2015

Overexpression of *SbMyb60* impacts phenylpropanoid biosynthesis and alters secondary cell wall composition in *Sorghum bicolor*

Erin D. Scully
University of Nebraska-Lincoln, erin.scully@ars.usda.gov

Tammy Gries
University of Nebraska-Lincoln, tgries2@unl.edu

Gautam Sarath
University of Nebraska-Lincoln, Gautam.sarath@ars.usda.gov

Nathan A. Palmer
Grain, Forage, and Bioenergy Research Unit, USDA-ARS, Lincoln, NE, nathan.palmer@ars.usda.gov

Lisa Baird
University of San Diego, Baird@sandiego.edu

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2016) **85,** 378–395 doi: 10.1111/tpj.13112

Overexpression of *SbMyb60* impacts phenylpropanoid biosynthesis and alters secondary cell wall composition in *Sorghum bicolor*

Erin D. Scully[1,2], Tammy Gries[1], Gautam Sarath[1], Nathan A. Palmer[1], Lisa Baird[3], Michelle J. Serapiglia[4], Bruce S. Dien[5], Akwasi A. Boateng[4], Zhengxiang Ge[6], Deanna L. Funnell-Harris[1,7], Paul Twigg[8], Thomas E. Clemente[6] and Scott E. Sattler[1,2,*]

[1]*Grain, Forage, and Bioenergy Research Unit, USDA-ARS, Lincoln, NE 68583, USA,*

[2]*Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583, USA,*

[3]*Department of Biology, Shiley Center for Science and Technology, University of San Diego, San Diego, CA 92110, USA,*

[4]*Agricultural Research Service, United States Department of Agriculture (USDA-ARS), Eastern Regional Research Center, 600 East Mermaid Lane, Wyndmoor, PA 19038, USA,*

[5]*National Center for Agricultural Utilization Research, USDA-ARS, 1815 North University Street, Peoria, IL 61604, USA,*

[6]*Department of Agronomy and Horticulture, Center for Plant Science Innovation, University of Nebraska, Lincoln, NE 68588, USA,*

[7]*Department of Plant Pathology, University of Nebraska-Lincoln, Lincoln, NE 68583, USA, and*

[8]*Biology Department, University of Nebraska-Kearney, Kearney, NE 68849, USA*

Received 10 July 2015; revised 9 December 2015; accepted 14 December 2015; published online 29 December 2015.
*For correspondence (e-mail Scott.Sattler@ars.usda.gov).

SUMMARY

The phenylpropanoid biosynthetic pathway that generates lignin subunits represents a significant target for altering the abundance and composition of lignin. The global regulators of phenylpropanoid metabolism may include MYB transcription factors, whose expression levels have been correlated with changes in secondary cell wall composition and the levels of several other aromatic compounds, including anthocyanins and flavonoids. While transcription factors correlated with downregulation of the phenylpropanoid biosynthesis pathway have been identified in several grass species, few transcription factors linked to activation of this pathway have been identified in C4 grasses, some of which are being developed as dedicated bioenergy feedstocks. In this study we investigated the role of *SbMyb60* in lignin biosynthesis in sorghum (*Sorghum bicolor*), which is a drought-tolerant, high-yielding biomass crop. Ectopic expression of this transcription factor in sorghum was associated with higher expression levels of genes involved in monolignol biosynthesis, and led to higher abundances of syringyl lignin, significant compositional changes to the lignin polymer and increased lignin concentration in biomass. Moreover, transgenic plants constitutively overexpressing *SbMyb60* also displayed ectopic lignification in leaf midribs and elevated concentrations of soluble phenolic compounds in biomass. Results indicate that overexpression of *SbMyb60* is associated with activation of monolignol biosynthesis in sorghum. *SbMyb60* represents a target for modification of plant cell wall composition, with the potential to improve biomass for renewable uses.

Keywords: lignin, monocot, monolignol, C4 grass, bioenergy, transgenic.

INTRODUCTION

Sorghum bicolor (sorghum) is a versatile heat- and drought-tolerant C4 grass that is rapidly emerging as a dedicated, short-rotation biofeedstock. Although traditionally grown in the United States for grain and forage production, sorghum has several advantages that make it an attractive crop for bioenergy purposes, including lower nitrogen and water requirements than corn. Importantly, it can be grown on low-quality and marginal lands (Rooney *et al.*, 2007) and biomass yields remain fairly consistent, even when grown under a range of environmental conditions (Miller and McBee, 1993). Its well-annotated, diploid genome makes the crop amenable to genetic improvement through traditional breeding practices and biotechnological approaches.

The biopolymer lignin strengthens and fortifies the plant cell wall and provides protection against pathogens and insect herbivores; however, its abundance and composition can greatly impact biomass and forage quality (Boerjan *et al.*, 2003). The phenylpropanoid biosynthesis pathway responsible for synthesizing the subunits used to build the lignin polymer is a major target for modifying biomass composition to improve the quality of biofeedstock and forages (Sattler *et al.*, 2010). Additionally, this pathway synthesizes an array of phenolic compounds which could serve as chemical precursors for the synthesis of high-value products. In flowering plants at least 11 enzymatic reactions are required to synthesize monolignols (sinapyl alcohol, coniferyl alcohol and *p*-coumaryl alcohol) from phenylalanine (Vanholme *et al.*, 2013) (Figure 1). Monolignols are subsequently transferred to the cell wall, oxidized by laccases/peroxidases and incorporated into lignin through radical oxidative polymerization (Boerjan *et al.*, 2003). This progression results in a complex polymer primarily comprising three phenolic subunits, guaiacyl (G-lignin), syringyl (S-lignin) and *p*-hydroxyphenyl (H-lignin), covalently linked by several bonds including highly resilient β-*O*-4 aryl ether linkages (Whetten and Sederoff, 1995). In C4 grasses (e.g. sorghum and maize), *brown midrib* (*bmr*) mutants have been useful for reducing lignin content (Cherney *et al.*, 1991) and altering secondary cell wall composition (Palmer *et al.*, 2008), which has led to enhanced saccharification and fermentation efficiencies and improvements in ethanol yields from lignocellulosic biomass (Dien *et al.*, 2009) and *in vitro* digestibility (Aydin *et al.*, 1999; Oliver *et al.*, 2004). Sorghum *bmr* mutants have also been shown to increase yields in emerging biofuel fermentation systems that utilize lignin as a substrate for lipid synthesis (Xie *et al.*, 2015). Together this research illustrates the plasticity of plants to tolerate changes in lignin composition and the utility of modifying biomass composition for industrial purposes.

Transcription factors (TF) that regulate phenylpropanoid biosynthesis have also been targeted to manipulate biomass composition. In the *Arabidopsis thaliana*, R2-R3 MYB family TFs regulate phenylpropanoid and secondary cell wall biosynthesis (Stracke *et al.*, 2001). Arabidopsis activation-tagged lines generated using the constitutive enhanced cauliflower mosaic virus 35S (E35S) promoter led to overexpression of several MYB TFs, which were often associated with elevated levels of monolignol pathway intermediates, flavonoids and anthocyanins (Borevitz, 2000). The *AtMYB46* TF and its homolog *AtMYB83* were shown to regulate the entire secondary cell wall biosynthetic network, including phenylpropanoid, cellulose and hemicellulose biosynthesis (McCarthy *et al.*, 2009). Overexpression of *AtMYB83* caused deposition of lignin, cellulose and xylan in leaf epidermal cells and induced the expression of several secondary cell wall biosynthetic genes, including cellulose synthase, 4-coumarate:CoA ligase (4CL) and caffeoyl-CoA-*O*-methyltransferase (CCoAOMT) (McCarthy *et al.*, 2009). Furthermore, two specific regulators of monolignol biosynthesis were identified in Arabidopsis, designated *AtMYB58* and *AtMYB63* (Zhou *et al.*, 2009), which activate lignin biosynthesis during secondary wall formation without affecting the biosynthesis of cellulose and hemicellulose. Overexpression of *AtMYB58* and *AtMYB63* resulted in activation of most genes encoding enzymes in the monolignol pathway, except for ferulate 5-hydroxylase (F5H), and caused ectopic deposition of lignin in leaves, while transgenic downregulation of these two TFs resulted in lower expression levels of the monolignol biosynthesis pathway and reduced secondary cell wall thickness (Zhou *et al.*, 2009). MYB TFs that positively regulate lignin biosynthesis have also been identified in several other dicots, including *Populus*, *Antirrhinum* and *Eucalyptus* (Tamagnone, 1998; Goicoechea *et al.*, 2005; McCarthy *et al.*, 2010).

Although sorghum and other C4 grasses are rapidly emerging as dedicated feedstocks for the bioeconomy, no TFs linked to the activation of lignin or phenylpropanoid biosynthesis have been identified; such TFs could be useful for optimizing cell wall composition. Altering the

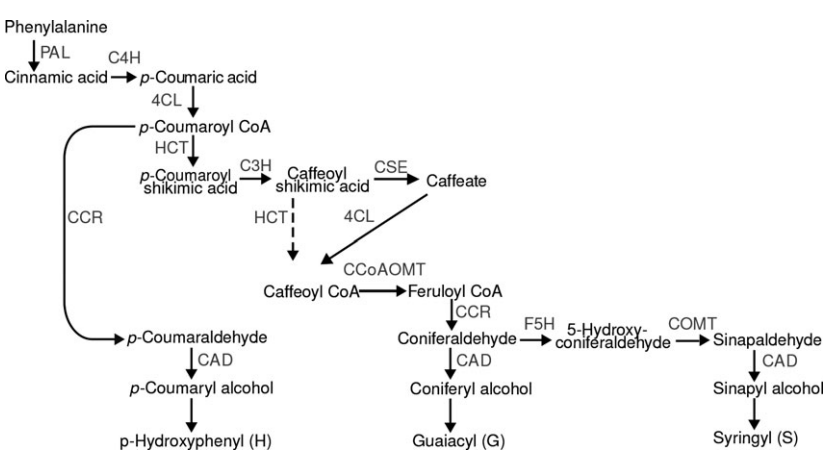

Figure 1. The monolignol biosynthetic pathway in sorghum based on consensus models from dicot and monocot plants (adapted from Vanholme *et al.*, 2013).
Enzyme abbreviations: PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; HCT, *p*-hydroxycinnamoyltransferase; C3H, 4-coumarate hydroxylase; CSE, caffeoyl shikimate esterase; CCoAOMT, caffeoyl-CoA-*O*-methyltransferase; CCR, cinnamoyl-CoA reductase; F5H, ferulate 5-hydroxylase; COMT, caffeic acid *O*-methyl transferase; CAD, cinnamyl alcohol dehydrogenase.

phenylpropanoid composition of herbaceous feedstocks and elevating lignin content may increase energy yields from thermal conversion processes and phenolic co-products from biomass would make useful precursors for green chemistry (McKendry, 2002; Yaman, 2004). Experiments reported herein demonstrate that overexpression of *SbMYB60*, a putative sorghum co-ortholog of *AtMyb58* and *AtMYB63*, is associated with higher expression levels of genes in the monolignol biosynthetic pathway. Overexpression of this TF in sorghum was associated with alterations in secondary wall composition, increased lignin content and increased abundances of several aromatic compounds. Changes in cell wall composition elevated total energy levels in sorghum biomass. *SbMyb60* represents a new target for engineering biomass for the bioeconomy in sorghum and other monocot grasses.

RESULTS

SbMyb60 is a potential co-ortholog of *AtMyb58* and *AtMyb63* in *Sorghum bicolor*

Approximately 113 myb family TFs and 80 myb-related family TFs have been previously identified and named in the sorghum genome (Yilmaz *et al.*, 2009) (http://grassius.org/grasstfdb.html). To identify myb TFs that positively regulate monolignol biosynthesis in sorghum, a maximum likelihood-based phylogenetic analysis revealed two co-orthologous genes in sorghum (Sobic.004G273800 and Sobic.006G199800) (Figure 2a and Figures S1 and S2 in the Supporting Information). Sobic.004G273800, designated as S*bMyb60* in the Grassius sorghum TF database (Yilmaz *et al.*, 2009), was selected for use in transformation experiments. The *SbMyb60* protein is 41% identical and 55% similar to *AtMYB58* and *AtMYB63* at the amino acid level.

