways and interactions between pathways, to better simulate the plant yield, and to design better strategies for plant genetic improvement (Hammer *et al.*, 2004). We can use transgenic or "cisgenic" strategies to quickly "stack" genes of interests from foreign or native genomic origins into switchgrass, and use synthetic biological approaches to engineer and move entire essential genetic components of a pathway into switchgrass (Benner & Sismour, 2005). For example, certain microbial metabolic pathways can be recruited and integrated into plant systems to make plants essentially bio-factories for

desirable products (Somleva *et al.*, 2008). All these approaches are emerging at an unprecedented speed. We can imagine that many genomics tools will be successfully applied to genetic improvement in the near future.

References

Abramson M, Shoseyov O, Shani Z. 2009. Plant cell wall reconstruction toward improved lignocellulosic production and processability. *Plant Science* **178**: 61-72.

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.

Anderson B, Ward J, Vogel K, Ward M, Gorz HJ, Haskins FA. 1988. Forage quality and performance of yearlings grazing switchgrass strains selected for differing digestibility. *Journal of Animal Science* **66**: 2239-2244.

Benner SA, Sismour AM. 2005. Synthetic biology. *Nature Reviews Genetics* **6**: 533-543.

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

Boateng AA, Mullen CA, Goldberg N, Hicks KB, Jung HJG, Lamb JAFS. 2008. Production of bio-oil from alfalfa stems by fluidized-bed fast pyrolysis. *Industrial & Engineering Chemistry Research* **47**: 4115-4122.

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

Boudet A. 2007. Evolution and current status of research in phenolic compounds. *Phytochemistry* **68**: 2722-2735.

Bouquin T, Meier C, Foster R, Nielsen ME, Mundy J. 2001. Control of specific gene expression by gibberellin and brassinosteroid. *Plant Physiology* **127**: 450-458.

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Rosas MO, Rocheford TR, Romay MC, Romero S, Salvo S, Villeda HS, Sofia da Silva H, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen MD. 2009. The genetic architecture of maize flowering time. *Science* **325**: 714-718.

Candela H, Johnston R, Gerhold A, Foster T, Hake S. 2008. The *milkweed pod1* gene encodes a KANADI protein that is required for abaxial/adaxial patterning in maize leaves. *The Plant Cell* **20**: 2073-2087.

Carroll A, Somerville C. 2009. Cellulosic biofuels. *Annual Review of Plant Biology* **60**: 165-182.

Casler MD, Stendal CA, Kapich L, Vogel KP. 2007. Genetic diversity, plant adaptation regions, and gene pools for switchgrass. *Crop Science* **47**: 2261-2273.

Chapple CCS, Vogt T, Ellis BE, Somerville CR. 1992. An *Arabidopsis* mutant defective in the general phenylpropanoid pathway. *The Plant Cell* **4**: 1413-1424.

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

Chuck G, Cigan AM, Saeteurn K, Hake S. 2007. The heterochronic maize mutant *Corngrass1* results from overexpression of a tandem microRNA. *Nature Genetics* **39**: 544-549.

Colasanti J, Yuan Z, Sundaresan V. 1998. The indeterminate gene encodes a zinc finger protein and regulates a leaf-generated signal required for the transition to flowering in maize. *Cell* **93**: 593-603.

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S. 2001. High-throughput screening for induced point mutations. *Plant Physiology* **126**: 480-484.

Compant S, Duffy B, Nowak J, Clement C, Barka EA. 2005. Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Applied and Environmental Microbiology* **71**: 4951-4959.

Corbesier L, Coupland G. 2006. The quest for florigen: a review of recent progress. *Journal of experimental botany* **57**: 3395-3403.

Distelfeld A, Li C, Dubcovsky J. 2009. Regulation of flowering in temperate cereals. *Current Opinion in Plant Biology* **12**: 178-184.

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A. 2004. *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT-like* gene expression independently of Hd1. *Genes & development* **18**: 926-936.

Eckardt NA. 2010. *YABBY* genes and the development and origin of seed plant leaves. *The Plant Cell* **22**: 2103.

Edwards GE, Franceschi VR, Voznesenskaya EV. 2004. Single-cell C4 photosynthesis versus the dual-cell (Kranz) paradigm. *Annuual Review Plant Biology* **55**: 173-196.

Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL. 2003. Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. *Current Biology* **13**: 1768-1774.

Endo A, Nakamura T, Ando A, Tokuyasu K, Shima J. 2008. Genome-wide screening of the genes required for tolerance to vanillin, which is a potential inhibitor of bioethanol fermentation, in *Saccharomyces cerevisiae*. *Biotechnology for Biofuels* **1**: 3-9.

Eshed Y, Izhaki A, Baum SF, Floyd SK, Bowman JL. 2004. Asymmetric leaf development and blade expansion in *Arabidopsis* are mediated by KANADI and YABBY activities. *Development* **131**: 2997-3006.

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

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

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science* **296**: 92-100.

Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N, Grima-Pettenati J. 2005. EgMYB2,

a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *The Plant Journal* **43**: 553-567.

Halpin C. 2004. Re-designing lignin for industry and agriculture. *Biotechnology & Genetic Engineering Reviews* **21**: 229-245.

Hammer GL, Sinclair TR, Chapman SC, Van Oosterom E. 2004. On systems thinking, systems biology, and the in silico plant. *Plant Physiology* **134**: 909-911.

Hatfield R, Ralph J, Grabber J. 1999. Cell wall cross-linking by ferulates and diferulates in grasses. *Journal of the Science of Food and Agriculture* **79**: 403-407.

Heaton EA, Flavell RB, Mascia PN, Thomas SR, Dohleman FG, Long SP. 2008. Herbaceous energy crop development: recent progress and future prospects. *Current Opinion in Biotechnology* **19**: 202-209.

Hong Z, Ueguchi-Tanaka M, Shimizu-Sato S, Inukai Y, Fujioka S, Shimada Y, Takatsuto S, Agetsuma M, Yoshida S, Watanabe Y. 2002. Loss-of-function of a rice brassinosteroid biosynthetic enzyme, C-6 oxidase, prevents the organized arrangement and polar elongation of cells in the leaves and stem. *The Plant Journal* **32**: 495-508.

Hopkins AA, Taliaferro CM, Murphy CD, Christian D. 1996. Chromosome number and nuclear DNA content of several switchgrass populations. *Crop Science* **36**: 1192-1195.

Izawa T, Oikawa T, Sugiyama N, Tanisaka T, Yano M, Shimamoto K. 2002. Phytochrome mediates the external light signal to repress *FT* orthologs in photoperiodic flowering of rice. *Genes & development* **16**: 2006-2020.

Kende H, van der Knaap E, Cho H-T. 1998. Deepwater rice: a model plant to study stem elongation. *Plant Physiology* **118**: 1105-1110.

Komiya R, Yokoi S, Shimamoto K. 2009. A gene network for long-day flowering activates *RFT1* encoding a mobile flowering signal in rice. *Development* **136**: 3443-3450.

Lawrence CJ, Walbot V. 2007. Translational genomics for bioenergy production from fuelstock grasses: maize as the model species. *The Plant Cell* **19**: 2091-2904.

Lee C, Teng Q, Huang W, Zhong R, Ye Z-H. 2009. Down-regulation of *PoGT47C* expression in poplar results in a reduced glucuronoxylan content and an increased wood digestibility by cellulase. *Plant and Cell Physiology* **50**: 1075-1089.

Lemus R, Brummer EC, Moore KJ, Molstad NE, Burras CL, Barker MF. 2002. Biomass yield and quality of 20 switchgrass populations in southern Iowa, USA. *Biomass and Bioenergy* **23**: 433-442.

Li L, Shi Z-Y, Li L, Shen G-Z, Wang X-Q, An L-S, Zhang J-L. 2010. Overexpression of *ACL1* (abaxially curled leaf 1) increased bulliform cells and induced abaxial curling of leaf blades in rice. *Molecular Plant* **3**: 807-817.

Li R, Qu R. 2011. High throughput *Agrobacterium*-mediated switchgrass transformation. *Biomass and Bioenergy* **35**: 1046-1054.

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**: 1-7.

Liang M, Davis E, Gardner D, Cai X, Wu Y. 2006. Involvement of *AtLAC15* in lignin synthesis in seeds and in root elongation of *Arabidopsis*. *Planta* **224**: 1185-1196.

Lionetti V, Francocci F, Ferrari S, Volpi C, Bellincampi D, Galletti R, D'Ovidio R, De Lorenzo G, Cervone F. 2010. Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. *Proceedings of the National Academy of Sciences* **107**: 616-621.

Lu K, Kaeppler SM, Vogel KP, Arumuganathan K, Lee DJ. 1998. Nuclear DNA content and chromosome numbers in switchgrass. *Great Plains Research* **8**: 269-280.

Maloof JN. 2010. Recent advances in regulation of flowering. *F1000 Biology Reports* **2**.

Mathé C, Barre A, Jourda C, Dunand C. 2010. Evolution and expression of class III peroxidases. *Archives of Biochemistry and Biophysics* **500**: 58-65.

Mathieu J, Yant LJ, Mürdter F, Küttner F, Schmid M. 2009. Repression of flowering by the miR172 target SMZ. *PLoS Biol* **7**: e1000148.

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.

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

Missaoui A, Paterson A, Bouton J. 2006. Molecular markers for the classification of switchgrass (*Panicum virgatum* L.) germplasm and to assess genetic diversity in three synthetic switchgrass populations. *Genetic Resources and Crop Evolution* **53**: 1291-1302.

Morinaka Y, Sakamoto T, Inukai Y, Agetsuma M, Kitano H, Ashikari M, Matsuoka M. 2006. Morphological alteration caused by brassinosteroid insensitivity increases the biomass and grain production of rice. *Plant Physiology* **141**: 924-931.

Moser L, Vogel K. 1995. *In: RF Barnes, DA Miller, CJ Nelson. Forages, vol.1: An introduction to grassland agriculture*: Iowa State University Press, Ames, Iowa, USA, pp. 409-420.

Narasimhamoorthy B, Saha M, Swaller T, Bouton J. 2008. Genetic diversity in switchgrass collections assessed by EST-SSR markers. *BioEnergy Research* **1**: 136-146.

O'Toole JC, Cruz RT. 1980. Response of leaf water potential, stomatal resistance, and leaf rolling to water stress. *Plant Physiology* **65**: 428-432.

Ohsugi R, Huber SC. 1987. Light modulation and localization of sucrose phosphate synthase activity between mesophyll cells and bundle sheath cells in C4 species. *Plant Physiology* **84**: 1096-1101.

Ohsugi R, Murata T. 1986. Variations in the leaf anatomy among some C4 Panicum species. *Annals of Botany* **58**: 443-453.

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A. 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* **457**: 551-556.

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.

Saathoff A, Tobias C, Sattler S, Haas E, Twigg P, Sarath G. 2011a. Switchgrass contains two cinnamyl alcohol dehydrogenases involved in lignin formation. *BioEnergy Research*: Online First.

Saathoff A, Sarath G, Chow E, Dien B, Tobias C,. 2011b. Downregulation of Cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. *Plos ONE* **6:** e16416.

Sakamoto T, Miura K, Itoh H, Tatsumi T, Ueguchi-Tanaka M, Ishiyama K, Kobayashi M, Agrawal GK, Takeda S, Abe K. 2004. An overview of gibberellin metabolism enzyme genes and their related mutants in rice. *Plant Physiology* **134**: 1642-1653.

Sakamoto T, Morinaka Y, Ohnishi T, Sunohara H, Fujioka S, Ueguchi-Tanaka M, Mizutani M, Sakata K, Takatsuto S, Yoshida S. 2005. Erect leaves caused by brassinosteroid deficiency increase biomass production and grain yield in rice. *Nature Biotechnology* **24**: 105-109.

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

Schmitt MR, Edwards GE. 1981. Photosynthetic capacity and nitrogen use efficiency of maize, wheat, and rice: a comparison between C3 and C4 photosynthesis. *Journal of experimental botany* **32**: 459-466.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA. 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**: 1112-1115.

Schubert C. 2006. Can biofuels finally take center stage? *Nature Biotechnology* **24**: 777-784.

Sharma MK, Sharma R, Cao P, Bartley LE, Ronald PC. 2011. Construction, characterization and screening of two switchgrass BAC libraries. *Plant & Animal Genome XIX Conference.* San Diego, CA

Shi ZY, Wang J, Wan XS, Shen GZ, Wang XQ, Zhang JL. 2007. Over-expression of rice *OsAGO7* gene induces upward curling of the leaf blade that enhanced erect-leaf habit. *Planta* **226**: 99-108.

Smita R, Nath D. 2008. Manipulation of lignin in plants with special reference to O-methyltransferase. *Plant Science* **174**: 264-277.

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.

Sowinski P, Szczepanik J, Minchin PEH. 2008. On the mechanism of C4 photosynthesis intermediate exchange between Kranz mesophyll and bundle sheath cells in grasses. *Journal of Experimental Botany* **59**: 1137-1147.

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science* **316**: 1033-1036.

Turck F, Fornara F, Coupland G. 2008. Regulation and identity of florigen: *FLOWERING LOCUS T* moves center stage. *Annual Review Plant Biology* **59**: 573-594.

Van Esbroeck GA, Hussey MA, Sanderson MA. 1997. Leaf appearance rate and final leaf number of switchgrass cultivars. *Crop Science* **37**: 864-870.

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

Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K, Lucas S, Harmon-Smith M, Lail K. 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**: 763-768.

Vogel K, Jung H. 2001. Genetic modification of herbaceous plants for feed and fuel. *Critical Reviews in Plant Sciences* **20**: 15-49.

Vogel K, Mitchell R. 2008. Heterosis in switchgrass: biomass yield in swards. *Crop Science* **48**: 2159-2164.

Wang L, Wang Z, Xu Y, Joo S-H, Kim S-K, Xue Z, Xu Z, Wang Z, Chong K. 2009. *OsGSR1* is involved in crosstalk between gibberellins and brassinosteroids in rice. *The Plant Journal* **57**: 498-510.

Wang L, Xu Y, Zhang C, Ma Q, Joo S-H, Kim S-K, Xu Z, Chong K. 2008. OsLIC, a novel CCCH-type zinc finger protein with transcription activation, mediates rice architecture via brassinosteroids signaling. *PLoS ONE* **3**: e3521.

Wang Y, Li J. 2008. Molecular basis of plant architecture. *Annual Review Plant Biology* **59**: 253-279.

Weber APM, von Caemmerer S. 2010. Plastid transport and metabolism of C3 and C4 plants: comparative analysis and possible biotechnological exploitation. *Current Opinion in Plant Biology* **13**: 256-264.

Weng J, Mo H, Chapple C. 2010. Over-expression of *F5H* in *COMT*-deficient *Arabidopsis* leads to enrichment of an unusual lignin and disruption of pollen wall formation. *The Plant Journal* **64**: 898-911.

Wu G, Park MY, Conway SR, Wang J-W, Weigel D, Poethig RS. 2009. The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell* **138**: 750-759.

Xie K, Wu C, Xiong L. 2006. Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant Physiology* **142**: 280-293.

Xu B, Huang L, Shen Z, Welbaum G, Zhang X, Zhao B. 2011a. Selection and characterization of a new switchgrass (*Panicum virgatum* L.) line with high somatic embryogenic capacity for genetic transformation**.** *Scientia Horticulturae* In press.

Xu B, Escamilla-Treviño L, Noppadon S, Shen Z, Shen H, Zhang P, Dixon P, Zhao B. 2011b. Silencing of *4-Coumarate: Coenzyme A ligase* in Switchgrass Leads to Reduced Lignin Content and Improved Fermentable Sugar Yields for Biofuel Production. *New Phytologist* In press.

Xu Z, Zhang D, Hu J, Zhou X, Ye X, Reichel K, Stewart N, Syrenne R, Yang X, Gao P. 2009. Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. *BMC Bioinformatics* **10**: S3.

Yamamuro C, Ihara Y, Wu X, Noguchi T, Fujioka S, Takatsuto S, Ashikari M, Kitano H, Matsuoka M. 2000. Loss of function of a rice *brassinosteroid*

insensitive1 homolog prevents internode elongation and bending of the lamina joint. *The Plant Cell* **12**: 1591-1606.

Yanovsky MJ, Kay SA. 2002. Molecular basis of seasonal time measurement in *Arabidopsis*. *Nature* **419**: 308-312.

Yu J, Hu S, Wang J, Wong GKS, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**: 79-92.

Zalapa J, Price D, Kaeppler S, Tobias C, Okada M, Casler M. 2011. Hierarchical classification of switchgrass genotypes using SSR and chloroplast sequences: ecotypes, ploidies, gene pools, and cultivars. *Theoretical and Applied Genetics* **122**: 805-817.

Zhang GH, Xu Q, Zhu XD, Qian Q, Xue HW. 2009. SHALLOT-LIKE1 is a KANADI transcription factor that modulates rice leaf rolling by regulating leaf abaxial cell development. *The Plant Cell* **21**: 719-735.

Zhang JZ, Creelman RA, Zhu JK. 2004. From laboratory to field. Using information from *Arabidopsis* to engineer salt, cold, and drought tolerance in crops. *Plant Physiology* **135**: 615-621.

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

Zhao Q, Gallego-Giraldo L, Wang H, Zeng Y, Ding S, Chen F, Dixon R. 2010. An NAC transcription factor orchestrates multiple features of cell wall development in *Medicago truncatula*. *The Plant Journal* **63**: 100-114.

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* **107**: 14496-14501.

Zhou J, Lee C, Zhong R, Ye Z. 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *The Plant Cell* **21**: 248-268.

Zhu XG, Long SP, Ort DR. 2008. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Current Opinion in Biotechnology* **19**: 153-159.

Chapter 2

Title: Selection and characterization of a new switchgrass (*Panicum virgatum* L.) line with high somatic embryogenic capacity for genetic transformation

Authors: Bin Xu[1], Linkai Huang[1, 2], Zhengxing Shen[1], Gregory E. Welbaum[1], Xunzhong Zhang[3], Bingyu Zhao[1*]

Affiliations: [1]Department of Horticulture, Virginia Polytechnic Institute and State University, Blacksburg, VA. [2]College of Animal Science and Technology, Si Chuan Agricultural Sciences, Ya-an, Si Chuan, China. [3]Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Corresponding Author: Bingyu Zhao, Department of Horticulture, Virginia Polytechnic Institute and State University, Blacksburg, VA. 24061; Phone: 540-231-1146; Fax: 540-231-3083; Email: bzhao07@vt.edu

ABSTRACT

Switchgrass (*Panicum virgatum* L.) is used horticulturally as an ornamental and agronomically as an animal feedstock and a putative bio-energy crop. Genetic transformation, using somatic embryogenic (SE) callus derived from mature seeds, is one strategy for improving switchgrass traits. A superior switchgrass line, HR8, was developed in this study using recurrent tissue culture selection from cv. Alamo. Eighty two percent of HR8 seeds germinated after harvest comparing to 27% for unselected 'Alamo'. Eighty five percent of germinated HR8 seeds produced SE callus. HR8 seeds had higher endogenous abscisic acid (ABA) contents and responded differently to exogenous additions of ABA in culture. Endophytes were isolated from switchgrass seeds and callus. HR8 callus had less endophytic contamination than 'Alamo' callus. HR8 SE calli were genetically transformable using *Agrobacterium*. Therefore, HR8 is a superior line for generating SE callus and *Agrobacterium*-mediated transformation.

Highlights

- An elite switchgrass line, HR8, was created through recurrent tissue culture selection.
- HR8 had exceptionally high somatic embryogenic callus production and lower seed dormancy.
- HR8 seeds had higher ABA content and responded differently to exogenous ABA.
- HR8 callus had less endophytic contamination.
- HR8 somatic embryogenic callus were transformable with *Agrobacterium tumefaciens*.

Keywords: abscisic acid, *Agrobacterium*-mediated transformation, endophyte, somatic embryogenic callus, seed germination

1. Introduction

Switchgrass (*Panicum virgatum* L.), a versatile C_4, warm-season, perennial grass native to North America, has been widely used as an ornamental and an animal feedstock for decades (McLaughlin and Adams Kszos, 2005). In recent years, switchgrass has been identified as a promising biofuel crop (McLaughlin and Adams Kszos, 2005). Genetic improvement using traditional plant breeding is difficult because switchgrass is an outcrossing, self-incompatible, polyploid monocot species (Bouton, 2008). Several molecular breeding programs have been established to improve switchgrass biomass yield and quality (Bouton, 2008).

One of the essential tools required for switchgrass molecular breeding is an efficient tissue culture and genetic transformation system. Several switchgrass transformation protocols have been published recently (Burris et al., 2009; Li and Qu, 2011; Richards et al., 2001; Somleva et al., 2002; Xi et al., 2009). However, further optimization of existing protocols is required to achieve high throughput transformation in switchgrass. One major hindrance to switchgrass transformation is the low efficiency (10-15%) of somatic embryogenic (SE) callus induction from mature switchgrass seeds (Burris et al., 2009; Li and Qu, 2011). Therefore, improving SE callus production from seeds will greatly improve switchgrass transformation efficiency.

Genotype is critical in determining tissue culture efficiency in various plant species (Landi et al., 2003). Because switchgrass cultivars are heterozygous, desirable traits can be selected from an existing cultivar (Burson et al., 2009).

Seeds of grass species are frequently contaminated with endophytes (Schardl et al., 2004). During tissue culture, endophyte contamination of callus is problematic, inhibiting callus growth and decreasing transformation efficiency (Bajaj et al., 2006). Therefore, selecting a switchgrass germplasm line free of endophytic contamination would be highly desirable for tissue culture and transformation research.

In this study, we report the development of a switchgrass germplasm line, HR8, from the commercial cultivar Alamo, which has improved SE callus production, less endophytic contamination, and can be stably transformed using *Agrobacterium tumefaciens*.

2. Materials and methods

2.1. *Plant material and tissue culture conditions*

All switchgrass plants were grown and maintained in a greenhouse or the Kentland Research Farm of Virginia Tech, Blacksburg, VA. Switchgrass seeds (cv. Alamo) were originally obtained from Ernst Conservation Seeds (Meadville, PA) in 2007 and stored at room temperature and ambient humidity. Switchgrass callus induction and regeneration conditions were followed according to the protocol described by Somleva et al. (2002). Compositions of the tissue culture media are listed in Table 1. All tissue culture chemicals were purchased from PhytoTechnology Lab., Shawnee Mission, KS. Two cycles of tissue culture selection were conducted as illustrated in Fig. 1. In the first cycle, a total of 500 SE calli was used to screen for high regeneration lines in RE Medium. Within 18 days, most SE calli had developed green shoot dots, but only 97 SE calli regenerated elongated shoots. Ten out of the 97 calli regenerated shoots faster than the remaining calli. One plant regenerated from each of the 10 selected calli was maintained, and the selected 10 plants were crossed to each other to produce seeds. Hundreds of seeds harvested from each selected plants were subjected to a second cycle of regeneration screening. The tissue culture-responsive SE calli were subsequently regenerated into plants, selected, and maintained in a greenhouse or the Kentland Farm of Virginia Tech. The tissue culture performance of one line, HR8, was further evaluated in 2009 and 2010. The numbers of germinated seeds, SE callus, and non-embryogenic callus were counted after 40 days on tissue culture medium. Seeds were germinated both in CI Medium and under conditions described by Shen et al. (2001), and similar results were observed. Percentages of SE and nonembryogenic callus were calculated by counting the numbers of each type and dividing by the number of germinated seeds. Callus was transferred to RE Medium and maintained in a growth chamber at 29^{o}C. Callus was induced in the dark, and regeneration was in the light (140 µmol/m^{2}/s cool white fluorescent irradiance) with a 16 h photoperiod.

2.2. *Microscopy*

Typical SE and non-embryogenic callus on CI Medium and regenerating SE callus on RE Medium were observed under a regular microscope (model SZXZ-RFL3, Olympus America, Melville, NY), and were also collected and prepared for scanning

electron microscopy according to the protocol described by Li et al. (2006). The mounted specimens were examined and digital images captured with a Scanning Electron Microscope (Phillips 505).

2.3. *ABA content*

Levels of ABA in switchgrass seeds were measured using indirect enzyme-linked immunoassays (ELISA) (Walker-Simmons and Sesing, 1990). The monoclonal antibody (Agdia, Elkhart, IN) was raised against cis-trans (+) ABA. Cis-trans (+) ABA (Sigma, St Louis, MO) was conjugated to ABA-C4'-BSA for preparing a standard curve that was used to calculate sample concentrations. Alkaline phosphatase-conjugated rabbit antimouse IgG was the secondary antibody (Sigma). The color reaction was measured at 405 nm using a microplate reader (Opsys MR, Thermo Labsystems, Chantilly, VA).

2.4. *Agrobacterium-mediated transformation*

A modified switchgrass transformation procedure described by Somleva et al. (2002) was used. In brief, SE callus were suspended in *Agrobacterium* Infection Medium (OD_{600} = 0.6) and vacuum infiltrated for 10 min at room temperature. The *A. tumefaciens* strain AGL1 was used harboring pCAMBIA1305.1-ubi: *BAR* or pSQ5 (Qu et al., 2008) binary vectors.

To construct vector pCAMBIA1305.1- ubi: *BAR*, a DNA fragment carrying *Ubiquitin* promoter: *BAR*-NosT was cloned from vector pATC17 (Christensen and Quail, 1996) and inserted into the *Hin*dIII site of pCAMBIA1305.1. The pCAMBIA1305.1-Ubi:*BAR* vector harbors *GUSPlus* as the reporter gene, which contains a catalase intron optimizing *GUS* expression in plant tissue. The pSQ5 vector harbors *GFP* as a reporter gene. Both binary vectors used the hygromycin B phosphotransferase gene (*HPTII*) as the selectable marker. After inoculation with *Agrobacterium*, calli were dried on a sterile paper towel before transfer to Co-Cultivation Medium. The calli were co-cultivated with *Agrobacterium* for 4 days at 23^{o}C in the dark. After co-cultivation, the infected calli were transferred without washing to CS Medium. After two months in CS Medium, vigorous callus was transferred to RS Medium. The calli were subcultured onto fresh medium every two weeks.

2.5. *Analysis of transgenic plants*

Calli and regenerated plants, which were transformed with pSQ5, were observed for GFP signals by fluorescent microscopy (model SZXZ-RFL3, Olympus America, Melville, NY) with GFP long-pass, narrow-band GFP, and FITC/TRITC filters (Chroma Technology Corp., Rockingham, VT). Plants transformed with pCAMIBA1305.1-Ubi: *BAR* were stained with GUS solution (Terada and Shimamoto, 1990) and tested for the resistance to the BASTA herbicide as previously described (Richards et al., 2001).

Plant genomic DNA was extracted from hygromycin-resistant plants as described by Missaoui et al., (2005). Southern blot analysis was performed as described previously (Missaoui et al., 2005). In brief, 10 µg switchgrass genomic DNA was digested with the restriction enzyme *Hin*dIII that is absent in the *HPTII* gene and only present in one side the T-DNA region. The digested DNA samples were electrophoresed in 0.8% agrose gels using 0.8% TBE buffer. The DNA was transferred to a nylon membrane (Schleicher and Schuell) using a TurboBlotter protocol (Whatman Ltd). The DNA filters were pre-hybridized in hybridization buffer [6xSSPE pH 7.0, 5x Denhardt Solution, and 0.5% sodium dodecyl sulfate (SDS)] containing 200 mg/ml of denatured Herring sperm DNA at 65 °C for 3 h. A fragment of the *HPT*II gene was amplified by using the primers mentioned above and used as a probe for hybridization at a concentration of $3.5x10^6$ dpm/ml in hybridization buffer. The hybridization was carried out at 68 ^oC for 24 h. After hybridization, the filters were washed with buffer three times for 90 min each in 2x SSC, 0.1% SDS, at 68°C, and exposed to X-ray film at –80 ^oC for 3 days.

2.6. *Identification of endophytes from switchgrass mature seeds and callus*

Mature switchgrass seeds were surface-sterilized with 70% ethanol (Fisher Scientific) for 10 s, washed twice with sterile water, then sanitized in a 50% dilution of 6% sodium hypochlorite (NaClO) with 0.1% Tween 20 for 2 min, and thrice washed with sterile water. Seeds were hand ground using a sterile mortar and pestle in 100 mM Tris-HCl (pH 8.0) and plated on Tryptone Glucose Yeast (TGY) medium (Mukhopadhyay et al., 1996). Genomic DNA was extracted for analysis from the bacterial colonies that grew on TGY medium.

Switchgrass callus exhibiting edema was considered to contain endophytes. Total DNA was extracted from edemic calli, healthy calli, and *in vitro* regenerated switchgrass

leaves. Primers for amplifying 16S and 18S rDNA to identify bacteria and fungi, respectively, were: 16S_FOR: 5'-AGAGTTTGATCCTGGCTCAG-3', 16S_REV: 5'-ACGGATACCTTGTTACGACTT-3', 18S_FOR: 5'-TTAGCATGGAATAATRRAATAGGA-3', 18S_REV: 5'-TCTGGACCTGGTGAGTTTCC-3'. PCR conditions were as described previously (Borneman and Hartin, 2000; Weisburg et al., 1991). The amplified fragments were gel-purified and cloned into the pENTR/D vector (Invitrogen, Carlsbad, CA) and transformed into *Escherichia coli* strain *DH5α*. Ten individual colonies from each transformation were randomly picked for plasmid mini-preparation and sequencing. The sequences were aligned using CLUSTALW and blasted against the Genebank database (www.ncbi.nlm.nih.gov) and 16S rDNA bacterial database (Maidak et al., 2001).

2.7. Statistical analysis

The effects of genotype and ABA on seed germination, as well as SE callus and non-embryogenic callus production, were analyzed using a complete factorial two-way ANOVA design with a minimum of three observations per treatment. Each Petri dish containing 30 seeds was considered as an observation. The comparison of treatment means was analyzed by Tukey HSD multiple comparison procedure using JMP software version 7 (SAS Inc., Cary NC).

3. Results

3.1. Selection of highly regeneration lines from switchgrass cultivar Alamo

During the initial tissue culture screening, SE calli were identified that regenerated many plantlets. The regenerated plants were grown in a greenhouse for seed production. These seeds were subjected to a second round of selection. In 2009, one particularly productive line, HR8, was identified. HR8 plants were cross pollinated with HR7, another productive selected line, for seed production in 2009 and 2010 (Fig. 1). HR lines and unselected 'Alamo' plants were grown in the same field but separate locations to prevent crossing.

3.2. HR8 exhibits high seed germination and SE callus production capacities

Newly harvested switchgrass seeds have deep seed dormancy that may take as long as 2 years of after-ripening to break (Shen et al., 2001). Untreated, fresh unselected 'Alamo' seeds germinated 26.8% (on average), while 82% of HR8 seeds germinated after

harvest. From the germinated seeds, unselected ‘Alamo’ produced 35.5% SE callus, while HR8 produced 84.9% SE callus (Fig. 2). The 83.6% regeneration of SE callus from HR8 was much higher than the 20.8% from unselected ‘Alamo’ on RE Medium (Fig. 2).

3.3. Characterization of callus

Switchgrass non-embryogenic callus had a distinct opaque gelatinous appearance (Fig. 3 A). In contrast, SE callus was pale yellow, granular and friable, or alternatively pale white, compact and dense (Fig. 3 B). Somatic embryo callus had smooth surfaces and granular structures when viewed under a microscope (Fig. 3 C). SE callus began to regenerate shoots within 10 days on RE Medium, while non-embryogenic callus failed to regenerate.

When viewed by electron microscopy, switchgrass non-embryogenic callus had a coarse surface structure with long, tubular cells (Fig. 4 A), while SE callus produced multiple granular structured SEs at different developmental stages (Fig. 4 B-C). After 10 days on RE Medium, shoots and leaves emerged from the coleoptiles of the SEs (Fig. 4 D).

3.4. Measurement of endogenous ABA and responses to exogenous ABA

The phytohormone ABA plays an important role in seed dormancy and somatic embryogenesis (Koornneef et al., 2002; Leung and Giraudat, 1998; von Arnold et al., 2002). To test effects of ABA, we germinated both HR8 and the unselected ‘Alamo’ seeds on CI Medium supplemented with various concentrations of ABA. There were significant interactions between ABA and genotype with respect to the production of SE callus and germination percentage (Table 2). We also observed that dehulling improved the germination of the unselected ‘Alamo’ seeds to 44.7% in comparison of the intact seeds that germinated 26.8%.

Exogenous ABA at concentrations from 1 to 100 μM increased the SE callus production for HR8 but not unselected ‘Alamo’ seeds. In terms of seed germination, exogenous application of greater than 10 μM ABA strongly inhibited seed germination of the unselected ‘Alamo’ but not HR8 seeds. We also measured endogenous ABA content of HR8 and unselected ‘Alamo’ seeds harvested from plants grown under the same conditions. ABA content of HR8 seeds [34.1 ±10.6 ng/g (means ± SE)] was significantly higher than that of ‘Alamo’ (11.2 ±1.6 ng/g). It is possible the high endogenous ABA

level contributes to the improved SE capacity of HR8, but further experiments are required to test whether/how the endogenous ABA is related to the production of SE in HR8.

3.5. SE callus derived from the HR8 line were transformable by Agrobacterium

HR8 SE calli were transformed with *Agrobacterium tumefaciens* strain AGL1 harboring expression vectors pCAMBIA1305.1-ubi: *BAR* or pSQ5. Thirty-two transgenic plants derived from pSQ5 and five plants derived from pCAMBIA1305.1-ubi: *BAR* were transplanted into soil and maintained in a greenhouse. Thirty days after transformation with pCAMBIA1305.1-ubi: *BAR,* 12.8% (15 of 117) of calli showed some degree of GUS-staining (Fig. 5 A). The 15 regenerated transgenic plants were maintained in greenhouse where leaves could also be GUS-stained and tested for BASTA resistance (Fig. 5 B, C). Thirty days following transformation with pSQ5, GFP was visualized in 11.3% (60 of 529) of calli (Fig. 5 D). GFP was also observed in the newly regenerated shoots and root tips of transgenic plants (Fig. 5 E-F). The integration of transgenes into the switchgrass genome was detected by southern blot analysis using the *HPTII* gene fragment as a probe, which demonstrated most lines contained one or two copies of the transgene (Fig. 5 J).

3.6. Identification of endophytes from switchgrass seeds

A high percentage of calli (both SE and non-SE) derived from the unselected ‘Alamo’ seeds, exhibited edema, typical of endophytic contamination (Fig. 3 F). Contaminated callus grew slowly and had lower regeneration percentages. The endophyte(s) was not eliminated by heating the seeds at 55 ℃ or treating seeds with Plant Preservative Mixture (Plant Cell Tech Inc., Washington, DC) (data not show).