Overexpression of *SbMyb60* in *Sorghum bicolor* (Tx430)

To evaluate the phenotype of ectopically overexpressing S*bMyb60* in sorghum, the genomic region of S*bMyb60* containing the coding region and a single intron was placed under the control of the E35S promoter with the CaMV 35S terminator to end transcription, and the cassette was subcloned into the binary vector pZP211 (Figure 2b). Sorghum plants were transformed with this construct using *Agrobacterium tumefaciens* as previously described (Howe *et al.*, 2006) (Methods S1). Three independent transformation events for sorghum overexpressing *SbMyb60* were characterized. These events were designated Myb ZG-129-4-10a, Myb ZG-129-4-15a and Myb ZG-124-1-2a. Using high-efficiency thermal asymmetric interlaced (hiTAIL) PCR, independent insertion points were confirmed for Myb ZG-129-4-10a and Myb ZG-124-1-2a; however, the insertion point for Myb ZG-129-4-15a could not be determined (Table S1, Methods S1). Overexpression of

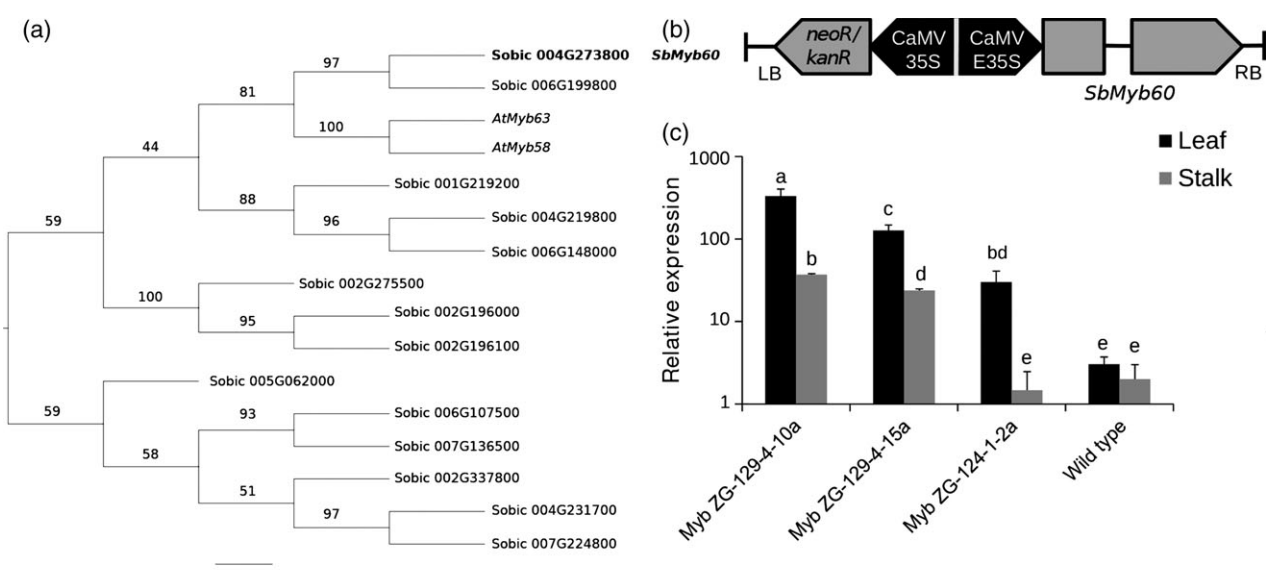

Figure 2. Overexpression of *SbMyb60 in planta*.
(a) Maximum likelihood-based phylogenetic analysis of homologs of *AtMYB58* and *AtMYB63* in sorghum. Bootstrap support values for $n = 500$ replicates are provided and the scale bar designates the number of substitutions per site.
(b) *SbMyb60* overexpression cassette used for genetic transformation of sorghum. An expression cassette containing *SbMyb60*, under the control of the E35S CAMV promoter with CaMV 35S terminator to end transcription, was genetically transformed into sorghum (Tx430).
(c) Relative expression levels of *SbMyb60* in transgenic and wild-type sorghum. Relative expression was determined via reverse transcription quantitative PCR using the ΔΔCt method and elf4A1 for normalization. Values displayed are normalized relative to ΔΔCt values from wild-type leaves or stalks. Error bars represent the standard error. Samples with different letters are statistically different at $\alpha = 0.05$ using Tukey's honestly significant difference test.

SbMyb60 in transgenic leaves and stalks, relative to the wild type, was confirmed using RT-qPCR (Figure 2c). Generally, *SbMyb60* expression levels were higher in leaves than in stalks for both wild-type and transgenic plants (Figure 2c). Myb ZG-129-4-10a had the highest levels of expression in leaves and stalks, approximately 750-fold and 30-fold higher than wild-type leaves and stalks, respectively, while Myb ZG-124-1-2a had the lowest expression levels. Although Myb ZG-124-1-2a expression levels in leaves were 30-fold higher than in the wild type, expression levels in stalks were not significantly different from the *SbMyb60* expression levels observed in wild-type stalks. Myb ZG-129-4-15a had intermediate expression levels of *SbMyb60* compared with the other two transgenic events, approximately 100- and 27-fold higher than expression levels in wild-type leaves and stalks, respectively.

SbMyb60 induces expression of monolignol biosynthetic genes

To evaluate the impact of *SbMyb60* on the expression of 10 monolignol biosynthetic pathway genes, transcript accumulations in leaves were quantified using RT-qPCR (Figure 3). Nine out of the 10 genes with involvement in monolignol biosynthesis in sorghum had elevated expression levels in plants overexpressing *SbMyb60*. The exception was a gene encoding cinnamate 4-hydroxylase (C4H), the second enzyme in the pathway (Figure 1). Relative expression levels of monolignol biosynthetic genes generally followed the same pattern as *SbMyb60* expression levels, with Myb ZG-129-4-10a events displaying the highest relative expression levels of the three transgenic events and Myb ZG-129-1-2a leaves displaying the lowest expression levels. The entire monolignol biosynthetic pathway, with the exception of C4H, was induced in Myb ZG-129-4-10a with the strongest induction of CCoAOMT (6.5-fold), cinnamyl alcohol dehydrogenase (CAD; 6.3-fold) and 4CL (5.8-fold). The expression levels of *p*-hydroxycinnamoyltransferase (HCT) and ferulate 5-hydroxylase (F5H) were only elevated in Myb ZG-129-4-10a plants. In the Myb ZG-129-4-15a and Myb ZG-124-1-2a events, seven of the ten monolignol biosynthesis pathway genes were more highly induced than in the wild type, while expression levels of C4H, HCT and F5H were not significantly different from the wild type. In Myb ZG-129-4-15a, the most strongly induced genes were 4CL (4.5-fold higher), C3H (4.3-fold higher) and CCoAOMT (4.5-fold higher) while, in Myb ZG-124-1-2a, the most strongly induced genes were phenylalanine ammonia lyase (PAL; 4.2-fold higher), 4CL (3.5-fold higher) and C3H (4-fold higher).

SbMyb60 overexpression induces accumulation of monolignol biosynthetic proteins

To determine if the induction of the monolignol biosynthetic genes resulted in higher protein abundances, Western blot analysis was performed using antibodies against several enzymes in the monolignol pathway, including 4CL, caffeic acid *O*-methyl transferase (COMT) and PAL (Figure 4a) and CAD and CCoAOMT (Figure 4b); antibodies against ascorbate peroxidase (APX) were used as a loading control. The intensities of the APX bands were consistent among all samples (Figure 4). In wild-

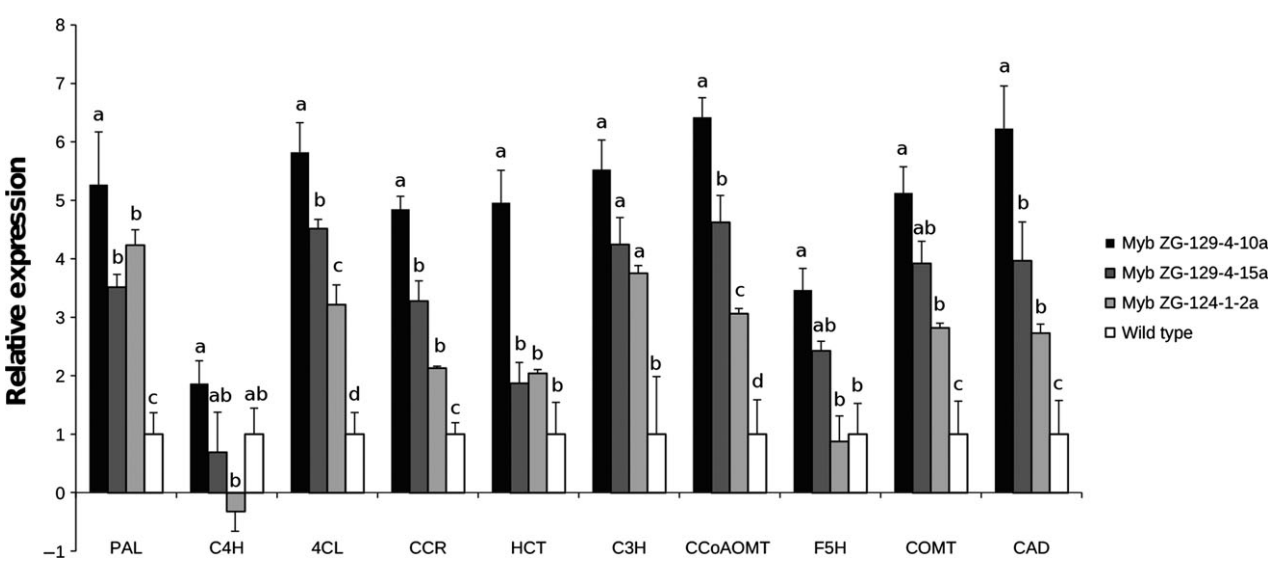

Figure 3. Relative expression levels of genes from the monolignol biosynthetic pathway determined via reverse transcription quantitative PCR.
Relative expression was determined using the ΔΔCt method with the α-tubulin gene (Sobic.001G107200.1) for normalization. Values displayed are normalized relative to ΔΔCt values from wild-type leaves or stalks. Error bars represent standard error. Samples with different letters are statistically different at α=0.05 using Tukey's honestly significant difference test. Gene identifiers for each monolignol pathway gene can be found in Table S1.

type leaves, bands corresponding to 4CL, COMT and CAD proteins were faint while bands corresponding to PAL and CCoAOMT were barely detectable, showing that these proteins were present at low levels. However, these bands in *SbMyb60* plants were much more intense, reflecting that overexpression of *SbMyb60* resulted in higher accumulations of these monolignol biosynthetic proteins in sorghum leaves. The intensities of the bands of the six monolignol biosynthesis proteins were significantly higher in Myb ZG-129-4-10a leaves compared with the wild type and the other two transgenic events. In contrast to the RT-qPCR expression data, which suggested that PAL, 4CL, COMT and CCoAOMT genes were more highly expressed in Myb ZG-129-4-15a leaves compared with Myb ZG-124-1-2a leaves, the bands corresponding to these proteins appeared to be more intense in Myb ZG-124-1-2a leaves.

SbMyb60 overexpression elevates enzymatic activities of phenylalanine ammonia lyase (PAL) and 4-coumarate:CoA ligase (4CL) in leaves

To determine whether the increased levels of monolignol-related biosynthetic proteins translated to increases in monolignol enzyme activity in *SbMyb60* plants, the activity levels of PAL (Figure 4c) and 4CL (Figure 4d) were measured. Enzyme activity levels of both PAL and 4CL were low in wild-type leaves, which is in agreement with previous results (Saballos *et al.*, 2012). On average, enzyme activity levels of PAL were approximately four- and three-fold higher in Myb ZG-129-4-10a and Myb ZG-129-4-15a plants, respectively, compared with the wild type (Figure 4c). However, PAL enzyme activity levels in Myb ZG-124-1-2a did not deviate significantly from wild-type values at alpha = 0.05. Activity levels of 4CL in leaves col-

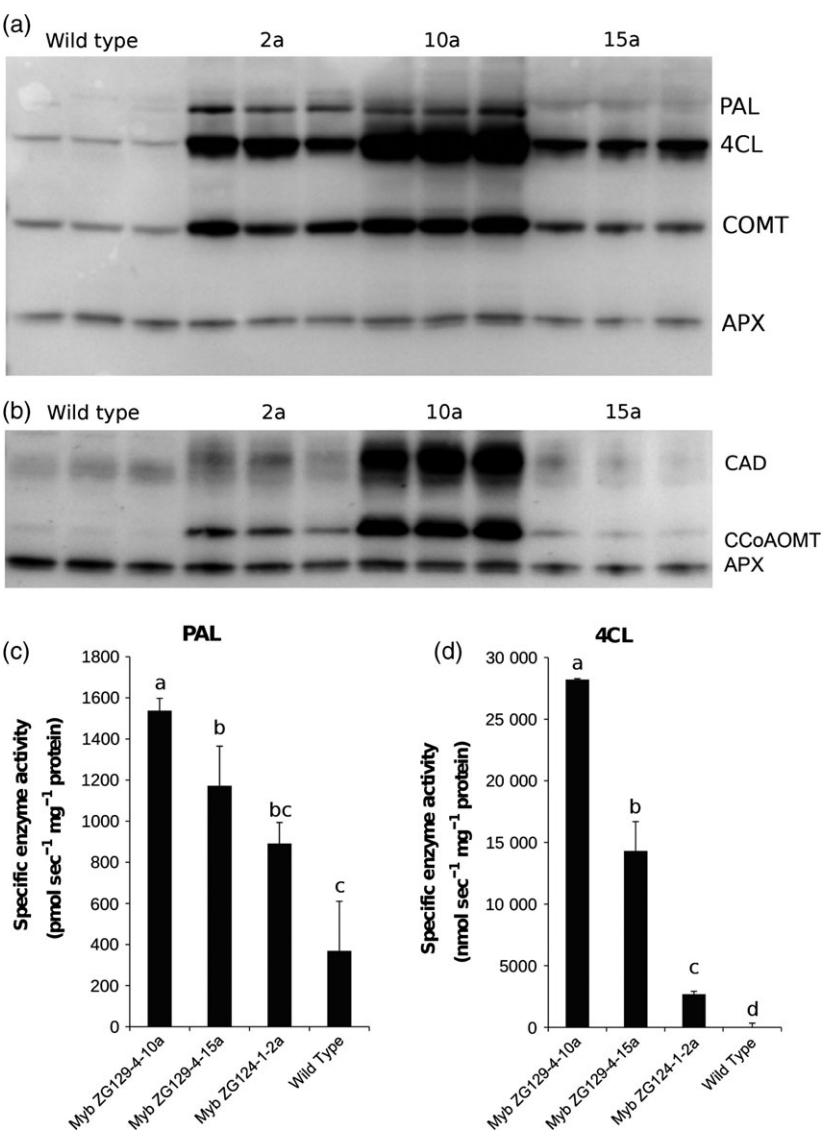

Figure 4. Abundances of proteins involved in monolignol biosynthesis and their enzymatic activities.

Immunoblot analysis of monolignol biosynthetic proteins (a) phenylalanine ammonia lyase (PAL), 4-coumarate:CoA ligase (4CL), and caffeic acid *O*-methyl transferase (COMT) and (b) cinnamyl alcohol dehydrogenase (CAD) and caffeoyl-CoA-*O*-methyltransferase (CCoAOMT). Ascorbate peroxidase (APX) was used as a loading control. Myb10a, Myb-ZG-129-4-10a; Myb15a, Myb-ZG-129-4-15a; Myb2a, Myb-ZG-124-1-2a

Enzyme activity levels of (c) PAL and (d) 4CL. Specific enzyme activity levels are displayed as pmol sec^{-1} mg^{-1} protein for PAL and nmol sec^{-1} mg^{-1} protein for 4CL. Error bars represent standard errors. Samples with different letters are statistically different at $\alpha = 0.05$ using an F-protected least significant difference test.

lected from all three transgenic events were significantly elevated compared with wild-type levels at alpha = 0.05 (Figure 4d). The specific activities for 4CL ranged from 28 000 nmol sec^{-1} mg^{-1} protein in Myb ZG-129-4-10a to approximately 2700 nmol sec^{-1} mg^{-1} in Myb ZG-124-1-2a.

SbMyb60 overexpression induces the accumulation of soluble monolignol pathway intermediates

To determine how *SbMyb60* overexpression affects flux through the phenylpropanoid biosynthetic pathway in plants grown under greenhouse conditions, soluble phenolics were extracted from stover (stalk and leaves) collected from mature *SbMyb60* and wild-type plants and the relative abundances of various monolignol pathway intermediates were measured by GC/MS (Figure 5). The phenolics *p*-coumaric acid (Figure 5a), caffeic acid (Figure 5b) and ferulic acid (Figure 5c) are common constituents of grass cell walls (Akin, 1982). Sinapic acid (Figure 5d) arises via methylation of the 3-hydroxyl group of 5-hydroxyferulate by COMT or via oxidation of sinapaldehyde (Nair, 2004). Syringic acid (Figure 5e) is probably formed through peroxisomal β-oxidation of sinapic acid (el-Basyouni *et al.*, 1964), and vanillic acid (Figure 5f) is derived from decomposition of ferulic acid or feruloyl CoA. Significantly higher levels of *p*-coumaric acid were detected in all three overexpression events relative to the wild type, ranging from 3.5-fold higher in Myb ZG-129-4-10a to 0.75-fold higher in Myb ZG-129-4-15a (Figure 5a). Levels of caffeic acid were elevated by 4.75-fold and 2-fold in the Myb ZG-124-1-2a and Myb ZG-129-4-15a events, respectively (Figure 5b), while levels in Myb ZG-129-4-10a did not deviate from those observed in the wild type at alpha = 0.05. Ferulic acid

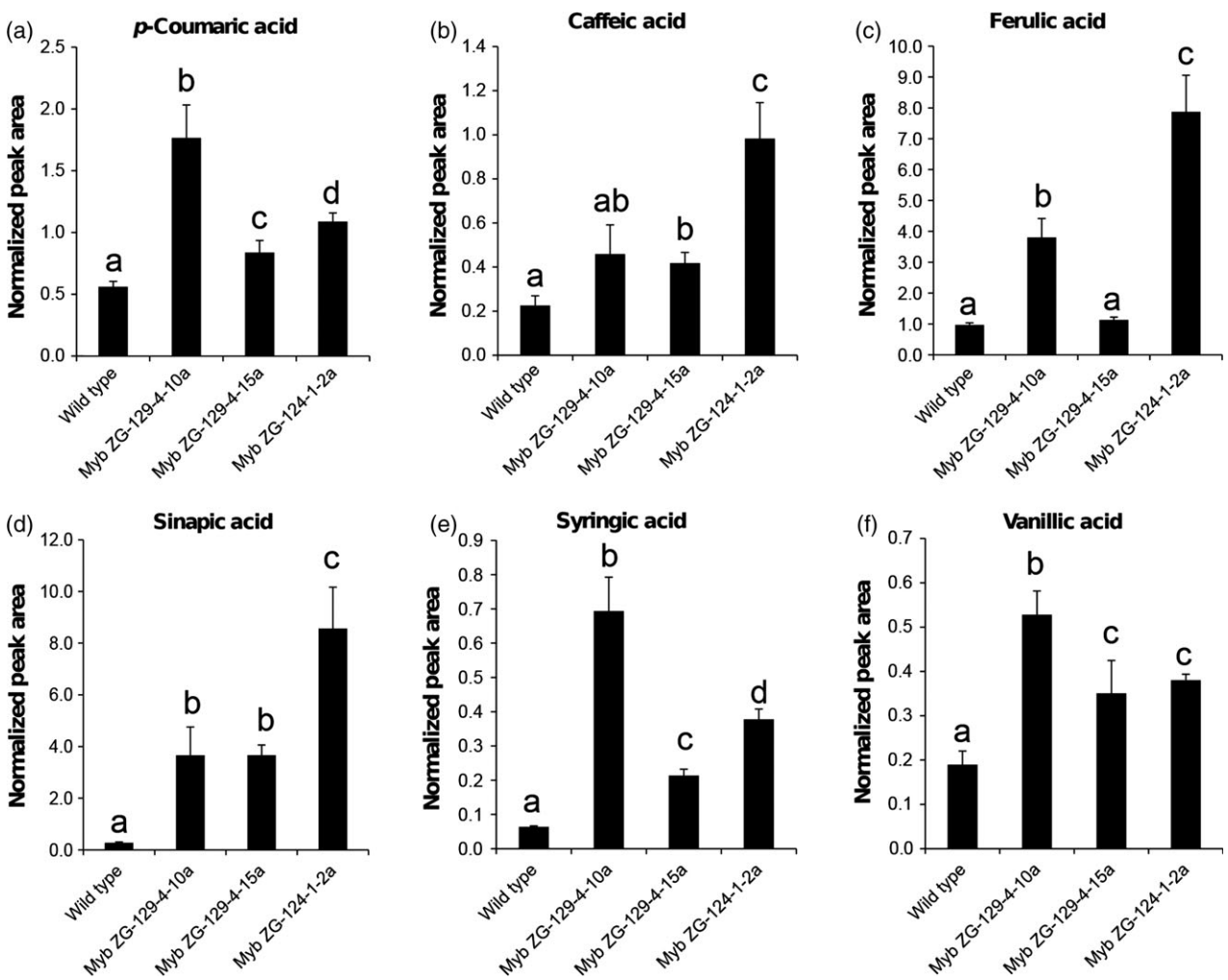

Figure 5. Relative abundances of intermediates of the monolignol biosynthetic pathway were determined by GC/MS analysis of soluble phenolics extracted from sorghum stover.

Peak area was normalized to the internal standard (toluic acid). Error bars represent standard errors and samples with different letters are statistically different at α = 0.05 using an F-protected least significant difference test. The relative abundances of (a) *p*-coumaric acid, (b) caffeic acid, (c) ferulic acid, (d) sinapic acid, (e) syringic acid, and (f) vanillic acid are presented.

levels were significantly elevated in the Myb ZG-129-4-10a and Myb ZG-124-1-2a events, respectively, but were not significantly different from the wild type in Myb ZG 129-4-15a plants (Figure 5c). The most striking difference was observed in the levels of free sinapic acid, which were elevated by approximately 39-fold over wild-type levels in the Myb ZG-129-4-10a and Myb ZG-129-4-15a events and over 80-fold in Myb ZG-124-1-2a plants (Figure 5d). Two other soluble phenolic compounds, syringic acid (Figure 5e) and vanillic acid (Figure 5f), were significantly elevated in all three events compared with the wild type. The highest levels of syringic and vanillic acid were observed in tissues harvested from Myb ZG-129-4-10a plants, which were approximately 9.3-fold and 2.9-fold higher than wild-type plants, respectively. In events Myb ZG-129-4-15a and Myb ZG-124-1-2a, syringic acid levels were 2.7-fold and 5.3 fold higher than in the wild type. Vanillin levels were approximately 1.5-fold higher compared with the wild type in both Myb ZG-129-4-15a and Myb ZG-124-1-2a events.

SbMyb60 overexpression alters lignin composition

Lignin composition in stover samples from mature wild-type and *SbMyb60* overexpressing plants was analyzed by thioacidolysis (Figure 6). Levels of S-lignin were significantly elevated in stover collected from all three overexpression events compared with the wild type, ranging from 1.25-fold higher in Myb ZG-124-1-2a to 2.5-fold higher in Myb ZG-129-4-10a (Figure 6a). The greater abundances of β-*O*-4-linked S-lignin observed in the transgenic events were associated with significant increases in the ratio of S- to G-lignin compared with wild-type plants. These increases in S/G ratios ranged from 1.73-fold higher in Myb ZG-129-4-10a to 1.23-fold higher in Myb ZG-124-1-2a compared with the wild type (Figure 6a). Levels of

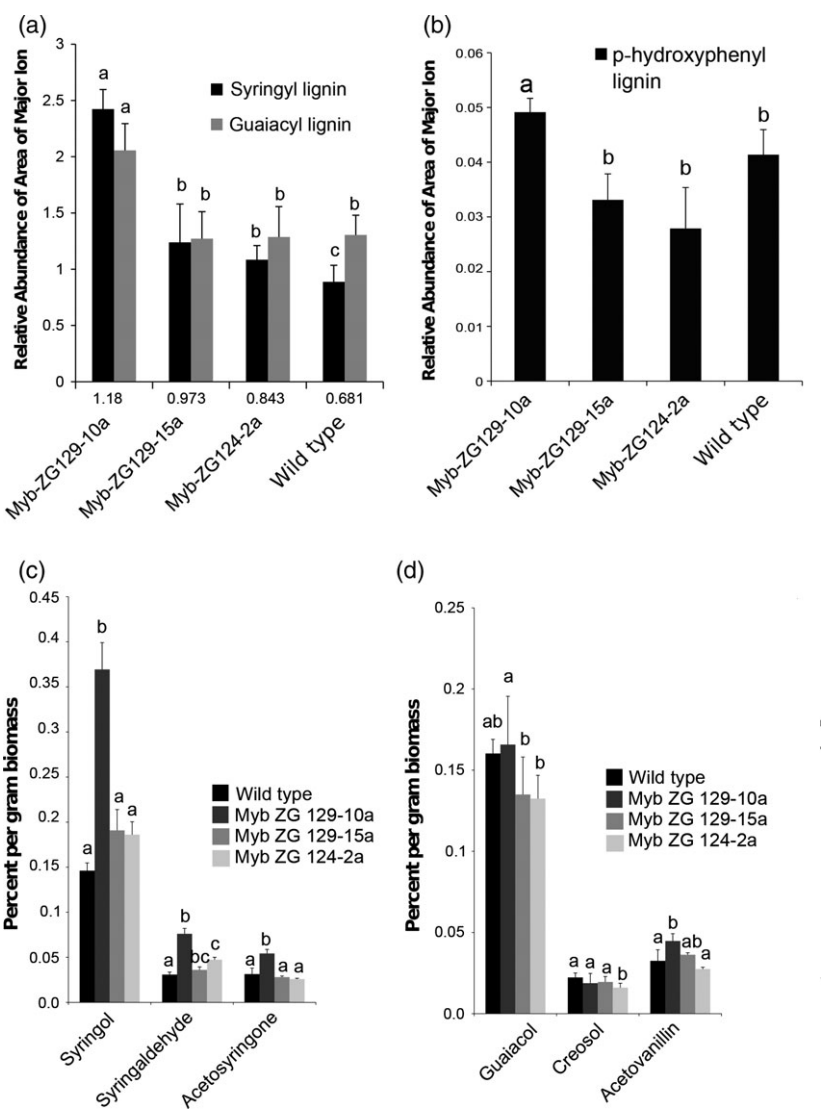

Figure 6. *SbMyb60* overexpression induces changes in lignin subunit composition. Samples with different letters are statistically different at α=0.05 using an F-protected least significant difference test.
(a) The abundance of β-*O*-4-linked syringyl (S) and guaiacyl (G) lignin subunits. The average S/G ratio for each line is displayed beneath the bar graph. All S/G ratios are statistically different from one another at α = 0.05.
(b) The abundance of β-*O*-4-linked *p*-hydroxyphenyl subunits.
(c) The abundances of S-lignin derivatives detected by pyrolysis-GC/MS.
(d) The abundances of G-lignin derivatives detected by pyrolysis-GC/MS.