Endophytes were isolated from seeds of switchgrass ‘Alamo’. Six bacterial colonies were randomly chosen based on their morphology for 16S rDNA sequencing and were classified as belonging to: the genus *Curtobacterium* (GeneBank ID: bankit1301295), the genus *Streptophyta* (GeneBank ID: bankit1301293), three were members of an unclassified genera in the family *Enterobacteriaceae* (GeneBank ID: bankit1301289, bankit1301290, bankit1301294), and one unclassified root (GeneBank ID: bankit1301292). However, no endophytic bacterium was isolated or detected by 16S rDNA PCR amplification from calli with or without edema, since 16S rDNA only exists

in prokaryotes. The dehulling and sterilization steps may have completely eliminated endophytic bacteria from switchgrass caryopses.

The calli with edema had a smooth shape and water-soaked appearance (Fig. 3 F). Attempts to culture the endophyte (s) from the contaminated calli on artificial media were unsuccessful, but we were able to detect the endophyte by PCR using 18S rDNA primers, indicating a eukaryotic identity (e.g. fungus). It is possible the edema calli was triggered by the infection of endophytes.

To confirm the identity of endophytes isolated from the edema calli, we amplified the 18S rDNA fragments by PCR using the DNA isolated from the edema calli and cloned into a pENTR/D vector for sequencing. From 10 randomly selected plasmid clones, the 18S rDNA sequences were identical, suggesting a single endophyte (GeneBank ID: bankit1301296). The 18S rDNA sequence of the endophyte shared 98% nucleotide identity with an unidentified eukaryote clone (GeneBank ID: AY897976.1), which exists in rice, ryegrass, and other grass species. The identical 18S rDNA fragment was also detected in surface-sterilized seeds collected from different locations in multiple years. Therefore, the switchgrass seeds and the callus derived from them were apparently all contaminated with the same fungal endophyte.

Seeds harvested from the HR8 line were less contaminated with the endophyte. There were 23.7 and 65.7% of calli of HR8 and 'Alamo' associated with endophytes, respectively, based on visual observation of callus derived from >200 seeds and PCR amplification with 18S rRNA primers.

4. Discussion

Recurrent selection for high *in vitro* regeneration capacity led to the identification of maize inbred lines especially suitable for SE callus production, anther culture and genetic transformation (Lowe et al., 2006; Machii et al., 1998; Marhic et al., 1998; Rosati et al., 1994). A similar strategy led us to identify the tissue culture-responsive switchgrass line HR8. Since switchgrass is a perennial grass, the HR8 lines can be maintained as a clone in a greenhouse or field to supply mature seeds to generate SE callus. Recently, SE callus was also generated from switchgrass inflorescence tissue (Burris et al., 2009). However, it is more convenient to produce SE callus for transformation from mature seeds, since

the availability of suitable explants is not limited (Xi et al., 2009). The selected SE callus of HR8 was sub-cultured for at least six months with no obvious loss of regeneration ability (data not shown), and few visible somatic mutations were observed in regenerated plants. It is possible that the polyploid nature of the switchgrass genome buffers genetic changes caused by tissue culture, explaining why phenotypic variation was rare.

HR8 SE calli were genetically transformed using *A. tumefaciens* by slightly modifying the protocol of Somleva et al. (2002). The modifications included: (1) after the co-cultivation with *Agrobacterium*, the calli were directly transferred onto CS Medium without washing to avoid the loss of fragile SE callus and to save labor. (2) Augmentin at 375 mg l^{-1} was added to the selection media to effectively inhibit growth of *Agrobacterium*. A stable transformation efficiency of 50% has been reported using a new transformation protocol and SE callus of 'Alamo' (Li and Qu, 2011). In the protocol developed by Li and Qu (2011), a high concentration of proline was included in tissue culture media, the co-cultured calli were desiccated and recovered on medium without hygromycin selection for 2 weeks. These modifications significantly improved the switchgrass transformation efficiency. In the report by Li and Qu, the cultivar 'Alamo' and 'Performer' produced less than 10 % SE calli, which is similar to our results. By using SE calli derived from HR8 and following the transformation protocol developed by Li and Qu (2011), switchgrass transformation efficiency could be further improved.

HR8 seeds also exhibited higher seed germination percentages than either commercial (data not shown) or unselected 'Alamo' seeds. Tissue culture selection for SE callus production capacity could favor this trait since only germinated seeds can produce calli (Lowe et al., 2006). Non-dormant switchgrass is highly valued because seed dormancy complicates successful field establishment for crop production (Shen et al., 2001).

Mature HR8 seeds also had significantly higher concentrations of endogenous ABA and responded to exogenous ABA differently than unselected 'Alamo'. Abscisic acid levels and sensitivity may contribute to the high SE production capacity of HR8. However, the role of ABA in switchgrass seed dormancy and SE production requires further investigation.

Many grass species are in symbiosis with endophytes (Schardl et al., 2004). However, in tissue culture, endophyte-free SE callus is desirable (Bajaj et al., 2006). The lower percentage of endophytic contamination in HR8 seeds was at least partially responsible for the improved production of SE callus. Since endophytes may enhance switchgrass growth by mitigating environmental stress in the field, further studies should focus on characterizing the causal endophyte as well as the interaction between it and switchgrass genotypes.

In summary, a new elite switchgrass line, HR8, was bred through recurrent tissue culture selection from 'Alamo'. HR8 seeds have higher ABA content and respond differently to exogenous ABA than 'Alamo'. HR8 has exceptionally high somatic embryogenic callus production with a lower percentage of callus contamination by endophyte and lower seed dormancy. HR8 somatic embryogenic callus were transformable with *Agrobacterium tumefaciens*.

5. Availability

The HR8 seeds will be deposited in the switchgrass collection at the USDA-ARS National Center for Genetic Resources Preservation (NCGRP) in Ft. Collins, Colorado USA. Small samples of HR8 seeds are available upon request from the corresponding author. Vegetative materials are maintained at the Virginia Tech, Kentland Research Farm, Blacksburg, VA.

Acknowledgements

The authors thank Dr. Richard Veilleux for comments on the manuscript. The AGL1 *Agrobacterium* strain was kindly provided by John Vogel (USDA-ARS, Western Regional Research Center, Albany, CA), and the pSQ5 binary vector by Venkatesan Sundaresan (University of California, Davis, CA). Some of the switchgrass seeds used in this study in the year 2007 were provided by Judy Fielder, Virginia Tech, Blacksburg, VA. This project was supported by a grant from the Virginia Agricultural Council (Project number 545), seed grants of Biodesign and Bioprocessing Research Center of The College of Agriculture and Life Sciences (BBRC) at Virginia Tech, and the Virginia Agricultural Experimental Station (VA135872).

Table 1. Tissue culture media used in this study.

Medium Name	Purpose	Compositions
CI Medium	SE callus induction	MS salts and vitamins, 3% maltose, 0.8% agar, 22.5 μM 2,4-D and 5 μM 6-BA
RE Medium	SE callus regeneration	MS salts and vitamins, 3% maltose, 0.8% agar, 1.44 μM GA_3
Infection Medium	Agrobacterium suspension and infection	1/2MS salts and vitamins, 3% maltose, 200 μM acetosyringone
Co-Cultivation Medium	Agrobacterium co-cultivation	1/2MS salts and vitamins, 3% maltose, 0.8% agar, 200 μM acetosyringone, 22.5 μM 2,4-D and 5 μM 6-BA
CS Medium	Transformed callus selection	CI Medium with 375 mg/l Augmentin, and 50 mg/l hygromycin
RS Medium	Transformed callus regeneration	RE Medium with 375 mg/l Augmentin, and 50 mg/l hygromycin

Table 2. Effects of ABA treatments on seed germination and SE callus formation of two switchgrass lines.

Switchgrass lines	**ABA conc.** μM	**Seed germination percentage (%)**	**SE callus percentage (%)**
HR8	0	81.9 ±10.2 [a]	87.1 ±4.4
HR8	1	79.0 ±8.3	91.6 ±4.4
HR8	10	57.6 ±5.5	98.9 ±4.9
HR8	100	51.1 ±6.4	96.8 ±5.7
Unselected	0	44.7 ±6.7	38.1 ±4.9
Unselected	1	43.6 ±10.7	27.4 ±4.9
Unselected	10	29.8 ±7.6	25.7 ±4.9
Unselected	100	7.3 ±6.8	0.0
		F(Lines) **[b]	F(Lines) **
		F(ABA) **	F(ABA) **
		F(Lines)(ABA) **	F(Lines)(ABA) **

[a] Means ± SE are based on at least three Petri plates of 30 seeds each.
[b] F of Fisher (from ANOVA). ** $p \leq 0.01$.

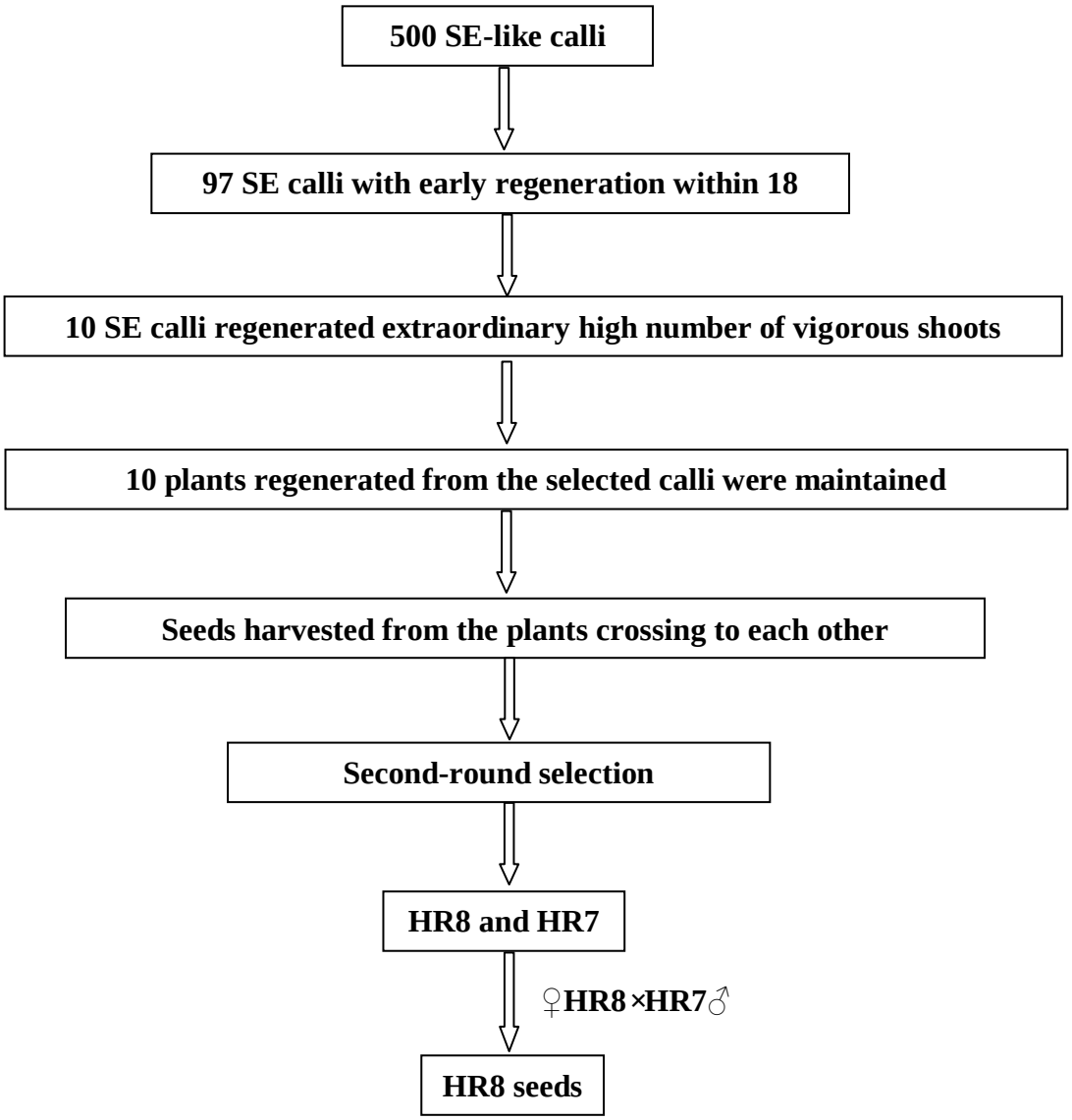


Fig. 1. Selection of the highly regenerable line HR8

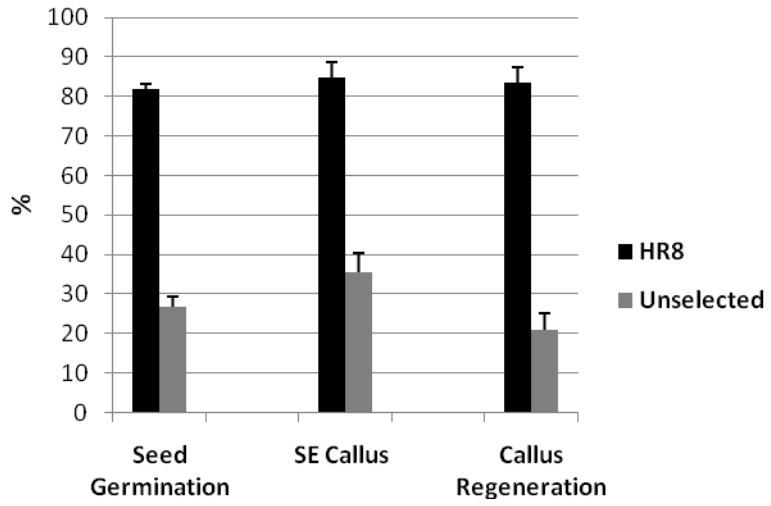


Fig. 2. HR8 has high seed germination, SE callus production percentage, and SE calli of HR8 have high regeneration percentage compared to unselected ‘Alamo’.

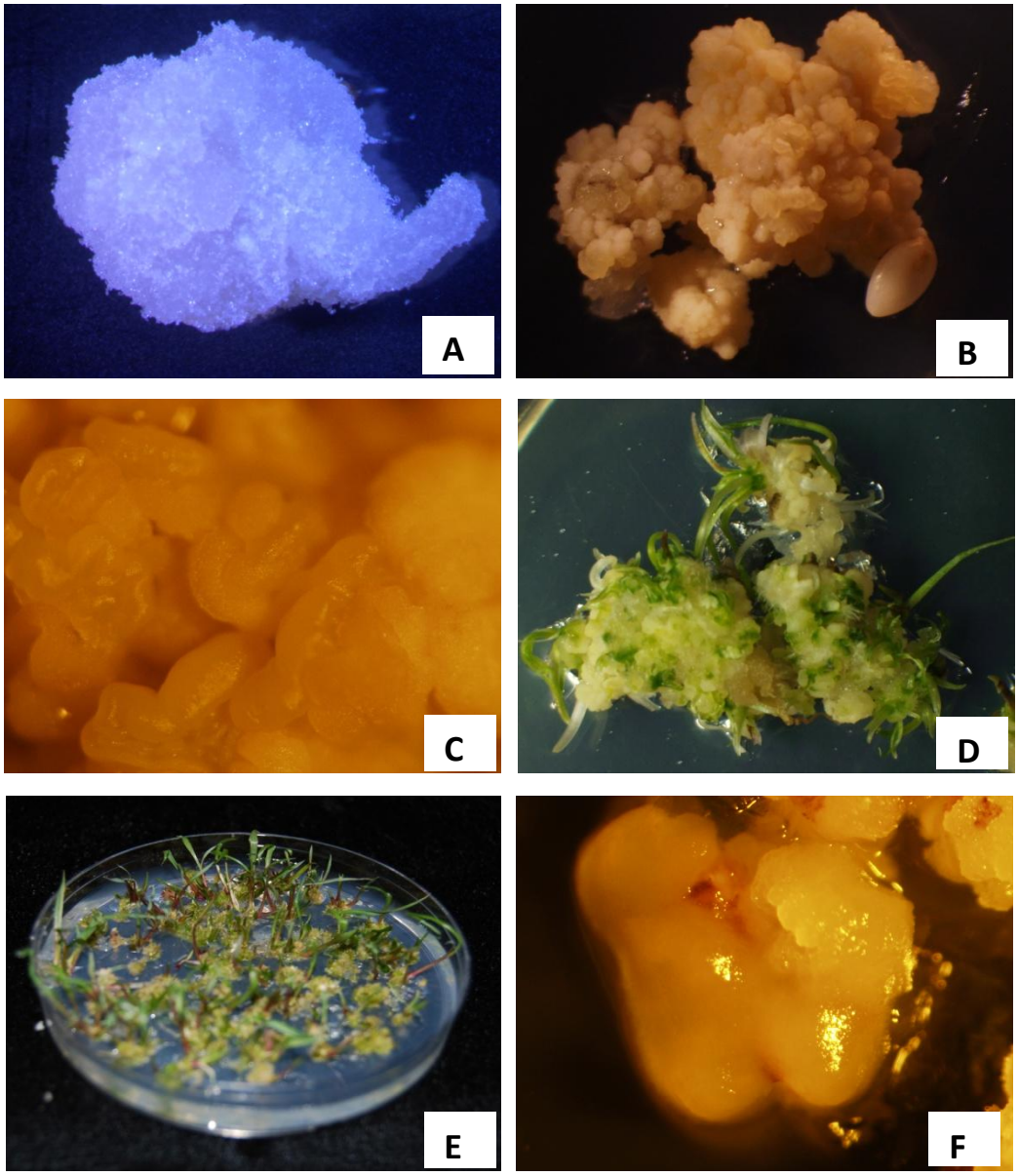


Fig. 3. Tissue culture of switchgrass line HR8. A. Non-embryogenic callus. B. SE callus. C. SE callus under microscope (20× magnification). D and E. Regeneration of SE callus in RE Medium after 10 (D) and 20 (E) days. F. Endophytic fungus-contaminated callus.

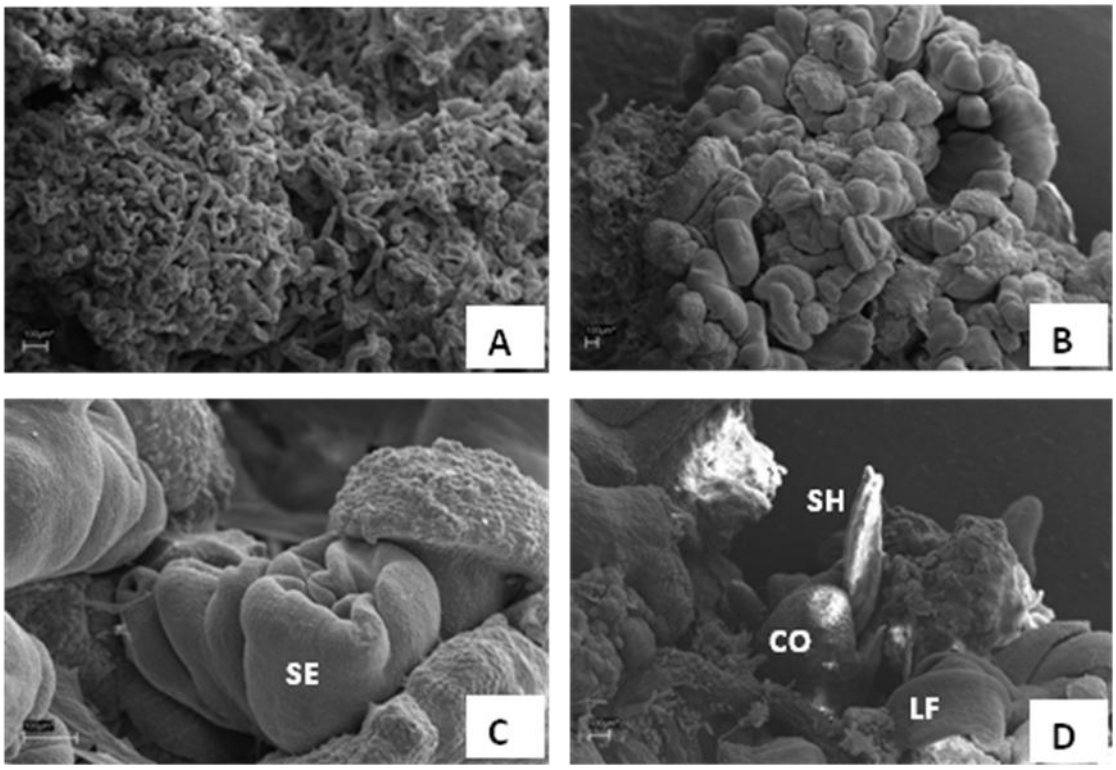


Fig. 4. Morphology of switchgrass callus under scanning electron microscope. The bar in each figure respresents 100 μm. A. Non-embryogenic callus showing irregularly interwined long, tubular cells on its surface. B. SE callus showing granular SEs on the surface. C. Two typical SEs. D. Regeneration of SE callus after 10 days in RE Medium showing shoot (SH) growing out of coleoptile (CO) and expanded leaves (LF) of shoots.

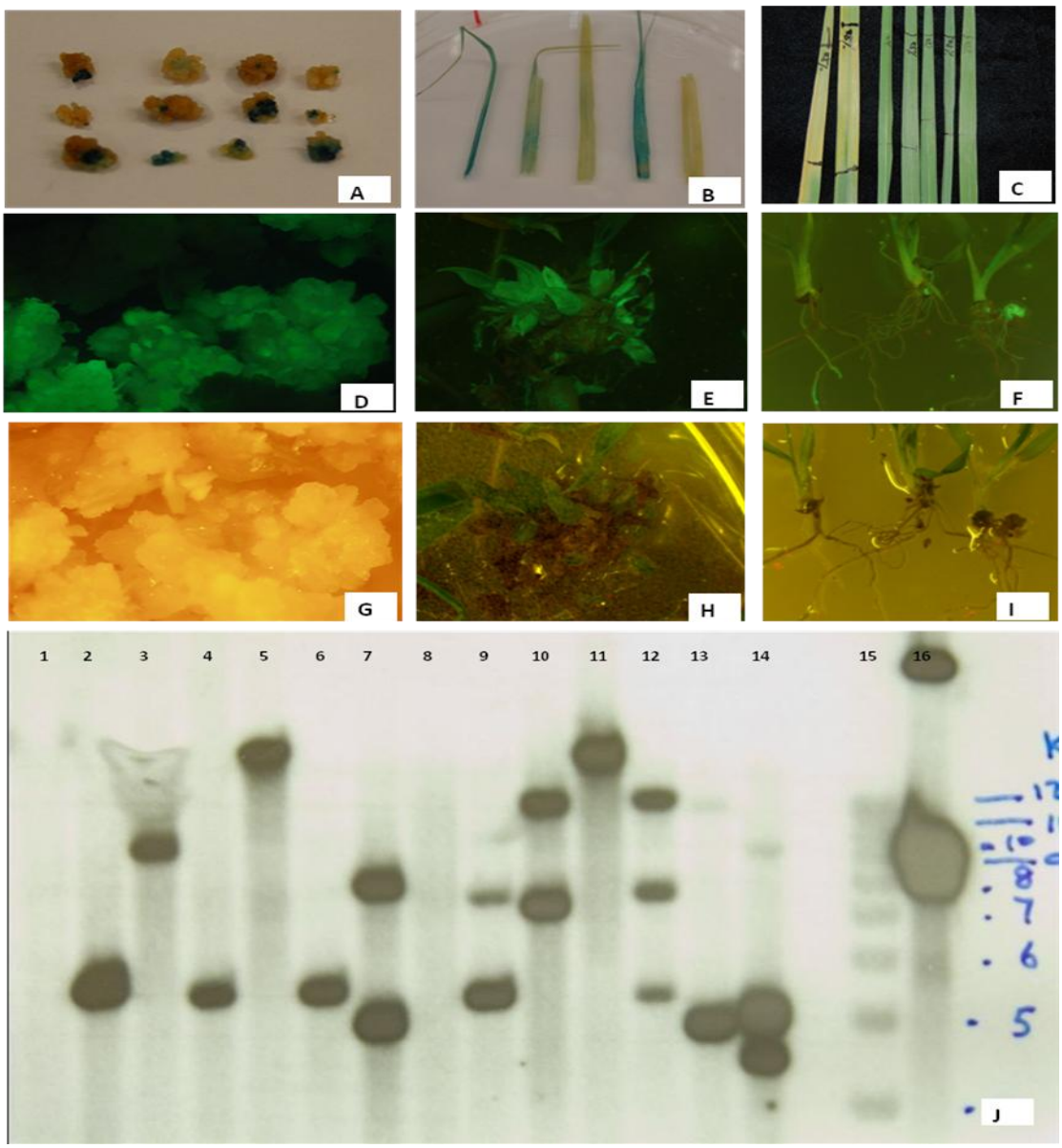


Fig. 5. *Agrobacterium*-mediated transformation of switchgrass line HR8 with two different plasmid constructs. For the transformation with pCAMBIA1305.1- Ubi-*BAR*, tranformed callus (A) and leaf sections were stained with GUS solution (B) (from left to right, 1-4 are from individual transgenic lines, and 5 is from a wildtype plant); The leaves of transgenic plants are resistant to the BASTA (C) (from left to right:1 and 2 are wildtype plants showing susceptbility to BASTA, 3-7 are individual transgenic lines that are resistant to BASTA). For the transformation with pSQ5, GFP flourencence was visualized under UV light in transformed callus (D), regenerated plant shoots (E), and new roots and shoots of transgenic seedlings (F). As control, the same tissue was visualized under regluar bright light (G to I). Southern blot analysis of transgenic plants with *Hpt* probe (J). Lane 1 is the negative control, lanes 2-14 are independent putative transgenic plants. Lane 15 is the 1kb ladder, lane 16 is the control plasmid DNA pCambia1305-Ubi-Bar.

References

Agrios, G.N., 2005. Plant Pathology, 5th ed. Elsevier, London, UK.

Bajaj, S., Ran, Y., Phillips, J., Kularajathevan, G., Pal, S., Cohen, D., Elborough, K., Puthigae, S., 2006. A high throughput *Agrobacterium tumefaciens*-mediated transformation method for functional genomics of perennial ryegrass (*Lolium perenne* L.). Plant Cell Rep. 25, 651-659.

Borneman, J., Hartin, R., 2000. PCR primers that amplify fungal rRNA genes from environmental samples. Appl. Env. Microb. 66, 4356.

Bouton, J., 2008. Improvement of switchgrass as a bioenergy crop, In: Vermerris, W. (Ed.), Genetic Improvement of Bioenergy Crops. Springer New York, pp. 309-345.

Burris, J., Mann, D., Joyce, B., Stewart, C., 2009. An improved tissue culture system for embryogenic callus production and plant regeneration in switchgrass (*Panicum virgatum* L.). BioEnerg. Res. 2, 267-274.

Burson, B., Tischler, C., Ocumpaugh, W., 2009. Breeding for reduced post-harvest seed dormancy in switchgrass: Registration of TEM-LoDorm switchgrass germplasm. J. Plant Regist. 3, 99.

Christensen, A.H., Quail, P.H., 1996. Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res. 5, 213-218.

Koornneef, M., Bentsink, L., Hilhorst, H., 2002. Seed dormancy and germination. Curr. Opin. Plant Biol. 5, 33-36.

Landi, P., Chiappetta, L., Sanguineti, M., Tuberosa, R., 2003. Bo21: A maize inbred line with high regeneration capacity from callus cultures of immature embryos. Maydica 48, 249-250.

Leung, J., Giraudat, J., 1998. Abscisic acid signal transduction. Ann. Rev. Plant Biol. 49, 199-222.

Li, R., Bruneau, A., Qu, R., 2006. Improved plant regeneration and *in vitro* somatic embryogenesis of St Augustine grass [*Stenotaphrum secundatum* (Walt.) Kuntze]. Plant Breeding 125, 52-56.

Li, R., Qu, R., 2011. High throughput Agrobacterium-mediated switchgrass transformation. Biomass Bioenerg. 35, 1046-1054.

Lowe, B., Way, M., Kumpf, J., Rout, J., Warner, D., Johnson, R., Armstrong, C., Spencer, M., Chomet, P., 2006. Marker assisted breeding for transformability in maize. Mol. Breeding 18, 229-239.

McLaughlin, S., Adams Kszos, L., 2005. Development of switchgrass (*Panicum virgatum*) as a bioenergy feedstock in the United States. Biomass Bioenerg. 28, 515-535.

Missaoui, A.M., Paterson, A.H., Bouton, J.H., 2005. Investigation of genomic organization in switchgrass (*Panicum virgatum* L.) using DNA markers. Theor. Appl. Genet. 110, 1372-1383.

Mukhopadhyay, K., Garrison, N., Hinton, D., Bacon, C., Khush, G., Peck, H., Datta, N., 1996. Identification and characterization of bacterial endophytes of rice. Mycopathologia 134, 151-159.

Qu, S., Desai, A., Wing, R., Sundaresan, V., 2008. A versatile transposon-based activation tag vector system for functional genomics in cereals and other monocot plants. Plant physiol. 146, 189-199.

Richards, H., Rudas, V., Sun, H., McDaniel, J., Tomaszewski, Z., Conger, B., 2001. Construction of a GFP-BAR plasmid and its use for switchgrass transformation. Plant Cell Rep. 20, 48-54.

Rosati, C., Landi, P., Tuberosa, R., 1994. Recurrent selection for regeneration capacity from immature embryo-derived calli in maize. Crop Sci. 34, 343-347.

Schardl, C.L., Leuchtmann, A., Spiering, M.J., 2004. Symbiosis of grass with seed borne fungal endophytes. Annu Rev Plant Biol. 55, 315-340.

Shen, Z., Parrish, D., Wolf, D., Welbaum, G., 2001. Stratification in switchgrass seeds is reversed and hastened by drying. Crop Sci. 41, 1546-1551.

Somleva, M., Tomaszewski, Z., Conger, B., 2002. Agrobacterium-mediated genetic transformation of switchgrass. Crop Sci. 42, 2080-2087.

Terada, R., Shimamoto, K., 1990. Expression of CaMV35S-GUS gene in transgenic rice plants. Mol. Gener. Genet. 220, 389-392.

von Arnold, S., Sabala, I., Bozhkov, P., Dyachok, J., Filonova, L., 2002. Developmental pathways of somatic embryogenesis. Plant Cell Tiss. Org. 69, 233-249.

Walker-Simmons, M., Sesing, J., 1990. Temperature effects on embryonic abscisic acid levels during development of wheat grain dormancy. J. Plant Growth Regul. 9, 51-56.

Weisburg, W., Barns, S., Pelletier, D., Lane, D., 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. 173, 697-703.

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. Bioenerg. Res. 2, 275-283.

Chapter 3

Title: Silencing of *4-Coumarate: Coenzyme A ligase* in Switchgrass Leads to Reduced Lignin Content and Improved Fermentable Sugar Yield for Biofuel Production

Authors: Bin Xu[1*], Luis L. Escamilla-Treviño[2,3*], Sathitsuksanoh Noppadon[4], Zhengxing Shen[1], Hui Shen[2], Y-H Percival Zhang[3,4], Richard A. Dixon[2,3], Bingyu Zhao[1,†]

[1]Department of Horticulture, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA; [2]Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA, [3]US Department of Energy BioEnergy Science Center (BESC); [4]Department of Biological System Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

* These authors contributed equally to this work.

† Corresponding author; email bzhao07@vt.edu; Tel: +1 540 231 1146; Fax: 540-231-3083

Summary

- The lignin content of feedstock has been proposed as one key agronomic trait impacting biofuel production from lignocellulosic biomass. 4-Coumarate: Coenzyme A ligase (4CL) is one of the key enzymes involved in the monolignol biosynthethic pathway.

- Two homologous *4CL* genes, *Pv4CL1* and *Pv4CL2*, were identified in switchgrass (*Panicum virgatum* L.) through phylogenetic analysis. Gene expression patterns and enzymatic activity assays suggested that Pv4CL1 is involved in monolignol biosynthesis. Stable transgenic plants were obtained with *Pv4CL1* downregulated.

- RNA interference of *Pv4CL1* reduced extractable 4CL activity by 80%, leading to a reduction in lignin content with decreased guaiacyl unit composition. Altered lignification patterns in the stems of RNAi transgenic plants were observed with phloroglucinol-HCl staining. The transgenic plants also had uncompromised biomass yields. After dilute acid pretreatment, the low lignin transgenic biomass had significantly increased cellulose hydrolysis (saccharification) efficiency.

- The results demonstrate that Pv4CL1, but not Pv4CL2, is the key 4CL isozyme involved in lignin biosynthesis, and reducing lignin content in switchgrass biomass by silencing *Pv4CL1* can remarkably increase the efficiency of fermentable sugar release for biofuel production.

Key words:

***Panicum virgatum*; monolignol pathway; *4-Coumarate: Coenzyme A ligase* (*4CL*); genetic transformation; saccharification; bioenergy.**

Introduction

The production of biofuels from renewable biomass could alleviate the dependence on fossil fuels, and this concept has led to a strong interest in developing biofuel feedstock crops and new biofuel conversion technologies (Carroll & Somerville, 2009). Switchgrass (*Panicum virgatum*), a warm-season perennial C_4 grass, has been considered as one prime candidate for lignocellulose-based feedstock production in the US (McLaughlin & Adams Kszos, 2005). One major breeding objective is to improve switchgrass feedstock quality for "transforming grass to gas" (Schubert, 2006). Feedstock quality essentially equates to the optimized cell wall composition of biomass, which impacts the efficiency of biofuel production through (bio)chemical conversion of sugars to fuels (Carroll & Somerville, 2009). Two major cell wall components, cellulose and hemicellulose, are the primary carbohydrate sources for lignocellulose-based bio-ethanol production through fermentation; while another cell wall component, lignin, adversely impacts bioconversion (Chen & Dixon, 2007). Lignin tightly binds to hemicellulose and cellulose, thereby blocking the access of hydrolytic enzymes, and also possibly inhibiting the activities of hydrolytic and fermentation enzymes during the bioconversion processes (Halpin, 2004; Keating *et al.*, 2006; Endo *et al.*, 2008; Abramson *et al.*, 2009). Therefore, it is hypothesized that switchgrass feedstock quality for bio-ethanol production can be improved by decreasing its lignin content (Carroll & Somerville, 2009). Comprehensive characterization of lignin biosynthetic pathways in switchgrass will enable us to manipulate the lignin content of switchgrass biomass through genetic engineering. Research on the molecular mechanisms regulating lignin biosynthesis in switchgrass has just started (Escamilla-Treviño *et al.*, 2009; Fu *et al.*, 2011a; Fu *et al.*, 2011b; Saathoff *et al.*, 2011a; Saathoff *et al.*, 2011b). One switchgrass lignin biosynthesis gene, *cinnamyl-alcohol dehydrogenase* (*CAD*), was recently identified, and the downregulation of switchgrass *CAD1* resulted in a decreased lignin content of switchgrass biomass that potentially enhances biofuel production (Fu *et al.*, 2011a; Saathoff *et al.*, 2011a). The overall biomass production of the low-lignin switchgrass plants was not characterized in these reports (Fu *et al.*, 2011b; Saathoff *et al.*, 2011a), and therefore an argument could not be made for the advantages of growing *CAD* downregulated switchgrass for feedstock production. In contrast, switchgrass plants

downregulated in the expression of another monolignol biosynthesis gene, *caffeic acid 3-O-methyltransferase* (*COMT*), were shown to have normal growth behavior and exhibit reduced recalcitrance for saccharification and fermentation to ethanol (Fu *et al.*, 2011b). In addition to providing proof of concept for lignin engineering in switchgrass, these results clearly confirm that the lignin biosynthetic pathways are evolutionarily conserved in different plant species including switchgrass (Xu *et al.*, 2009; Weng & Chapple, 2010). In contrast to switchgrass, the monolignol biosynthetic pathways have been well studied in model plant species, such as Arabidopsis, alfalfa (*Medicago sativa*), and poplar (*Populus trichocarpa x Populus deltoids*, *P. tremuloides*, or *P. tomentos*) (Smita & Nath, 2008; Carroll & Somerville, 2009). The knowldege of lignin synthesis from model plant species enables us to identify lignin-related genes in switchgrass, and therefore to manipulate the lignin content in switchgrass biomass at different stages in the pathway to optimize processing efficiency.

The phenolic polymer lignin is derived from ρ-hydroxycinnamic alcohols (monolignols) via combinatorial radical coupling reactions (Boudet, 2007; Umezawa, 2010). Approximately ten key enzymes are involved in the monolignol biosynthetic pathway in model plant species (Hisano *et al.*, 2009), and most of these gene-homologues could be identified from the switchgrass EST database (Tobias *et al.*, 2008). Among the monolignol biosynthesis enzymes, 4-(hydroxy) cinnamoyl CoA ligase (4CL) is a key enzyme involved in early steps of the monolignol biosynthetic pathway. 4CL catalyzes the formation of activated thioesters to convert ferrulic acid and sinapic acid to their corresponding alcohols, which may act as substrates for entry into different branch pathways of phenylpropanoid metabolism (Lee *et al.*, 1997). *4CL* genes normally belong to a small gene family. In Arabidopsis, three 4CL isozymes, At4CL1, At4CL2, and At4CL3, with different substrate preferences and gene expression patterns, have been identified. At4CL1 and At4CL2 are involved in the monolignol biosynthetic pathway, while At4CL3 participates in flavonoid and other non-lignin biosynthetic pathways (Ehlting *et al.*, 1999; Cukovica *et al.*, 2001). In poplar, two functionally divergent 4CLs were identified. Ptr4CL1 is devoted to lignin biosynthesis in developing xylem tissues; whereas Ptr4CL2 is possibly involved in flavonoid biosynthesis in epidermal cells (Hu *et al.*, 1998). Downregulaton of *At4CL1* in Arabidopsis or *Ptr4CL1* in poplar resulted in

reduced lignin content (Hu *et al.*, 1999; Sanchez *et al.*, 2006; Voelker *et al.*, 2010) and little changed biomass production (Sanchez *et al.*, 2006). The *4CL* genes were not co-localized within the quantitative trait loci regulating biomass production in *Eucalyptus* (Kirst *et al.*, 2004). Therefore, based on the characterization of 4CLs in other plant species, we hypothesize that identifying the switchgrass 4CL isozyme involved in the monolignol biosynthetic pathway, and downregulating this specific *4CL* gene, could reduce switchgrass lignin content without significantly adverse effects on biomass production.

In this report we identified two switchgrass *4CL* genes through phylogenic analysis of different 4CL homologues. The enzyme activities and substrate preferences of the two switchgrass 4CL isoforms were determined. One gene, *Pv4CL1*, was silenced by RNA interference (RNAi). The phenotypes of the transgenic plants including biomass yield, cell wall composition, and cellulose hydrolysis efficiency were characterized in detail. Our results indicated that Pv4CL1, but not Pv4CL2, was the key 4CL isozyme involved in the monolignol biosynthetic pathway, and reducing lignin content in switchgrass biomass by silencing *Pv4CL1* can significantly increase the efficiency of fermentable sugar release for biofuel production.

MATERIALS AND METHODS

Cloning *Pv4CL1* and *Pv4CL2* cDNAs

4CL sequences of *Zea mays* were used as Query for BLAST searches against the available switchgrass sequences in public databases. Full length consensus sequences from multiple cDNA alignments were used for primer design. TRIzol Reagent (Invitrogen, Carlsbad, CA**)** was used for RNA extraction. DNA contamination was eliminated by treating total RNA with UltraPure DNase I (Invitrogen). The integrity and quantity of total RNA were checked by running through a 0.8% agarose gel and through a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE). cDNA synthesis was performed using the SuperScript III First-Strand System for RT-PCR Kit (Invitrogen) with an oligo-dT primer. The full length cDNA was amplified by PCR using KOD DNA polymerase (EMD, San Diego, CA), and cloned into the vector p-

ENTR /D-TOPO. Sequences of all primers used in this study are listed in Table S1 (Supporting information). The primers used for *Pv4CL1* and *Pv4CL2* cloning were *Pv4CL1*_ORF_For and *Pv4CL1*_ORF_Rev, and *Pv4CL2*_ORF_For and *Pv4CL2*_ORF_Rev, respectively.

RT-PCR and qRT-PCR

For qRT-PCR, total RNA was isolated from young switchgrass plants [E4 stage (elongation stage with 4 internodes) internodes, leaves, nodes, leaf sheaths, R1 (Reproductive Stage 1) inflorescences, and from fully elongated flower stalks, leaves, and internodes]. For qRT-PCR, PRIMER EXPRESS_ software (version 3.0; Applied Biosystems, Foster City, CA) was used to design primer sets for *Pv4CL1*, *Pv4CL2* and the reference genes [*Pv_UBIQUITIN* (FL955474.1) and *Pv_ACTIN2* (FL724919.1)] (Table S1). The qRT-PCR was performed with ABsolute Blue QPCR Sybr Green ROX mix (ABgene) in the ABI 7500 Real-Time PCR System or ABI PRISM 7900HT Sequence Detection System (Applied Biosystems Inc.) in a 25 µl or 10 µl reaction volume, respectively, according to the manufacturer's instructions. Each sample had 3 replicates, and the data were normalized against the reference genes. There was no amplification of the primer pairs without the cDNA templates. RT-PCR was also used to detect the transcript levels of *Pv4CL1* in different transgenic lines using RNA isolated from the third internodes of each plant.

Expression of Switchgrass Pv4CL1 and Pv4CL2 in *E. coli*

Pv4CL1 and *Pv4CL2* were sub-cloned into the expression vector pDEST17 using Gateway technology (Invitrogen). *E. coli* strain Rosetta cells harboring the *Pv4CL1* or *Pv4CL2* constructs were cultured at 37 ℃ until OD_{600} reached 0.6-0.7, and protein expression was then induced by adding isopropyl 1-thio β-galactopyranoside (IPTG) at a final concentration of 0.5 mM, followed by incubation at 16 ℃ for 18-20 h. Frozen cell pellets from 25 ml of induced culture were thawed at room temperature and resuspended in 1.2 ml of extraction-washing buffer (10 mM imidazole, 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol, and 10 mM β-mercaptoethanol). The extracts were sonicated three times for 20 sec, and the supernatants recovered after centrifugation at 16,000 x *g* were mixed with equilibrated Ni-NTA beads (Qiagen, Germantown, MD) and incubated

at 4 ℃ for 30 min under constant inversion to allow the His-tag proteins to bind to the beads. The beads were washed three times with 1 ml of extraction-washing buffer. Target proteins were then eluted off from the beads with 250 µl of elution solution (300 mM imidazole, 50 mM Tris-HCl buffer pH 8.0, 500 mM NaCl, 10% glycerol, and 10 mM β-mercaptoethanol). The purity of eluted target proteins was verified by SDS-PAGE and protein concentrations were determined using the BioRad protein assay (BioRad, Hercules, CA).

Enzyme Activity Assays and Kinetics

Pure recombinant enzymes (10 to 400 ng) were incubated at 30 ℃ (10 or 30 min) with 50 mM Bis-Tris propane buffer pH 7.5, 2.5 mM $MgCl_2$, 5 mM ATP, 1 mM Coenzyme A, and 2 to 100 µM substrate (cinnamic acid, 4-coumaric acid, caffeic acid, ferulic acid or sinapic acid) in a final volume of 100 µl. The reactions were stopped by adding 10 µl of glacial acetic acid. Reaction products were analyzed by reverse-phase HPLC on a C18 column (Spherisorb 5µ ODS2, Waters, Milford, MA) in a step gradient using 1% phosphoric acid in water as solvent A and acetonitrile as solvent B. Calibration curves were constructed with authentic standards of each product. The 4CL test substrates cinnamic acid, 4-coumaric acid, caffeic acid, ferulic acid and sinapic acid were purchased from Sigma-Aldrich (St. Louis, MO), while the 4CL products for calibration curves, 4-coumaroyl CoA, caffeoyl CoA, and feruloyl CoA, were synthesized as described previously (Stockigt & Zenk, 1975).

Construction of Gateway Compatible Vectors

The pCAMBIA1305.2 vector was modified to be a Gateway compatible binary vector for switchgrass transformation. The pUC19 vector was first digested with *Eco*RI and *Sph*I and blunt-ended with Klenow DNA polymerase. Re-ligation of the treated pUC19 vector led to the new plasmid pUC19-Δ*Eco*RI-*Sph*I that had a unique *Hin*dIII site. A *Hin*dIII DNA fragment from pAHC27 that carried the maize *Ubi* promoter and the *uid*A (*GUS)* gene was subcloned into pUC19-Δ*Eco*RI-*Sph*I. The *uid* A (*GUS)* gene was replaced with a *Bam*HI-*Eco*RV-HA-*Sac*I linker (5'-GGATCCGATATCTATCCATACGATGTGCCAGATTACGCATAGGAGCTC-3') to generate pUC19-Ubi-HA-NosT. The ccdB(B) cassette frame A was then inserted into the

*Eco*RV site of pUC19-Ubi-HA-NosT to generate pUC19-Ubi-DesA-HA-NosT. The *Hin*dIII fragment from pUC19-Ubi-DesA-HA-NosT was subcloned into the *Hin*dIII site of pCAMBIA1305.2, which resulted in pZHAO1629 (Fig. S2, see Supporting Information). This vector allowed us to either overexpress a target gene or silence a gene in the grass species.

An Entry vector, pEntry/D-Kannibal, for gene silencing was also constructed (Fig. S2b). The pEntry-1A vector (Invitrogen) was modified for cloning fragments of both antisense and sense strands of *Pv4CL1*. In brief, a *Sal*I-*Xba*I DNA fragment carrying the PDK intron from pKannibal (Wesley et al., 2001) was cloned into pEntry-1A to generate pEntry/D-Kannibal.

A 203-bp cDNA fragment of *Pv4CL1*, spanning part of the putative Box I domain (Stuible & Kombrink, 2001) (Fig. S1, see Supporting Information), was amplified from the cDNA of switchgrass cv. Alamo using a nested RT-PCR method. The first pair of primers was Pv4CL_1st Round_For and Pv4CL_1st Round_Rev, and the nested PCR primers were Pv4CL_H3RI_For and Pv4CL SalXba_Rev. The *Pv4CL1* fragments from pEN-Pv4CL1 were sequentially cloned into the *Sal*I/*Eco*RI and *Hin*dIII/*Xba*I sites. This cloning step generated the RNAi entry vector pEntry/D-Kannibal-2x*Pv4CL1*. The Kannibal-2x*Pv4CL1* was cloned into pZHAO1629 by LR Gateway cloning reaction to generate pZHAO1629-2×Pv4CL. The binary vector was transformed into *Agrobacterium tumefaciens* strain C58C1 by electroporation.

Switchgrass Genetic Transformation

Mature seeds of switchgrass line HR8 selected from cv. Alamo were used for all tissue culture and genetic transformations in this study. A modified *Agrobacterium*-mediated transformation protocol was used to transform switchgrass with the RNAi binary vector (Somleva *et al.*, 2002). In brief, somatic embryogenic calli were suspended in *Agrobacterium* solution (OD_{600}=0.6) and vacuum infiltrated for 10 min with occasional shaking. After *Agrobacterium*-inoculation, the calli were blotted on sterile paper towels and then transferred to the co-cultivation medium for 4 days at 23^{o}C in the dark. After co-cultivation, the calli were transferred onto callus and then regeneration media selected

under 50 mg/l hygromycin B (Sigma Chemical Co., St. Louis, MO). The regenerated plants were verified by PCR, Southern blot, and GUS staining.

The verified transgenic plants were grown in the Horticulture greenhouse at Virginia Tech, with temperatures set at 22^{o}C night and 28^{o}C day with 12-14 hour light. The plants were grown in Miracle-Gro Potting Mix (Miracle-Gro Lawn Products, Inc., Marysville, OH) in 1.1×10^{-2} m^3 pots and watered about twice a week. Wild type (WT) plants regenerated from non-transformed calli were also grown in the same greenhouse under the same conditions. Each transgenic line was multiplied by splitting tillers and maintained in the greenhouse. Plant samples were harvested when 50% of the tillers had flowered.

Switchgrass is gametophytically self-incompatible. Therefore, we obtained T_1 plants by crossing the T_0 transgenic line-115 with WT plants. The T_1 plants segregated in a 1:1 ratio according to the presence of the *HPTII* gene detected by PCR (Table 3), and were grown and harvested under the same conditions as mentioned above.

4CL Activity Assays in Plant Protein Extracts

Ground stem tissue (1 g), harvested from stems at the same growth stage, was suspended in 2.7 ml of extraction buffer (100 mM Tris-Cl pH=7.5, 10% glycerol, 1mM PMSF and 0.5 mM of PMSF) and 0.1 g of polyvinylpolypyrrolidone was then added. The suspension was kept on ice for 45 min with occasional vortexing. The supernatant was recovered after centrifugation (12,000 x *g* for 5 min), and desalted by passing it through a PD-10 column (GE-Healthcare) according to the manufacturer's instructions. The crude protein extracts (3 to 4 µg) were incubated at 30 ℃ for 10 to 30 min with 50 mM Bis-Tris propane buffer pH 7.5, 2.5 mM $MgCl_2$, 5 mM ATP, 1mM Coenzyme A and 60 µM 4-coumaric acid in a final volume of 100 µl. The reactions were stopped by adding 10 µl of glacial acetic acid. Reaction products were analyzed by reverse-phase HPLC on a C18 column (Spherisorb 5µ ODS2, Waters, Milford, MA) in a step gradient using 1% phosphoric acid in water as solvent A and acetonitrile as solvent B. Calibration curves were constructed with authentic standard of the product 4-coumaroyl CoA.

Carbohydrate and Lignin Assays

Whole stems (from the first internode and above) of RNAi transgenic and WT control plants were collected and dried for cell wall composition analysis. The structural carbohydrate composition of switchgrass biomass was determined using a modified quantitative saccharification (QS) procedure (Moxley & Zhang, 2007). Monomeric sugars were measured with a Shimadzu HPLC equipped with a Bio-Rad Aminex HPX-87P column (Richmond, CA). Lignin and ash were measured according to the standard National Renewable Energy Laboratory (NREL) biomass protocol (Sluiter *et al.*, 2004). The concentrations of glucose and xylose in the enzymatic hydrolysates were measured with a Shimadzu HPLC equipped with a Bio-Rad Aminex HPX-87H chromatography column. Furfural and HMF were not observed in the hydrolysates (<0.001%, w/v).

Determination of Monolignol Composition by Thioacidolysis/GC-MS

Whole stems of different plants were dried and treated for thioacidolysis followed by GC-MS to measure the monolignol composition. Extractive-free lignin was made by acetone extraction in a Soxhlet apparatus for 24 h (Rolando *et al.*, 1992). The dried lignin of each sample was processed through a recently revised thioacidolysis method (Robinson & Mansfield, 2009). The silylated sample was injected into the GC column (Restek RTX5-MS, 1 uM film thickness, 30 M x 3.2 mM i.d., Thames Restek UK Ltd, Windsor, UK). The GC-MS analysis was modified from a previous method (Rolando *et al.*, 1992) and performed on a VG 70SE double focusing magnetic sector instrument, interfaced to a HP5790 GC.

Histology and Microscopy

The internodes of the T_1 segregating plants were embedded in 2.5% agarose and cut with a Leica VT1200 vibrating blade microtome (Bannockburn, US) into 50 µm thick sections. Phloroglucinol and Mäule reagents were used to stain the 50 µm thick stem sections to analyze the lignin deposition patterns by visualization under an Olympus SZXZ-RFL3 fluorescence microscope (Olympus America, Melville, NY, USA) (Coleman *et al.*, 2008; Pomar *et al.*, 2002).

Dilute Acid (DA) Pretreatment and Enzymatic Hydrolysis

The dried switchgrass materials were ground and sieved through a size 40-60 mesh. The switchgrass samples were pretreated with dilute acid (DA), using 1.3% (w/w) sulfuric acid at a solid loading of 10% (w/w) at 130 ℃, 15 psi (autoclave) for 40 min. After DA, the hydrolysates were separated by centrifugation. The switchgrass residues were washed with water prior to enzymatic hydrolysis. The DA-pretreated switchgrass samples were diluted to 20 g biomass per L in 50 mM sodium citrate buffer (pH 4.8) with supplementary addition of 0.1% (w/v) NaN_3, as described previously (Moxley & Zhang, 2007; Zhu *et al.*, 2009). All hydrolysis experiments were carried out in a rotary shaker at 250 rpm and 50 ℃. The enzyme loadings were five filter paper units (FPUs) of cellulase (Novozymes Inc., Bagsvaerd, Denmark) and 10 units of β-glucosidase (Novozymes) per gram of biomass. The cellulose and β-glucosidase enzyme activities were confirmed with standard protocols (Adney and Baker, 1996). The protein content was determined by Bicinohoninic acid (BCA) assay using bovine serum albumin as a protein standard. The estimated protein contents of cellulase and b-glucosidase were ~ 143 mg/mL. After enzymatic hydrolysis, glucan digestibility was calculated as described previously (Zhang *et al.*, 2009). The mass balance of dilute acid pretreatment and enzymatic hydrolysis is shown in Fig. S3 (supporting information).

RESULTS

Isolation and Characterization of Switchgrass *4CL* Genes and Proteins

Using both a full length sequence and conserved domains of a maize (*Zea mays*) *4CL* gene, *Zm4CL* (AY566301), as a query to BLAST against the switchgrass nucleotide and EST databases, we identified six ESTs annotated as *4CL*-like genes. Only two genes were classified with other characterized *4CLs* by phylogenetic analysis (Fig. 1). We therefore named the switchgrass *4CL* gene (EU491511.1) as *Pv4CL1*, and another *4CL* homologue (JF414903) as *Pv4CL2*.

Pv4CL1 has an open reading frame (ORF) of 1629 nucleotides encoding a 542 amino acid protein with a predicted molecular mass of 58.35 kDa and a pI of 5.38. *Pv4CL2* has an ORF of 1728 nucleotides encoding a protein of 575 amino acids (61.07 kDa) with a calculated pI of 5.37. The protein sequences deduced from the *Pv4CL1* and *Pv4CL2* cDNAs show 60% identity (Fig. S1, see Supporting Information), which suggests they

are homologues rather than two alleles although the tetraploid switchgrass cv. Alamo may have multiple alleles of *Pv4CL* genes. Both sequences have the AMP binding domain (*PFSSGTTGLPKGV* for 4CL1 and *PYSSGTTGLPKGV* for 4CL2), the *GEICIRGR* motif (Stuible & Kombrink, 2001) and the conserved *VPP* and *PVL* domains (Schneider *et al.*, 2003), all characteristics of 4CL enzymes.

A phylogenetic tree of Pv4CL1, Pv4CL2 and most other 4CL proteins was constructed, and showed similar phylogenetic patterns to those reported previously (Fig. 1), in which all 4CLs could be classified into two major classes (Ehlting *et al.*, 1999; Cukovica *et al.*, 2001). Pv4CL1 was classified in the Class I group, along with the characterized 4CL enzymes, such as Arabidopsis At4CL1, At4CL2, aspen Ptr4CL1, and pine Pt4CL1, that are devoted to the monolignol biosynthetic pathway (Hu *et al.*, 1998; Ehlting *et al.*, 1999; Wagner *et al.*, 2009). Pv4CL2 was classified in the Class II group, in which the characterized 4CL enzymes, such as Arabidopsis At4CL3 and aspen Ptr4CL2, mainly participate in the flavonoid biosynthetic pathway (Hu *et al.*, 1998; Ehlting *et al.*, 1999).

The expression patterns of *Pv4CL1* and *Pv4CL2* were analyzed by real-time PCR (qRT-PCR). *Pv4CL1* transcripts were more abundant in the highly lignified internodes than in leaves and other tissues with relatively lower lignin contents (Fig. 2). In the internodes, the *Pv4CL1* transcript level is about 7 times higher than that of *Pv4CL2* (Fig. 2). In switchgrass internodes, the lignin content increases with increasing distance from the peduncle (Sarath *et al.*, 2007; Shen *et al.*, 2009). The transcript abundance of *Pv4CL1* in different organs largely correlates with the cell wall lignification pattern. Based on the expression pattern and phylogenetic analysis, we hypothesize that *Pv4CL1* is the functional 4CL enzyme involved in the monolignol biosynthetic pathway in switchgrass.

E. coli expressed His-tagged Pv4CL1 and Pv4CL2 fusion proteins were purified to homogeneity (Fig. 3a). The enzymatic activities of the purified proteins were initially screened by determining their ability to ligate Coenzyme A (CoA) to form the respective CoA esters. Both enzymes were active with 4-coumaric, caffeic, and ferulic acids, but cinnamic and sinapic acids were not substrates.

Kinetic parameters of both recombinant enzymes were determined for all three substrates using a fixed concentration of Coenzyme A (Table 1). Chromatograms and curves of reaction velocity versus substrate concentration for the three substrates are shown in Fig.

3b-c. Kinetic parameters for Pv4CL2 using ferulic acid were not determined because the efficiency of the reaction was low in comparison with 4-coumaric acid or caffeic acid. The preferred substrate for both enzymes was 4-coumaric acid with similar efficiencies (Kcat/Km), but the Kcat value for Pv4CL2 was lower than that for Pv4CL1 (Table 1).

Downregulating *Pv4CL1* Expression by RNAi

A 203-bp fragment of *Pv4CL1* (Fig. S1) that is specific to this gene was used to generate the RNAi construct pZHAO1629-2×Pv4CL. One set of Gateway-compatible entry and destination vectors (Fig. S1) were constructed for generating the RNAi vector. *Agrobacterium*-mediated transformation of switchgrass with the RNAi vector yielded more than 100 putative transgenic plants. The transgenic lines were verified by detecting the presence of the *HPTII* gene by PCR and Southern blot, and the *uidA* gene by GUS staining (Data not shown).

We selected seven T_0 generation transgenic plants to monitor *Pv4CL1* transcript levels by RT-PCR and qRT-PCR. The RT-PCR (Fig. 4a) and qRT-PCR results were consistent showing that the transcript levels of *Pv4CL1* in the transgenic lines ranged from 0.05- to 0.73- fold that of the WT control plants (Fig. 4b).

Segregated T_1 plants were further studied. In the transgenic T_1 plants, levels of *Pv4CL1* transcripts, but not *Pv4CL2* transcripts, were greatly reduced (Fig. 4c), confirming that the RNAi construct specifically targeted *Pv4CL1*. Protein extracts from pooled stem tissues of three T_1 transgenic plants were assayed for 4CL activities with 4-coumaric acid and Coenzyme A as substrates under optimal conditions. The result showed that transgenic T_1 plants exhibited on average an 80% reduction in 4CL activity (Fig. 4d-e).

Suppression of *Pv4CL1* Results in Phenotypic Alterations and Reduced Lignin Content

Different T_0 transgenic switchgrass lines with low *Pv4CL1* transcript abundances showed browning on parts of the leaf mid-vein (Fig. 5a), and sporadically exhibited brown patches in stem internodes (Fig. 5b), similar to low lignin *brown-midrib* (*bm*) maize mutants (Cherney *et al.*, 1991). The inner sides of the basal stems (e.g., the stems below the fourth internodes) became reddish-brown (Fig. 5c). With decreasing distance from the flower stalks, the number of dark brown patches on the outside of the stems decreased,

and the reddish-brown color on the inner side of the stems gradually reduced to that of WT plants. The mature roots of the transgenic lines turned reddish brown to various degrees. However, the newly elongated roots and the root tips were still white, similar to the WT plants (Fig. 5d).

The aboveground biomass yields of four independent T_0 lines and three tissue culture regenerated WT plants were measured. As shown in Table 2, silencing *Pv4CL1* did not affect the biomass yields in T_0 transgenic lines. To confirm the effect of silencing *Pv4CL1* in switchgrass, we further measured the biomass yields and other phenotypes in T_1 plants. The reddish-brown color in mature roots and basal stems co-segregated with the RNAi transgene in T_1 plants. The biomass yield and other agronomic traits related to biomass production (e.g., tiller number and plant height) were not significantly different between T_1 plants with or without the transgene (Table 3), suggesting that silencing of *Pv4CL1* did not significantly affect the biomass yields of switchgrass plants grown under greenhouse conditions.

We measured the cell wall composition of four individual T_0 plants (Table 4-5) and the pooled T_1 plants with or without the transgene (Table 6). The T_0 plants had 23-34% less acid-insoluble lignin or 17- 32% less total lignin than WT plants, and varied levels of cellulose (referred to as glucan) and hemicellulose (referred to as xylan, the predominant component of hemicellulose in switchgrass) contents (Table 4). Monolignol composition [hydroxyphenyl (H), guaiacyl (G), and syringyl (S)] of four T_0 transgenic lines and the WT control plants were measured. As shown in Table 5, the T_0 transgenic lines had similar S, but less G and higher H contents than WT plants.

Between the T_1 plants with or without the RNAi transgene, changes of cell wall composition were also observed. The pooled transgenic T_1 plants had 22% acid-insoluble lignin or 22% total lignin reduction compared to WT plants. The transgenic T_1 plants had similar S content, but 47% less G content and 45% more H content than non-transgenic control T_1 plants (Table 6). Compared to cell wall composition of T_0 plants, segregated T_1 plants all had relatively high lignin content but low cellulose content at harvest time, possibly because T_0 and T_1 plants were grown and measured at different times. Therefore, cell wall composition of T_0 and T_1 plants was only compared to the corresponding control plants that were grown under the same conditions. Nevertheless, consistent trends of low

lignin content and altered monolignol composition were observed in multiple T_0 and T_1 plants.

The lignin deposition patterns in T_0 and T_1 plants were characterized (Fig. 6). Phloroglucinol staining, which detects hydroxycinnamaldehyde end groups in native lignin (Pomar *et al.*, 2002), showed that there was reduced lignin deposition in the collenchyma, sclerenchyma, and even in the parenchyma cells of the transgenic T_1 plants. Mäule reagents which specifically stain S lignin, showed no difference between transgenic lines and WT control plants (Fig. 6a and b) (Coleman *et al.*, 2008).

4CL-Downregualted Switchgrass Biomass has Improved Yields of Fermentable Sugars

Segregating T_1 plant material was subjected to enzymatic hydrolysis with or without acid pre-treatment. All non-pretreated samples exhibited comparatively low enzymatic digestibility. The dilute acid (DA) pretreated samples exhibited enhanced enzymatic digestibility for glucan yield but not for xylan yield. The low lignin transgenic plant materials yielded 57% more fermentable sugar than the WT material with DA pretreatment, suggesting the central importance of reducing lignin content for biofuel production (Fig. 7). Decreased lignin content in the RNAi:*Pv4CL1* transgenic plants improved biomass saccharification efficiency and therefore may improve the economics of switchgrass production and the utilization of bioenergy.

DISCUSSION

Plants with reduced lignin content can be identified through breeding and selection [such as the maize *bm* and sorghum *bmr* mutants (Vignols *et al.*, 1995; Halpin *et al.*, 1998)] or purposely altered through genetic engineering (Hisano *et al.*, 2009). Genetic engineering can directly manipulate the genes involved in the lignin biosynthetic pathway by silencing (e.g., RNAi), blocking the expression of the genes by artificial zinc finger chimeras (Sanchez *et al.*, 2006), or by manipulating transcription factors that regulate the expression of single to multiple lignin synthesis gene(s) (Zhou *et al.*, 2009; Zhao *et al.*, 2010; Zhong *et al.*, 2010). Since limited information about lignin synthesis and its regulatory mechanism is available for grass species including switchgrass (Carroll & Somerville, 2009; Escamilla-Treviño *et al.*, 2009; Saathoff *et al.*, 2011b), silencing the

gene(s) involved in the monolignol biosyntheic pathway is currently the most straightforward way to reduce lignin content.

Despite the importance of grass species for forage and future biofuel production, only a few reports are available on the functional characterization of lignin-synthesis related genes in grasses (besides the maize *bm* and sorghum *bmr* mutants) (Fu *et al.*, 2011b; Chen *et al.*, 2003; Bell *et al.*, 2004; Escamilla-Treviño *et al.*, 2009; Tu *et al.*, 2010; Saathoff *et al.*, 2011b), possibly because of the difficulty in identifying the potential target gene(s) and the time-consuming genetic transformation processes.

Pv4CL1 Is a Key Functional 4CL Isozyme in the Monolignol Biosynthetic Pathway

Lignin biosynthetic pathways are conserved in most plant species (Boerjan *et al.*, 2003; Umezawa, 2010). Both the recombinant Pv4CL1 and Pv4CL2 proteins showed 4CL enzyme activity *in vitro*, since both proteins possess the representative domains of 4CL enzymes (Stuible et al., 2000; Schneider et al., 2003). Both enzymes had similar efficiencies (Kcat/Km), but Pv4CL1 had the higher Km and Kcat values. It is possible that the substrate availability or concentration in a specific cell or tissue type could be a factor in determining which of the two 4CL forms is responsible for 4-coumaroyl CoA formation. Furthermore, our transgenic plants downregulated for Pv4CL1, but not for Pv4CL2, showed more than 80% reduction of 4CL activity in stems, indicating that Pv4CL1 provides most of the activity in a tissue with active lignification.

Correlation between *Pv4CL1* Transcript Levels and Cell Wall Composition Change

The T_0 generation transgenic plants had different levels of *Pv4CL1* transcript abundance, which is a rather common phenomenon for RNAi transgenic plants. The lignin content of different T_0 transgenic lines largely correlated with their *Pv4CL1* transcript abundance, although inconsistencies were sometimes observed. The inconsistencies might be caused by the heterozygous genetic background, even though all the T_0 plants are half-siblings, and by the independent T-DNA inserted loci. Therefore, it is important both to analyze the T_1 segregating plants with pooled samples to minimize the effects of differences in genetic background, and also to analyze several independent T_0 RNAi transgenic lines to minimize the effects of the T-DNA insertion sites in the genome.

RNAi:*Pv4CL1* transgenic plants exhibited a substantial decrease in G units, and slightly increased H units in the lignin polymer, as consistently illustrated in T_0 and segregating T_1 plants (Table 5-6). Similar results were observed following downregulation of 4CL orthologs in Arabidopsis (Lee *et al.*, 1997), tobacco (Kajita *et al.*, 1997), and pine (Wagner *et al.*, 2009). In the monolignol biosynthetic pathway, Arabidopsis ferulate 5-hydroxylase (F5H) has a Km of 3.06 μM for catalyzing coniferaldehyde to 5-OH-coniferaldehyde, or a Km of 1.76 μM for catalyzing coniferyl alcohol to 5-OH-coniferyl alcohol, respectively, which produces precursors for sinapyl alcohol (S lignin) (Weng, J-K *et al.*, 2010). However, the best Arabidopsis cinnamyl alcohol dehydrogenase (CAD), AtCAD5, which can catalyze coniferaldehyde to coniferyl alcohol has a Km of 35 μM (Kim *et al.*, 2004). In switchgrass, F5H homologs have not been characterized yet. Switchgrass has at least two *CAD* genes. The expression level of *PvCAD1* is more than 10 times higher than *PvCAD2*, and PvCAD1 has a relatively high Km of 10.9 μM for catalyzing coniferaldehyde (Saathoff *et al.*, 2011).Therefore, lower amounts of coniferaldehyde resulted by the diminished substrates pool due to the downregulation of 4CL1, favor F5H but not CAD because of its 10 times lower Km, and thereby decrease the formation of G but not S lignin, which could have happened in 4CL downregulated Arabidopsis and possibly in switchgrass and other plant species.

Growth Performance of RNAi:*Pv4CL1* Transgenic Lines

4CL participates in an early step of the general phenylpropanoid pathway by producing the monolignol precursor ρ-coumaroyl-CoA. This metabolic intermediate is also a precursor for the production of many secondary metabolites, such as stilbenes and flavonoids (Boudet, 2007). Therefore, the downregulation of *4CL* could have pleiotropic effects, such as color changes in leaf mid-ribs, mature stems and roots. These color changes are common phenomena when downregulating gene(s) in the monolignol biosynthetic pathway (Kajita *et al.*, 1996; Wagner *et al.*, 2009; Voelker *et al.*, 2010), which often leads to ectopic accumulation of flavonoids (Besseau *et al.*, 2007). However, the ectopic accumulation of flavonoids caused by silencing monolignol biosynthetic pathway genes does not in itself directly impact plant growth, at least in Arabidopsis (Li *et al.*, 2010).