β-*O*-4-linked G-lignin subunits were not substantially different from wild-type G-lignin levels in ZG-129-4-15a and Myb ZG-124-1-2a events (Figure 6a), but levels in Myb ZG-129-4-10a were approximately 1.6-fold higher than the levels of G-lignin in wild-type plants (Figure 6a). Levels of β-*O*-4-linked H-lignin were not significantly affected in Myb ZG-129-4-15a and Myb ZG-124-1-2a plants, but were approximately 20% higher in Myb ZG-129-4-10a plants compared with the wild type (Figure 6b).

To more thoroughly assess the impact of *SbMyb60* overexpression on lignin polymer composition, pyrolysis-GC/MS (py-GC/MS) was performed on tissue collected from mature plants (Table S2). Compared with thioacidolysis, which measures the relative abundance of β-*O*-4-linked subunits (Rolando *et al.*, 1992), pyrolysis breaks a variety of different linkages in lignin and other secondary cell wall polymers (Mu *et al.*, 2013) and provides a broader evaluation of lignin and secondary cell wall composition. Higher levels of syringyl derivatives in *SbMyb60* plants were also detected via py-GC/MS in the three transgenic events (Figure 6c). Specifically, syringol and acetosyringone were significantly elevated in Myb ZG-129-4-10a while levels of syringaldehyde were elevated in all three transgenic events. In general, levels of syringaldehyde were highest in Myb ZG-129-4-10a at approximately 2-fold higher than wild-type, while levels were 0.5-fold higher in Myb ZG-129-4-15a and Myb ZG-124-1-2a events. Additionally, levels of acetosyringone and syringol were elevated by approximately 2-fold over the wild type in Myb ZG-129-4-10a plants. In contrast to the thioacidolysis analysis, which indicated that levels of β-*O*-4-linked G-lignin subunits were elevated in Myb ZG-129-4-10a plants, levels of several guaiacyl derivatives measured by py-GC/MS, such as guaiacol, did not appear to be elevated in Myb ZG 129-4-10a plants or in any of the other transgenic events (Figure 6d). While creosol (a guaiacyl derivative) was also not elevated in any of the transgenic events, levels of acetovanillone were slightly elevated in Myb ZG-129-4-10a plants compared with both the wild type and the other transgenic events. Taken together, the results of thioacidolysis show that *SbMyb60* overexpression results primarily in elevations of β-*O*-4-linked S-lignin subunits in all three transgenic events and greater abundances of β-*O*-4 G-lignin subunits in Myb ZG-129-4-10a.

SbMyb60 overexpression alters cell wall composition

To determine whether *SbMyb60* overexpression affected cell wall composition or the abundance of lignin, stover from mature wild-type and *SbMyb60* overexpression events was analyzed using the National Renewable Energy Laboratory (NREL) analytical procedure (Sluiter *et al.*, 2008). Levels of Klason lignin were elevated by approximately 10% over wild-type levels in biomass collected from Myb ZG-129-4-10a (Table 1). However, levels of Klason lignin in biomass collected from Myb ZG-129-4-15a plants were not significantly different from the wild type and Klason lignin levels in Myb ZG-124-1-2a plants were approximately 15% lower than levels detected in wild-type plants. In addition to affecting levels of lignin in biomass, *SbMyb60* overexpression also appears to have an impact on the abundances of various structural carbohydrates in stover (Table 1). Levels of two major grass structural cell wall polysaccharides, cellulose and xylan, were reduced in Myb ZG-129-4-10a plants by approximately 9% and 15%, respectively, compared with the wild type (Table 1). However, cellulose and xylan levels were not substantially different in biomass collected from Myb ZG-129-4-15a and Myb ZG-124-1-2a plants compared with the wild type. Levels of other carbohydrates, such as arabinan and starch, were not substantially different in any of the transgenic events compared with the wild type; however, galactan levels were slightly elevated at approximately 28% and 37% higher in materials collected from Myb ZG-129-4-10a and Myb ZG-129-4-15a plants compared with the wild type (Table 1). The ash content, which represents the incombustible fraction of biomass, was elevated by over 1.5-fold in all three transgenic events compared with the wild type.

To further evaluate changes in biomass composition caused by overexpression of *SbMyb60*, fiber analysis was also performed on biomass collected from *SbMyb60* and

Table 1 Plant cell wall composition. The abundances of cell wall components were measured according to the analytical procedure of the National Renewable Energy Laboratory. All values below are represented in terms of g kg^{-1} of biomass. Means and standard errors are shown. Samples with different letters are statistically different at $\alpha = 0.05$

	Wild type	Myb ZG-129-4-10a	Myb ZG-129-4-15a	Myb ZG-124-1-2a
Structural carbohydrates	644.15 ± 46.44(a)	549.78 ± 6.31(b)	572.67 ± 21.84(ab)	582.01 ± 8.36(ab)
Cellulose	356.75 ± 21.57(a)	301.63 ± 3.06(b)	320.45 ± 13.44(ab)	323.02 ± 5.98(ab)
Xylan	263.32 ± 23.86(a)	223.93 ± 3.12(b)	225.08 ± 6.78(ab)	233.25 ± 2.35(ab)
Arabinan	24.08 ± 1.73(a)	26.40 ± 0.67(a)	26.15 ± 1.75(a)	25.49 ± 0.27(a)
Galactan	6.71 ± 1.80(a)	8.59 ± 0.49(b)	9.16 ± 1.37(b)	7.73 ± 0.09(a)
Starch	1.26 ± 0.62(a)	0.24 ± 0.24(a)	1.30 ± 0.40(a)	1.20 ± 0.47(a)
Klason lignin	129.90 ± 3.49(a)	143.94 ± 4.82(b)	127.73 ± 5.57(a)	111.46 ± 2.59(c)
Ash	14.33 ± 2.69(a)	23.39 ± 3.88(b)	25.05 ± 4.68(b)	24.07 ± 0.82(b)

wild-type plants to measure the levels of neutral detergent fiber (NDF), acid detergent fiber (ADF) and acid detergent lignin (ADL). While NDF values were only lower in biomass collected from Myb ZG-129-4-10a and Myb ZG-129-4-15a plants, ADF values were lower in biomass collected from all three transgenic events compared with the wild type (Figure 7a, b). Levels of ADL were elevated in Myb ZG-129-4-10a by approximately 10% over levels observed in wild-type plants (Figure 7c). The ADL levels in Myb ZG-129-4-15a and Myb ZG-124-1-2a plants did not deviate substantially from the levels observed in wild-type plants. Detergent cellulose, estimated as the difference between ADF and ADL, was approximately 16.0%, 15.5% and 18.9% lower than wild-type levels in Myb ZG-129-4-10a, Myb ZG-129-4-15a and Myb ZG-124-1-2a, respectively, while detergent hemicellulose, estimated as the difference between NDF and ADF, was approximately 0.8%, 0.4% and 3.8% lower in Myb ZG-129-4-10a, Myb ZG-129-4-15a and Myb ZG-124-1-2a compared with the wild type.

Lower levels of several major carbohydrate derivatives were also detected by py-GC/MS in plants overexpressing *SbMyb60*. These reductions were most prominent in the Myb ZG-129-4-10a line, which displayed significantly lower levels of the cellulose derivatives furfural and 2 (5H)-furanone compared with the wild type and the other two transgenic events (Figure 7d). Levels of furfural were reduced by approximately 30% and 38% in Myb ZG-129-4-10a compared with the wild type and Myb ZG-124-1-2a, respectively, while levels of 2(5H)-furanone were reduced by approximately 19% compared with both the wild type and the other two transgenic events. The Myb ZG-129-4-15a and Myb ZG-124-1-2a events particularly affected levels of 2,3-butanedione and propanoic acid: 2,3-butanedione levels were reduced by approximately 19% in all three transgenic events compared with the wild type, while levels of propanoic acid were approximately 37% lower than wild-type in all events. The major pyrolysis products detected from decomposition of carbohydrates are acetic acid and acetol (Demirbas, 2000), all of which were significantly lower than wild-type levels in all three transgenic events (Figure 7e), providing further evidence of impacts of *SbMyb60* overexpression on cell wall carbohydrates. The lowest levels of acetic acid were detected in Myb ZG-129-4-10a, which were reduced by 26% compared with the wild type, while levels in Myb ZG-124-1-2a were reduced by 14% compared with the wild type. Levels of acetol were reduced by approximately 33% in all three transgenic events. Notably, the carbohydrate levoglucasan, which is proposed to originate as the primary decomposition product of cellulose during pyrolysis, was elevated 2-fold in Myb ZG-124-1-2a, but did not deviate from wild-type levels in the other two transgenic events (Figure 7e). Levels of additional pyrolysis products are listed in Table S2.

SbMyb60 overexpression increases energy content of biomass

To determine the impact of changes in the secondary cell wall composition associated with overexpression of *SbMyb60*, the total biomass energy was measured from stover collected from mature wild-type and *SbMyb60* plants using bomb calorimetry (Figure 7f). Total energy levels were elevated in all three events compared with the wild type, indicating that biomass harvested from *SbMyb60* plants was more energy dense. Myb ZG-124-1-2a biomass had the highest total energy levels (154 cal g^{-1} higher than wild-type biomass) while energy levels in Myb ZG-129-4-10a and Myb ZG-129-4-15a (77–97 cal g^{-1} higher than wild-type biomass) did not differ from one another. On average, the increases to total energy observed in biomass from the three *SbMyb60* overexpression lines exceeded 100 cal g^{-1}.

Overexpression of *SbMyb60* is associated with ectopic lignin deposition in leaves

To investigate ectopic effects of *SbMyb60* overexpression on cell wall architecture, the fifth leaf from the bottom of each plant was collected, sectioned and stained with phloroglucinol to detect lignin (Figure 8). On the adaxial surface of the leaves (Figure 8a–d), lignification extended further inward towards the vascular bundles in the overexpression events. Higher levels of lignification were also observed around the vascular bundles and around the xylem and phloem elements in the transgenic lines. The layer of lignified cells on the abaxial regions in Myb ZG-129-4-10a plants was much thicker relative to the wild type (Figure 8e, f) and to the other two transgenic events (Figure 8g, h). The intensity of the phloroglucinol stain along the midrib was similar to the wild type in the Myb ZG-129-4-15a and Myb ZG-124-1-2a events.

Overexpression of *SbMyb60* affects growth and development

Growth and development of *SbMyb60* plants were altered compared with wild-type plants (Figure 8i–l). Plant height at maturity was reduced by approximately 38.6 cm in Myb ZG-129-4-10a, 45 cm in Myb ZG-129-4-15a and 33.5 cm in Myb ZG-124-1-2a events compared with the wild type (Table S3). Additionally, anthesis was significantly delayed in *SbMyb60* plants, ranging from 14 days later than the wild type in Myb ZG-129-1-2a plants to almost 20 days later in Myb ZG-129-4-10a plants (Table S3). Tillering was not significantly affected in Myb ZG-129-4-15a and Myb ZG-124-1-2a events; however, slightly less tillering was observed in Myb ZG-129-4-10a events (Table S3). Plant architecture was significantly affected in the Myb ZG-129-4-10a overexpression line (Figures 8j). The internodes of this line were substantially shorter compared with the wild