We measured biomass yields and cell wall compositions in both T_0 and T_1 transgenic lines, and we did not see any significant change in plant growth among four T_0 and T_1 segregating plants (Table 2-3). Downregulation of *At4CL1* in Arabidopsis did not result in compromised biomass production (Sanchez *et al.*, 2006). However, the silencing of 4CLs in tobacco (Kajita *et al.*, 1997), pine (Wagner *et al.*, 2009) and poplar (Voelker *et al.*, 2010) resulted in stunted plant growth in some transgenic lines, primarily caused by deformation of xylem tissue, and deposition of tyloses and phenolics in xylem vessels of poplar, thus blocking water transport (Kitin *et al.*, 2010). Possibly because of the anatomical difference between grasses and trees, or because of differences in their tolerance to lignin modification, the present RNAi:*Pv4CL1* switchgrass plants did not show any obvious growth abnormalities under greenhouse conditions. Likewise, RNAi:*CCR1* and RNAi:*COMT1* transgenic ryegrass (Tu *et al.*, 2010), and RNAi:*COMT* transgenic switchgrass (Fu *et al.*, 2011b) did not show compromised biomass production. Because lignin deposition is influenced by abiotic and biotic stress (Halpin, 2004; Boudet, 2007), it will be interesting to further confirm if RNAi: *Pv4CL1* transgenic switchgrass plants have low lignin content under both greenhouse and field conditions. A recent study showed that low lignin transgenic poplar and WT poplar differed in their field performance possibly because of reduced wood strength and stiffness in transgenic plants (Voelker *et al.*, 2011). It will also be interesting to observe the stand integrity (e.g lodging) of the low lignin switchgrass in field conditions.

Reducing Biofuel Production Costs by Downregulating Feedstock Lignin Content

Reduced lignin content of biomass could improve saccharification efficiency through enzyme hydrolysis, and therefore reduce the cost of biofuel production (Fu *et al.*, 2011a), though in some bioenergy processes such as pyrolysis and combustion, the higher lignin feedstock could be desirable because of its high energy contents (Boateng *et al.*, 2008). In this study, the T_0 transgenic plants have varied S:G ratios (Table 5). Therefore, it is still inconclusive whether the increased S:G ratio together with lower lignin contents contribute to the increased cellulose hydrolysis efficiency. F5H overexpressing transgenic poplar has a higher S lignin with unchanged lignin content, and this has lead to a significant improvement in pulping and bleaching efficiencies (Huntley *et al.*, 2003),

presumably because the S units form fewer crosslink bonds (Huntley *et al.*, 2003; Grabber *et al.*, 2004). However, no clear correlations between sugar-releasing efficiency and monolignol unit composition have been shown from studies on biomass derived from *Medicago truncatula* (Chen & Dixon, 2007).

Reducing lignin content has increased cellulose hydrolysis efficiency (saccharification efficiency) in model plants such as *M. truncatula* and Arabidopsis (Chen & Dixon, 2007; Weng, J *et al.*, 2010). The same is true in switchgrass by comparing the processing ability of tissues at different developmental stages, and more recently by the analysis of genetically modified switchgrass with reduced expression of *CCOMT* or *CAD* (Fu *et al.* 2011a; Fu *et al.*, 2011b; Saathoff *et al.*, 2011a). Similar results with the present low lignin plant materials were observed (Fig. 7). Further improvements of switchgrass feedstock quality by genetic engineering along with efficient bioprocessing and conversion technologies will lead to economical biofuel production in the future.

ACKNOWLEDGMENTS

We thank Mr. Kim Harich for GC-MS analysis, Dr. Jim Tokuhisa for providing some equipment for thioacidolysis, Dr. Amy Brunner for providing the fluorescent microscope for histology study, Dr. Qiang Cheng for technical support, and Dr. Richard Veilleux for critical reading of the manuscript. This work was supported by grants from the Virginia Agricultural Council (Project number 545), the Biodesign and Bioprocessing Research Center of the College of Agriculture and Life Sciences at Virginia Tech, the Virginia Agricultural Research Station (VA135872), and the U.S. Department of Energy Bioenergy Research Centers, through the Office of Biological and Environmental Research in the DOE Office of Science.

Table 1. Kinetic parameters of Pv4CL1 and Pv4CL2.

Substrate	Km (μM)		Kcat (s^{-1})		Kcat/Km (s^{-1} μM^{-1})	
	4CL1	4CL2	4CL1	4CL2	4CL1	4CL2
4-Coumaric acid	18.8	6.3	15.9	4.0	0.84	0.63
Caffeic acid	13.5	17.4	5.8	2.2	0.43	0.13
Ferulic acid	12.9	nd	7	nd	0.54	nd
Sinapic acid	No conversion					
Cinnamic acid	No conversion					

nd, not determined because of inefficient conversion.

Table 2. Growth performance of wild type and T_0 lines. Second year biomass production of four T_0 lines and three wild type plants were measured.

Switchgrass lines	Mature root color	Basal stems color	2nd year yield			
			Tiller number	Dry weight (g)	Averaged dry biomass (g)	
Wild type-1	White	White	19	76.5		
Wild type-2	White	White	17	96.0	WT	96.6 (27.6) [a]
Wild type-3	White	White	23	131.0		
T_0-4	Light brown	Light brown	17	84.6		
T_0-6	Reddish brown	Reddish brown	18	86.5	T_0 transgenic plants	93.2 (15.8)
T_0-7	Reddish brown	Reddish brown	20	84.8		
T_0-115	Reddish brown	Reddish brown	24	117.0		

[a] No statistically significant difference was detected between wild type and T_0 transgenic plants. SE is in parenthesis.

Table 3. Growth performance of segregating T_1 plants.

T1 Plants	Number of plants	Mature root color	Basal stems color	Tiller number	Tiller height (cm)	Above-ground DW [c] (g)	Below-ground DW (g)	Total DW (g)
Wild type [a]	12	White	White	75.3 [d] (14.1) [b]	148.0 [d] (9.3)	149.1 [d] (29.9)	106.7 [d] (27.3)	255.7 [d] (50.3)
Transgenic	13	Reddish brown	Reddish brown	83.9 (21.4)	147.8 (13.4)	141.2 (21.8)	118.8 (25.9)	259.4 (45.1)

[a] Detected by PCR with *HPTII* gene primers, [b] Standard Error (SE). [c] Dry Weight (DW). [d] No statistically significant difference was detected between wild type and transgenic plants.

Table 4. T_0 Transgenic plants have reduced lignin contents. Three wild type control plants grown under the same condition with transgenic plants were pooled together for the analysis. The carbohydrate data were averaged from the results of two separate experiments.

Switchgrass Lines	**Carbohydrates mg / 100mg**					**Lignin mg / 100mg**		
	Glucan	**Xylan**	**Galactan**	**Arabinan**	**Mannan**	**Acid-insoluble**	**Acid-soluble**	**Total Lignin**
Wild type	34.8 (0.4)[a]	17.6 (0.3)	1.6 (0.0)	2.2 (0.0)	0.1 (0.2)	18.5 (0.0)	0.7 (0.2)	19.2 (0.2) **
Line -4	35.4 (1.2)	17.6 (0.9)	2.0 (0.2)	2.4 (0.3)	0.2 (0.0)	14.2 (0.4)	1.7 (0.3)	15.9 (0.5)
Line -6	36.9(0.5)	19.3 (0.0)	1.9 (0.2)	2.5 (0.2)	0.0 (0.0)	14.2 (0.9)	1.4 (0.2)	15.6 (0.9)
Line -7	34.7 (0.8)	18.8 (0.6)	1.7 (0.1)	2.0 (0.2)	0.2 (0.2)	13.8 (0.2)	1.3 (0.4)	15.1 (0.4)
Line -115	39.2 (0.1)	19.5 (0.3)	1.2 (0.3)	1.8 (0.3)	0.0 (0.3)	12.2 (0.2)	0.9 (0.3)	13.1 (0.4)

[a] SE is in parenthesis.
** Statistically significant differences ($p<0.01$) between wild type and transgenic plants.

Table 5. T_0 Transgenic plants have altered monolignol composition. Three wild type control plants grown under the same condition with transgenic plants were pooled together for the analysis.

Switchgrass Lines	Monolignol composition mg / 100mg			G:S
	H	G	S	
Wild type	0.23 (0.09)[a]	11.60 (0.18)**	7.37 (0.27)	1.57**
T_0-4	0.30 (0.23)	8.58 (1.04)	7.01 (0.95)	1.22
T_0-6	1.48 (0.23)	7.20 (0.69)	6.92 (0.92)	1.04
T_0-7	1.41 (0.35)	7.08 (0.08)	6.61 (0.39)	1.07
T_0-115	1.11 (0.19)	3.60 (0.22)	8.39 (0.03)	0.43

[a] SE is in parenthesis. ** Statistically significant differences ($p<0.01$) between wild type and transgenic plants.

Table 6. Cell wall composition of segregating T_1 plants. Stems of the segregating plants were independently pooled (transgenic and wild type) together, and used for cell wall composition analysis. Three experimental repeats were conducted with the pooled plant material.

Switchgrass T_1 Lines	Carbohydrates mg/100mg		Lignin mg/100mg			Monolignol composition mg/100mg			G:S
	Glucan	Xylan	Acid-insoluble	Acid-soluble	Total Lignin	H	G	S	
Wild type	27.7 (1.0)[a]	15.5 (0.5)	22.4 (0.5)	1.2 (0.5)	23.6 (0.7) **	0.39 (0.03) **	11.27 (0.43) **	10.74 (0.46)	1.05
Transgenic	29.9 (1.6)	17.4 (1.1)	17.4 (1.2)	1.0 (0.4)	18.4 (1.3)	0.71 (0.04)	5.98 (0.11)	10.70 (0.15)	0.56

[a] SE is in parenthesis.
** Statistically significant differences ($p<0.01$) between wild type and transgenic plants.

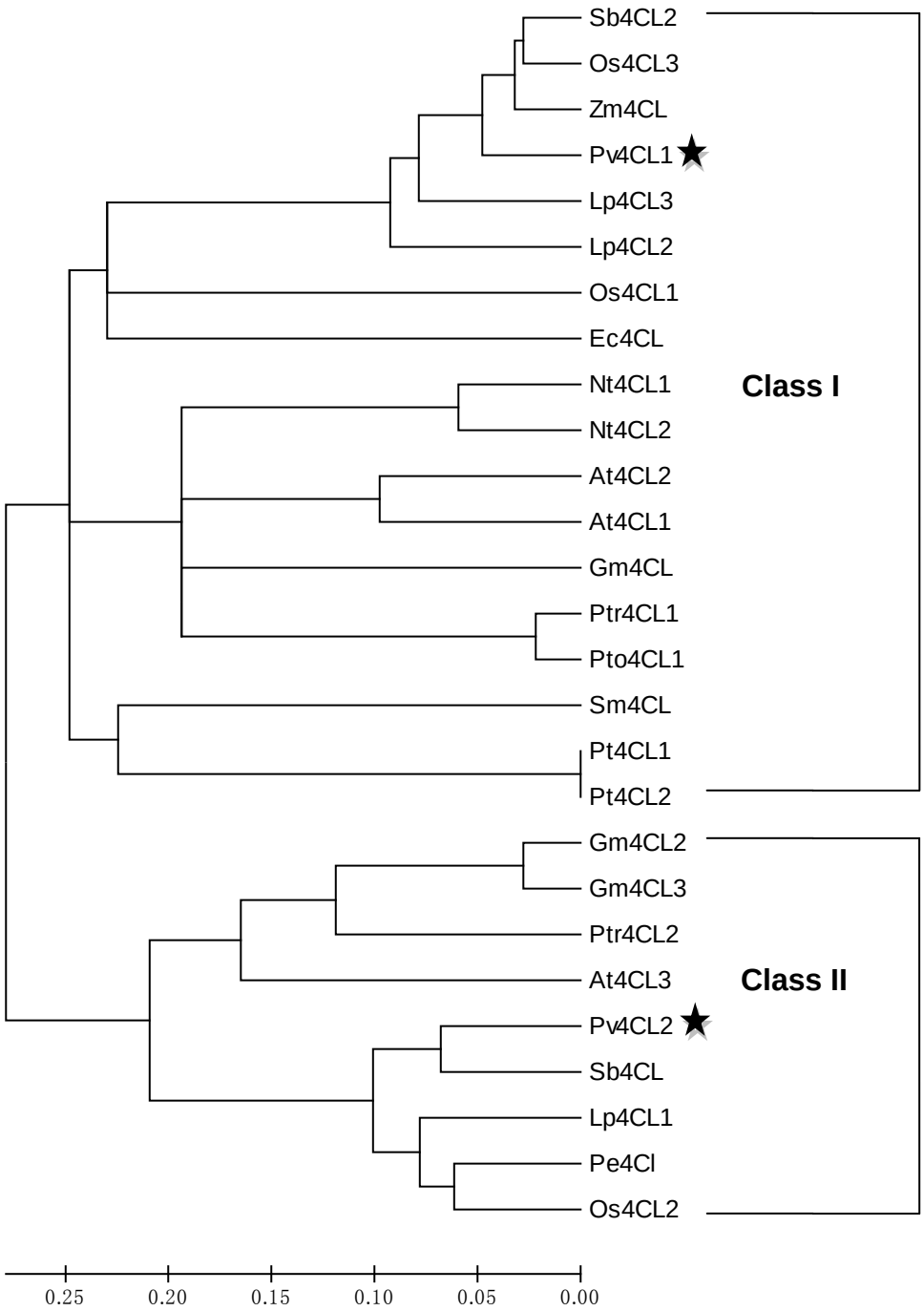
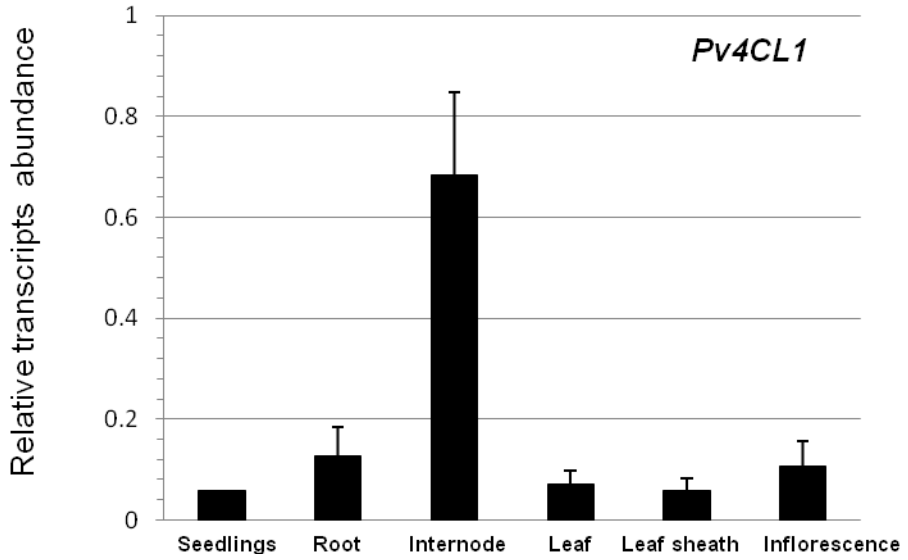


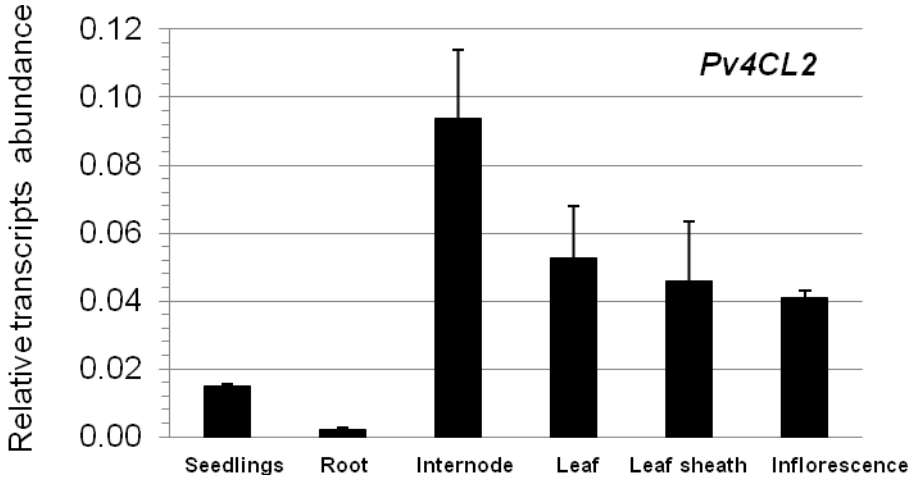
Fig. 1 Phylogenetic tree of Pv4CL1 and its homologues in some model plants. Pv4CL1 is classified in the Class I group, whereas Pv4CL2 is in the Class II group. The multiple alignments were done with ClustalW, and the Neighbor Joining (NJ) tree was built using Mega4 software (Tamura *et al.*, 2007). The optimal tree with the sum of branch length = 3.65635142 is shown. The phylogenetic tree was linearized assuming equal evolutionary rates in all lineages. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). Scale indicates amino acid substitutions per position. The NCBI accession numbers of the protein sequences are as follows: Arabidopsis At4CL1 (NP_175579), At4CL2 (NP_188761), At4CL3 (NP_176686); Aspen (*Populus tremuloides)* Ptr4CL1 (AAC24503.1), Ptr4CL2, (AAC24504.1); Bamboo (*Phyllostachys edulis*) putative Pe4CL (FP101648.1); Eucalyptus (*Eucalyptus camaldulensis*) Ec4CL (ACX68559.1); Maize Zm4CL (NM_001111788); Poplar (*P. tomentosa*) Pto4CL1 (AAL02145.1); Pine Pt4CL1 (PTU12012), Pt4CL2 (PTU12013); Rice (*Oryza sativa*) Os4CL1 (BAD05189), Os4CL2 (Q42982), Os4CL3 (AB234050); Ryegrass (*Lolium perenne*) Lp4CL1 (AAF37732.1), Lp4CL2 (AAF37733.1), Lp4CL3 (AAAF37734.1); *Selaginella moellendorffii* Sm4CL (XP_002985214.1); Sorghum (*Sorghum bicolor*) Sb4CL (AAA64913.1), putative Sb4CL2 (XP_002451647.1); Soybean (*Glycine max*) Gm4CL (AAL98709), Gm4CL2 (P31687), Gm4CL3 (AAC97389); Tobacco (*Nicotiana tabacum*) Nt4CL1 (O24145), Nt4CL2 (O24146).

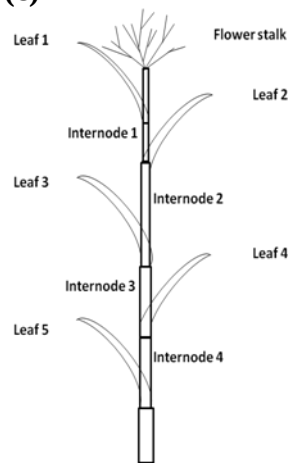
(a)


Pv4CL1
Relative transcripts abundance
1
0.8
0.6
0.4
0.2
0
Seedlings
Root
Internode
Leaf
Leaf sheath
Inflorescence

(b)


Pv4CL2
Relative transcripts abundance
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Seedlings
Root
Internode
Leaf
Leaf sheath
Inflorescence

(c)



(d)

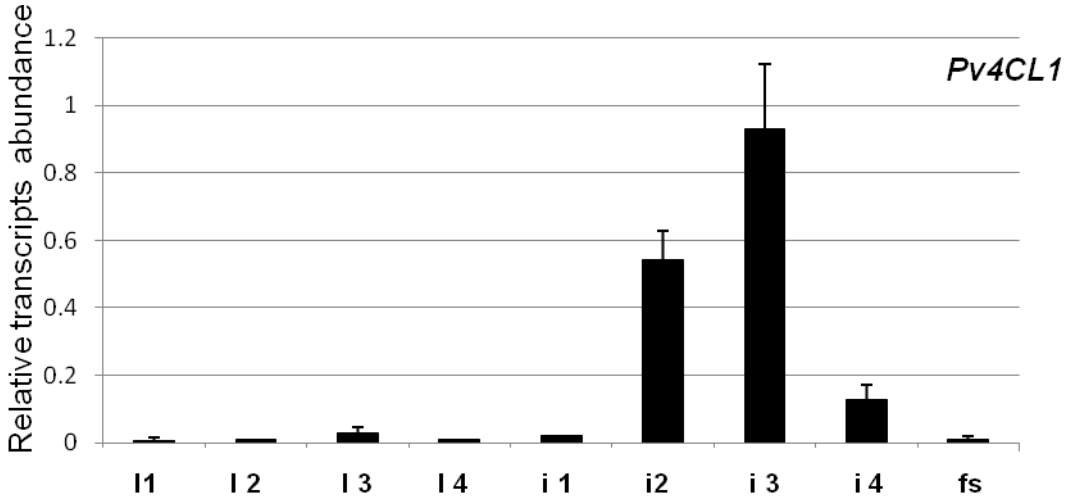
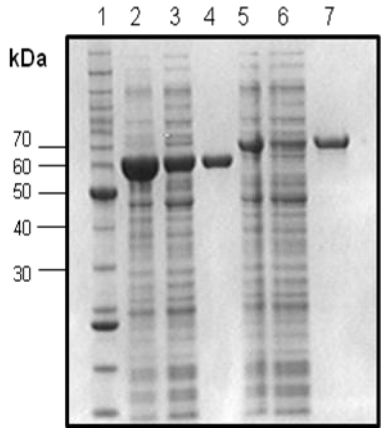
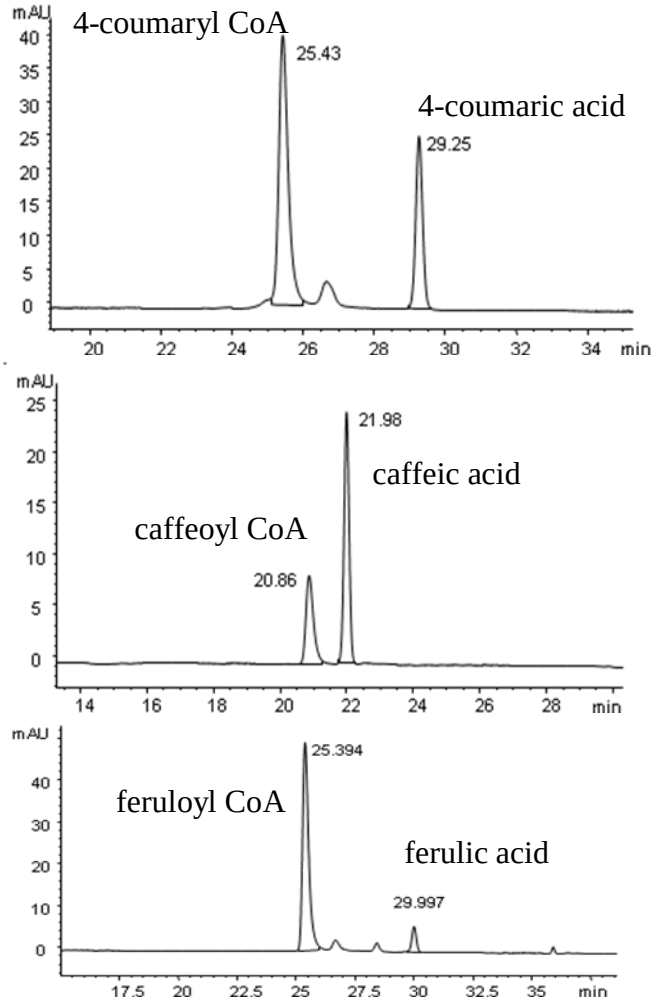


Fig. 2 Transcript abundance of *Pv4CL1* and *Pv4CL2* in different organs. Quantitative real-time PCR using *ACTIN* and *UBIQUITIN* as reference genes was used to determine transcript abundance of Pv4CL1 (a) and Pv4CL2 (b) in six switchgrass organs. (c) Graphical representation of a switchgrass plant. (d) Transcript abundance of *Pv4CL1* in different internodes and leaves. L1 to l4, leaf 1 to leaf 4; i1 to i4, internode 1 to internode 4; fs, flower stalk. Representative data are shown from two biological repeats. The dissociation curve for the qRT-PCR products showed that the qRT-PCR primers were gene specific

(a)


1 2 3 4 5 6 7
kDa
70
60
50
40
30

(b)


mAU
4-coumaryl CoA
25.43
4-coumaric acid
29.25
min
mAU
21.98
caffeic acid
caffeoyl CoA
20.86
min
mAU
25.394
feruloyl CoA
ferulic acid
29.997
min

(c)

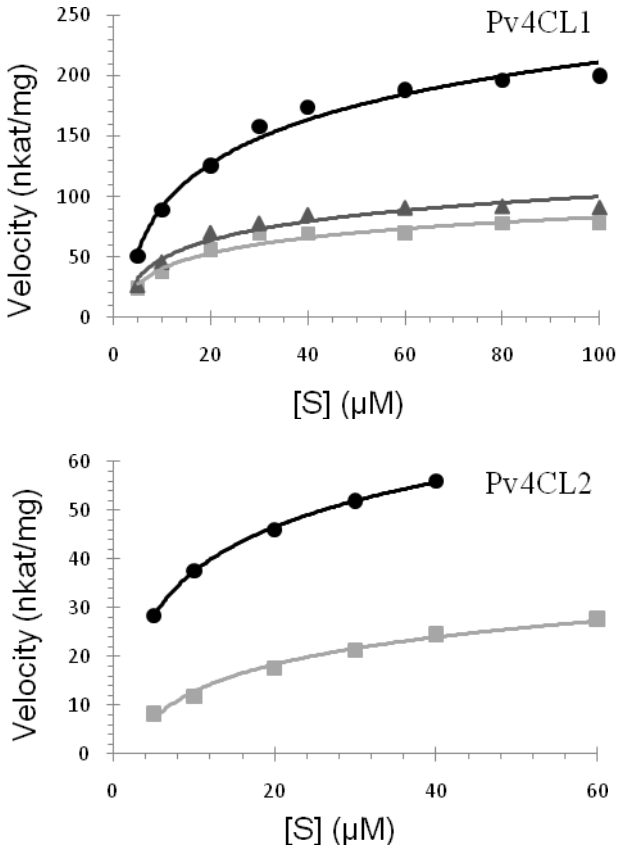
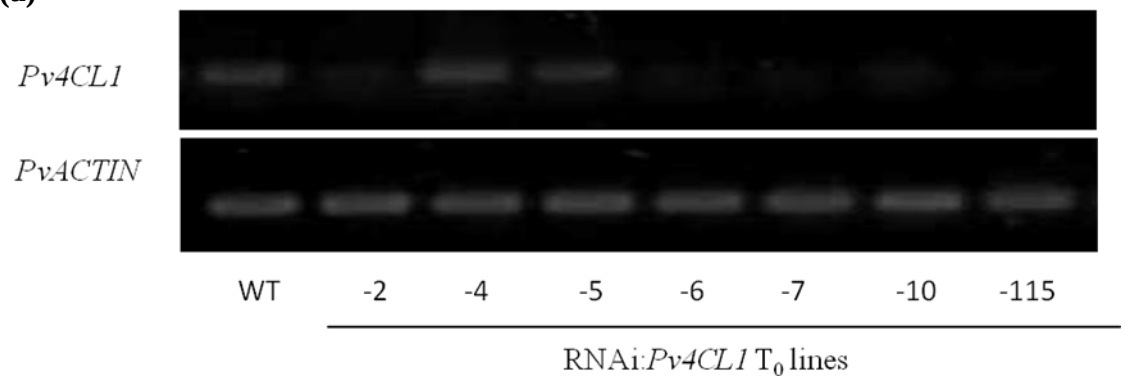
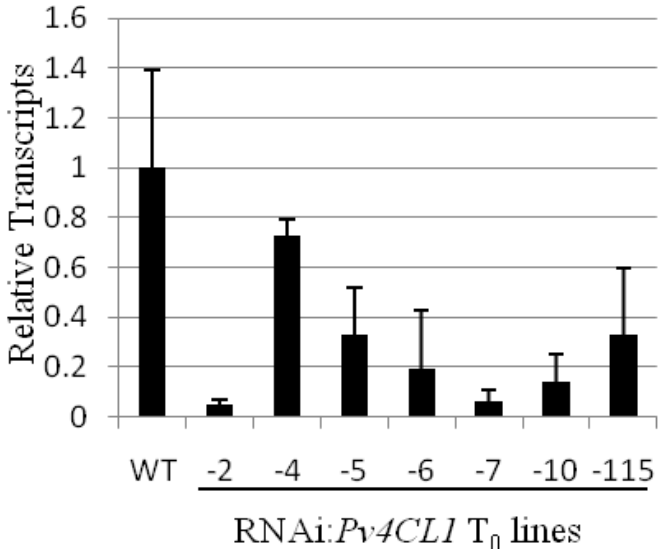


Fig. 3 Functional characterization of Pv4CL1 and Pv4CL2. (a) Protein gel blot analysis of recombinant enzymes expressed in *E. coli*; 1, Molecular weight markers; 2, 3, 4; Pv4CL1; 5, 6 and 7; Pv4CL2; 2 and 5 total proteins from induced cells; 3 and 6 soluble proteins from induced cells; 4 and 7 pure recombinant protein. (b) HPLC chromatograms and retention times of substrates and products after 4CL reaction. (c) Curves of reaction velocity versus substrate concentration for Pv4CL1 and Pv4CL2 towards 4-coumaroyl CoA (dots), caffeoyl CoA (squares) or feruloyl CoA (triangles).

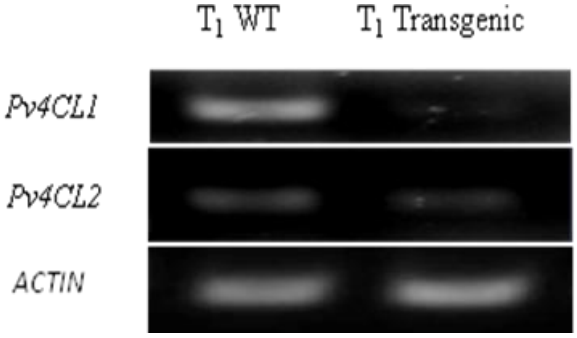
(a)


Pv4CL1
PvACTIN
WT
-2
-4
-5
-6
-7
-10
-115
RNAi:Pv4CL1 T0 lines

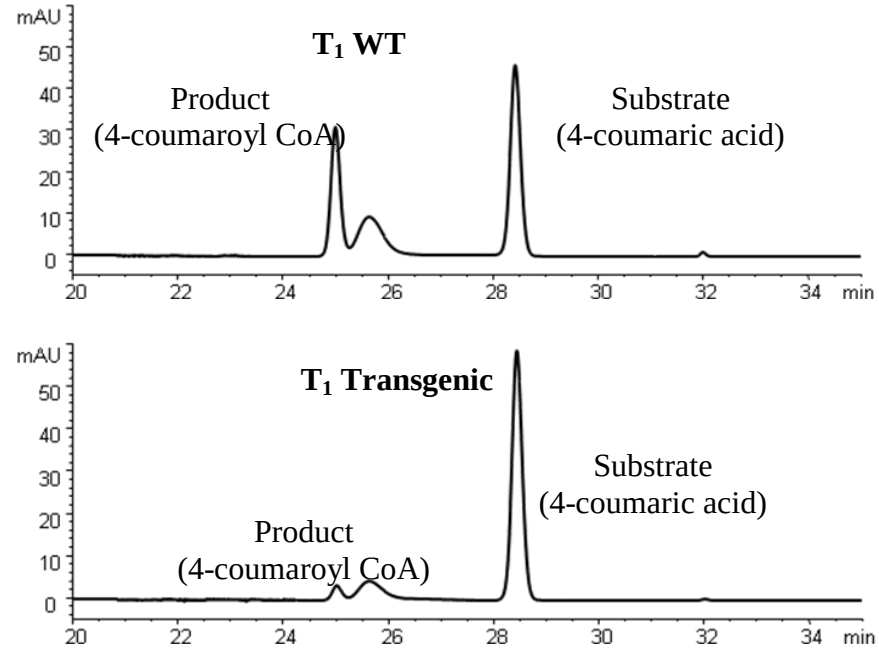
(b)


Relative Transcripts
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
WT
-2
-4
-5
-6
-7
-10
-115
RNAi:Pv4CL1 T0 lines

(c)


T1 WT
T1 Transgenic
Pv4CL1
Pv4CL2
ACTIN

(d)



(e)

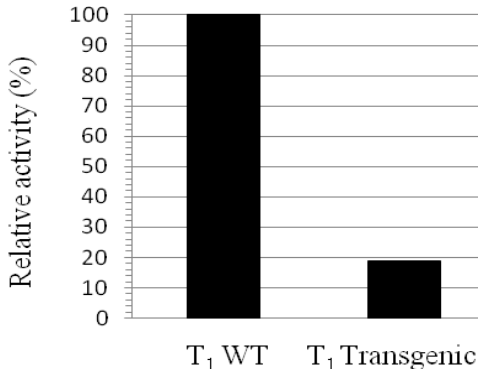


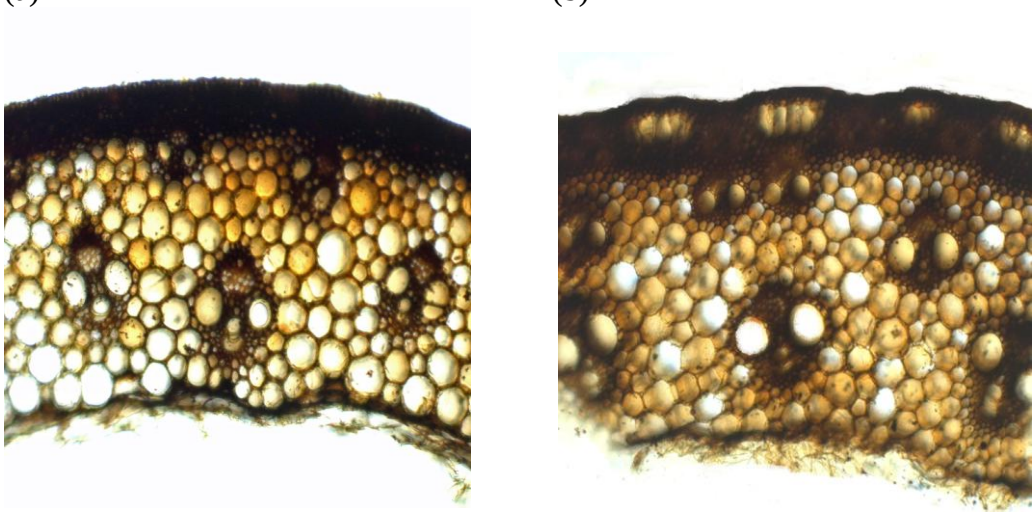
Fig. 4 RNAi:Pv4CL1 lines have reduced *Pv4CL1* transcript abundances and 4CL enzymatic activity. RT-PCR (a) and qRT-PCR (b) showing decreased transcript abundance in individual T_0 transgenic lines. In the T_1 lines, which segregated as transgenic or non-transgenic (wild type), the pooled transgenic lines have decreased transcript abundance detected by RT-PCR (c), and lower 4CL enzymatic activity (d-e). All RNA was extracted from the 3rd internode of switchgrass lines at the flowering stage. For RT-PCR, the reference gene was *Pv_ACTIN2* amplified with primers Pv_ACTIN_for and Pv_ACTIN_rev. A fragment of *Pv4CL1* was amplified with the primers Pv4CL_1st Round_For and Pv4CL SalXba_Rev; and a fragment of *Pv4CL2* was amplified with F_Pv4CL2 and R_Pv4CL2. The RT-PCRs were run with 28 cycles for *Pv4CL1*, *Pv4CL2* and *PvACTIN* with the same amount of cDNA loading in the PCR reaction mix. Two different sets of RNA of the samples were extracted to run two biological replicates of the qRT-PCR, and both *Pv_ACTIN2* and *Pv-UBIQUITIN* were used as reference genes.