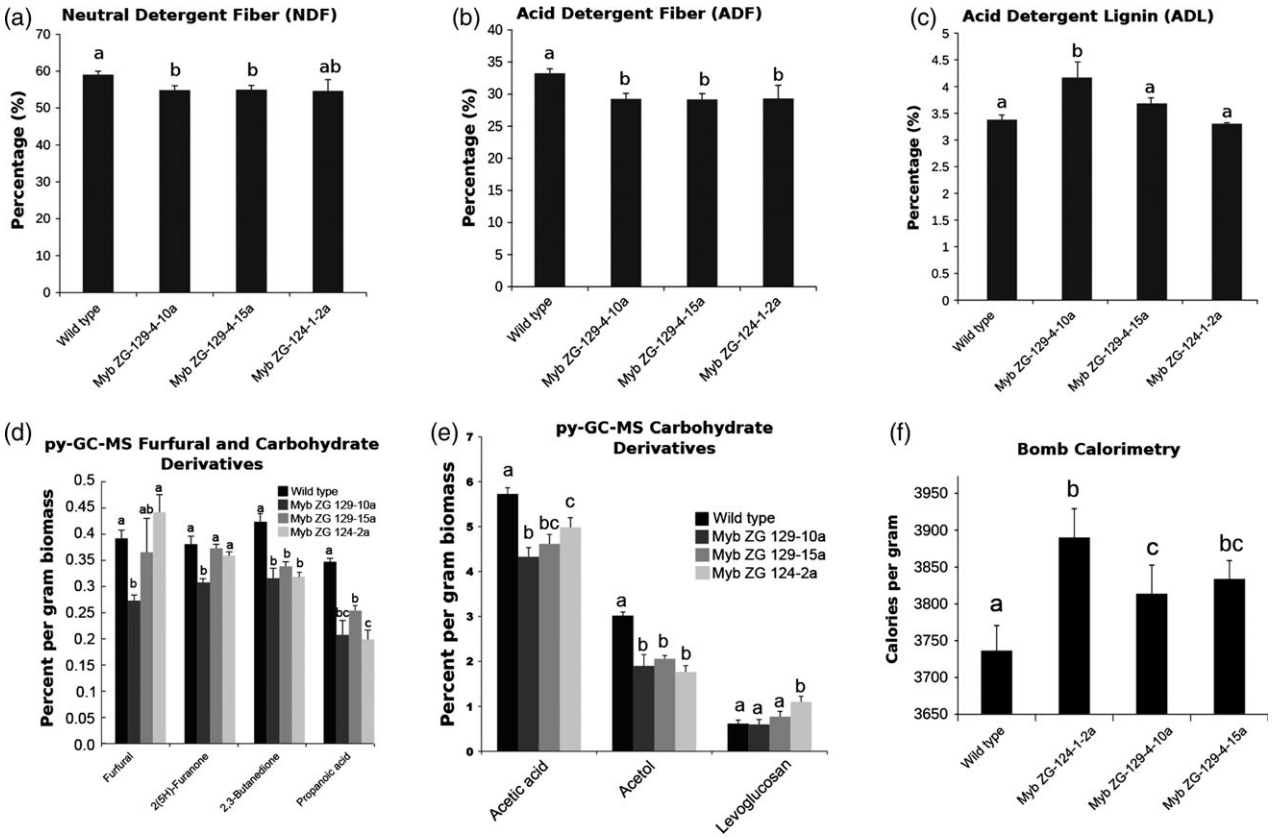

Figure 7. *SbMyb60* overexpression changes the structural composition of the secondary cell walls of mature plants.
Neutral detergent fiber (NDF), acid detergent fiber (ADF) and acid detergent lignin (ADL) from mature plants were measured using an ANKOM fiber analyzer. The relative abundances of carbohydrates in mature plants were also measured using pyrolysis-GC/MS. Error bars represent standard errors. Samples with different letters are statistically different at $\alpha = 0.05$ using an F-protected least significant difference test.
Abundance of (a) NDF, (b) ADF, and (c) ADL. Abundance of (d) furfural and other carbohydrates and (e) carbohydrate derivatives detected via pyrolysis-GC/MS. Additional carbohydrate derivatives detected via pyrolysis-GC/MS are presented in Table S2. (f)Total energy content of *SbMyb60* and wild-type sorghum biomass.

type or the other two events, and its leaves were more erect than in wild-type plants as a potential consequence. While impacts on plant architecture were more subtle in Myb ZG-129-4-15a and Myb ZG-124-1-2a events compared with Myb ZG-129-4-10a, differences were still apparent compared with the wild-type. In Myb ZG-129-4-15a, the leaves tended to grow less erect than in the wild-type and the other two transgenic events (Figure 8k). Myb ZG-124-1-2a (Figure 8l) had similar growth patterns to Myb ZG-129-1-15a, but the leaves were more erect compared with Myb ZG-129-4-15a.

DISCUSSION

C4 grasses are rapidly emerging as renewable feedstocks for the bioeconomy, providing renewable biomass for biofuels, chemical precursors for green chemistry and other industrial applications (Rooney *et al.*, 2007; Byrt *et al.*, 2011). The ability to manipulate cell wall composition may allow biomass content to be tailored to specific industrial applications (Boudet *et al.*, 2003). The phenylpropanoid pathway synthesizes an array of aromatic compounds with key roles in plant physiology, including cell wall biosynthesis, defense against herbivory and pathogenesis, and stress response (Dixon and Paiva, 1995; Boerjan *et al.*, 2003). Therefore, this pathway represents a useful target for manipulating plant cell wall and phytochemical compositions (Li *et al.*, 2008) for industrial purposes or to create plants that are more tolerant to a variety of abiotic (Dai *et al.*, 2012; Yang *et al.*, 2012) and biotic stressors (Dixon *et al.*, 1996; Ibraheem *et al.*, 2015). While this pathway has been well studied in dicot species (e.g. poplar and Arabidopsis) comparatively little is known about its regulation in monocots, such as sorghum, which is rapidly emerging as a biofeedstock and can be grown in drought-prone areas. The MYB TFs are one of the major regulators of phenylpropanoid biosynthesis, and specific MYB proteins have been shown to induce the biosynthesis of various monolignols (Zhong and Ye, 2009; Zhou *et al.*, 2009), flavo-

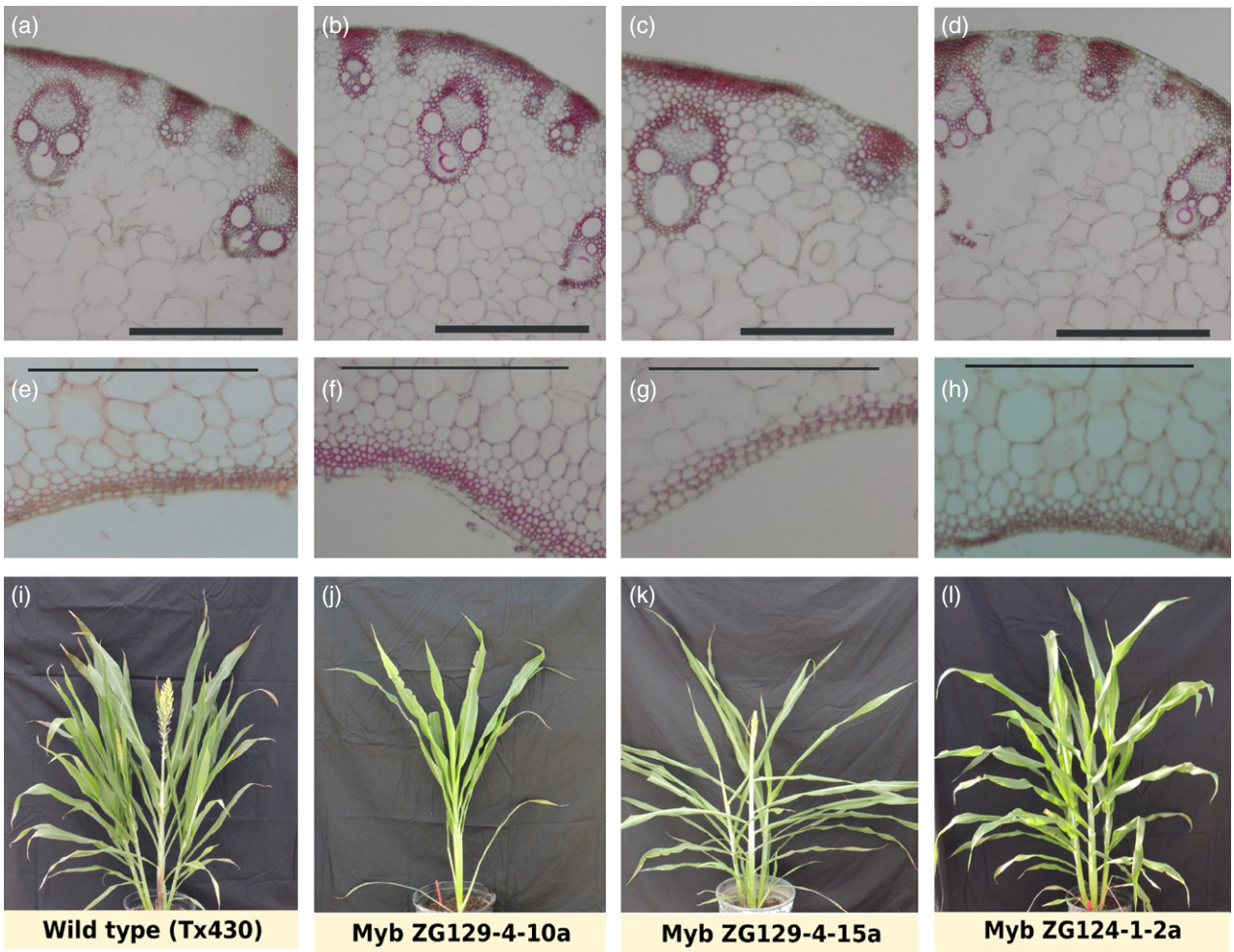

Figure 8. *SbMyb60* overexpression induces changes in growth and lignin deposition in leaves.
(a)–(h) Phloroglucinol-stained leaf (a–d) and midrib (e–h) sections from *SbMyb60* and wild-type plants. Scale bar = 500 μm. (i)–(l) Phenotype of the wild type and *SbMyb60* at 8 weeks.

noids (Grotewold *et al.*, 1994; Moyano *et al.*, 1996), anthocyanins (Albert *et al.*, 2015; Ibraheem *et al.*, 2015) and other aromatic compounds (Borevitz, 2000). In this study, the function of the *SbMyb60* TF was investigated *in planta* in the emerging bioenergy grass sorghum. Many of the genes described in the monolignol biosynthetic pathway were activated in plants overexpressing *SbMyb60*, demonstrating that this TF is associated with activation of monolignol biosynthesis in sorghum. Transcriptional induction of these monolignol biosynthetic genes in leaves translated into greater accumulations of corresponding proteins and increased enzyme activity levels for two key pathway enzymes. Ectopic lignin deposition in cortical cells adjacent to the leaf epidermis and increased lignification around the vascular bundles and xylem elements were observed in all three *SbMyb60* overexpression events. Taken together, these data indicated that overexpression of *SbMyb60* is sufficient to induce lignin biosynthesis in sorghum. The abundances of β-*O*-4-linked S-lignin subunits in stover collected from mature plants were substantially elevated in all three overexpression events, which suggested that overexpression of this TF predominantly drives the flux of the monolignol pathway towards the synthesis of S-lignin in sorghum. In the canonical monolignol biosynthetic pathway, F5H and COMT are specifically committed to the synthesis of S-lignin (Grand, 1984). Previous studies have shown that modifying the expression levels of these genes significantly influences the abundance of S subunits in the lignin polymer (Meyer *et al.*, 1998; Marita *et al.*, 1999) and, in the case of COMT, underexpression also results in the incorporation of 5-hydroxyconiferyl alcohol into lignin instead of S-lignin subunits (Piquemal *et al.*, 2002. The levels of expression of COMT were significantly upregulated in all three *SbMyb60* overexpression lines, which could explain the higher abundances of S-lignin detected in these lines compared with in the wild type. Additionally,

F5H was exclusively upregulated in event Myb ZG-129-4-10a, which displayed the highest abundance of S-lignin of the three transgenic lines. The elevated expression level of this gene could explain why the abundance of S-lignin was highest in Myb ZG-129-4-10a biomass.

Phenotypic effects were largely correlated with expression levels of *SbMyb60*, which varied significantly among the three transgenic events included in this study. Elevated levels of Klason lignin and ADL were observed in Myb ZG-129-4-10a, which had the highest level of *SbMyb60* expression. Myb ZG-129-4-10a was also the only line that had elevated levels of β-*O*-4-linked G- and H-lignin compared with wild-type plants. Expression of the HCT and F5H genes was also only induced in this line (Figure 3). The elevated levels of HCT could contribute to the increased lignin content observed in this line by directing the flux of *p*-coumaryl CoA toward monolignol biosynthesis and away from other aromatic biosynthetic pathways that share *p*-coumaroyl CoA as a substrate (Boerjan *et al.*, 2003). The accumulation of flavonoids and other aromatics has been previously observed in HCT-downregulated plants (Besseau *et al.*, 2007). This hypothesis also explains why monolignol pathway intermediates (soluble phenolics) were relatively lower in Myb ZG-129-4-10a compared with the other two transgenic events, which suggests that flux was directed toward the synthesis of monolignols and incorporation into lignin. In addition, the growth and development of Myb ZG-129-4-10a plants were more severely altered compared with the other two transgenic events and wild-type plants, possibly due to high levels of lignin impeding cell wall expansion during growth, as has been observed in other high-lignin plants (Caño-Delgado *et al.*, 2000). While the growth of this plant was reduced by approximately 29% compared with the wild type, the 10% increases in ADL and Klason lignin were substantial. For example, similar elevations in lignin content in switchgrass were previously shown to have dramatic negative impacts on digestibility by ruminants (Sarath *et al.*, 2011; Vogel *et al.*, 2013), demonstrating that small changes in the abundance of lignin can have large implications for end usage in herbaceous plants. Further, the total energy levels in Myb ZG-129-4-10a biomass were elevated by 77 cal g^{-1} of biomass, which represents a substantial increase in total energy per gram relative to wild-type biomass.