(a) (b)

(c) (d)

Fig. 5 RNAi:*Pv4CL1* lines have altered phenotypes. Some RNAi T_0 lines with low lignin content had brown coloration in parts of their leaf veins (a); brown patches in stems (b); reddish-brown coloration on the inner sides of basal stems (c); and brownish color in the mature roots (d). Note: In each picture, the left plant materials are from an RNAi transgenic plant, and the right ones are from wild type.

(a) (b)



(c) (d)

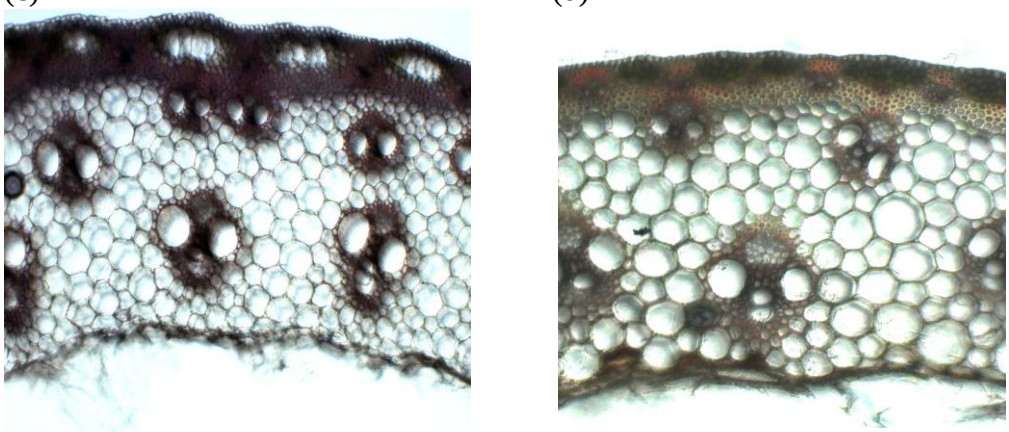


Fig. 6 Histological staining for lignin deposition. Compared to the wild type plants (a, c), the transgenic plants did not show altered staining by Mäule reagent (b). However, lignin staining by phloroglucinol-HCl reagent was clearly reduced (d). The third internodes were used to make the 50 μm thick sections. Five tillers from the transgenic or non-transgenic (wild-type) lines were taken for the sections, and consistent observations were obtained.

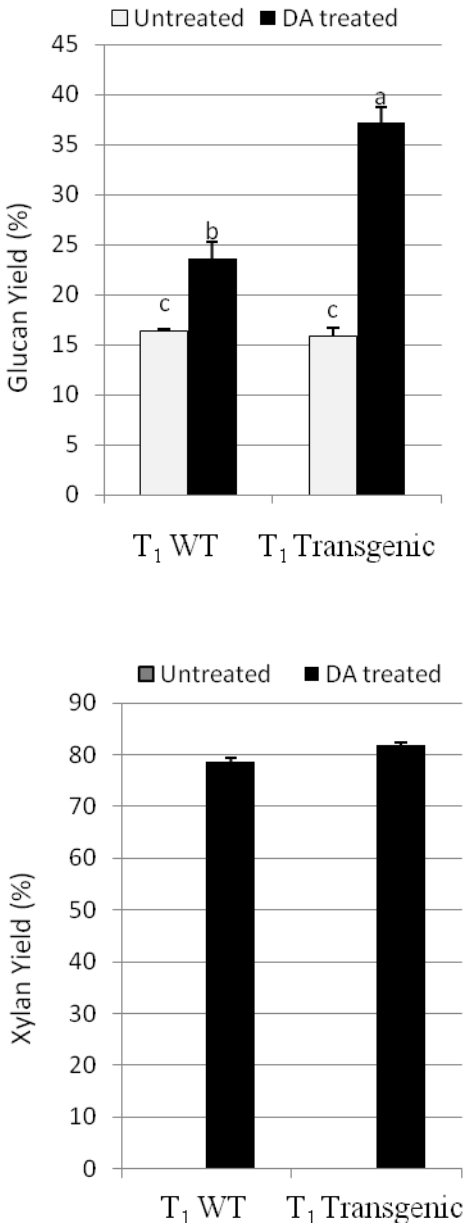


Fig. 7 RNAi:Pv4CL1 switchgrass biomass has improved saccharification efficiency. Biomasses of segregating T_1 plants was either treated with diluted acid (DA) or analyzed without pre-treatment (untreated). Without DA treatment, the xylan yields of both wild-type and transgenic lines were zero. The bars are standard errors. Different letters above the bars indicate statistically significant difference at the level of α=0.01.

References

Abramson M, Shoseyov O, Shani Z. 2009. Plant cell wall reconstruction toward improved lignocellulosic production and processability. *Plant Science* **178**: 61-72.

Adney B, Baker J. 1996. Measurement of cellulase activities. LAP-006 NREL Analytical Procedure, National Renewable Energy Laboratory, Golden, CO (1996).

Bell J, Lehmann D, Wang Z, Chen L, Auh C, 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. *The Plant Cell* **19**: 148-162.

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

Boudet A. 2007. Evolution and current status of research in phenolic compounds. *Phytochemistry* **68**: 2722-2735.

Carroll A, Somerville C. 2009. Cellulosic biofuels. *Annual Review of Plant Biology* **60**: 165-182.

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

Chen L, Auh C, Dowling P, Bell J, Chen F, Hopkins A, Dixon R, Wang Z. 2003. Improved forage digestibility of tall fescue (*Festuca arundinacea*) by transgenic down-regulation of cinnamyl alcohol dehydrogenase. *Plant Biotechnology Journal* **1**: 437-449.

Cherney J, Cherney D, Akin D, Axtell J. 1991. Potential of brown-midrib, low-lignin mutants for improving forage quality. *Advances in Agronomy* **46**: 157-198.

Coleman H, Park J, Nair R, Chapple C, Mansfield S. 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.

Cukovica D, Ehlting J, Ziffle J, Douglas C. 2001. Structure and evolution of *4-coumarate: coenzyme A ligase* (*4CL*) gene families. *Biological Chemistry* **382**: 645-654.

Ehlting J, Buttner D, Wang Q, Douglas C, Somssich I, Kombrink E. 1999. Three 4-coumarate: coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *The Plant Journal* **19**: 9-20.

Endo A, Nakamura T, Ando A, Tokuyasu K, Shima J. 2008. Genome-wide screening of the genes required for tolerance to vanillin, which is a potential inhibitor of bioethanol fermentation, in *Saccharomyces cerevisiae*. *Biotechnology for Biofuels* **1**: 3-9.

Escamilla-Treviño L, Shen H, Uppalapati S, Ray T, Tang Y, Hernandez T, Yin Y, Xu Y, Dixon R. 2009. Switchgrass (*Panicum virgatum*) possesses a divergent family of cinnamoyl CoA reductases with distinct biochemical properties. *The New phytologist* **185**: 143-155.

Fu C, Mielenz JR, Xiao X, Ge Y, Hamilton CY, Rodriguez M, Chen F, Foston M, Ragauskas A, Bouton J, Dixon RA, Wang Z-Y. 2011a. 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, Xiao X, Xi Y, Ge Y, Chen F, Bouton J, Dixon R, Wang Z-Y. 2011b. Downregulation of cinnamyl alcohol dehydrogenase (CAD) leads to improved saccharification efficiency in switchgrass. *BioEnergy Research* **4**: 1-12.

Grabber J, Ralph J, Lapierre C, Barrière Y. 2004. Genetic and molecular basis of grass cell-wall degradability. I. Lignin-cell wall matrix interactions. *Comptes Rendus Biologies* **327**: 455-465.

Halpin C. 2004. Re-designing lignin for industry and agriculture. *Biotechnology & genetic engineering reviews* **21**: 229-245.

Halpin C, Holt K, Chojecki J, Oliver D, Chabbert B, Monties B, Edwards K, Barakate A, Foxon GA. 1998. *Brown-midrib* maize (*bm1*)-a mutation affecting the cinnamyl alcohol dehydrogenase gene. *The Plant Journal* **14**: 545-553.

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

Hu WJ, Lung J, Harding S, Popko J, Ralph J, Stokke D, Tsai C, Chiang VL. 1999. Repression of lignin biosynthesis in transgenic trees promotes cellulose accumulation and growth. *Nature Biotechnology* **17**: 808-812.

Hu WJ, Kawaoka A, Tsai CJ, Lung J, Osakabe K, Ebinuma H, Chiang VL. 1998. Compartmentalized expression of two structurally and functionally distinct *4-coumarate: CoA ligase* genes in aspen (*Populus tremuloides*). *Proceedings of the National Academy of Sciences, USA* **95**: 5407-5412.

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

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.

Kajita S, Katayama Y, Omori S. 1996. Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for *4-coumarate: coenzyme A ligase*. *Plant and Cell Physiology* **37**: 957-965.

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

Kim SJ, Kim MR, Bedgar DL, Moinuddin SGA, Cardenas CL, Davin LB, Kang C, Lewis NG. 2004. Functional reclassification of the putative *cinnamyl alcohol dehydrogenase* multigene family in Arabidopsis. *Proceedings of the National Academy of Sciences, USA* **101**: 1455-1460.

Kirst M, Myburg A, De Leon J, Kirst M, Scott J, Sederoff R. 2004. Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus

analysis of cDNA microarray data in an interspecific backcross of eucalyptus. *Plant Physiology* **135**: 2368-2378.

Kitin P, Voelker S, Meinzer F, Beeckman H, Strauss S, Lachenbruch B. 2010. Tyloses and phenolic deposits in xylem vessels impede water transport in low-lignin transgenic poplars: a study by cryo-fluorescence microscopy. *Plant Physiology* **154**: 887-898.

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

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

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

Moxley G, Zhang YHP. 2007. More accurate determination of acid-labile carbohydrates in lignocellulose by modified quantitative saccharification. *Energy & Fuels* **21**: 3684-3688.

Pomar F, Merino F, Ros Barceló A. 2002. *O*-4-Linked coniferyl and sinapyl aldehydes in lignifying cell walls are the main targets of the Wiesner (phloroglucinol-HCl) reaction. *Protoplasma* **220**: 17–28.

Robinson A, Mansfield S. 2009. Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and NIR-based prediction modeling. *The Plant Journal* **58**: 706-714.

Rolando C, Monties B, Lapierre C. 1992. Thioacidolysis. In Methods in lignin chemistry; Dence, C. W., Lin, S. Y., Eds.;Springer-Verlag: Berlin, pp 334-349.

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

Saathoff AJ, Tobias C, Sattler S, Haas E, Twigg P, Sarath G. 2011b. Switchgrass contains two cinnamyl alcohol dehydrogenases involved in lignin formation. *BioEnergy Research*: **4**:120-133.

Sanchez J, Ullman C, Moore M, Choo Y, Chua N. 2006. Regulation of *Arabidopsis thaliana 4-coumarate: coenzyme-A ligase-1* expression by artificial zinc finger chimeras. *Plant Biotechnology Journal* **4**: 103-114.

Sarath G, Baird L, Vogel K, Mitchell R. 2007. Internode structure and cell wall composition in maturing tillers of switchgrass (*Panicum virgatum*. L). *Bioresource technology* **98**: 2985-2992.

Schneider K, Hövel K, Witzel K, Hamberger B, Schomburg D, Kombrink E, Stuible HP. 2003. The substrate specificity-determining amino acid code of 4-coumarate:CoA ligase. *Proceedings of the National Academy of Sciences, USA* **100**: 8601-8606.

Schubert C. 2006. Can biofuels finally take center stage? *Nature Biotechnology* **24**: 777-784.

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 Research* **2**: 233-245.

Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D. 2004. Determination of structural carbohydrates and lignin in biomass.In. *NREL, Golden, CO*.

Smita R, Nath D. 2008. Manipulation of lignin in plants with special reference to O-methyltransferase. *Plant Science* **174**: 264-277.

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

Stockigt J, Zenk MH. 1975. Chemical syntheses and properties of hydroxycinnamoyl-coenzyme A derivatives. *Z Naturforsch [C]* **30**: 352-358.

Stuible H, Kombrink E. 2001. Identification of the substrate specificity-conferring amino acid residues of 4-coumarate: coenzyme A ligase allows the rational design of mutant enzymes with new catalytic properties. *Journal of Biological Chemistry* **276**: 26893-26897.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular biology and evolution* **24**: 1596-1599.

Tobias C, Sarath G, Twigg P, Lindquist E, Pangilinan J, Penning B, Barry K, McCann M, Carpita N, Lazo G. 2008. Comparative genomics in switchgrass using 61,585 high-quality expressed sequence tags. *The Plant Genome* **1**: 111-124.

Tu Y, Rochfort S, Liu Z, Ran Y, Griffith M, Badenhorst P, Louie G, Bowman M, Smith K, Noel J. 2010. Functional analyses of *caffeic acid O-methyltransferase* and *cinnamoyl-CoA-reductase* genes from perennial ryegrass (*Lolium perenne*). *The Plant Cell* **22**: 3357-3373.

Umezawa T. 2010. The cinnamate/monolignol pathway. *Phytochemistry Reviews* **9**(1): 1-17.

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

Voelker S, Lachenbruch B, Meinzer F, Jourdes M, Ki C, Patten A, Davin L, Lewis N, Tuskan G, Gunter L. 2010. Antisense down-regulation of *4CL* expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiology* **154**: 874-886.

Voelker SL, Lachenbruch B, Meinzer FC, Strauss SH. 2011. Reduced wood stiffness and strength, and altered stem form, in young antisense *4CL* transgenic poplars with reduced lignin contents. *New Phytologist* **189**: 1096-1109.

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.

Weng JK, Akiyama T, Bonawitz ND, Li X, Ralph J, Chapple C. 2010a. Convergent evolution of syringyl lignin biosynthesis via distinct pathways in the lycophyte *Selaginella* and flowering plants. *The Plant Cell* **22**: 1033-1045.

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

Weng JK, Mo H, Chapple C. 2010b. Over-expression of *F5H* in *COMT*-deficient Arabidopsis leads to enrichment of an unusual lignin and disruption of pollen wall formation. *The Plant Journal* **64**: 898-911.

Xu Z, Zhang D, Hu J, Zhou X, Ye X, Reichel K, Stewart N, Syrenne R, Yang X, Gao P. 2009. Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. *BMC Bioinformatics* **10**: S3.

Zhang YHP, Berson E, Sarkanen S, Dale B. 2009. Sessions 3 and 8: Pretreatment and biomass recalcitrance: fundamentals and progress. *Applied Biochemistry and Biotechnology* **153**: 80-83.

Zhao Q, Gallego Giraldo L, Wang H, Zeng Y, Ding S, Chen F, Dixon R. 2010. An NAC transcription factor orchestrates multiple features of cell wall development in *Medicago truncatula*. *The Plant Journal* **63**: 100-114.

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

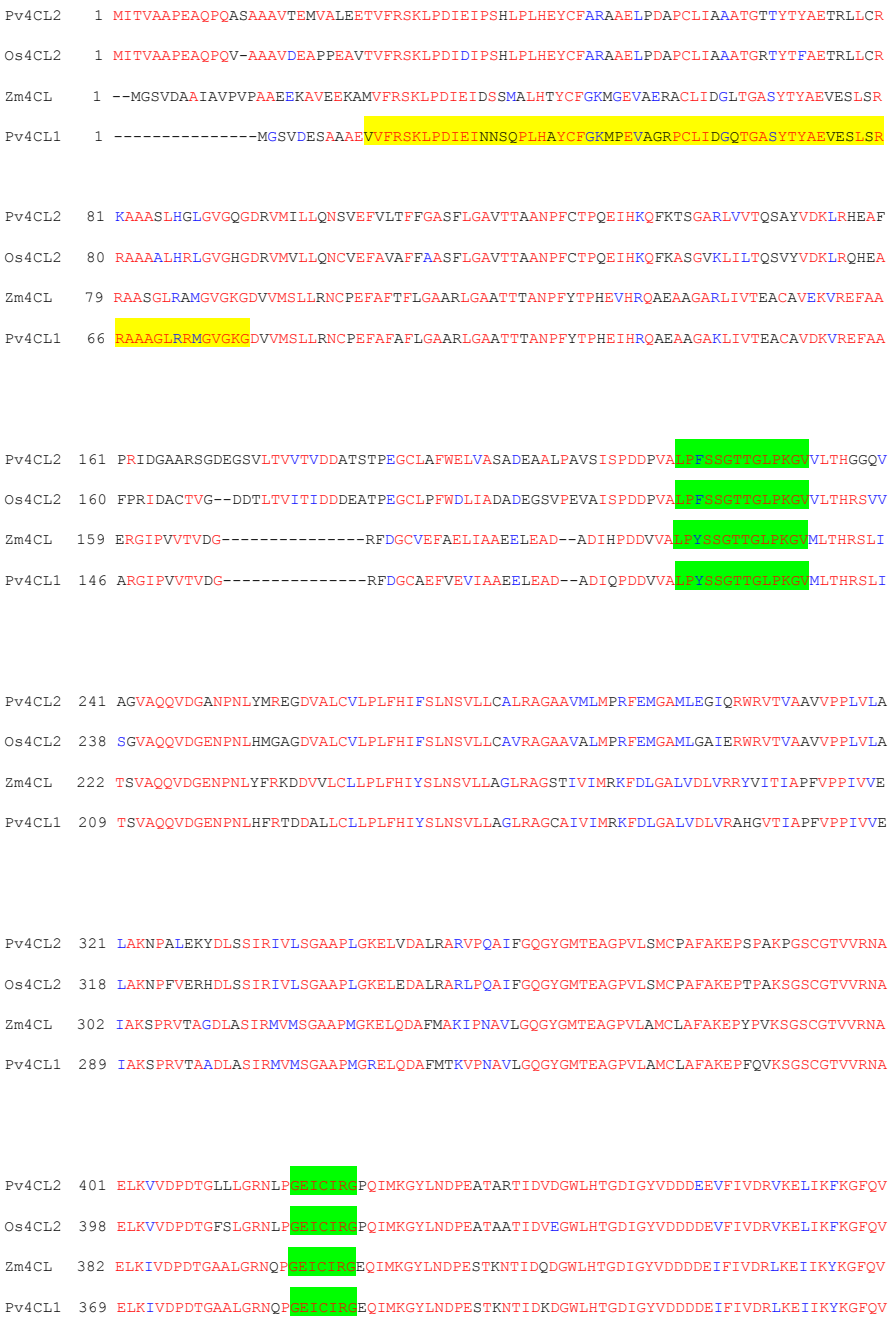
Zhou J, Lee C, Zhong R, Ye Z. 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *The Plant Cell* **21**: 248-266.

Zhu Z, Sathitsuksanoh N, Vinzant T, Schell DJ, McMillan JD, Zhang YHP. 2009. Comparative study of corn stover pretreated by dilute acid and cellulose solvent-based lignocellulose fractionation: Enzymatic hydrolysis, supramolecular structure, and substrate accessibility. *Biotechnology and Bioengineering* **103**: 715-724.

Supplementary data

Fig. S1 Alignment of two switchgrass 4CL amino acid sequences with their rice and maize orthologs (a) and alignment of two switchgrass *4CL* nucleotide sequences (b). The yellow-color highlighted amino acid or nucleotide sequences were used for the construct of the RNAi:*Pv4CL1* vector. The green-color highlighted amino acids sequences were the conserved Box I and Box II domains of 4CL (Stuible and Kombrink, 2001).

(a)



(b)

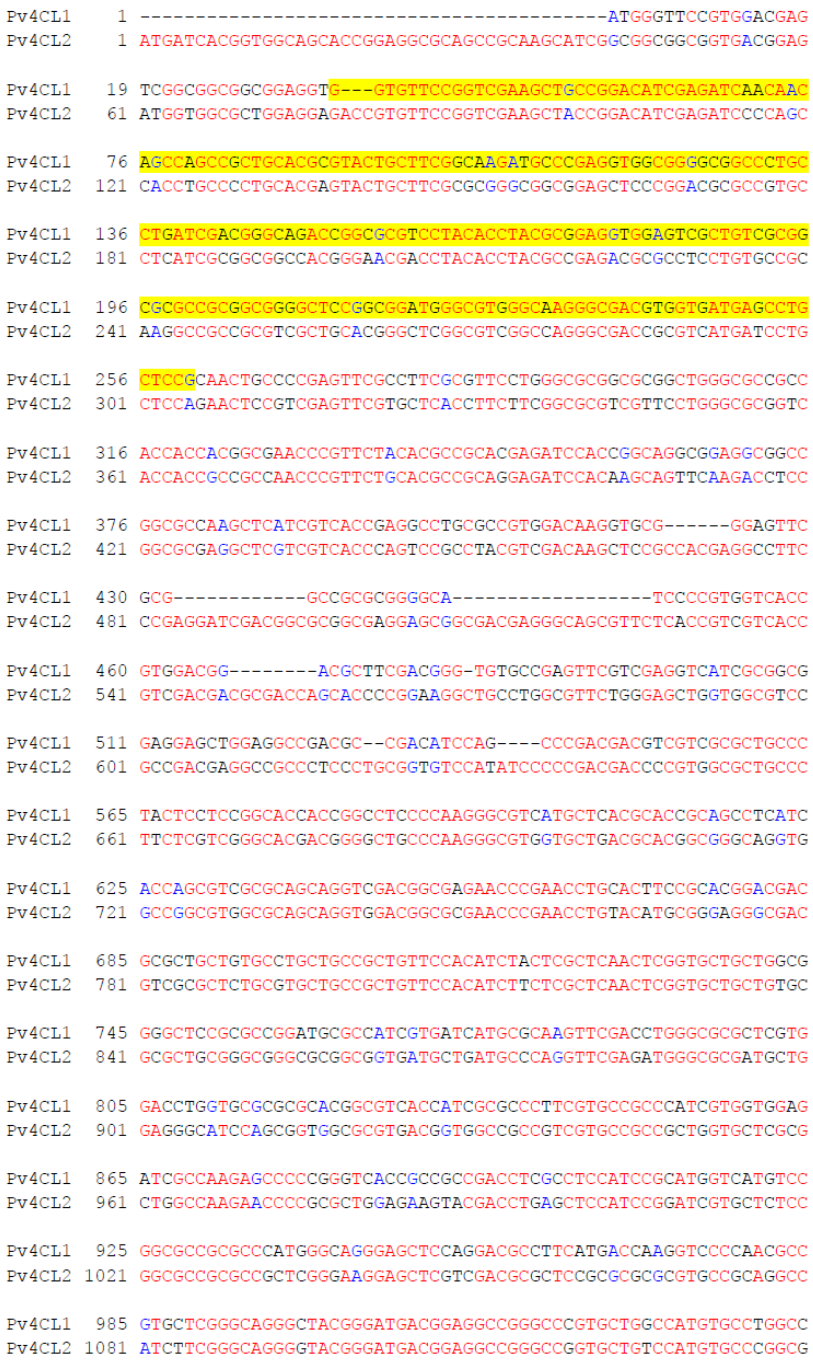

Pv4CL1 1 ----------ATGGGTTCCGTGGACGAG
Pv4CL2 1 ATGATCACGGTGGCAGCACCGGAGGCGCAGCCGCAAGCATCGGCGGCGGCGGTGACGGAG
Pv4CL1 19 TCGGCGGCGGCGGAGGTG---GTGTTCCGGTCGAAGCTGCCGGACATCGAGATCAACAAC
Pv4CL2 61 ATGGTGGCGCTGGAGGAGACCGTGTTCCGGTCGAAGCTACCGGACATCGAGATCCCCAGC
Pv4CL1 76 AGCCAGCCGCTGCACGCGTACTGCTTCGGCAAGATGCCCGAGGTGGCGGGGCGGCCCTGC
Pv4CL2 121 CACCTGCCCCTGCACGAGTACTGCTTCGCGCGGCGGCGGAGCTCCCGGACGCGCCGTGC
Pv4CL1 136 CTGATCGACGGGCAGACCGGCGCGTCCTACACCTACGCGGAGGTGGAGTCGCTGTCGCGG
Pv4CL2 181 CTCATCGCGGCGGCCACGGGAACGACCTACACCTACGCCGAGACGCGCCTCCTGTGCCGC
Pv4CL1 196 CGCGCCGCGGCGGGGCTCCGGCGGATGGGCGTGGGCAAGGGCGACGTGGTGATGAGCCTG
Pv4CL2 241 AAGGCCGCCGCGTCGCTGCACGGGCTCGGCGTCGGCCAGGGCGACCGCGTCATGATCCTG
Pv4CL1 256 CTCCGCAACTGCCCCGAGTTCGCCTTCGCGTTCCTGGGCGCGGCGCGGCTGGGCGCCGCC
Pv4CL2 301 CTCCAGAACTCCGTCGAGTTCGTGCTCACCTTCTTCGGCGCGTCGTTCCTGGGCGCGGTC
Pv4CL1 316 ACCACCACGGCGAACCCGTTCTACACGCCGCACGAGATCCACCGGCAGGCGGAGGCGGCC
Pv4CL2 361 ACCACCGCCGCCAACCCGTTCTGCACGCCGCAGGAGATCCACAAGCAGTTCAAGACCTCC
Pv4CL1 376 GGCGCCAAGCTCATCGTCACCGAGGCCTGCGCCGTGGACAAGGTGCG------GGAGTTC
Pv4CL2 421 GGCGCGAGGCTCGTCGTCACCCAGTCCGCCTACGTCGACAAGCTCCGCCACGAGGCCTTC
Pv4CL1 430 GCG----------GCCGCGCGGGGCA----------TCCCCGTGGTCACC
Pv4CL2 481 CCGAGGATCGACGGCGCGGCGAGGAGCGGCGACGAGGGCAGCGTTCTCACCGTCGTCACC
Pv4CL1 460 GTGGACGG--------ACGCTTCGACGGG-TGTGCCGAGTTCGTCGAGGTCATCGCGGCG
Pv4CL2 541 GTCGACGACGCGACCAGCACCCCGGAAGGCTGCCTGGCGTTCTGGGAGCTGGTGGCGTCC
Pv4CL1 511 GAGGAGCTGGAGGCCGACGC--CGACATCCAG----CCCGACGACGTCGTCGCGCTGCCC
Pv4CL2 601 GCCGACGAGGCCGCCCTCCCTGCGGTGTCCATATCCCCCGACGACCCCGTGGCGCTGCCC
Pv4CL1 565 TACTCCTCCGGCACCACCGGCCTCCCCAAGGGCGTCATGCTCACGCACCGCAGCCTCATC
Pv4CL2 661 TTCTCGTCGGGCACGACGGGGCTGCCCAAGGGCGTGGTGCTGACGCACGGCGGGCAGGTG
Pv4CL1 625 ACCAGCGTCGCGCAGCAGGTCGACGGCGAGAACCCGAACCTGCACTTCCGCACGGACGAC
Pv4CL2 721 GCCGGCGTGGCGCAGCAGGTGGACGGCGCGAACCCGAACCTGTACATGCGGGAGGGCGAC
Pv4CL1 685 GCGCTGCTGTGCCTGCTGCCGCTGTTCCACATCTACTCGCTCAACTCGGTGCTGCTGGCG
Pv4CL2 781 GTCGCGCTCTGCGTGCTGCCGCTGTTCCACATCTTCTCGCTCAACTCGGTGCTGCTGTGC
Pv4CL1 745 GGGCTCCGCGCCGGATGCGCCATCGTGATCATGCGCAAGTTCGACCTGGGCGCGCTCGTG
Pv4CL2 841 GCGCTGCGGGCGGGCGCGGCGGTGATGCTGATGCCCAGGTTCGAGATGGGCGCGATGCTG
Pv4CL1 805 GACCTGGTGCGCGCGCACGGCGTCACCATCGCGCCCTTCGTGCCGCCCATCGTGGTGGAG
Pv4CL2 901 GAGGGCATCCAGCGGTGGCGCGTGACGGTGGCCGCCGTCGTGCCGCCGCTGGTGCTCGCG
Pv4CL1 865 ATCGCCAAGAGCCCCCGGGTCACCGCCGCCGACCTCGCCTCCATCCGCATGGTCATGTCC
Pv4CL2 961 CTGGCCAAGAACCCCGCGCTGGAGAAGTACGACCTGAGCTCCATCCGGATCGTGCTCTCC
Pv4CL1 925 GGCGCCGCGCCCATGGGCAGGGAGCTCCAGGACGCCTTCATGACCAAGGTCCCCAACGCC
Pv4CL2 1021 GGCGCCGCGCCCGCTCGGGAAGGAGCTCGTCGACGCGCTCCGCGCGCGCGTGCCGCAGGCC
Pv4CL1 985 GTGCTCGGGCAGGGCTACGGGATGACGGAGGCCGGGCCCGTGCTGGCCATGTGCCTGGCC
Pv4CL2 1081 ATCTTCGGGCAGGGGTACGGGATGACGGAGGCCGGGCCGGTGCTGTCCATGTGCCCGGCG

Fig. S2 Key features of the vectors used in this study. The pZHAO1629 (a) vector was constructed as a Gateway-compatible destination vector for monocot transformation. The pEntry/D-Kannibal (b) was used to construct the palindromic structure of a target gene separated by a PDK intron. LB: the left T-DNA border; RB: the right T-DNA border; UbiP: maize *Ubi* promoter and 5' UTR; ccdB (B): the Invitrogen destination cassette frame B; HA: hemoglobin epitope tag; NosT: Nos terminator. Other features are identical to the original pCAMBIA1305.2 vector; AttL1 and attL2 are the flanking sites allowing easy site-specific recombination to the destination vector. The two pairs of unique restriction sites flanking the PDK intron allow cloning of the antisense and sense strands of the targeted genes.

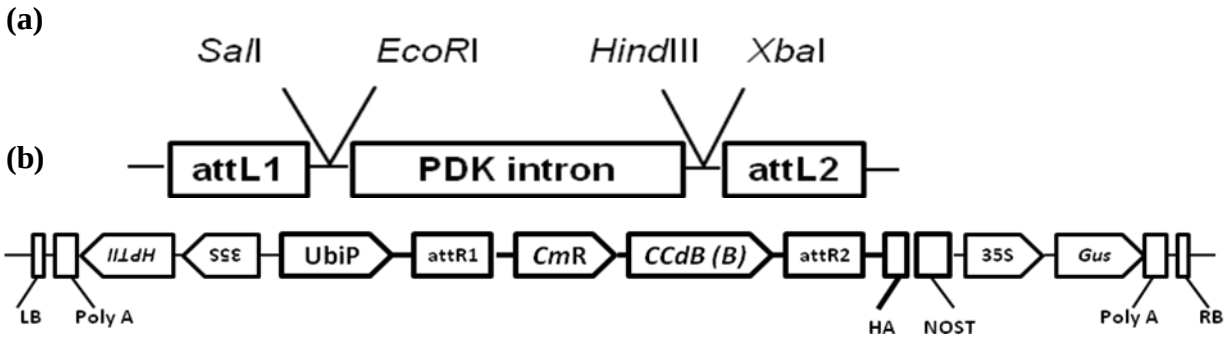
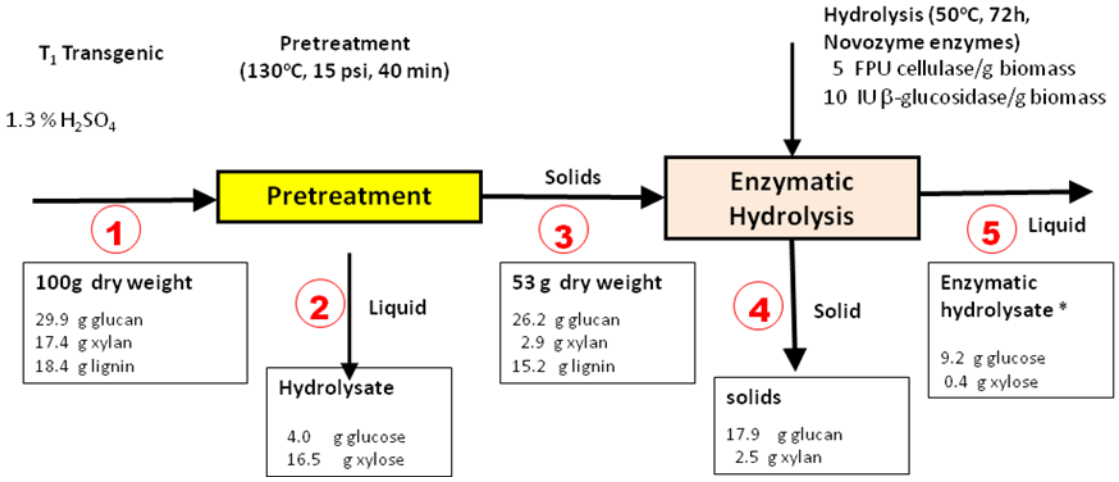


Fig. S3 Mass balance of Dilute Acid Pretreatment and Enzymatic Hydrolysis of T_1 transgenic (a) and wild type (b) biomass.

(a)



(b)

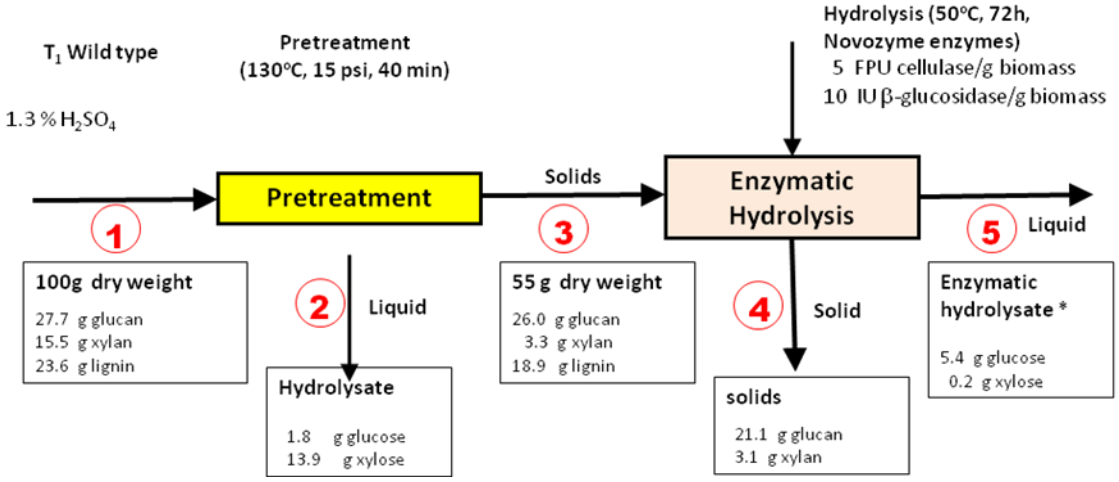


Table S1 Primers used in the study.

Primer Sets	Primer Sequences (from 5' to 3')	PCR Targets (NCBI Accession Number)
Pv4CL1_ORF_For	CACCATGGGTTCCGTGGACGAGTC	*Pv4CL1*(EU491511)
Pv4CL1_ORF_Rev	TCAGTGGACACCGGCGGCGAG	*Pv4CL1*(EU491511)
Pv4CL2_ORF_For	CACCATGATCACGGTGGCAGCACC	*Pv4CL2* (JF414903)
Pv4CL2_ORF_Rev	TGAGGCTGCAGCTACTCAGG	*Pv4CL2* (JF414903)
Pv4CL_1st Round_For	TAGTTGAACCAATCGCCTACTGCAC	*Pv4CL1*(EU491511)
Pv4CL_1st Round_Rev	GAACTCCCGCACCTTGTCCACGGCGC	*Pv4CL1*EU491511)
Pv4CL_H3RI_For	CACCAAGCTTGAATTCGGTGTTCCGGTCGAAGCTG	*Pv4CL1*(EU491511)
Pv4CL SalXba_Rev	GCTCTAGAGTCGACGAACTCCCGCACCTTGTCC	*Pv4CL1*(EU491511)
Pv_ACTIN_for	CACTGGAATGGTCAAGGCAG	*PvACTIN2* (FL724919.1)
Pv_ACTIN_Rev	CTCCATGTCATCCCAGTTG	*PvACTIN2* (FL724919.1)
F_Pv4CL1	TCGAAGCTGCCGGACATC	*Pv4CL1*(EU491511)
R_Pv4CL1	GGCATCTTGCCGAAGCAGTA	*Pv4CL1*(EU491511)
F_Pv4CL2	GAAGTCTCCGGTTCGTCGTCGG	*Pv4CL2* (JF414903)
R_Pv4CL2	CTTGCGGCACAGGAGGCGCGT	*Pv4CL2* (JF414903)
F_PvUBIQUITIN	CAGCGAGGGCTCAATAATTCCA	*PvUBIQUITIN* (FL955474.1)
R_PvUBIQUITIN	TCTGGCGGACTACAATATCCA	*PvUBIQUITIN* (FL955474.1)
F_PvACTIN2	GCGAGCTTCCCTGTAGGTA	*PvACTIN2* (FL724919.1)
R_PvACTIN2	CGAACCCAGCCTTCACCATAC	*PvACTIN2* (FL724919.1)
F_HPTII	CAAACTGTGATGGACGACACCG	*HPTII (AF354046.1)*
R_HPTII	TATATGCTCAACACATGAGCG	*HPTII (AF354046.1)*

SUPPLEMENTAL LITERATURE CITED

Stuible H, Kombrink E (2001) Identification of the substrate specificity-conferring amino acid residues of 4-coumarate: coenzyme A ligase allows the rational design of mutant enzymes with new catalytic properties. J Bioll Chem **276:**26893-26897

Chapter 4

Title: Ectopic expression of *AtLOV1* in switchgrass altered plant architecture, lignin content, and flowering time

Authors: Bin Xu[1], Sathitsuksanoh Noppadon[2], Yuhong Tang[3], Michael Udvardi[3], Jiyi Zhang[3], Zhengxing Shen[1], Maria Balota[4], Kim Harich[5], Percival Y-H Zhang[2], Bingyu Zhao[1†]

1. Department of Horticulture, Virginia Tech, Blacksburg, VA. 24061; 2. Department of Biological Systems Engineering, Virginia Tech, Blacksburg, VA. 24061; 3. Noble Foundation, Ardmore, OK. 73401; 4. Department of Plant Pathology, Plant Physiology and Weed Science, Virginia Tech, Blacksburg, VA. 24061; 5. Department of Biochemistry, Virginia Tech, Blacksburg, VA. 24061.

[†] Corresponding author; email bzhao07@vt.edu; Tel: +1 540 231 1146; Fax: 540-231-3083

Abstract

Switchgrass (*Panicum virgatum* L.) is one prime candidate crop for biofuel feedstock production in the United States. As a self-incompatible, polyploid perennial plant species, it is challenging to breed elite and stable switchgrass cultivars with traditional breeding methods. Translational genomics is a contributory strategy for switchgrass genetic improvement, especially for generating elite traits absent in natural switchgrass populations. Erect leaf phenotype can improve plant architecture through increased shade avoidance and thereby maximize plant biomass yield in a dense field population. In this study, we constitutively expressed an *Arabidopsis* NAC transcriptional factor gene, *LONG VEGETATIVE PHASE ONE* (*AtLOV1*), in switchgrass. Ectopic expression of *AtLOV1* causes smaller leaf angles by changing morphologies of epidermal cells in the leaf collar region, affected lignin content and monolignol composition, and also delayed flowering time in switchgrass. Global gene-expression analysis of the transgenic plants revealed an array of responding genes with predicted functions in plant development, cell wall biosynthesis and flowering time. To our knowledge, this is the first report of altering leaf angle, cell wall composition and flowering time by ectopically-expressing a single gene, which further suggests the potential of applying translational genomics to the genetic improvement of crops.

Key words: *Panicum virgatum* L., leaf angle, lignin, flowering, NAC

Introduction

Switchgrass (*Panicum virgatum* L.) is an attractive biomass crop for the US, because it is a North American native perennial C4 grass rich in natural diversity, and it is naturally adapted over a large portion of the continent (McLaughlin & Adams Kszos, 2005). Genetic improvement of switchgrass for the purpose of large scale biomass feedstock production started recently in the US. Currently, all switchgrass cultivars are synthetic and were directly selected from their original habitats, where average biomass yields range from four to ten tons/acre/year dependent upon cultivar and field condition (Lemus *et al.*, 2002). Due to its polyploid genome and gametophytic self-incompatibility, switchgrass breeding is a challenge (Vogel & Mitchell, 2008).

By combining strategies of traditional breeding and genetic engineering, it could be possible to significantly hasten the process of selection in elite switchgrass cultivars. The knowledge gained from the extensive study of model plant species, such as *Arabidopsis*, rice, *Medicago truncatula*, and *Brachypodium distachyon*, could also significantly contribute to overcoming important agronomic problems for major food crops and non-food crops such as switchgrass (Zhang *et al.*, 2004; Lawrence & Walbot, 2007). Candidate genes isolated from model plant species could be directly expressed in crop plants to improve agronomic traits (Takeda & Matsuoka, 2008) through efficient tissue culture based genetic transformation systems (Li & Qu, 2011; Xu, *et al.*, 2011a). In certain cases, novel phenotypes could also be triggered by ectopically expressing a candidate gene across different plant species. For example, overexpression of an *Arabidopsis* transcription factor gene *SHINE1* (*SHN1*) in rice caused an increased cellulose and decreased lignin content in the cell wall biomass, where similar phenotypes were not observed in the *Arabidopsis* genetic background (Ambavaram *et al.*, 2011).

An erect leaf phenotype can improve plant architecture by decreasing the shade effects and thereby maximizing plant biomass yield in a dense field population. Small leaf angles against vertical stems (erect leaves) are required to sustain a high leaf area index (LAI, ratio of upper leaf surface area per unit of land area) and to increase the light perception for photosynthesis in dense plantings (Sakamoto *et al.*, 2005). A progression of decreased

leaf angles is associated with higher yields of newly developed rice cultivars (Sinclair & Sheehy, 1999). Mutations in several brassinosteroid (BR)-related genes, overexpression of one *ARGONAUTE* gene (*OsAGO4*), and mutation of one *KANADI* gene (*SLL1*) induced smaller leaf angles in rice plants (Shi *et al.*, 2007; Wang *et al.*, 2008; Yang & Hwa, 2008; Zhang *et al.*, 2009). Altered leaf angle could be a result of developmental changes in the leaf collar region as in the case of some BR-related mutants (Wang *et al.*, 2008), or caused by the leaf rolling (upward curved leaf) as in the case of *OsAGO4* and *SLL1* mutants (Shi *et al.*, 2007; Zhang *et al.*, 2009). In switchgrass, a linear relationship between leaf area index (LAI) and seasonal biomass yield was observed (Madakadze *et al.*, 1998). Decreasing leaf angles by genetic engineering could increase switchgrass biomass yield in dense field plantings.

Higher lignin content in plant leaf and stem tissues could also contribute to an erect leaf phenotype (Zhang *et al.*, 2009). Although low lignin content in switchgrass biomass is preferred for biological conversion of lignocellulosic feedstock to biofuel (Chen & Dixon, 2007), biomass feedstock with high energy content and low mineral residues is highly desirable in other bioenergy-production processes such as pyrolysis and combustion (Boateng *et al.*, 2008).

The accumulation of aboveground biomass ceases when plants begin flowering (Van Esbroeck, *et al.*, 1997). Plants with delayed flowering time have extended vegetative growth and usually produce more biomass. In switchgrass, the lowland ecotypes flower late in high latitude areas and therefore produce higher yields than the upland ecotypes (Lemus *et al.*, 2002). Although genetic components controlling switchgrass flowering time haven't been characterized, it is speculated that switchgrass flowering genes should be similar to its closest outcrossing relative, maize (Demura & Ye, 2010). Interestingly, maize flowering time, unlike *Arabidopsis* and rice, is controlled by a collection of quantitative loci with smaller effects (Buckler *et al.*, 2009), though some conserved flowering genes, such as *flowering locus T* (*FT*, *ZCN8* in maize) and *Indeterminate1* (*id1*), have also been identified in the maize genome (Matsubara *et al.*, 2008; Meng *et al.*, 2011). It is intriguing to control flowering time by identifying and manipulating the flowering genes in switchgrass.

Transcription factor (TF) proteins bind conserved cis-elements to transactivate or suppress the expression of a group of genes which might involve in multiple cellular processes (Jedlicka *et al.*, 1997). The expression of plant TF is normally tightly controlled by developmental stages of specific organs and tissues or regulated by various environmental stimuli (Century *et al.*, 2008). Therefore, knockout of a transcription factor often results in pleiotropic effects. On the other hand, constitutive expression of a transcription factor frequently led to gain-of-function traits. Therefore, it is possible to regulate multiple traits such as: erect leaf, altered lignin content, and delayed flowering time by manipulating one or a few plant TF genes.

NAC (NAM, ATAF1/2 and CUC2) domain genes constitute the largest plant-specific TF family (105 NAC genes in Arabidopsis, and 75 in rice) (Ooka *et al.*, 2003). Despite the fact that NAC transcription factors have diverse roles in plant development and stress adaptation, thus far, none of the characterized NAC transcription factors functions in the regulation of leaf angle (Jensen, *et al.*, 2010).

Recently, it was reported that an Arabidopsis NAC domain TF gene, *Long Vegetative Phase 1* (*AtLOV1*, At2g02450) contributes to cold tolerance and delayed flowering time under the long day photoperiod when it was overexpressed in Arabidopsis (Yoo *et al.*, 2007). Here, we demonstrate that switchgrass plants with ectopically overexpressed *AtLOV1* have unique phenotypic changes, including erect leaves, increased lignin content, and delayed flowering time. Further breeding selection by incorporating the *AtLOV1* transgene in elite switchgrass cultivars could improve switchgrass biomass feedstock production especially in densely planted fields.

MATERIALS AND METHODS

Gene Cloning and Switchgrass Genetic Transformation

A Gateway compatible binary vector pVT1629 (Xu, *et al.*, 2011b), modified from pCAMBIA1305.2, was developed for switchgrass genetic transformation, in which a ccdB(B) cassette can be replaced with the gene of interest through LR reaction

(Invitrogen, Carlsbad, CA). The maize *Ubi* promoter (Christensen & Quail, 1996) was used in the vector to drive the gene of interest.

The full length *AtLOV1* gene was amplified from genomic DNA of *Arabidopsis thaliana* ecotype Columbia-0 with primers LOV1_BamHFor and LOV1_SalRev (supplementary table S1), cloned into the vector p-ENTR/D-TOPO, and sequenced. The *AtLOV1* was then cloned into pVT1629 from the Entry vector through LR reaction. The resultant pVT1629-LOV1 vector was transformed into *Agrobacterium tumefaciens* C58C1 through electroporation. Somatic embryogenic calli induced from seeds of a selected line, HR8, were used for genetic transformation (Xu, *et al.*, 2011a). The regenerated plants were selected under 50 mg/L of hygromycin B (Sigma Chemical Co., St. Louis, MO). Regenerated plants from independent callus were further verified by PCR, GUS staining (data not shown) and Southern blot analysis. Six independent transgenic lines were transplanted in Miracle-Gro Potting Mix (Miracle-Gro Lawn Products, Marysville, OH) in 1.1×10^{-2} m^3 pots in the Horticulture greenhouse at Virginia Tech (Blacksburg, VA) with temperatures set at 22^{o}C/night and 28^{o}C/day with 12-14 light. The plants were well watered by maintaining the soil water content at ~80%. Wild type (WT) plants regenerated from non-transformed calli were also grown under the same condition. T_1 generation transgenic plants were obtained by crossing independent T_0 transgenic plants with the WT controls. The presence or absence of the *AtLOV1* transgene in T_1 plants was detected by PCR with primers LOV1_BamHFor and LOV1_SalRev (data not shown). T_1 segregated plants were maintained in the greenhouse under the same condition as described above.

RT-PCR and Southern Blot

Total RNA was extracted from leaves of transgenic and WT plants by using the TRIzol Reagent (Invitrogen). cDNAs were generated by reverse transcription with SuperScript III (Invitrogen). The expression of *AtLOV1* was detected by RT-PCR using LOV1.2_FOR and LOV1_SalRev (supplementary table S1). Genomic DNA was extracted from four T_0 transgenic *AtLOV1* plants for Southern blot analysis using a DNA fragment of the *HPTII* gene as a probe (Xu *et al*., 2011a). In breif, 10 μg of switchgrass

genomic DNA was digested with restriction enzyme *Hin*dIII which is absent in the *HPTII* gene but only present in one side the T-DNA region. The digested DNA samples were electrophoresed in 0.8% agrose gel, and then transferred to a nylon membrane (Whatman Schleicher and Schuell, Keene, NH). After pre-hybridization in hybridization buffer [6xSSPE pH 7.0, 5x Denhardt Solution, and 0.5% sodium dodecyl sulfate (SDS)] containing 200 mg/ml of denatured Herring sperm DNA at 65 °C for 3 h, hybridization with the probe was carried out at 68 ^{o}C for 24 hrs. After hybridization, the filters were washed three times for 90 min each in a buffer (2 x SSC, 0.1% SDS), at 68°C, and exposed to X-ray film at –80 ^{o}C for 3 days.

Scanning Electron Microscopy

Leaf collars of T_1 generation *AtLOV1* transgenic and WT plants were collected and prepared for scanning electron microscopy according to the protocol described by Li *et al.* (2006). The mounted specimens were examined with a scanning electron microscope (Phillips 505) and digital images were captured at the core facility of the Maryland-Virginia veterinary school at Virginia Tech.

Lignin Content and Monolignol Composition Measurement

Lignin and ash were measured according to the standard biomass protocol developed by the National Renewable Energy Laboratory (NREL) (Sluiter *et al.*, 2004). Monolignol composition was analyzed with a recently modified thioacidolysis protocol (Robinson & Mansfield, 2009). The silylated sample was injected into the GC column (Restek RTX5-MS, 1 uM film thickness, 30 M x 3.2 mM i.d., Thames Restek UK Ltd, Windsor, UK), and the GC-MS was performed on a VG 70SE double focusing magnetic sector instrument, interfaced to a HP5790 GC, and the results were analyzed according to a previous reported method (Xu, *et al.*, 2011b).

Plant Growth Measurement and Statistical Analysis

T_1 generation plants segregating with or without the *AtLOV1* transgene were used to determine plant growth measurement. The growth conditions of the T_1 plants are

described as above. Flowering time and biomass yields were recorded from 9 transgenic and 9 WT plants grown at the same stages and harvested at mid-September.

In other experiments, at least three technical repeats for each plant sample were conducted for statistical analysis. The comparison of treatment means were analyzed by the Tukey HSD multiple comparison procedure using JMP software version 7 (SAS Inc., Cary NC).

Microarray Analysis

Total RNA of *AtLOV1* transgenic and WT control plants was extracted from the joint region between the leaf and leaf sheath, including a 2 cm long leaf base and a 2 cm long leaf sheath close to the leaf collar, using TRIzol Reagent (Invitrogen) and purified by Qiagene RNeasy columns (QIAGEN, Valencia, CA). Plant materials were from newly emerged tillers (with two fully expanded leaves) of plants under well watered conditions (80% soil water content). In order to minimize the effect of the heterogeneous genetic background of switchgrass, we extracted each RNA from mixed samples of three T_1 generation segregated plants. RNA was quantified and evaluated for purity using a Nanodrop Spectrophotometer ND-100 (NanoDrop Technologies, Willington, DE) and Bioanalyzer 2100 (Agilent, Santa Clara, CA).

Three RNAs from pooled plants were used as three biological repeats in the microarray. For each sample, 500 ng of total RNA was used for the expression analysis using a custom designed switchgrass cDNA chip Pvi_cDNAa520831 (Affymetrix, Santa Clara, CA). Probe labeling, chip hybridization and scanning were performed according to the manufacturer's instructions from the IVT Express Kit (Affymetrix). Data normalization between chips was conducted using RMA (Robust Multichip Average) (Irizarry *et al.*, 2003). Gene selections based on Associative T-test (Dozmorov & Centola, 2003) were made using Matlab software (MathWorks, Natick, MA). In this method, the background noise presented between replicates and technical noise during microarray experiments were measured by the residual presented among a group of genes whose residuals are homoscedastic. Genes whose residuals between the compared sample pairs that were

significantly higher than the measured background noise level were considered to be differentially expressed. A selection threshold of 2 for transcript ratios and a Bonferroni-corrected P-value threshold of 4.07977E-07 were used. The Bonferroni-corrected P-value threshold was derived from 0.05/N in these analyses (Abdi, 2007) , where N is the number of probes sets (122,556) on the chip. False discovery rate of all significant genes was monitored with Q-Values obtained by EDGE software (Storey & Tibshirani, 2003; Leek *et al.*, 2005).

For hierarchical clustering, the z-scores for expression values of the selected genes were calculated and imported into MeV open software (http://www.tm4.org/mev/). Pearson correlation was used to calculate the similarity in expression patterns between genes.

A total of 125 microarray probes detected genes with altered expressions. The probe sequences were used as Query for BLAST searching against the EST (Expressed Sequence Tag) databases in both NCBI (National Center for Biotechnology Information) and in the Samuel Roberts Noble Foundation (Ardmore, OK). Probes detecting different parts of the same gene showed similar expression level changes. A few probes detected genes of virus origins, which are possibly caused by viral contamination in plants. After eliminating the repeated probes and probes against virus sequences, a total of 104 switchgrass genes was shown to have altered expression levels. The longest EST or probe sequences were used to perform a BLAST search against the maize, rice and *Arabidopsis* genome databases to manually annotate the putative functions of the switchgrass genes.

Real-time RT-PCR

Microarray data were validated with real-time RT-PCR (qRT-PCR) on 10 selected genes with two biological repeats. The RNA samples were collected from pooled T_1 transgenic and wild type control plants as described above in the microarray analysis section. The qRT-PCR was performed with ABsolute™ Blue QPCR SYBR® Green plus ROX mix kit (Thermo Fisher Scientific, Inc., Waltham, MA) using the ABI 7500 Real-Time PCR System. The PCR reaction was set in a 25 µl reaction volume following the manufacturer's instructions. Each sample had three replicates, and the data was

normalized against the expression of the reference gene, switchgrass *ACTIN2*. Dissociation curves showed that the primers used for qRT-PCR were gene-specific. Primers used in qRT-PCR are listed in supplementary Table S2.

RESULTS

Overexpression of *AtLOV1* in switchgrass causes an erect leaf phenotype

The full length *AtLOV1* gene containing its original introns was transformed into switchgrass cv. 'Alamo' driven by the maize *Ubiquitin* promoter. After confirming the integration of T-DNA in the plant genome by PCR (data not shown), we grew six transgenic lines in a greenhouse for further analysis. Most AtLOV1 transgenic lines have smaller leaf angles than that of wild type (WT) plants (Fig. 1). The small leaf angle makes plants more "erect". The leaves of transgenic plants also appeared greener than those of WT plants.

AtLOV1 can be alternatively spliced into two transcripts in Arabidopsis (supplementary Fig. S1). In the ectopic switchgrass genetic background, *AtLOV1* is still faithfully spliced into two transcripts (Fig. 1 c). The erect leaf phenotype of transgenic plants was correlated with the expression of *AtLOV1*, but not only with integration of the transgene (Fig. 1 b-c). For simplification, AtLOV1 transgenic plants mentioned elsewhere refer to transgenic plants with the *AtLOV1* gene expressed.

T_1 generation plants were also obtained by crossing two independent T_0 generation AtLOV1 transgenic plants with WT plants. The phenotype of T_1 plants cosegregated with the transgene, further confirming that *AtLOV1* caused the erect leaf phenotype (data now shown).

Alterations of the cell morphology in the leaf collar region resulted in erect leaves

The erect leaves of transgenic plants were investigated by examining the leaf collar (or lamina joint) region. As shown in Fig. 2, in comparison with WT plants, the transgenic plants have increased edge-collar height but decreased center-collar height. Cell shape

and size in the leaf collar region were altered in the transgenic plants: epidermal cells close to the collar center of transgenic plants (Fig. 2 g) are more stretched than those of WT plants (Fig. 2 h); and the reverse is true with epidermal cells close to the collar edges (Fig. 2 i-k). Our observation suggests the alteration of cell shape, size and organization in the leaf collar region changed the Abaxial: Adaxial collar height ratio, and resulted in a smaller leaf angles in transgenic switchgrass plants.

Altered lignin content and monolignol composition of AtLOV1 transgenic plants

It is known that higher lignin content in leaf and stem tissue could increase the mechanical strength of plant, and therefore could contribute to the erect leaf phenotype. To investigate whether the transgenic plants also have higher lignin content, we monitored the native lignin deposition in the stem of transgenic plants using phloroglucinol staining, which detects coniferaylaldehyde structures of lignin (Dence, 1992). The stem of transgenic plants was stained in a deeper color than that of WT plants (supplementary Fig. S2), suggesting that the transgenic plants have higher lignin contents. We also quantitatively measured lignin content and monolignol compositions of independent transgenic lines: the transgenic plants have about 6-10% higher total lignin content than that of WT plants (table 2); the transgenic plants also have altered monolignol compositions that decreased ratios between guaiacyl and syringyl (G: S) were detected in the transgenic plants (table 2).

AtLOV1 transgenic switchgrass plants have delayed flowering time and uncompromised overall biomass yield

Nine independent T_1 generation transgenic and WT plants were evaluated for their growth performance. The first emergence of flowering of the transgenic plants was about four to six days later than WT controls. When harvested in mid-September, the transgenic plants had a lower percentage of flowering tillers, and there are more leaves of flowering tillers of the transgenic plants (table 3), showing that overexpression of *AtLOV1* delayed switchgrass flowering time.

The transgenic and WT plants were harvested at mid-September after ~four months growth in the greenhouse. Compared to WT plants, the transgenic plants have comparable tiller numbers, similar tiller stem width, but slightly reduced plant height (table 1). Individual transgenic plants have slightly less aboveground plant biomass, but similar belowground and whole plant biomass (table 1).

Global gene expression analysis of AtLOV1 transgenic switchgrass

To understand molecular mechanisms contributing to the unique phenotype of the transgenic plants, we performed global gene expression profiling using the switchgrass microarray chip. The reliability of microarray assay was validated by qRT-PCR on selected genes (Fig. 3). Overexpression *AtLOV1* induced altered expressions of a total of 104 switchgrass genes, among which 55 genes were upregulated and 49 genes downregulated with more than two-fold expression changes (Fig. 3). Among the 104 genes, 37 genes encode unknown proteins, 12 are grass (Poaceae family) unique genes. Through BLAST and phylogenetic analysis, we manually annotated the putative biological functions of some switchgrass genes according to their rice or *Arabidopsis* orthologs (supplementary table S3). Through the global gene-expression analysis, we did not identify any known abaxial-adaxial cell fate-determining genes (e.g., *KANADI* family genes or BR-synthesis and signaling pathway genes). However, genes potentially related to cell wall biosynthesis (e.g., KanlowCTG40909_s_at, microarray probe IDs), cell fate determination (e.g., AP13ITG40821_s_at), and flowering (e.g., AP13CTG21204_at) were identified. The expression of ten genes was validated by qRT-PCR analysis (Fig. 3). Based on the preliminary gene expression data, we proposed a simplified model to explain the phenotypes caused by the overexpression of *AtLOV1* in switchgrass (Fig. 4). Potential functional roles of some genes are further discussed in the Discussion section.

DISCUSSION

AtLOV1 produces two alternatively spliced isoforms in both *Arabidopsis* and switchgrass

AtLOV1 gene is alternatively spliced generating two isoforms of transcript (ANAC034/035, or LOV1.1/1.2) in *Arabidopsis* and switchgrass, and both transcripts localized in the nucleus (supplementary Fig. S3). It was confirmed that the phenotypes in *Arabidopsis* were mainly due to the expression of the longer transcript isoforms (*AtLOV1.2*) (Yoo *et al.*, 2007). In our study, we used the genomic DNA of the *AtLOV1* gene that could produce both isoforms of the transgenic switchgrass (Fig. 1 b). We also generated switchgrass lines only expressing the longer transcript isoform (*AtLOV1.2*), which conferred exactly the same phenotype (data not shown), which is consistent with the result in Arabidopsis (Yoo *et al*., 2007). However, it is unclear what the roles/functions of the shorter transcript are in either *Arabidopsis* or switchgrass.

According to Ooka *et al*., (2003), NAC proteins can be divided into two major groups with 18 subgroups based on their sequence similarity. NAC proteins within the subgroup often share similar biological functions (Ooka *et al.*, 2003). AtLOV1 (ANAC034/035) belongs to the subgroup of ONAC022, and its closely related rice orthologs are ONAC075 and ONAC049 (Ooka *et al.*, 2003). However, none of these genes within the subgroup of *ONAC022* were functionally characterized except *AtLOV1*. According to the expression data deposited in the NCBI EST database, *ONAC075* (Os01g66490) is exclusively expressed in the panicle, and *ONAC049* (Os08g0113500) is mainly expressed in flower, leaf, and stem tissues. It will be interesting to characterize the functional roles of ONAC075 (Os01g66490) and ONAC049 (Os08g0113500) to see if they have similar functions in common with *AtLOV1*.

NAC TF genes have a conserved N-terminal NAC domain but a diversified C-terminal transcriptional activation region (TAR). Since the genome sequence of switchgrass is not yet available, we BLAST searched the current switchgrass EST database with the full length, NAC, or TAR sequences of *AtLOV1*. However, no sequence with significant similarity was identified in the switchgrass EST database.

The ectopic expression of *AtLOV1* confers unique phenotypes in switchgrass

NAC proteins comprise one of the largest families of plant-specific transcription factors in plants (Ooka *et al.*, 2003). Many NAC transcription factors were confirmed to have roles in regulation of plant development and stress perception (Jensen *et al.*, 2010). However, *AtLOV1* is the only reported NAC domain family gene involved in the regulation of cell morphology in the leaf collar region in an ectopic genetic background. *AtLOV1* also induced pleiotropic changes in switchgrass, such as increased lignin content and delayed flowering time. Such phenotypes were controlled by different genes in previous reports (Torii *et al.*, 1996; Masle *et al.*, 2005; Zhang *et al.*, 2009). However, it is interesting that the ectopic overexpression of *AtLOV1* alone causes all of these distinct phenotypes in switchgrass.

Small leaf angles of AtLOV1 transgenic switchgrass

Rice mutants with reduced brassinosteroid (BR) content or reduced BR-sensitivity are often severely dwarf, but have erect and dark green leaves (Yamamuro *et al.*, 2000; Hong *et al.*, 2002; Sakamoto *et al.*, 2005; Morinaka *et al.*, 2006; Wang *et al.*, 2008). Interestingly, the AtLOV1 transgenic switchgrass also have erect and dark green leaves, and their plant heights are slightly shorter than those of WT controls. However, global gene expression analysis did not identify genes that are significantly homologous to any known genes involved in BR-related pathway (supplementary table S3), suggesting that either an unidentified genetic pathway is contributable to the erect leaf phenotype or *AtLOV1* directly targets downstream genes controlling the leaf angle.

Leaf curling also impacts the leaf angles, which is often directly resulted by changes of leaf abaxial: adaxial cell fates (Zhang *et al.*, 2009). In *Arabidopsis*, the density of abaxial epidermal cells is regulated by a group of genes within the *YABBY* and *KANADI* families (Emery *et al.*, 2003; Eshed *et al.*, 2004; Eckardt, 2010), while the adaxial cell identity is controlled by genes that belong to the HD-ZIP III family (e.g., *PHB, PHV*) (Emery *et al.*, 2003). Maize and rice *KANADI* family genes also function in defining the abaxial cell identity (Candela *et al.*, 2008; Zhang *et al.*, 2009). However, the leaves of AtLOV1 transgenic switchgrass lines are not severely curved (data not shown), suggesting that leaf curling is not a major factor contributing to the small leaf angle in *AtLOV1* transgenic

switchgrass plants. And no genes homologous to those involved in abaxial: adaxial cell fates was identified in the microarray analysis.

We have observed a wide range of natural variation of leaf angles in our collected switchgrass germplasm (data not shown). Switchgrass leaf angle is also developmentally regulated, where the leaves are more erect in young plants emerging in the late spring than those developed later in tall tillers in the summer season (data not shown). Therefore, we hypothesize that the key gene(s) regulating leaf angles are under natural selection depending on their original inhabits and are developmentally regulated. The ectopic overexpression of the *Arabidopsis* gene *AtLOV1* by using a strong constitutive promoter, maize polyubiquitin-1 (Ubi-1) promoter, may constitutively regulate the leaf angle gene(s) across all switchgrass developmental stages in different leaf positions, which results in smaller leaf angles and a erect leaf phenotype.

AtLOV1 confers different phenotypes in switchgrass and *Arabidopsis*

In Arabidopsis, the overexpression of *AtLOV1* resulted in delayed flowering time under long-day photoperiod and increased cold tolerance by negatively regulating the expression of *FT* and *CONSTANS* (*CO*) and positively regulating the expression of C-repeat-binding factor/dehydration responsive element-binding factor 1(*CBF/DREB1*) gene family (Yoo *et al.*, 2007). Interestingly, *AtLOV1* transgenic switchgrass plants also have slightly delayed flowering time, but an obviously increased cold resistance was not observed by freezing cold-acclimatized plants under 0 ^{o}C, -7 ^{o}C or -14 ^{o}C (data not shown). However, we cannot rule out the possibility that *AtLOV1* transgenic plants may have altered cold tolerance under certain low temperatures in fields. It is also possible that switchgrass is insensitive to cold hardness in general, and therefore the cold-resistant effect caused by the overexpression of *AtLOV1* is not easy to be determined under our testing conditions.

AtLOV1 may regulate the expression of different genes in different plant species if those genes have the conserved cis-elements in their promoter regions. Overexpression of the *Arabidopsis SHN1* gene in rice also triggered unexpected but highly desirable traits

(Ambavaram *et al.*, 2011). In Arabidopsis, *SHN1* regulates cuticular wax content on leaf surfaces (Aharoni *et al.*, 2004); however, *SHN1* overexpression rice plants have higher cellulose but less lignin content in their cell walls (Ambavaram *et al.*, 2011).

It would be interesting to find the switchgrass genes which are directly regulated by the TF protein, AtLOV1. However, switchgrass genome sequence is not yet available that full length genes and their corresponding promoters cannot be easily analyzed. The microarray analysis provided valuable preliminary data for the further study. As all known NAC domain TFs bind certain conserved cis-elements (Ooka *et al.*, 2003), it is possible to identify these cis-elements in promoters of the differentially expressed switchgrass genes once the whole genome information is available.

Global gene expression analysis identified candidate genes that may control phenotypes of the AtLOV1 transgenic switchgrass plants

We performed switchgrass microarray for global gene expression analysis of the transgenic switchgrass plants. Because the phenotype was most dramatic in the leaf and leaf collar region, RNA was extracted from the leaf base, leaf collar and leaf sheath region for the microarray analysis. Our results showed that about 104 switchgrass genes have significantly altered expression levels (>2 fold change) in response to the expression of *AtLOV1*. Putative biological functions of these genes were manually annotated based on their homologues in *Arabidopsis* or rice (supplementary table S3). Candidate genes which might contribute to the phenotype of transgenic plants are discussed as follows. Further characterization of these *AtLOV1* responsive genes in switchgrass may help unravel genetic pathways for small leaf angle, cell wall biosynthesis and flowering time in switchgrass.

In the transgenic plants, 16 differentially expressed genes are putatively related to plant development. One upregulated gene (probe ID: AP13ITG40821_s_at; corresponding EST No. HO250279.1) is an ortholog of human *Lateral Signaling Target Protein*, which may also serve as a negative regulator of epidermal growth factor receptor signaling in plants as well as in animals (Mosesson *et al.*, 2009). Therefore, it may regulate the

developmental process of leaf collar cells as well. Another downregulated gene (probe ID: AP13ITG61387_at; no corresponding EST deposited at GeneBank yet) is the rice and Arabidopsis orthologous *NAC1*, which may be involved in auxin signaling and plant height (Xie *et al.*, 2000).