Conversely, Myb ZG-124-1-2a displayed both the lowest levels of *SbMyb60* overexpression and the lowest expression levels of monolignol biosynthesis genes of the three transgenic events. Likewise, this line also displayed the lowest PAL and 4CL enzyme activity levels. This line accumulated the highest levels of several monolignol pathway intermediates (caffeic, ferulic, and sinapic acid) compared with the wild type and the two other events. The higher levels of these compounds in Myb ZG-124-1-2a is probably because only a subset of the monolignol biosynthesis genes were upregulated, potentially increasing flux through only some steps of the pathway and leading to the accumulation of pathway intermediates but not to increased synthesis of monolignols or their incorporation into lignin. In contrast, these pathway intermediates may have accumulated to lower levels in biomass harvested from Myb ZG-129-4-10a and Myb ZG-129-4-15a due to higher expression levels of monolignol biosynthesis genes compared with Myb ZG-124-1-2a, which probably drives the flux through the monolignol biosynthesis pathway towards the synthesis of monolignols, preventing the accumulation of these intermediates. This hypothesis is supported by enzyme activity levels of PAL and 4CL, which are lower in Myb ZG-124-1-2a compared with the other two *SbMyb60* overexpression lines, and analysis of Klason and ADL lignin, which demonstrates that lignin levels are lowest in Myb ZG-124-1-2a biomass. Similarly, the sorghum *bmr6* and *bmr12* mutants, which are impaired in monolignol biosynthesis, also accumulate these pathway intermediates (Palmer *et al.*, 2008). Despite having the lowest levels of lignin of the three transgenic events and lower levels of structural carbohydrates than wild-type plants, total energy levels were highest in Myb ZG-124-1-2a plants, which suggests that phenolic intermediates and other phenolic polymers make significant contributions to total energy in biomass. For example, grass cell walls, unlike dicot cell walls, contain significant amounts of ester-linked phenolic compounds such as *p*-coumaric and ferulic acid, which are separate from the lignin polymers and may contribute to energy levels in biomass (Grabber *et al.*, 1991; Boerjan *et al.*, 2003; Grabber and Lu, 2007). The phenotypic variation observed among the transgenic events suggests that expression levels of *SbMyb60* can be potentially manipulated to fine-tune secondary cell wall and phytochemical composition of plant biomass.

Although *SbMyb60* is a co-ortholog of *AtMYB58* and *AtMYB63*, several phenotypic differences were observed between *SbMyb60* and Arabidopsis *AtMYB58* and *AtMYB63* overexpression events (Zhou *et al.*, 2009). First, unlike *AtMYB58* and *AtMYB63*, overexpression of *SbMyb60* did not induce C4H even though the promoter of this gene contains several canonical AC elements for myb binding. However, *p*-coumaric acid was elevated in all three transgenic events. Although Sobic.002G126600 was the most highly expressed C4H in lignified sorghum tissues (Shakoor *et al.*, 2014), *SbMyb60* overexpression could have activated another related cytochrome P450 hydroxylase gene whose protein catalyzes this reaction. Alternatively, there is evidence that the monocot PAL enzymes have the ability to convert tyrosine directly into *p*-coumaric acid, thus bypassing the C4H step (Rosler, 1997). Second, expression of F5H was induced in the Myb ZG-129-4-10a line, which has not previously been observed in plants overexpressing either *AtMYB58* or *AtMYB63* (Zhou *et al.*,

2009). The Arabidopsis F5H gene lacks canonical AC elements required for MYB binding and is instead regulated by SND1, an NAC domain TF, (Zhou *et al.*, 2009). The sorghum F5H homolog contains only degenerate AC motifs (Figure S3); however, the high levels of *SbMyb60* overexpression in Myb ZG-129-4-10a might induce expression of F5H through these motifs or by activating other TFs. Lastly, *AtMyb58* and *AtMyb63* overexpression did not affect levels of structural carbohydrates in biomass, suggesting that they specifically regulated monolignol biosynthesis (Zhou *et al.*, 2009). However, lower levels of structural carbohydrates were observed in all three *SbMyb60* events. Previously, downregulation of 4CL in poplar was associated with reductions in Klason lignin and elevations in cellulose, suggesting a balancing mechanism between cellulose and phenylpropanoid biosynthesis (Hu *et al.*, 1999). While the underlying mechanism(s) for this observation has yet to be elucidated, it is possible that carbon flux was being redirected away from carbohydrate biosynthesis to the shikimate and phenylpropanoid pathways in the *SbMyb60* overexpression events. In addition, the overexpression of *SbMyb60* could have activated other myb family TFs as well as TFs from other families (i.e. bZIP, NAC and WKRY) (reviewed in Zhong and Ye, 2009) that regulate lignin, cellulose and hemicellulose biosynthesis. It is possible that the stimulation of other TFs in the *SbMyb60* overexpression lines could be responsible for the reduced carbohydrate levels as well as some of the other differences in secondary cell wall composition associated with the three transgenic lines. Further investigations will elucidate whether *SbMyb60* can trigger activation and/or repression of other TFs involved in secondary cell wall biosynthesis. Despite these phenotypic differences, several similarities were noted between the *SbMyb60*, *AtMYB58* and *AtMYB63* overexpression lines which demonstrate that these orthologous TFs may have common effects on biomass composition and plant growth and development in sorghum and Arabidopsis. For example, plants overexpressing *AtMYB58* and *AtMYB63* at levels approximately 10 000- fold and 80-fold higher than wild-type levels induced genes encoding nine monolignol biosynthesis enzymes by levels that were 5-fold to 20-fold higher than in wild-type plants (Zhou *et al.*, 2009). The levels of MYB TF overexpression and induction of monolignol biosynthesis genes in the Arabidopsis overexpression lines are similar to the levels observed in the *SbMyb60* overexpression lines presented in this study. In addition, transgenic downregulation of both *AtMYB58* and *AtMYB63* was correlated with downregulation of the same nine monolignol biosynthesis pathway genes, supporting the hypothesis that these two TFs likely function as regulators of monolignol biosynthesis. While changes to total lignin content were not documented in the Arabidopsis study, similar elevations in S- and G-lignin levels and higher S/G ratios were documented in both the sorghum and Arabidopsis overexpression lines. Like *SbMyb60* plants, *AtMYB58* and *AtMYB63* plants were also substantially stunted in growth comparison with the wild type (Zhou *et al.*, 2009), suggesting that excess lignification also impedes normal growth and development in Arabidopsis.

In conclusion, this study demonstrates that *SbMyb60* is associated with elevated expression of genes involved in monolignol biosynthesis in sorghum and that its expression levels can be manipulated to modify cell wall composition and soluble phenolic profiles. Moreover, this study also indicates that despite the low amino acid similarity between *SbMyb60* and its Arabidopsis homologs, this element of the monolignol biosynthesis network is conserved between monocots and dicots, although there are some subtle phenotypic differences between the sorghum *SbMyb60* overexpression lines and the Arabidopsis *AtMYB58* and *AtMYB63* lines. Certainly, functional characterization of additional myb TFs in sorghum and other bioenergy grasses will lead to the identification of other mechanisms that could be exploited to engineer the phenylpropanoid biosynthesis pathway to produce aromatic compounds of interest.

EXPERIMENTAL PROCEDURES

Identification of putative *AtMYB58* and *AtMYB63* orthologs associated with lignin biosynthesis in *Sorghum bicolor*

To identify orthologous sequences in *Sorghum bicolor*, protein homologs of *AtMYB58* and *AtMYB63* were identified using the InParanoid-based ortholog and paralog assignments (Remm *et al.*, 2001; Ostlund *et al.*, 2010) available through Phytozome (http://phytozome.jgi.doe.gov/pz/portal.html#). Fourteen putative homologous genes were identified in sorghum. Multiple amino acid sequence alignments generated using MUSCLE (Edgar, 2004) and ProtTest (Darriba *et al.*, 2011) were used to predict optimal evolutionary models. Unrooted maximum likelihood trees were constructed using Garli (v.2.0) (Zwickl, 2006). The WAG + G + I evolutionary model was selected and 500 non-parametric bootstrap replicates were performed.

Construction of *SbMyb60* expression cassette and genetic transformation of *Sorghum bicolor* (Tx430)

SbMyb60 (Sobic.004G273800) was amplified using turbo *Pfu* DNA polymerase (Agilent, http://www.agilent.com/) from genomic DNA isolated from *S. bicolor* (BTx623) plants and subcloned between the E35S CaMV promoter and the 35S CaMV terminator. The E35S promoter is a variant of the CaMV 35S promoter that contains a tandem duplication of the 250-bp upstream region. The E35S promoter has been previously shown to increase the transcription levels of transferred DNA by over 10-fold compared with the unmodified CaMV 35S promoter (Kay *et al.*, 1987). The cassette was introduced into the binary vector pPZP211 (Hajdukiewicz *et al.*, 1994) (Methods S2). The recombinant vector was introduced into *A. tumefaciens* strain NTL4/pTiKPSF2 (Palanichelvam *et al.*, 2000) via tri-parental mating. *Sorghum bicolor* (RTx430) plants

were transformed as previously described (Howe *et al.*, 2006). Three transformation events, designated Myb ZG-129-4-10a, Myb ZG-129-4-15a and Myb ZG-124-2a, were selected for further characterization.

Cultivation of plants and collection of materials

Seeds from T_2 plants were planted a soil mixture consisting of a 1:2:1:1 ratio of soil:peat moss:vermiculite:sand and were arranged in the greenhouse in a randomized complete block design. Seedlings were culled to one plant per pot after 2 weeks and a total of six plants per line were grown at two different time points. Growth conditions were augmented with high-pressure sodium lights on a 16-h daylight cycle. Approximately 0.20 ml of fertilizer (Dyna Green All Purpose 12-12-12, Hummert, https://www.hummert.com/) were applied to each pot every 7 days. Temperatures in the greenhouse were maintained at 29–30°C during daylight hours and 26–27°C at night. Leaf and stalk tissues were harvested separately from 5 to 6-week-old plants grown during the first time point, frozen immediately in liquid N, ground using a freezer mill (SPEX SamplePrep, http://www.spexsampleprep.com/), and used for reverse transcription quantitative PCR (RT-qPCR). Specifically, the fifth leaf from the base of each plant and the stalk segment 10 cm above the soil for were collected for leaf and stalk tissue, respectively. Leaf materials were also used for Western blots, enzyme activity assays and microscopy. Biomass from whole mature plants (stalks and leaves) grown during the second time point was dried at 50°C, ground in a Wiley mill fitted with a 2-mm mesh screen (Arthur H. Thomas Co., http://www.ahthomas.com/), ground in a cyclone mill fitted with a 1-mm mesh screen (UDY Corp., http://www.udyone.com/), then used for lignin and cell wall composition analysis and soluble phenolic profiling.

Reverse transcription PCR to confirm *SbMyb60* overexpression

Total RNA was extracted from leaves and stalks harvested from 5 to 6-week-old greenhouse-grown plants from four individual plants per line. Approximately 100 mg of homogenized leaf or stalk material was added to 1 ml of TriPure Isolation Reagent (Roche, http://www.roche.com/). The RNA was recovered and purified using the RNA Clean and Concentrator Kit (Zymo, https://www.zymoresearch.com/) and on-column DNase treatment was performed. One microgram of RNA was used for cDNA synthesis with the iScript Reverse Transcription Supermix for qRT-PCR (Bio-Rad, http://www.bio-rad.com/). Twenty-microliter qPCR reactions were performed in triplicate with 0.6 μM SbMyb60F primer, 0.6 μM SbMyb60R primer (Table S4), 1 μl 1:5 diluted cDNA and 10 μl SsoAdvanced SYBR Green Supermix (Bio-Rad) using the Bio-Rad CFX Connect Real Time System. Thermal cycling parameters were as follows: initial denaturation for 30 sec at 95°C, 40 cycles of denaturation at 95°C for 5 sec and annealing at 62°C for 10 sec. Primer specificities were confirmed by dissociation curve analysis, consisting of denaturation at 95°C for 5 sec, cooling to 65°C for 15 sec, and gradual heating at 0.2°C sec^{-1} to a final temperature of 95°C. Relative expression was computed using the ΔΔCt method with elf4A1 for normalization. No-template and no reverse transcription controls were also analyzed to verify the absence of DNA contamination. Relative expression values were statistically analyzed using ANOVA. Pairwise comparisons among genotypes were performed using Tukey's honestly significant difference (HSD) test with an experiment-wise error rate of $\alpha = 0.05$ using R (v.3.1.1 for Linux). Three biological replicates were analyzed for each line in triplicate.

Quantitative RT-PCR to quantify relative expression of genes in the monolignol biosynthesis pathway

Relative expression levels of monolignol biosynthetic genes were measured using RT-qPCR on RNA collected from leaf tissue of *SbMyb60* overexpression plants. The RNA extractions, cDNA synthesis and qPCR assays were performed as described above. A list of qPCR primers specific for monolignol biosynthetic genes can be found in Table S4. α-Tubulin was used for normalization. Three biological replicates were analyzed in triplicate for each line and statistical analysis was performed as described above.