One downregulated gene (microarray probe ID: AP13CTG21204_at; corresponding EST GeneBank No. HO298851.1) in *AtLOV1* switchgrass plants is the floral inducer *FT* gene ortholog of rice and *Arabidopsis* (Tamaki *et al.*, 2007), which may contribute to the delayed flowering time phenotype in *AtLOV1* transgenic plants. As reported previously, different switchgrass cultivars could produce similar final leaf numbers under the same growth conditions (Van Esbroeck *et al.*, 1997). We observed different switchgrass cultivars (both upland and lowland) flowering at the time when they only have 2-3 leaves in the greenhouse during the winter time (temperature set at 22-28 ℃, under natural light) at Blacksburg, Virginia, USA. These observations suggest that switchgrass flowering time is regulated by photoperiod and autonomous pathways. Similar to maize, switchgrass is an outcrossing plant species and flowers in response to short-day photoperiods (Lawrence & Walbot, 2007). Therefore, switchgrass may share similar flowering regulatory pathways with maize. It would be interesting to test whether the switchgrass *FT* homolog identified in our microarray analysis has similar functions as the maize *FT gene* (Meng *et al*., 2011).

In *AtLOV1* transgenic switchgrass plants, we identified nine differentially expressed genes that are related to cell wall biosynthesis, among which five genes are upregulated and four are downregulated. One upregulated gene (probe ID: AP13ITG62738RC_at; EST No. FL891887.1) has a significant homology to the *GP1* which involves in plant cell wall extension and biosynthesis (Keskiaho *et al.*, 2007). Another upregulated gene (probe ID: AP13ITG76341_at; EST No. FL759004.1) that may encode proteins has conserved domains of cellulose synthase, though it was not clustered with other dicot cellulose synthase (data not shown). It is possible that this gene encodes a grass-specific cellulose synthase and upregulation of the gene (EST No. FL759004.1) may alter the cell wall content in LOV1 transgenic plants.

Nine proteins that are presumably localized in the chloroplast, among which seven were upregulated and two were downregulated. One downregulated gene (probe ID: KanlowCTG40909_s_at; EST No. HO321666.1) is homologous to *AT5G47860* whose protein localizes in chloroplast and is exclusively expressed in guard cells of *Arabidopsis* (Schelbert *et al.*, 2009).

SUMMARY

Multiple elite traits of switchgrass for bioenergy production were obtained by overexpressing a single NAC domain gene, *AtLOV1*. Some phenotypic changes were not observed in transgenic *Arabidopsis* plants. The tetraploid genetic background of switchgrass may have rich genetic components required for the pleiotropic phenotype change. This study further implies the potential of applying translational genomics as a new strategy to assist the molecular breeding of switchgrass.

ACKNOWLEDGEMENTS

We thank Dr. Richard Veilleux and Ms. Kerri Mills for critical reading of the manuscript. This work was supported by grants from the Biodesign and Bioprocessing Research Center of the College of Agriculture and Life Sciences at Virginia Tech, the Virginia Agricultural Experimental Station (VA135872).

Table 1. Length of leaf collar region of AtLOV1 transgenic plants compared to wild type plants.

T_1 segregated plants	Length of collar edge (mm)	Length of collar center (mm)	Ratio between lengths of collar edge and collar center
AtLOV1	2.1 (±0.1) **	2.0 (±0.1) **	1.04 (±0.01) **
Wild type	3.2(±0.2)	1.3 (±0.3)	2.65 (±0.24)

[a] SE is in parenthesis.** Statistically significant differences ($p<0.01$) between wild type and transgenic plants.

Table 2. Lignin content and monolignol composition of independent T_0 AtLOV1 transgenic plants compared to wild type plants. Three wild type plants were pooled for the analysis.

Switchgrass T_0 Lines	Total lignin mg/100mg	Monolignol composition (%)			G:S
		Hydroxyphenyl (H)	Guaiacyl (G)	Syringyl (S)	
Wild type	20.6 (±0.2)[a]**	1.2 (±1.0) **	60.4 (±3.4) **	38.4 (±2.4)	0.64 (±0.01) **
AtLOV1-2	21.8 (±0.3)	3.9 (±0.6)	51.7 (±4.5)	44.4 (±4.0)	0.86 (±0.01)
AtLOV1-16	21.8 (±0.1)	n/a	n/a	n/a	n/a
AtLOV1-130	22.7 (±0.2)	5.2 (±1.2)	51.8 (±1.3)	43.0 (1.7±)	0.83 (±0.01)

[a] SE is in parenthesis.** Statistically significant differences ($p<0.01$) between wild type and transgenic plants.

Table 3. AtLOV1 transgenic plants have delayed flowering time compared to wild type plants.

T_1 plants	Number of plant	Fully flowered tillers (%)	Leaf number of flowered tillers
AtLOV1	9	42.7 (±5.1)[a]	6.31 (±0.07) **
Wild type	9	48.6 (±3.6)	5.68 (±0.13)

[a] SE is in parenthesis. ** Statistically significant differences ($p<0.01$) between wild type and transgenic plants.

Table 4. AtLOV1 transgenic plants have uncompromised whole plant biomass yield but decreased aboveground plant biomass yield.

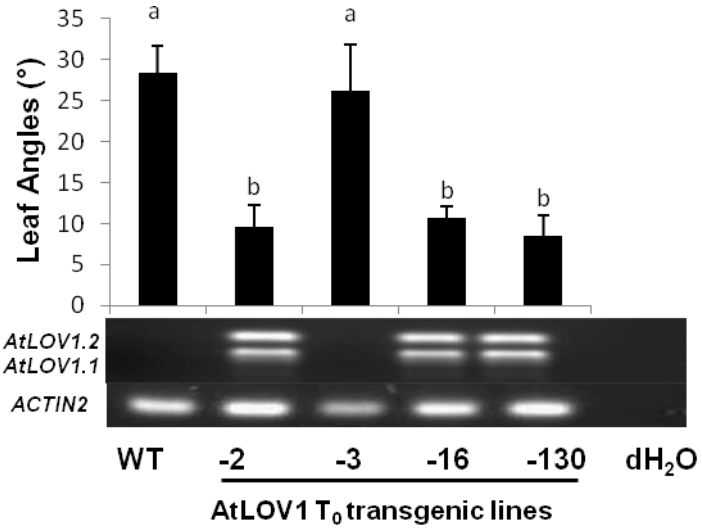
T_1 Plants	Number of plants	Tiller number	Plant height (cm)	Stem width (mm)	Biomass yields (g)		
					AP biomass	BP biomass	WP biomass
WT	9	38.1 (±2.4) [a]	146.8 (±1.0) **	4.95 (±0.07)	140.9 (±3.6) **	98.5 (±7.8)	239.4 (±10.8)
AtLOV1	9	37.0 (±2.1)	138.0 (±1.2)	4.82 (±0.07)	111.6 (±3.8)	104.4 (±6.1)	216 (±8.2)

AP: Aboveground plant; BP (belowground plant); and WP (whole plant).
[a] SE is in parenthesis. ** Statistically significant differences ($p<0.01$) between wild type and transgenic plants.

a


AtLOV1
WT
AtLOV1
WT

b


Leaf Angles (°)
35
30
25
20
15
10
5
0
a
b
a
b
b
AtLOV1.2
AtLOV1.1
ACTIN2
WT
-2
-3
-16
-130
dH2O
AtLOV1 T0 transgenic lines

c

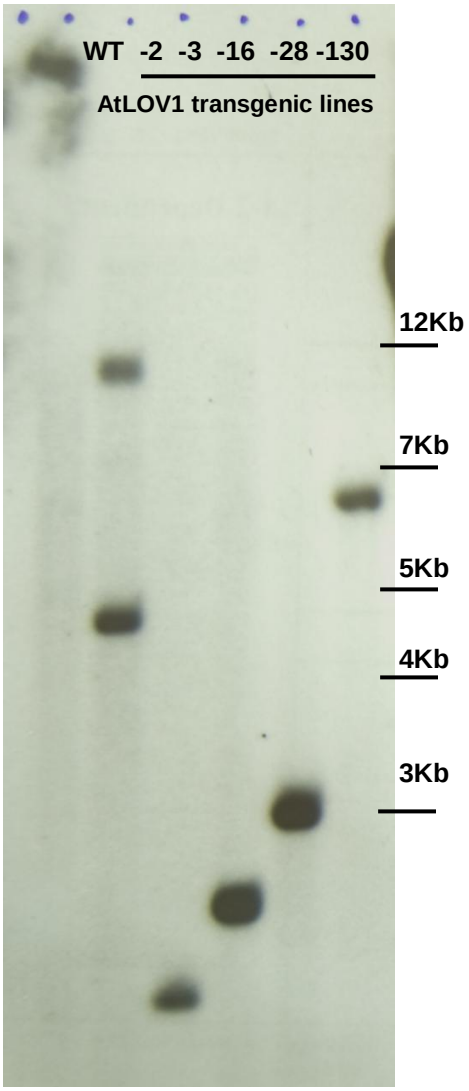


Fig. 1 Overexpression of *AtLOV1* caused erect leaf phenotype in switchgrass. (**a**) Transgenic plants have smaller leaf angles compared to wild type (WT) plants; (**b**) The leaf angles of four independent T_0 transgenic lines are correlated with the expression of *AtLOV1* in plants as detected by RT-PCR (noting that *AtLOV1* pre-mRNA can be alternatively spliced in switchgrass yielding two transcripts *AtLOV1.1* and *AtLOV1.2*); (**c**) Southern blot analysis with a *hph* probe showed the T-DNA integrations in four T_0 transgenic line. Different letters above bars indicate the statistically significant difference (α=0.01).

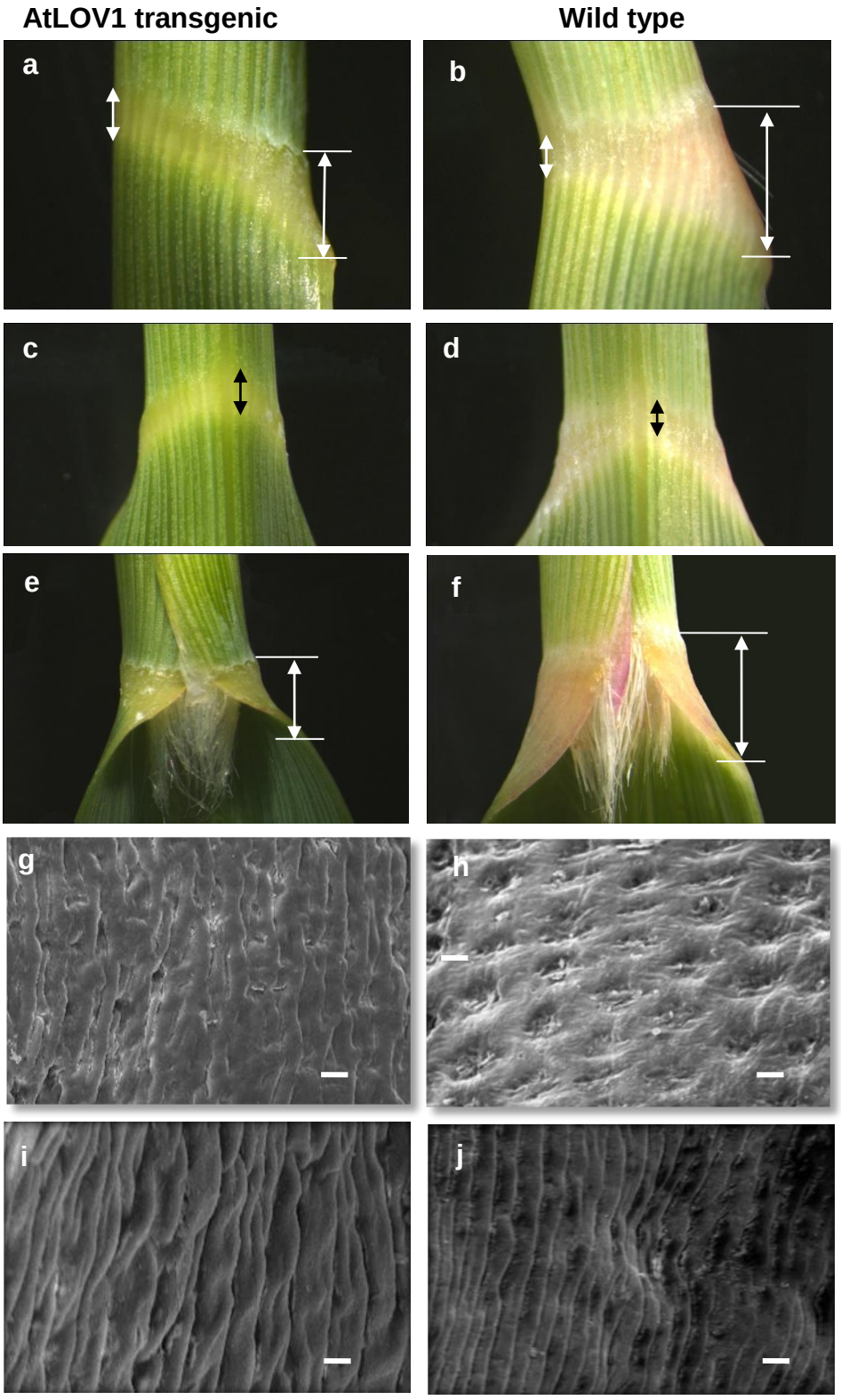
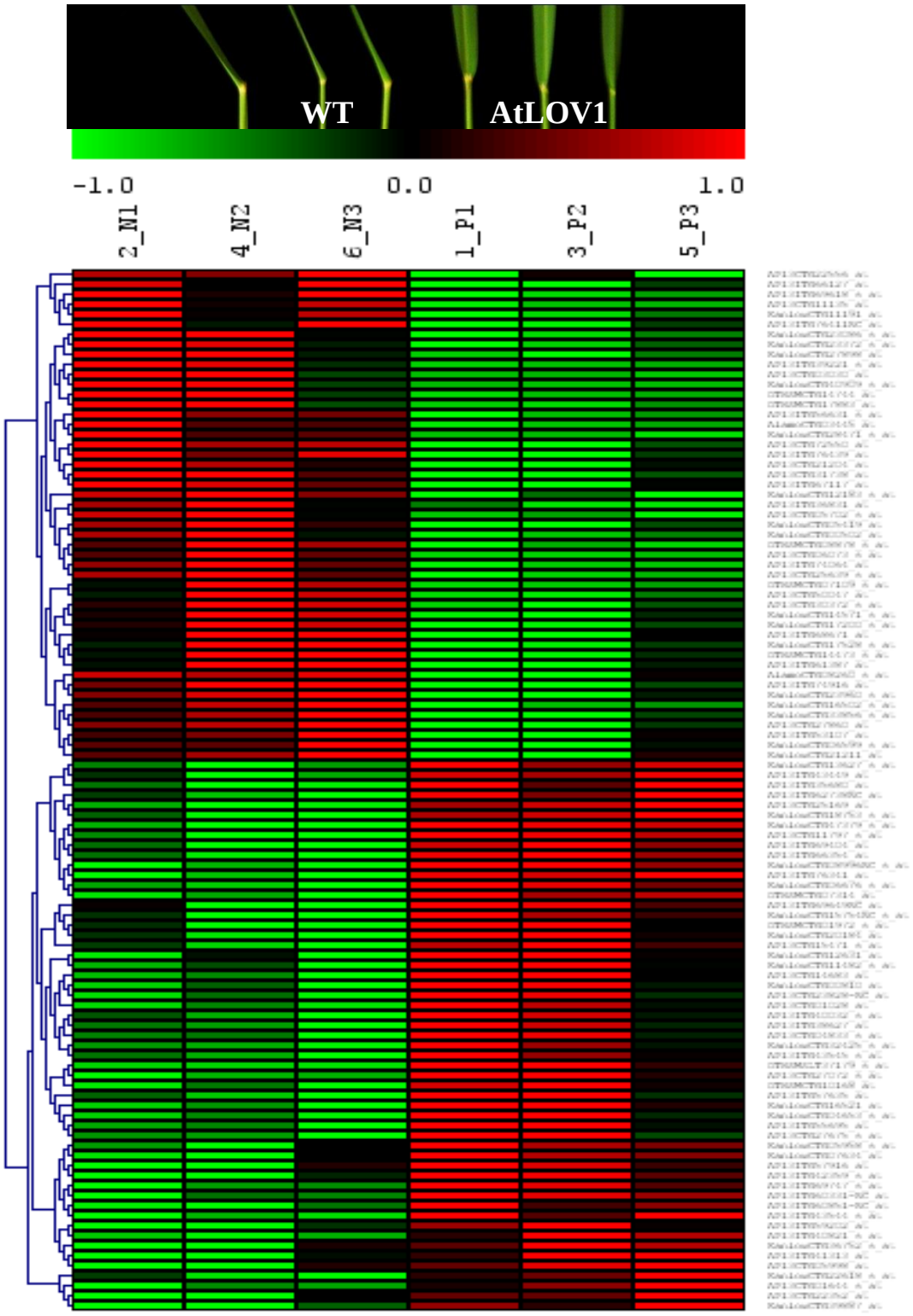


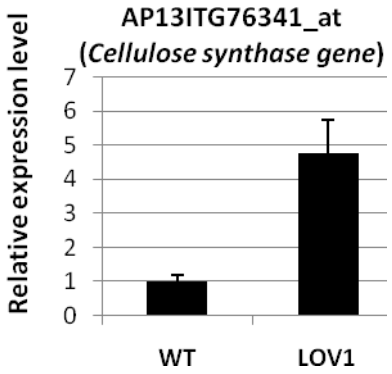
Fig. 2 Plant morphology at the leaf collar region in AtLOV1 transgenic and wild type (WT) control plants. Compared to WT plants, AtLOV1 transgenic plants have decreased

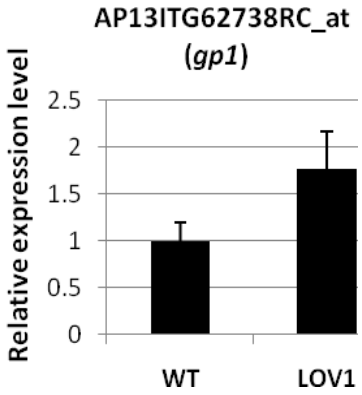
length of the collar center (c, d) and increased length of the collar edge (e, f). Epidermal cells under scanning electron microscope are observed (g-j). Epidermal cells in the collar center region are more stretched in the transgenic plants (g) than in WT plants (h); and epidermal cells in the collar edge region are less stretched in the transgenic plants (i) than in WT plants (j). Bars represent 10 μm.

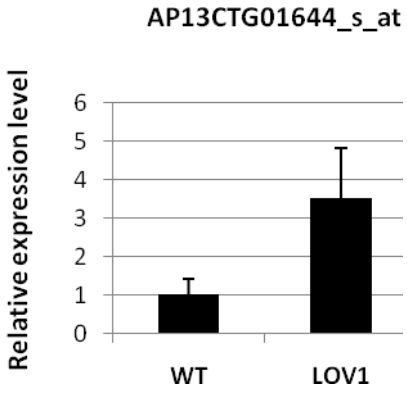
a

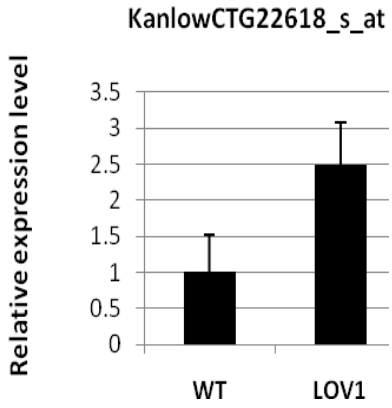


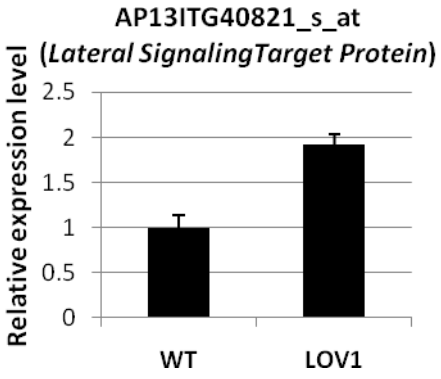
b

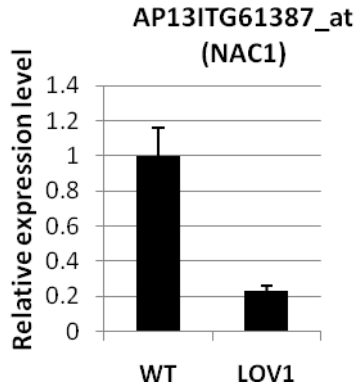

AP13ITG76341_at
(Cellulose synthase gene)
Relative expression level
7
6
5
4
3
2
1
0
WT
LOV1

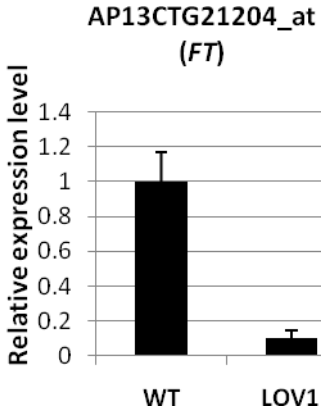

AP13ITG62738RC_at
(gp1)
Relative expression level
2.5
2
1.5
1
0.5
0
WT
LOV1

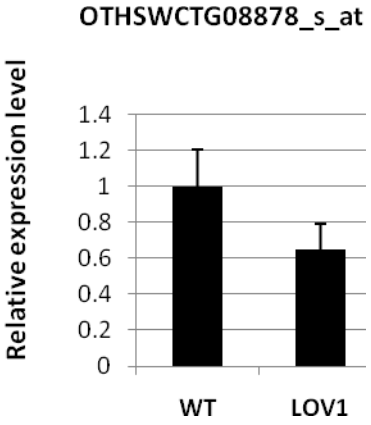

AP13CTG01644_s_at
Relative expression level
6
5
4
3
2
1
0
WT
LOV1

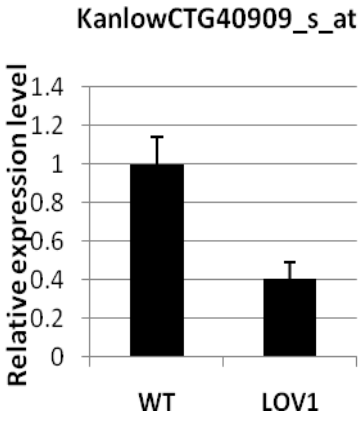

KanlowCTG22618_s_at
Relative expression level
3.5
3
2.5
2
1.5
1
0.5
0
WT
LOV1

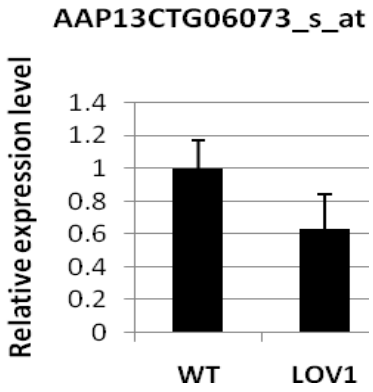

AP13ITG40821_s_at
(Lateral SignalingTarget Protein)
Relative expression level
2.5
2
1.5
1
0.5
0
WT
LOV1


AP13ITG61387_at
(NAC1)
Relative expression level
1.4
1.2
1
0.8
0.6
0.4
0.2
0
WT
LOV1


AP13CTG21204_at
(*FT*)
Relative expression level
1.4
1.2
1
0.8
0.6
0.4
0.2
0
WT
LOV1


OTHSWCTG08878_s_at
Relative expression level
1.4
1.2
1
0.8
0.6
0.4
0.2
0
WT
LOV1


KanlowCTG40909_s_at
Relative expression level
1.4
1.2
1
0.8
0.6
0.4
0.2
0
WT
LOV1


AAP13CTG06073_s_at
Relative expression level
1.4
1.2
1
0.8
0.6
0.4
0.2
0
WT
LOV1

c

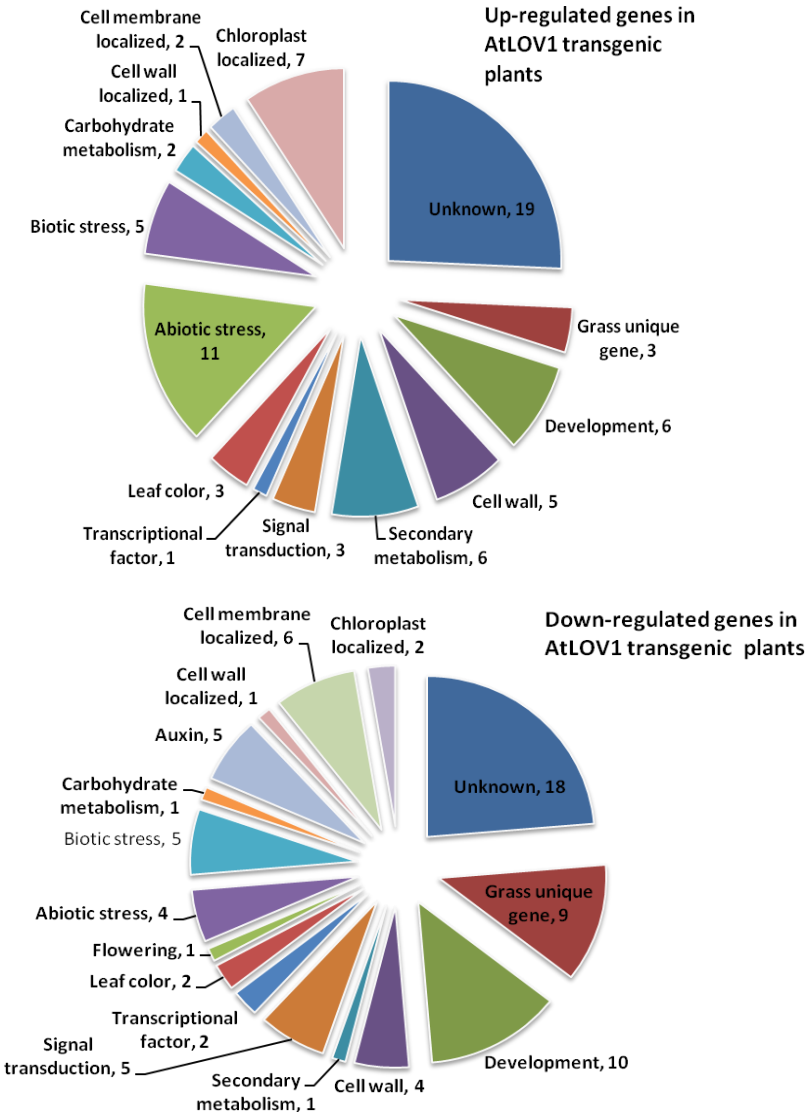


Fig. 3 Microarray analyses on differentially expressed genes in AtLOV1 transgenic plants. Heat map of 104 differentially expressed genes showing a total of 55 genes is upregulated and 49 genes downregulated above two folds in the transgenic plants (a). Expression levels of ten selected genes ware validated with Realtime-PCR, the detected genes were indicated as microarray probe ID with corresponding orthologous gene's name in parenthesis when it is available (b). The differentially expressed genes are annotated and classified in different categories (c). Three biological repeats of microarray and two biological repeats of QPCR were conducted.

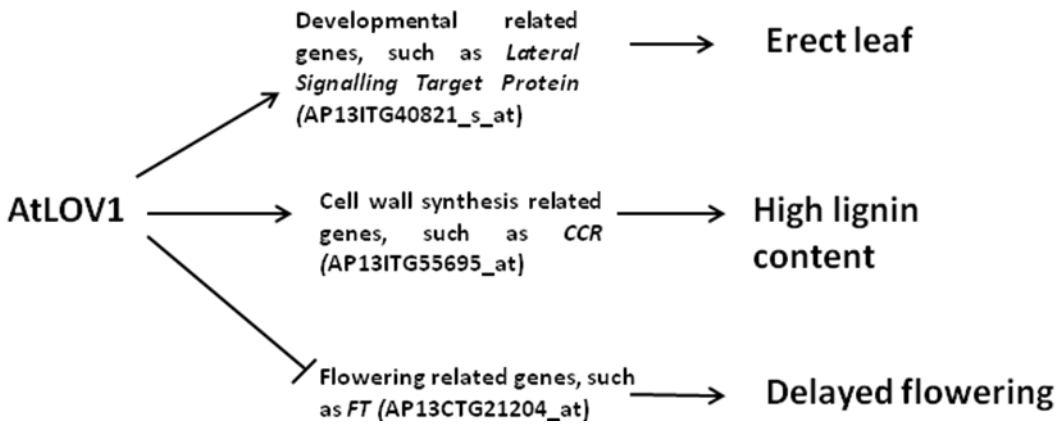


Fig. 4 Ectopic overexpression of *AtLOV1* in switchgrass induced distinct phenotypes by regulating gene expression in different signaling pathways.

References

Abdi H 2007. Bonferroni and Sidak correction for multiple comparisons. In: Sage ed. *Encyclopedia of Measurement and Statistics*. Salkind, NJ: Thousand Oaks,CA, pp103-106.

Aharoni A, Dixit S, Jetter R, Thoenes E, Van Arkel G, Pereira A. 2004. The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in Arabidopsis. *The Plant Cell* **16**: 2463-2480.

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.

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Rosas MO, Rocheford TR, Romay MC, Romero S, Salvo S, Villeda HS, Sofia da Silva H, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen MD. 2009. The genetic architecture of maize flowering time. *Science* **325**: 714-718.

Candela H, Johnston R, Gerhold A, Foster T, Hake S. 2008. The milkweed *pod1* gene encodes a KANADI protein that is required for abaxial/adaxial patterning in maize leaves. *The Plant Cell* **20**: 2073-2087.

Century K, Reuber TL, Ratcliffe OJ. 2008. Regulating the regulators: the future prospects for transcription-factor-based agricultural biotechnology products. *Plant physiology* **147**: 20-29.

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

Christensen AH, Quail PH. 1996. Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic research* **5**: 213-218.

Demura T, Ye ZH. 2010. Regulation of plant biomass production. *Current opinion in plant biology* **13**: 298-303.

Dence C 1992. The determination of lignin.In: Methods in lignin chemistry. Springer-Verlag: Heidelberg. 33-61.

Dozmorov I, Centola M. 2003. An associative analysis of gene expression array data. *Bioinformatics* **19**: 204-211.

Eckardt NA. 2010. *YABBY* genes and the development and origin of seed plant leaves. *The Plant Cell* **22**: 2103.

Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL. 2003. Radial patterning of Arabidopsis shoots by *class III HD-ZIP* and *KANADI* genes. *Current Biology* **13**: 1768-1774.

Eshed Y, Izhaki A, Baum SF, Floyd SK, Bowman JL. 2004. Asymmetric leaf development and blade expansion in Arabidopsis are mediated by KANADI and YABBY activities. *Development* **131**: 2997-3006.

Hackbusch J, Richter K, Müller J, Salamini F, Uhrig JF. 2005. A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 4908-4912.

Hong Z, Ueguchi-Tanaka M, Shimizu-Sato S, Inukai Y, Fujioka S, Shimada Y, Takatsuto S, Agetsuma M, Yoshida S, Watanabe Y. 2002. Loss-of-function of a rice brassinosteroid biosynthetic enzyme, C-6 oxidase, prevents the organized arrangement and polar elongation of cells in the leaves and stem. *The Plant Journal* **32**: 495-508.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**: 249-264.

Jedlicka P, Mortin MA, Wu C. 1997. Multiple functions of *Drosophila* heat shock transcription factor *in vivo*. *The EMBO Journal* **16**: 2452-2462.

Jensen MK, Kjaersgaard T, Nielsen MM, Galberg P, Petersen K, O'Shea C, Skriver K. 2010. The *Arabidopsis thaliana* NAC transcription factor family: structure-function relationships and determinants of *ANAC019* stress signalling. *Biochemical Journal* **426**: 183-196.

Jensen MK, Kjaersgaard T, Petersen K, Skriver K. 2010. NAC genes: Time-specific regulators of hormonal signaling in Arabidopsis. *Plant signaling & behavior* **5**: 907-910.

Keskiaho K, Hieta R, Sormunen R, Myllyharju J. 2007. *Chlamydomonas reinhardtii* has multiple prolyl 4-hydroxylases, one of which is essential for proper cell wall assembly. *The Plant Cell* **19**: 256-269.

Lawrence CJ, Walbot V. 2007. Translational genomics for bioenergy production from fuelstock grasses: maize as the model species. *The Plant Cell* **19**: 2091-2904.

Leek JT, Monsen E, Dabney AR, Storey JD. 2005. EDGE: extraction and analysis of differential gene expression. *Bioinformatics* **22**: 507-508.

Lemus R, Brummer EC, Moore KJ, Molstad NE, Burras CL, Barker MF. 2002. Biomass yield and quality of 20 switchgrass populations in southern Iowa, USA. *Biomass and Bioenergy* **23**: 433-442.

Li R, Bruneau A, Qu R. 2006. Improved plant regeneration and *in vitro* somatic embryogenesis of St Augustinegrass [*Stenotaphrum secundatum* (Walt.) Kuntze]. *Plant breeding* **125**: 52-56.

Li R, Qu R. 2011. High throughput *Agrobacterium*-mediated switchgrass transformation. *Biomass and Bioenergy* **35**: 1046-1054.

Madakadze I, Stewart K, Peterson P, Coulman B, Samson R, Smith D. 1998. Light interception, use-efficiency and energy yield of switchgrass (*Panicum virgatum* L.) grown in a short season area. *Biomass and Bioenergy* **15**: 475-482.

Masle J, Gilmore SR, Farquhar GD. 2005. The *ERECTA* gene regulates plant transpiration efficiency in Arabidopsis. *Nature* **436**: 866-870.

Matsubara K, Yamanouchi U, Wang ZX, Minobe Y, Izawa T, Yano M. 2008. *Ehd2*, a rice ortholog of the maize *INDETERMINATE1* gene, promotes flowering by up-regulating *Ehd1*. *Plant physiology* **148**: 1425-1435.

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.

Meng X, Muszynski MG, Danilevskaya ON. 2011. The *FT*-like *ZCN8* gene functions as a floral activator and is involved in photoperiod sensitivity in maize. *The Plant Cell* **23**: 942-960.

Morinaka Y, Sakamoto T, Inukai Y, Agetsuma M, Kitano H, Ashikari M, Matsuoka M. 2006. Morphological alteration caused by brassinosteroid insensitivity increases the biomass and grain production of rice. *Plant Physiology* **141**: 924-931.

Mosesson Y, Chetrit D, Schley L, Berghoff J, Ziv T, Carvalho S, Milanezi F, Admon A, Schmitt F, Ehrlich M. 2009. Monoubiquitinylation regulates endosomal localization of Lst2, a negative regulator of EGF receptor signaling. *Developmental cell* **16**: 687-698.

Ooka H, Satoh K, Doi K, Nagata T, Otomo Y, Murakami K, Matsubara K, Osato N, Kawai J, Carninci P. 2003. Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA research* **10**: 239-247.

Robinson AR, Mansfield SD. 2009. Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near infrared reflectance based prediction modeling. *The Plant Journal* **58**: 706-714.

Sakamoto T, Morinaka Y, Ohnishi T, Sunohara H, Fujioka S, Ueguchi-Tanaka M, Mizutani M, Sakata K, Takatsuto S, Yoshida S. 2005. Erect leaves caused by brassinosteroid deficiency increase biomass production and grain yield in rice. *Nature Biotechnology* **24**: 105-109.

Schelbert S, Aubry S, Burla B, Agne B, Kessler F, Krupinska K, H rtensteiner S. 2009. Pheophytin pheophorbide hydrolase (pheophytinase) is involved in chlorophyll breakdown during leaf senescence in Arabidopsis. *The Plant Cell* **21**: 767-785.

Shi ZY, Wang J, Wan XS, Shen GZ, Wang XQ, Zhang JL. 2007. Over-expression of rice *OsAGO7* gene induces upward curling of the leaf blade that enhanced erect-leaf habit. *Planta* **226**: 99-108.

Sinclair TR, Sheehy JE. 1999. Erect leaves and photosynthesis in rice. *Science* **283**: 1455.

Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D. 2004. Determination of structural carbohydrates and lignin in biomass. *NREL, Golden, Co.*

Storey JD, Tibshirani R. 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 9440-9445.

Takeda S, Matsuoka M. 2008. Genetic approaches to crop improvement: Responding to environmental and population changes. *Nature Reviews Genetics* **9**: 444-457.

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science* **316**: 1033-1036.

Torii KU, Mitsukawa N, Oosumi T, Matsuura Y, Yokoyama R, Whittier RF, Komeda Y. 1996. The Arabidopsis *ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *The Plant Cell* **8**: 735-746.

Van Esbroeck G, Hussey M, Sanderson M. 1997. Leaf appearance rate and final leaf number of switchgrass cultivars. *Crop Science* **37**: 864-870.

Vogel KP, Mitchell RB. 2008. Heterosis in switchgrass: biomass yield in swards. *Crop Sci.* **48**: 2159-2164.

Wang L, Xu Y, Zhang C, Ma Q, Joo S-H, Kim S-K, Xu Z, Chong K. 2008. OsLIC, a novel CCCH-type zinc finger protein with transcription activation, mediates rice architecture via brassinosteroids signaling. *PLoS ONE* **3**: e3521.

Xie Q, Frugis G, Colgan D, Chua NH. 2000. Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes & Development* **14**: 3024-3036.

Xu B, Huang L, Shen Z, Welbaum G, Zhang X, Zhao B. 2011a. Selection and characterization of a new switchgrass (*Panicum virgatum* L.) line with high somatic embryogenic capacity for genetic transformation**.** *Scientia Horticulturae* In press.

Xu B, Escamilla-Treviño L, Noppadon S, Shen Z, Shen H, Zhang P, Dixon P, Zhao B. 2011b. Silencing of *4-Coumarate: Coenzyme A ligase* in Switchgrass Leads to Reduced Lignin Content and Improved Fermentable Sugar Yields for Biofuel Production. *New Phytologist* In press.

Yamamuro C, Ihara Y, Wu X, Noguchi T, Fujioka S, Takatsuto S, Ashikari M, Kitano H, Matsuoka M. 2000. Loss of function of a rice brassinosteroid insensitive1 homolog prevents internode elongation and bending of the lamina joint. *The Plant Cell* **12**: 1591-1606.

Yang X, Hwa C. 2008. Genetic modification of plant architecture and variety improvement in rice. *Heredity* **101**: 396-404.

Yoo SY, Kim Y, Kim SY, Lee JS, Ahn JH. 2007. Control of flowering time and cold response by a NAC-domain protein in Arabidopsis. *PLoS ONE* **2**: e642.

Zhang GH, Xu Q, Zhu XD, Qian Q, Xue HW. 2009. SHALLOT-LIKE1 is a KANADI transcription factor that modulates rice leaf rolling by regulating leaf abaxial cell development. *The Plant Cell* **21**: 719-735.

Supplementary Materials and Methods

Cellular Localization of AtLOV1.1 and AtLOV1.2

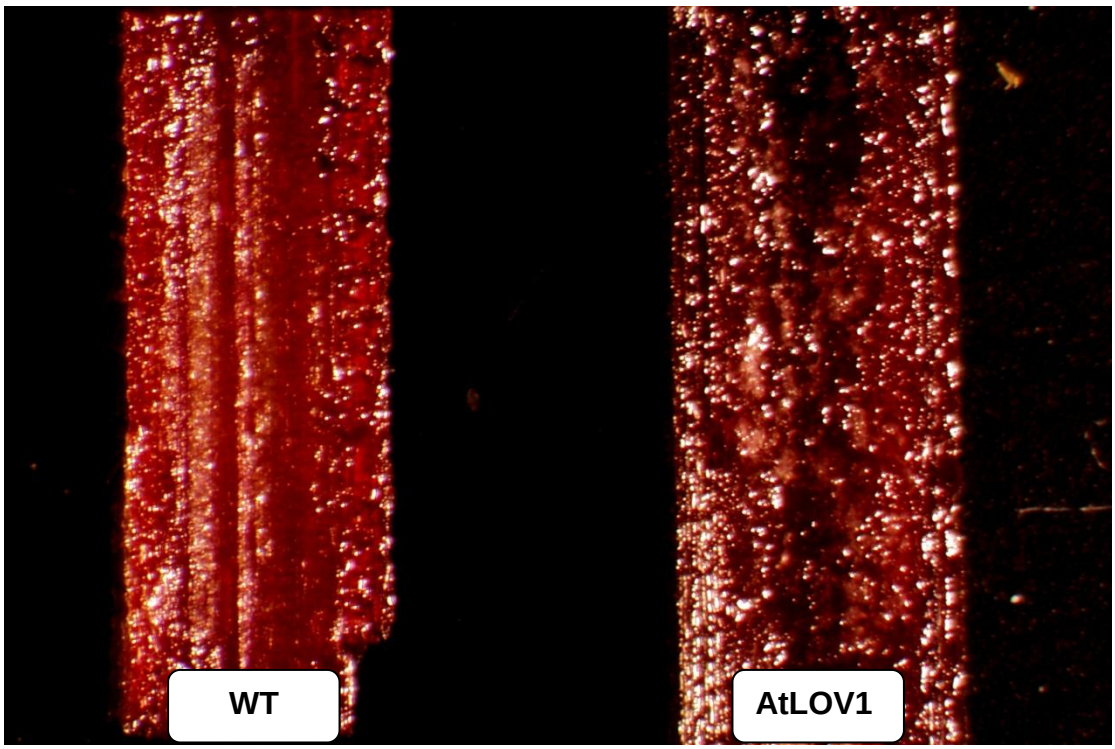
The two transcripts of *AtLOV1*, *AtLOV1.1* and *AtLOV1.2*, were amplified from cDNA of Arabidopsis ecotype Columbia-0, cloned into pENTR/D vector and sequenced. The genes were subcloned into pEarlyGate101 by LR reaction (Invitrogen, Carlsbad, CA). The targeted genes cloned in pEarlyGate101 is driven by 35S promoter and fused with a C-terminal GFP gene (Earley *et al.*, 2006). The resultant binary vectors were electro-transformed into *Agrobacterium* strain GV3101. A transient expression assay was conducted to observe the subcellular localization of AtLOV1.1:GFP, AtLOV1.2:GFP and the control GFP:GFP by injecting the *Agrobacterium* carrying corresponding vectors into *Nicotiana benthamiana* according to a protocol described by Hackbusch *et al.* (2005). The GFP signal was observed under a SZXZ-RFL3 fluorescence microscope (Olympus America, Melville, NY, USA).

Histology and Microscopy

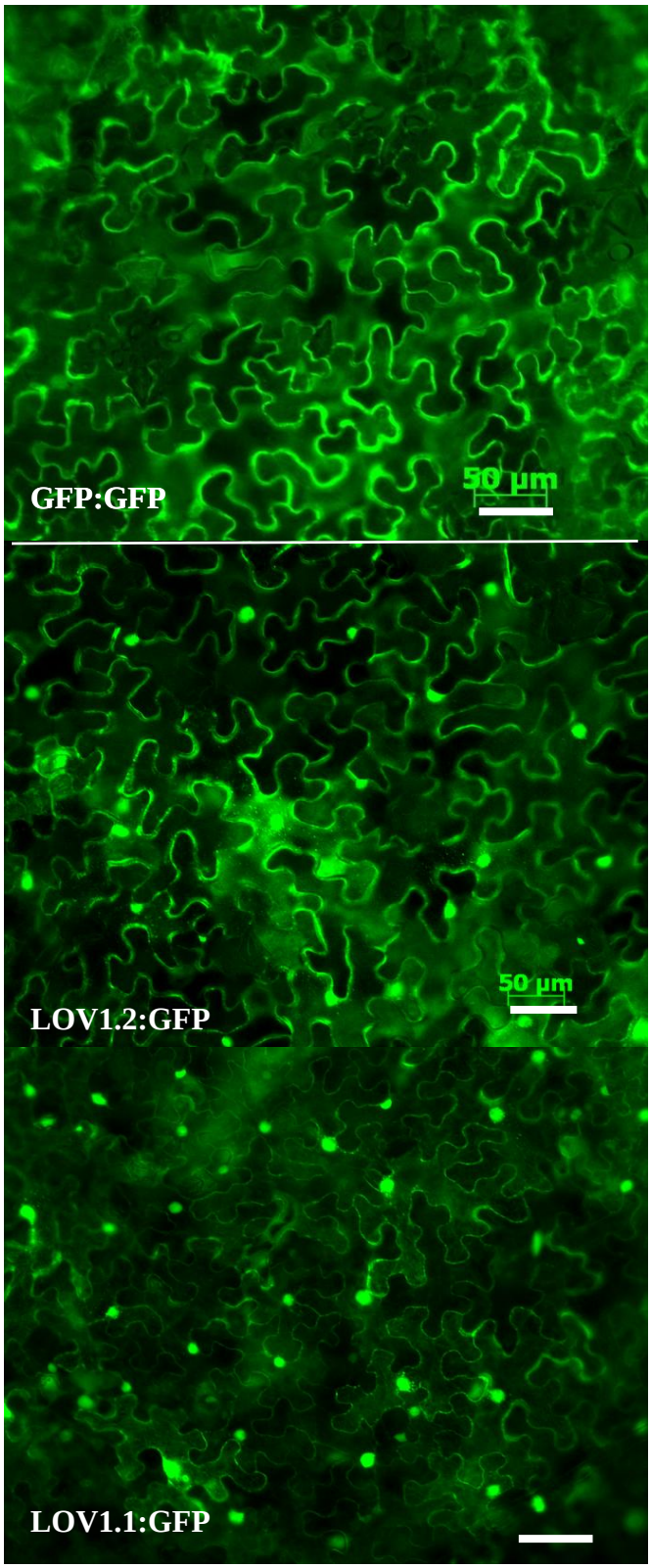
The internodes of the AtLOV1 transgenic and wild type control plants were stained with phloroglucinol reagents (Pomar *et al.*, 2002) to analyze the lignin deposition patterns and visualized under an Olympus SZXZ-RFL3 fluorescence microscope (Olympus America, Melville, NY, USA).

Supplementary Data

Supplementary Fig. S1 *AtLOV1* has two transcripts due to alternative splicing. *AtLOV1.1* has four introns and exons and *AtLOV1.2* has three introns and exons. The filled boxes represent exons and lines represent the introns. The numbers indicate the nucleotide position.



Supplementary fig. S2 Phloroglucinol staining of lignin deposition patterns in AtLOV1 transgenic and wild type (WT) plants. Darker red color in AtLOV1 suggests it has higher lignin content than the WT control.



Supplementary fig. S3. The subcellular localization of AtLOV1.1 and AtLOV1.2 fused with C-terminal GFP proteins indicate that both of them localized into plant nucleus. a.

GFP:GFP control, b. AtLOV1.1:GFP, c. AtLOV1.2:GFP. The GFP signal was observed under fluorescence microscope. The green foci represent the plant nuclei. The bars represent 50 μm.

Supplementary Table S1. Primers used for PCR and RT-PCR.

Primer Sets	Primer Sequences (from 5' to 3')
LOV1_SalRev	GTCGACCTTAAAAGGAATATTAGTATAG
LOV1_BamHFor	CACCGGATCCATGGCAATTGTATCCTCCACAAC
HYGRO_FOR	CAAACTGTGATGGACGACACCG
HYGRO_REV	TATATGCTCAACACATGAGCG
LOV1.2_FOR	CAAGACGATGATGAAACTGCC
AD_lov_276aa_RIFOR	ACGAATTCTCTGACGTTACCATTGCTCTAG
AD_lov_414aa_BamREV	ACGTCGACCTTAAAAGGAATATTAGTATAG
AD_lov_313aa_RIFOR	ACGAATTC ACTGCCATTGTTGACGATCTTC
AD_325aa_BamREV	ACGGATCCGTAGTTAACTAGTCTTTGAAG

Supplementary Table S2. Primers used in qRT-PCR for some microarray data validation

probe set id	primers NAME	primer sequence
AP13CTG21204_at	21204_F	ACATCGCAGCCACAGGTTTG
AP13CTG21204_at	21204_R	CGGTGTTTGTACCGCAAGTG
OTHSWCTG08878_s_at	08878_F	GGCATCCATGCACTCTTCCT
OTHSWCTG08878_s_at	08878_R	CGACTCCTCATCCTCAGCATCTTAC
AP13CTG06073_s_at	06073_F	GAAGTGCATCTGCTTCAATCTGA
AP13CTG06073_s_at	06073_R	GTGGAAGTGATGACAGGATTGATC
AP13ITG61387_at	61387_F	TCGCGTATTCGCTTGTATAGTGA
AP13ITG61387_at	61387_R	ACGGCAGAGATCACACTAACCA
KanlowCTG40909_s_at	40909_F	ACCGCGGAGGTGAACTTCTG
KanlowCTG40909_s_at	40909_R	ATGCCCGAGTCAGTACCCAAAG
AP13CTG01644_s_at	01644_F	TCGCCACCTTCACATAACCA
AP13CTG01644_s_at	01644_R	GGGAGCACATCCTTGGATGA
AP13ITG62738RC_at	62738_F	CCTTCGGCATAATCGAAATCAAGA
AP13ITG62738RC_at	62738_R	CGGCTCGACAATCAGACATC
AP13ITG40821_s_at	40821_F	GCTGCAATGACGTTTACAGACA
AP13ITG40821_s_at	40821_R	CAGACGCAAACACTCACAAATCG
AP13ITG76341_at	76341_F	GCTGTGGTTACTGATGAATGGAAT
AP13ITG76341_at	76341_R	GGTCGATGGGACTGGTAGAAG
KanlowCTG22618_s_at	22618_F	ATCCGCCCGCGGAACCAT
KanlowCTG22618_s_at	22618_R	GGCCGACCGATTCTTGATGTAC

Supplementary Table S3. Differentially expressed genes in AtLOV1 transgenic switchgrass.
||||, no information available; NA, cannot be identified.

Probeset_id	Ratio of gene expression (Transgenic: WT)	Rice homolog	Arabidopsis homolog	Description
OTHSWCTG07109_s_at	0.18	NA	AT3G23637	AT3G23637 is DVL family protein involves in plant development.) Over-expression of some of this family proteins (eg.DVL1), results in pleiotropic phenotypes featured by shortened stature, rounder rosette leaves, clustered inflorescences, shortened pedicles, and siliques with pronged tips; DVL genes have overlapping function in the DVL gene family.
AP13CTG50047_at	0.27	NA	AT1G77131	AT1G77131, is a pseudogene of PGSIP, glycogenin glucosyltransferase; and OsGGT relates to stress response
KanlowCTG05419_at	0.33	Os02g32469	NA	uncharacterized; PRK12678,transcription termination factor Rho;
KanlowCTG17528_s_at	0.33	Os05g28740	NA	pfam00582, Universal stress protein family. The universal stress protein UspA is a small cytoplasmic bacterial protein whose expression is enhanced when the cell is exposed to stress agents.
KanlowCTG00502_at	0.34	Os02g01590	AT1G62660	Glycosyl hydrolases family 32 protein; FUNCTIONS IN: hydrolase activity, hydrolyzing O-glycosyl compounds; INVOLVED IN: sucrose catabolic process, using beta-fructofuranosidase, carbohydrate metabolic process; LOCATED IN: vacuole; smart00640,Glycosyl hydrolases family 32.,3E-164||AT1G62660,Galactose metabolism; Starch and sucrose metabolism; Metabolic pathways; Galactose metabolism; Starch and sucrose metabolism; Metabolic pathways
AP13ITG74916_at	0.35	Os03g19470	AT5G06700	AT5G06700 encodes a member of the TBL (TRICHOME BIREFRINGENCE-LIKE) gene family containing a plant-specific DUF231 (domain of unknown function) domain. TBL gene family has 46 members, two of which (TBR/AT5G06700 and TBL3/AT5G01360) have been shown to be involved in the synthesis and deposition of secondary wall cellulose, presumably by influencing the esterification state of pectic polymers.
AP13ITG56631_s_at	0.37	Os08g04890	AT5G44460	AT5G44460 encodes calmodulin like protein; functions in calcium ion binding; involved in cadmium ion toxicity
AP13CTG30372_s_at	0.38	Os11g34920	AT3G07040	AT3G07040 is RPS3 (RPM1), an NB-LRR gene, confering disease resistance to Pst against avrB and avrRpm1
AP13CTG21204_at	0.38	Os06g35940	AT1G65480	The FT homolog; FT, together with LFY, promotes flowering and is antagonistic with its homologous gene. late flowering phenotype under long-day conditions (40.6 leaves versus 15.0 leaves for wild type); late flowering; strong influence of short days on flowering time; small but significant effect of vernalization on flowering time; narrow leaves, increased number of cauline leaves; flowering occurs about 13 days later than wild type.; PLN00169,CETS family protein; Provisional,7E-77|FAMILY NOT NAMED,4.3E-75|
AP13ITG36831_at	0.39	NA	NA	||||

KanlowCTG23086_s_at	0.40	Os11g12740	AT1G33440	has transporter activity; At-homolog has oligopeptide transport function, and respond to nematode--stress?; pfam00854,POT family. The POT (proton-dependent oligopeptide transport) family all appear to be proton dependent transporters.,0.000000003|SUBFAMILY NOT NAMED
KanlowCTG16502_s_at	0.42	Os11g26790	NA	possibly a grass-specific dehydrin; pfam00257,Dehydrin
KanlowCTG12183_s_at	0.42		AT1G20450	AT1G20450 involves in cold acclimation, regulation of seed germination, response to abscisic acid stimulus, response to cold, response to stress, response to water, response to water deprivation; colocalizes with actin cytoskeleton; functions in actin binding
OTHSWCTG08878_s_at	0.42	Os05g38680	AT3G55240	Over-expression of AT3G55240 leads to PEL (Pseudo-Etiolation in Light) phenotype.; pfam09713,Plant protein 1589 of unknown function (A_thal_3526).
KanlowCTG33856_s_at	0.42	Os03g61150	AT1G19530	AT1G19530, involves in N-terminal protein myristoylation, anaerobic F16respiration; expressed in leaf apex, inflorescence meristerm, hypocotyl, root, flower; PHA03307,transcriptional regulator ICP4
AP13CTG72550_at	0.42	Os03g28300	AT1G79570	At-homologs involved in pollen germination. pfam07714,Protein tyrosine kinase.,7E-62|SERINE/THREONINE PROTEIN KINASE,8.1E-184| |
AP13ITG53107_at	0.43	Os03g03510	AT1G01140	cd05580,Catalytic domain of the Protein Serine/Threonine Kinase, STKs catalyze the transfer of the gamma-phosphoryl group from ATP to potassium ion cellular response to potassium ion starvation, regulation of potassium ion transport, response to response to water deprivation, somatal movement, potassium ion starvation, cold, mannitol stimulus, salt and wounding; serine/threonine kinase activity; involved in signal transduction pathway; locates at cytoplasm, cytosol, nucleus, plasma membrane;serine/threonine residues on protein substrates. PKA is present ubiquitously in cells and interacts with many different downstream targets. It plays a role in the regulation of diverse processes such as growth, development, memory, metabolism, gene expression, immunity, and lipolysis
AP13ITG74064_at	0.43	Os09g35810	AT1G01225	unknown function; |NC DOMAIN-CONTAINING PROTEIN
AP13ITG39221_s_at	0.44	Os01g36580	AT5G64700	protein auxin-induced protein 5NG4, putative; response to auxin stimulus; At-homolog is close to Walls are thin 1 (WAT1), an arabidopsis homolog of Medicago truncatula NODULIN21, is a tonoplast-localized protein required for secondary wall formation in fibers.Analysis of wat1 mutants revealed two cell wall-related phenotypes in stems: a defect in cell elongation, resulting in a dwarfed habit and little to no secondary cell walls in fibers;
AP13CTG27860_at	0.45	Os06g48020	AT4G30170	cd00693,Horseradish peroxidase and related secretory plant peroxidases. Class III peroxidases are found in the extracellular space or in the vacuole in plants where they have been implicated in hydrogen peroxide detoxification, auxin catabolism and lignin biosynthesis, and stress response. Class III peroxidases contain four conserved disulphide bridges and two conserved calcium binding sites.,5E-104| |AT4G30170, Phenylpropanoid biosynthesis;|Os06g48020.1 Phenylpropanoid biosynthesis;

OTHSWCTG14473_s_at	0.45	NA	NA	||||
KanlowCTG14571_s_at	0.46	Os01g48444	AT2G22670	pfam02309,AUX/IAA family. at-homolog is IAA8; Transcription of the AUX/IAA family of genes is rapidly induced by the plant hormone auxin. Some members of this family are longer and contain an N terminal DNA binding domain. The function of this region is uncertain.
KanlowCTG11191_at	0.48	NA	NA	||||
AP13CTG03030_at	0.48	Os02g05660	AT1G55150	COG0513,Superfamily II DNA and RNA helicases [DNA replication, recombination, and repair / Transcription / Translation, ribosomal structure and biogenesis] |AT1G55150,Spliceosome|Os02g05660.1,Spliceosome
AP13ITG66127_at	0.48	Os03g55030	AT5G05870	|UDP-GLUCOSYLTRANSFERASE|AT5G05870,Zeatin biosynthesis|
KanlowCTG40909_s_at	0.48	Os04g33280	AT5G47860	Unknown function
AP13ITG76411RC_at	0.48	Os05g43610	AT4G21070	BRCA1 C Terminus (BRCT) domain. The BRCT domain is found predominantly in proteins involved in cell cycle checkpoint functions responsive to DNA damage. T
OTHSWCTG14744_at	0.49	NA	NA	||||
AlamoCTG08260_s_at	0.49	Os04g17100	At5g66110.1	Unknown function
AP13ITG76439_at	0.49	Os01g20830	NA	Uncharacterized protein
OTHSWCTG17883_at	0.49	Os05g40820	AT3G53020	Ribosomal protein L24e.
AP13ITG67117_at	0.50	Os06g35060	NA	|COPPER TRANSPORT PROTEIN ATOX1 (METAL TRANSPORT PROTEIN ATX1)
AP13CTG25639_s_at	0.50	Os11g41500	NA	Unknown function
KanlowCTG23960_s_at	0.50	Os07g44290	AT3G17510	smart00220,Serine/Threonine protein kinases, catalytic domain. Phosphotransferases. Serine or threonine-specific kinase subfamily.
KanlowCTG21211_at	0.50	Os01g01340	AT3G26740.1	Light regulated protein Lir1. Lir1 mRNA accumulates in the light, reaching maximum and minimum steady-state levels at the end of the light and dark period, respectively. Plants germinated in the dark have very low levels of lir1 mRNA, whereas plants germinated in continuous light express lir1 at an intermediate but constant level. It is thought that lir1 expression is controlled by light and a circadian clock. The exact function of this family is unclear
KanlowCTG13627_s_at	2.01	Os03g56270	AT5G65700	leucine-rich repeat receptor-like protein kinase;

KanlowCTG05958_s_at	2.01	Os07g28480	AT1G17180	GST_C family, Class Tau subfamily; The plant-specific class Tau GST subfamily has undergone extensive gene duplication. The Arabidopsis and Oryza genomes contain 28 and 40 Tau GSTs, respectively. They are primarily responsible for herbicide detoxification together with class Phi GSTs, showing class specificity in substrate preference. Tau enzymes are highly efficient in detoxifying diphenylether and aryloxyphenoxypropionate herbicides. In addition, Tau GSTs play important roles in intracellular signalling, biosynthesis of anthocyanin, responses to soil stresses and responses to auxin and cytokinin hormones.
AP13ITG69649RC_at	2.01	Os06g30640	AT2G45570	flavonoid 3'-monooxygenase,
AP13ITG60851-RC_at	2.02	Os10g38140	AT1G10370	GST_C family, Class Tau subfamily; The plant-specific class Tau GST subfamily has undergone extensive gene duplication. The Arabidopsis and Oryza genomes contain 28 and 40 Tau GSTs, respectively. They are primarily responsible for herbicide detoxification together with class Phi GSTs, showing class specificity in substrate preference. Tau enzymes are highly efficient in detoxifying diphenylether and aryloxyphenoxypropionate herbicides. In addition, Tau GSTs play important roles in intracellular signalling, biosynthesis of anthocyanin, responses to soil stresses and responses to auxin and cytokinin hormones.
AP13ITG40032_s_at	2.03	Os03g09270	AT3G49600	Unknown function
KanlowCTG15754RC_s_at	2.04	Os03g52010	AT1G27480	Lecithin:cholesterol acyltransferase. Lecithin:cholesterol acyltransferase (LACT) is involved in extracellular metabolism of plasma lipoproteins, including cholesterol.
AP13ITG57916_at	2.04	NA	NA	||||
OTHSWCTG01972_s_at	2.05	Os01g29330	AT4G24310	Protein of unknown function (DUF679). This family contains several uncharacterized plant proteins.
AP13ITG38627_at	2.05	Os01g50200	AT3G16520	PLN00164,glucosyltransferase; Provisional,9E-119|UDP-GLUCOSYLTRANSFERASE, |AT3G50740,Phenylpropanoid biosynthesis
KanlowCTG06676_s_at	2.05	Os11g03240	AT1G33110	putative efflux protein, MATE family. The MATE family consists of probable efflux proteins including a functionally characterized multi drug efflux system from Vibrio parahaemolyticus, a putative ethionine resistance protein of Saccharomyces cerevisiae, and the functionally uncharacterized DNA damage-inducible protein F (DinF) of E. coli.
AP13CTG14683_at	2.05	Os04g44530	AT1G73390	BRO1-like domain. This domain is found in a number proteins including Rhophilin and BRO1.
KanlowCTG00810_at	2.06	Os07g26630	AT3G53420	Major intrinsic protein. MIP (Major Intrinsic Protein) family proteins exhibit essentially two distinct types of channel properties: (1) specific water transport by the aquaporins, and (2) small neutral solutes transport, such as glycerol by the glycerol facilitators.
KanlowCTG16521_at	2.09	Os06g08830	AT2G36750	UDP-glucosyltransferase family protein |AT2G36750, Zeatin biosynthesis

KanlowCTG04653_s_at	2.10	Os03g08330	AT1G19180	tify domain. This short possible domain is found in a variety of plant transcription factors that contain GATA domains as well as other motifs. Although previously known as the Zim domain this is now called the tify domain after its most conserved amino acids. TIFY proteins can be further classified into two groups depending on the presence (group I) or absence (group II) of a C2C2-GATA domain. Functional annotation of these proteins is still poor, but several screens revealed a link between TIFY proteins of group II and jasmonic acid-related stress response.
AP13CTG01644_s_at	2.11	Os01g60700	AT3G08760	Catalytic domain of Protein Tyrosine Kinases. Protein Tyrosine Kinase (PTK) family, catalytic domain.
AP13CTG25169_at	2.11	Os05g02310	AT3G53620	PLN02373,inorganic pyrophosphatase, | |AT3G53620,Oxidative phosphorylation|Os05g02310.1|SORBI_09g001530,Oxidative phosphorylation
KanlowCTG07634_at	2.11	Os12g31000	AT1G78160	Pumilio-family RNA binding domain. These proteins function as translational repressors in early embryonic development.
AP13ITG42359_s_at	2.12	08g015360	AT5G20850	|DNA REPAIR PROTEIN RAD51 HOMOLOG 3, R51H3,8.9E-130|AT5G20850,Homologous recombination|Os01g39630.1; |SORBI_08g015360,Homologous recombination
KanlowCTG47379_s_at	2.13	Os04g33590	AT2G36290	alpha/beta hydrolase fold. This catalytic domain is found in a very wide range of enzymes.,
OTHSWSLT37179_s_at	2.13	NA	AT5G55240	| | | |
AP13ITG62738RC_at	2.14	Os07g04930	At4g38080.1	DNA polymerase III subunits gamma and tau;
AP13ITG69404_at	2.17	Os06g09240	AT5G17050	UDP-glucoronosyl and UDP-glucosyl transferase., |OS06G0192100 |AT5G17050, Flavone and flavonol biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
AP13ITG43449_at	2.17	NA	NA	| | | |
AP13CTG04833_s_at	2.18	Os08g14570	AT4G30210	NADPH cytochrome p450 reductase (CYPOR) serves as an electron donor in several oxygenase systems and is a component of nitric oxide synthases and methionine synthase reductases.
AP13ITG59202_at	2.19	Os06g29730	AT1G64450.	Unknown function
KanlowCTG08996RC_s_at	2.20	Os03g22790	At5g18670.1	beta-amylase| |AT4G17090,Starch and sucrose metabolism; Metabolic pathways|Os03g22790.1|
KanlowCTG20184_at	2.23	Os11g38810	AT1G67070	mannose-6-phosphate isomerase, |Os11g38810.1;13111|SORBI_05g023560,Fructose and mannose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
AP13ITG69747_s_at	2.26	Os06g42560	AT5G38530	tryptophan synthase subunit beta; |AT5G38530,Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites|Os06g42560.2, Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
KanlowCTG18753_s_at	2.27	Os02g49870	At4g01150.1	PLN02777,photosystem I
AP13CTG11797_s_at	2.27	Os01g60830	At3g56290.1	Unknown function
AP13CTG23828-RC_at	2.30	Os04g46440	NA	DNA-binding domain in plant proteins such as APETALA2 and EREBPs.

AP13ITG66354_at	2.32	Os12g04424	NA	Strictosidine synthase. Strictosidine synthase is a key enzyme in alkaloid biosynthesis.
KanlowCTG12631_at	2.33	Os04g20230	AT5G55940	Unknown function
AP13CTG27072_s_at	2.33	Os05g08770	NA	Unknown function
AP13ITG55695_at	2.33	Os01g45200	AT1G15950	cinnamoyl-CoA reductase,2E-109|CINNAMOYL-COA REDUCTASE,2.5E-88|AT1G15950,Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites|Os01g45200.1|SORBI_07g021680,Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
AP13ITG60331-RC_at	2.34	NA	NA	| | | |
AP13ITG35680_at	2.36	Os01g60830	At3g56290.1	Unknown function
AP13ITG40821_s_at	2.39	Os01g4273	At1g28600.1	SGNH_plant_lipase_like, a plant specific subfamily of the SGNH-family of hydrolases, a diverse family of lipases and esterases. |LATERAL SIGNALING TARGET PROTEIN
AP13ITG76341_at	2.43	Os04g35020.1	At2g32540.1	cellulose synthase, |CELLULOSE SYNTHASE-LIKE PROTEIN
AP13CTG15471_s_at	2.51	Os01g16030.4	At1g10630.1	ARF family protein; Provisional, |FAMILY NOT NAMED
OTHSWCTG10168_at	2.53	NA	NA	| | | |
AP13CTG01028_at	2.55	Os06g14324.1	NA	Caleosin related protein. This family contains plant proteins related to caleosin. Caleosins contain calcium-binding domains and have an oleosin-like association with lipid bodies. Caleosins are present at relatively low levels and are mainly bound to microsomal membrane fractions at the early stages of seed development. As the seeds mature, overall levels of caleosins increased dramatically and they were associated almost exclusively with storage lipid bodies. T
KanlowCTG32425_s_at	2.55	Os06g39960.1	NA	basic region leucin zipper. Unknown function
KanlowCTG11482_s_at	2.56	Os02g44200.1	NA	Predicted thioesterase [General function prediction only]
AP13ITG43545_s_at	2.59	Os02g10860.1	At5g11260.1	basic region leucin zipper, Unknown function
KanlowCTG36752_s_at	2.77	Os10g40030.8	NA	oxidoreductase; Unknown function
OTHSWCTG07314_at	2.78	Os06g39960.1	NA	basic region leucin zipper. Unknown function
AP13ITG57635_at	2.92	Os08g08850.1	At4g27030.1	Kua-ubiquitin conjugating enzyme hybrid localisation domain.
AP13CTG27675_s_at	3.03	Os03g04470.1	At2g36885.1	transcription termination factor Rho
KanlowCTG22618_s_at	3.11	Os02g43640.1	NA	alpha-ketoglutarate decarboxylase
AP13ITG43544_s_at	3.14	Os02g10860.1	At5g11260.1	basic region leucin zipper., basic region leucin zipper. Unknown function

AP13ITG41313_at	3.68	NA	NA	||||
AP13CTG05998_at	3.81	Os10g40020.3	NA	oxidoreductase; Provisional,1E-67|FAMILY NOT NAMED
AP13CTG22352_at	3.97	NA	NA	||||
KanlowCTG39887_at	5.01	Os05g33310.1	AT4G02940.1	AT4G02940.1 belongs to pfam03171,2OG-Fe(II) oxygenase superfamily.

SUPPLEMENTAL LITERATURE CITED

Pomar F, Merino F, Ros Barceló A. 2002. *O*-4-Linked coniferyl and sinapyl aldehydes in lignifying cell walls are the main targets of the Wiesner (phloroglucinol-HCl) reaction. *Protoplasma* **220**: 17–28.

Hackbusch J, Richter K, Muller J, Salamini F, Uhrig JF. 2005. A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 4908-4912.

Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song KM, Pikaard CS. 2006. Gateway-compatible vectors for plant functional genomics and proteomics. *Plant Journal* **45**: 616-629.