Western blot and immunodetection

Proteins from *SbMyb60* overexpression events and wild-type plants were isolated from leaf tissue of 5–6-week-old sorghum plants using an extraction buffer containing protease inhibitor (Sigma-Aldrich, http://www.sigmaaldrich.com/) (Sattler *et al.*, 2009). Protein concentrations were measured with a dye-binding assay (Pierce 660 Protein assay kit, https://www.thermofisher.com/) using BSA as a standard. Western blot analyses were performed as described previously (Sattler *et al.*, 2009, 2012; Saballos *et al.*, 2012). The first blot was probed with 1:14 000 dilutions of PAL, 1:12 000 dilutions of 4CL and 1:14 000 dilutions of COMT. A second blot was probed using 1:1000 dilutions of cinnamyl alcohol dehydrogenase (CAD) and 1:1000 dilutions of CCoAOMT. As loading controls we used 1:4000 dilutions of APX primary antibodies. The secondary antibody was detected using chemiluminescence with Amersham ECL Western blotting reagent (GE Healthcare, http://www.gehealthcare.com/). Each lane on the gels represents protein extract collected from one individual plant.

Phenylalanine ammonia lyase and 4-coumarate:CoA ligase enzyme assays

The 4CL assays were performed as described previously using *p*-coumarate as the substrate (Knobloch and Hahlbrock, 1977; Saballos *et al.*, 2012). To assay for PAL activity, proteins were extracted in 0.1 M sodium borate buffer (pH 8.8) and the homogenates were passed through 2.0-ml Zeba Spin Desalting Columns (Thermo-Scientific, http://www.thermoscientific.com/). The reaction was initiated by adding 0.02 M of L-phenylalanine to 20 μl desalted protein extract (Gómez-Vásquez *et al.*, 2004). The conversion of phenylalanine to cinnamic acid was assayed by monitoring the change in absorbance at 290 nm every minute for a period of 30 min. No-substrate and no-extract controls lacked activity. Specific enzyme activity was calculated by normalizing by the amount of protein added to each reaction. Four biological replicates from each line were analyzed in triplicate.

Analysis of soluble phenolic intermediates of the monolignol biosynthesis pathway

Soluble phenolic compounds were extracted from ground sorghum as described previously (Palmer *et al.*, 2008). Ten microliters of each extract was derivatized by adding 50 μl of pyridine containing toluic acid (internal standard) and 80 μl of *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA) (Thermo Fisher). Compounds of interest were analyzed as silyl derivatives (Palmer *et al.*, 2008). The derivatization reactions were carried out at 60°C for 90 min and were analyzed on an Agilent G2570A integrated GC/MS system equipped with a G2913A autoinjector module, 6850

Series II GC, and a 5973 Network Mass spectrometer (Agilent). Relative abundances of soluble phenolic compounds were determined by the peak areas of major ions. Between-sample normalization was performed using the peak area of the toluic acid internal standard. Four biological replicates were analyzed as technical duplicates.

Thioacidolysis

Lignin subunit composition of material collected from mature sorghum plants was determined by thioacidolysis followed by GC/MS. Whole plant samples were washed, derivatized and analyzed as described previously (Palmer *et al.*, 2008) and 50 μl 1.0 mg ml^{-1} 4-4′-ethylidenebisphenol in dichloromethane was added to each sample as an internal standard. Analysis was performed in duplicate on four individual plants per line.

Pyrolysis by py-GC/MS

Micropyrolysis of all biomass samples was performed according to Mihalcik *et al.* (2011) using a Frontier Lab (http://www.frontier-lab.com/) Double-Shot Micro Pyrolyzer PY-3030iD with the Frontier Lab Auto-Shot Sampler AS-1020E attached to a gas chromatograph, Shimadzu GC-2010 (Columbia, MD) coupled with a Shimadzu GCMS-QP2010S mass spectrometer (http://www.shimadzu.com/). A sample size of 350–450 μg was subjected to single-shot pyrolysis at 500°C for 30 sec using stainless steel cups (disposable eco-cup LF; Frontier Lab) followed by GC separation. Quantitative analysis of individual chemical products was done by the external standard method, using authentic samples to generate calibration curves (Mihalcik *et al.*, 2011). Materials from four biological replicates per line were analyzed in triplicate.

Cell wall composition and fiber analysis

The abundances of various structural carbohydrates (cellulose, xylan, arabinin), Klason lignin and ash were measured using the NREL analytical procedure (Sluiter *et al.*, 2008). Fiber analysis was performed on ground plant material using an ANKOM fiber analyzer as previously described (Vogel *et al.*, 1999). Analysis was performed on four biological replicates per line in triplicate.

Sample fixation and histochemical staining of lignin

Midribs from the fifth leaf from the bottom of each plant were taken from 5 to 6-week-old sorghum plants. The portion of the leaf 1.5 cm from the petiole was fixed in 3.7% formaldehyde and embedded in paraffin. Twenty-micrometer sections were deparaffinized, hydrated for 30 min and stained for 15 sec in phloroglucinol in 20% HCl. Sections were examined and photographed using an Olympus BX-51 light microscope (http://www.olympus-ims.com/).

Bomb calorimetry

The total energy content of biomass was determined using a Parr 6400 bomb calorimeter (Parr Instrument Company, http://www.parrinst.com/). Approximately 200 mg of plant material was added to the sample holder and 600 mg of light mineral oil was added to the biomass as a combustion aid. The total energy of each biomass sample was calculated by subtracting the total energy released from combustion of mineral oil alone from the total energy of the combined mineral oil and biomass sample and standardizing by sample weight. Total energy was measured in duplicate for four biological replicates.

Statistical analysis

Statistical analysis of results from enzyme activity level assays, soluble phenolics analysis, fiber analysis, chemical composition analysis, py-GC/MS analysis and bomb calorimetry analysis was performed using SAS (https://www.sas.com/) PROC MIXED v.9.2 (SAS, 2002–2008). Heterogeneous variances were detected using Levene's test for homogeneity and were addressed with PROC MIXED. Pairwise comparisons among genotypes were performed using F-protected least significant differences with an experiment-wise error rate of $\alpha = 0.05$ (SAS). Information about the number of biological and technical replicates used for each experiment can be found in the detailed Experimental Procedures.

ACKNOWLEDGEMENTS

We thank John Toy and Nathan Peroutka-Bigus for technical assistance and Pat O'Neill for statistical support. This research was supported by the United States Department of Agriculture – National Institute of Food and Agriculture AFRI grant number 2011-67009-30026 (SES, TEC, and DLF-H) and additional funding from USDA-ARS, CRIS projects 3042-21220-032-00D (SES DLF-H) and 3042-21000-030-00D (GS, NAP). The US Department of Agriculture, Agricultural Research Service, is an equal opportunity/affirmative action employer and all agency services are available without discrimination. Mention of commercial products and organizations in this manuscript is solely to provide specific information. It does not constitute endorsement by USDA-ARS over other products and organizations not mentioned.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Neighbor-joining analysis of sorghum MYB genes and *AtMYB58* and *AtMYB63*.

Figure S2. Neighbor-joining analysis of Arabidopsis MYB genes and *SbMyb60* and Sobic.006G199800.

Figure S3. Promoter analysis for sorghum F5H (Sobic.001G196300).

Table S1. Genomic insertion sites of E35S::*SbMyb60* constructs in three transgenic sorghum lines.

Table S2. Major products detected via pyrolysis-GC/MS in *SbMyb60* and wild-type plants.

Table S3. Agronomic data from wild-type and *SbMyb60* plants.

Table S4. Real-time-qPCR primers used to analyze expression levels of genes from the monolignol biosynthesis pathway.

Methods S1. High-efficiency thermal asymmetric interlaced PCR to determine location of transgene insertion site.

Methods S2. Construction of *SbMyb60* T-DNA.

REFERENCES

Akin, D.E. (1982) Forage cell wall degradation and ?-coumaric, ferulic, and sinapic acids. *Agron. J.* **74**, 424. Available at: https://dl.sciencesocieties.org/publications/aj/abstracts/74/3/AJ0740030424 [Accessed May 7, 2015].

Albert, N.W., Griffiths, A.G., Cousins, G.R., Verry, I.M. and Williams, W.M. (2015) Anthocyanin leaf markings are regulated by a family of R2R3-MYB genes in the genus *Trifolium*. *New Phytol.* **205**, 882–893. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25329638 [Accessed May 24, 2015].

Aydin, G., Grant, R.J. and O'Rear, J. (1999) Brown midrib sorghum in diets for lactating dairy cows. *J. Dairy Sci.* **82**, 2127–2135. Available at: http://www.sciencedirect.com/science/article/pii/S0022030299754561 [Accessed October 14, 2014].

el-Basyouni, S.Z., Chen, D., Ibrahim, R.K., Neish, A.C. and Towers, G.H.N. (1964) The biosynthesis of hydroxybenzoic acids in higher plants. *Phytochemistry*, **3**, 485–492. Available at: http://www.sciencedirect.com/science/article/pii/S0031942200880255 [Accessed April 7, 2015].

Besseau, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B. and Legrand, M. (2007) Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. *Plant Cell*, **19**, 148–162. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1820963&tool=pmcentrez&rendertype=abstract [Accessed May 8, 2015].

Boerjan, W., Ralph, J. and Baucher, M. (2003) Lignin biosynthesis. *Annu. Rev. Plant Biol.* **54**, 519–546. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14503002 [Accessed July 10, 2014].

Borevitz, J.O. (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell*, **12**, 2383–2394. Available at: http://www.plantcell.org/content/12/12/2383.full [Accessed December 8, 2014].

Boudet, A.M., Kajita, S., Grima-Pettenati, J. and Goffner, D. (2003) Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends Plant Sci.* **8**, 576–581. Available at: http://www.sciencedirect.com/science/article/pii/S1360138503002486 [Accessed May 12, 2015].

Byrt, C.S., Grof, C.P.L. and Furbank, R.T. (2011) C4 plants as biofuel feedstocks: optimising biomass production and feedstock quality from a lignocellulosic perspective. *J. Integr. Plant Biol.* **53**, 120–135. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21205189 [Accessed February 17, 2015].

Caño-Delgado, A.I., Metzlaff, K. and Bevan, M.W. (2000) The eli1 mutation reveals a link between cell expansion and secondary cell wall formation in *Arabidopsis thaliana*. *Development*, **127**, 3395–3405. Available at: http://dev.biologists.org/content/127/15/3395.short [Accessed October 5, 2015].

Cherney, J.H., Cherney, D.J.R., Akin, D.E. and Axtell, J.D. (1991) *Advances in Agronomy Volume 46*, Elsevier. Available at: http://www.sciencedirect.com/science/article/pii/S0065211308605805 [Accessed November 5, 2014].

Dai, X., Wang, Y., Yang, A. and Zhang, W.-H. (2012) OsMYB2P-1, an R2R3 MYB transcription factor, is involved in the regulation of phosphate-starvation responses and root architecture in rice. *Plant Physiol.* **159**, 169–183. Available at: http://www.plantphysiol.org/content/159/1/169.abstract [Accessed April 20, 2015].

Darriba, D., Taboada, G.L., Doallo, R. and Posada, D. (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*, **27**, 1164–1165. Available at: http://bioinformatics.oxfordjournals.org/content/27/8/1164.short [Accessed January 7, 2015].

Demirbas, A. (2000) Mechanisms of liquefaction and pyrolysis reactions of biomass. *Energy Convers. Manage.* **41**, 633–646. Available at: http://www.sciencedirect.com/science/article/pii/S0196890499001302 [Accessed January 21, 2015].

Dien, B.S., Sarath, G., Pedersen, J.F., Sattler, S.E., Chen, H., Funnell-Harris, D.L., Nichols, N.N. and Cotta, M.A. (2009) Improved sugar conversion and ethanol yield for forage sorghum (*Sorghum bicolor L. Moench*) lines with reduced lignin contents. *Bioenergy Res.* **2**, 153–164. Available at: http://link.springer.com/10.1007/s12155-009-9041-2 [Accessed September 2, 2014].

Dixon, R.A. and Paiva, N.L. (1995) Stress-induced phenylpropanoid metabolism. *Plant Cell*, **7**, 1085–1097. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=160915&tool=pmcentrez&rendertype=abstract [Accessed February 16, 2015].

Dixon, R.A., Lamb, C.J., Masoud, S., Sewalt, V.J.H. and Paiva, N.L. (1996) Metabolic engineering: prospects for crop improvement through the genetic manipulation of phenylpropanoid biosynthesis and defense responses — a review. *Gene*, **179**, 61–71. Available at: http://www.sciencedirect.com/science/article/pii/S0378111996003277 [Accessed May 24, 2015].

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797. Available at: http://nar.oxfordjournals.org/content/32/5/1792.short [Accessed July 11, 2014].

Goicoechea, M., Lacombe, E., Legay, S. *et al.* (2005) EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J.* **43**, 553–567. Available at: http://doi.wiley.com/10.1111/j.1365-313X.2005.02480.x [Accessed March 18, 2015].

Gómez-Vásquez, R., Day, R., Buschmann, H., Randles, S., Beeching, J.R. and Cooper, R.M. (2004) Phenylpropanoids, phenylalanine ammonia lyase and peroxidases in elicitor-challenged cassava (*Manihot esculenta*) suspension cells and leaves. *Ann. Bot.* **94**, 87–97. Available at: http://aob.oxfordjournals.org/content/94/1/87.short [Accessed January 31, 2015].

Grabber, J.H. and Lu, F. (2007) Formation of syringyl-rich lignins in maize as influenced by feruloylated xylans and p-coumaroylated monolignols. *Planta*, **226**, 741–751. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17457604 [Accessed May 24, 2015].

Grabber, J.H., Jung, G.A. and Hill, R.R. (1991) Chemical composition of parenchyma and sclerenchyma cell walls isolated from Orchardgrass and Switchgrass. *Crop Sci.* **31**, 1058. Available at: https://dl.sciencesocieties.org/publications/cs/abstracts/31/4/CS0310041058 [Accessed May 24, 2015].

Grand, C. (1984). Ferulic acid 5-hydroxylase: a new cytochrome P-450-dependent enzyme from higher plant microsomes involved in lignin synthesis. *FEBS Lett.* **169**, 7–11. Available at: http://www.sciencedirect.com/science/article/pii/0014579384802781 [Accessed October 5, 2015].

Grotewold, E., Drummond, B.J., Bowen, B. and Peterson, T. (1994) The myb-homologous P gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell*, **76**, 543–553. Available at: http://www.sciencedirect.com/science/article/pii/0092867494901171 [Accessed May 18, 2015].

Hajdukiewicz, P., Svab, Z. and Maliga, P. (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol. Biol.* **25**, 989–994. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7919218 [Accessed February 9, 2015].

Howe, A., Sato, S., Dweikat, I., Fromm, M. and Clemente, T. (2006) Rapid and reproducible *Agrobacterium*-mediated transformation of sorghum. *Plant Cell Rep.* **25**, 784–791. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16528567 [Accessed February 9, 2015].

Hu, W.J., Harding, S.A., Lung, J., Popko, J.L., Ralph, J., Stokke, D.D., Tsai, C.J. and Chiang, V.L. (1999) Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nat. Biotechnol.* **17**, 808–812. Available at: http://dx.doi.org/10.1038/11758 [Accessed August 19, 2014].

Ibraheem, F., Gaffoor, I., Tan, Q., Shyu, C.-R. and Chopra, S. (2015) A sorghum MYB transcription factor induces 3-deoxyanthocyanidins and enhances resistance against leaf blights in maize. *Molecules*, **20**, 2388–2404. Available at: http://www.mdpi.com/1420-3049/20/2/2388/htm [Accessed May 24, 2015].

Kay, R., Chan, A., Daly, M. and McPherson, J. (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science*, **236**, 1299–1302. Available at: http://www.sciencemag.org/content/236/4806/1299.short [Accessed October 5, 2015].

Knobloch, K.-H. and Hahlbrock, K. (1977) 4-Coumarate:CoA ligase from cell suspension cultures of *Petroselinum hortense* Hoffm. *Arch. Biochem. Biophys.* **184**, 237–248. Available at: http://www.sciencedirect.com/science/article/pii/0003986177903472 [Accessed February 10, 2015].

Li, X., Weng, J.-K. and Chapple, C. (2008) Improvement of biomass through lignin modification. *Plant J.* **54**, 569–581. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18476864 [Accessed May 24, 2015].

Marita, J.M., Ralph, J., Hatfield, R.D. and Chapple, C. (1999) NMR characterization of lignins in *Arabidopsis* altered in the activity of ferulate 5-hydroxylase. *Proc. Natl Acad. Sci. USA*, **96**, 12328–12332. Available at: http://www.pnas.org/content/96/22/12328.short [Accessed October 6, 2015].

McCarthy, R.L., Zhong, R. and Ye, Z.-H. (2009) *MYB83* is a direct target of SND1 and acts redundantly with *MYB46* in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell Physiol.* **50**, 1950–1964. Available at: http://pcp.oxfordjournals.org/content/50/11/1950.long [Accessed January 15, 2015].

McCarthy, R.L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., Spicer, C. and Ye, Z.-H. (2010) The poplar MYB transcription factors, *PtrMYB3* and *PtrMYB20*, are involved in the regulation of secondary wall biosynthesis. *Plant Cell Physiol.* **51**, 1084–1090. Available at: http://pcp.oxfordjournals.org/content/51/6/1084.full [Accessed March 18, 2015].

McKendry, P. (2002) Energy production from biomass (part 1): overview of biomass. *Bioresour. Technol.* **83**, 37–46. Available at: http://www.sci-

encedirect.com/science/article/pii/S0960852401001183 [Accessed November 20, 2014].

Meyer, K., Shirley, A.M., Cusumano, J.C., Bell-Lelong, D.A. and Chapple, C. (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. *Proc. Natl Acad. Sci. USA*, **95**, 6619–6623. Available at: http://www.pnas.org/content/95/12/6619.short [Accessed October 6, 2015].

Mihalcik, D.J., Mullen, C.A. and Boateng, A.A. (2011) Screening acidic zeolites for catalytic fast pyrolysis of biomass and its components. *J. Anal. Appl. Pyrolysis*, **92**, 224–232. Available at: http://www.sciencedirect.com/science/article/pii/S0165237011001094 [Accessed January 11, 2015].

Miller, F.R. and McBee, G.G. (1993) Genetics and management of physiologic systems of sorghum for biomass production. *Biomass Bioenergy*, **5**, 41–49. Available at: http://www.sciencedirect.com/science/article/pii/096195349390006P [Accessed April 2, 2015].

Moyano, E., Martínez-Garcia, J.F. and Martin, C. (1996) Apparent redundancy in myb gene function provides gearing for the control of flavonoid biosynthesis in antirrhinum flowers. *Plant Cell*, **8**, 1519–1532. Available at: http://www.plantcell.org/content/8/9/1519.abstract [Accessed May 24, 2015].

Mu, W., Ben, H., Ragauskas, A. and Deng, Y. (2013) Lignin pyrolysis components and upgrading—technology review. *Bioenergy Res.* **6**, 1183–1204. Available at: http://link.springer.com/article/10.1007/s12155-013-9314-7 [Accessed October 5, 2015].

Nair, R.B. (2004) The *Arabidopsis thaliana* REDUCED EPIDERMAL FLUORESCENCE1 gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. *Plant Cell*, **16**, 544–554. Available at: http://www.plantcell.org/content/16/2/544.long [Accessed April 7, 2015].

Oliver, A.L., Grant, R.J., Pedersen, J.F. and O'Rear, J. (2004) Comparison of brown midrib-6 and -18 forage sorghum with conventional sorghum and corn silage in diets of lactating dairy cows. *J. Dairy Sci.* **87**, 637–644. Available at: http://www.sciencedirect.com/science/article/pii/S0022030204732063 [Accessed October 8, 2014].

Ostlund, G., Schmitt, T., Forslund, K., Köstler, T., Messina, D.N., Roopra, S., Frings, O. and Sonnhammer, E.L.L. (2010) InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res.* **38**, D196–D203. Available at: http://nar.oxfordjournals.org/content/38/suppl_1/D196.short [Accessed January 27, 2015].

Palanichelvam, K., Oger, P., Clough, S.J., Cha, C., Bent, A.F. and Farrand, S.K. (2000) A second T-region of the soybean-supervirulent chrysopine-type Ti plasmid pTiChry5, and construction of a fully disarmed vir helper plasmid. *Mol. Plant Microbe Interact.* **13**, 1081–1091. Available at: http://apsjournals.apsnet.org/doi/abs/10.1094/MPMI.2000.13.10.1081 [Accessed February 9, 2015].

Palmer, N.A., Sattler, S.E., Saathoff, A.J., Funnell, D., Pedersen, J.F. and Sarath, G. (2008) Genetic background impacts soluble and cell wall-bound aromatics in brown midrib mutants of sorghum. *Planta*, **229**, 115–127. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18795321 [Accessed June 10, 2014].

Piquemal, J., Chamayou, S., Nadaud, I. *et al.* (2002) Down-regulation of caffeic acid o-methyltransferase in maize revisited using a transgenic approach. *Plant Physiol.* **130**, 1675–1685. Available at: http://www.plantphysiol.org/content/130/4/1675.long [Accessed October 6, 2015].

Remm, M., Storm, C.E. and Sonnhammer, E.L. (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* **314**, 1041–1052. Available at: http://www.sciencedirect.com/science/article/pii/S0022283600951970 [Accessed January 5, 2015].

Rolando, C., Monties, B. and Lapierre, C. (1992). Thioacidolysis. In *Methods in Lignin Chemistry* (Stephen Y. Lin and Carlton W ed). Berlin: Springer, pp. 334–349. Available at: http://link.springer.com/chapter/10.1007%2F978-3-642-74065-7_23# [Accessed October 5, 2015].

Rooney, W.L., Blumenthal, J., Bean, B. and Mullet, J.E. (2007) Designing sorghum as a dedicated bioenergy feedstock. *Biofuels, Bioprod. Biorefin.* **1**, 147–157. Available at: http://doi.wiley.com/10.1002/bbb.15 [Accessed January 20, 2015].

Rosler, J. (1997) Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. *Plant Physiol.* **113**, 175–179. Available at: http://www.plantphysiol.org/content/113/1/175.abstract [Accessed May 24, 2015].

Saballos, A., Sattler, S.E., Sanchez, E., Foster, T.P., Xin, Z., Kang, C., Pedersen, J.F. and Vermerris, W. (2012) Brown midrib2 (Bmr2) encodes the major 4-coumarate:coenzyme A ligase involved in lignin biosynthesis in sorghum (*Sorghum bicolor (L.) Moench*). *Plant J.* **70**, 818–830. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22313236 [Accessed September 14, 2014].

Sarath, G., Dien, B., Saathoff, A.J., Vogel, K.P., Mitchell, R.B. and Chen, H. (2011). Ethanol yields and cell wall properties in divergently bred switchgrass genotypes. *Bioresour. Technol.* **102**, 9579–9585. Available at: www.sciencedirect.com/science/article/pii/S0960852411010376 [Accessed October 5, 2015].

Sattler, S.E., Saathoff, A.J., Haas, E.J., Palmer, N.A., Funnell-Harris, D.L., Sarath, G. and Pedersen, J.F. (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the sorghum brown midrib6 phenotype. *Plant Physiol.* **150**, 584–595. Available at: http://www.plantphysiol.org/content/150/2/584.full [Accessed June 10, 2014].

Sattler, S.E., Funnell-Harris, D.L. and Pedersen, J.F. (2010) Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. *Plant Sci.* **178**, 229–238. Available at: http://www.sciencedirect.com/science/article/pii/S0168945210000051 [Accessed September 10, 2014].

Sattler, S.E., Palmer, N.A., Saballos, A., Greene, A.M., Xin, Z., Sarath, G., Vermerris, W. and Pedersen, J.F. (2012) Identification and characterization of four missense mutations in Brown midrib 12 (Bmr12), the caffeic *o*-methyltranferase (COMT) of sorghum. *Bioenergy Res.* **5**, 855–865. Available at: http://link.springer.com/10.1007/s12155-012-9197-z [Accessed September 29, 2014].

Shakoor, N., Nair, R., Crasta, O., Morris, G., Feltus, A. and Kresovich, S. (2014) A *Sorghum bicolor* expression atlas reveals dynamic genotype-specific expression profiles for vegetative tissues of grain, sweet and bioenergy sorghums. *BMC Plant Biol.* **14**, 35. Available at: http://www.biomedcentral.com/1471-2229/14/35 [Accessed May 11, 2015].

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, D., Templeton, D. and Crocker, D. (2008) *Determination of Structural Carbohydrates and Lignin in Biomass*. Golden, CO: National Renewable Energy Laboratory.

Stracke, R., Werber, M. and Weisshaar, B. (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* **4**, 447–456. Available at: http://www.sciencedirect.com/science/article/pii/S1369526600001990 [Accessed January 21, 2015].

Tamagnone, L. (1998) The *AmMYB308* and *AmMYB330* transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell*, **10**, 135–154. Available at: http://www.plantcell.org/content/10/2/135.short [Accessed April 2, 2015].

Vanholme, R., Cesarino, I., Rataj, K. *et al.* (2013) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science*, **341**, 1103–1106. Available at: http://www.sciencemag.org/content/341/6150/1103.abstract [Accessed January 21, 2015].

Vogel, K.P., Pedersen, J.F., Masterson, S.D. and Toy, J.J. (1999) Evaluation of a filter bag system for NDF, ADF, and IVDMD forage analysis. *Crop Sci.* **39**, 276. Available at: https://dl.sciencesocieties.org/publications/cs/abstracts/39/1/CS0390010276 [Accessed February 14, 2015].

Vogel, K.P., Mitchell, R.B., Sarath, G., Jung, H.G., Dien, B.S. and Casler, M.D. (2013) Switchgrass biomass composition altered by six generations of divergent breeding for digestibility. *Crop Sci.* **53**, 853–862. Available at: https://dl.sciencesocieties.org/publications/cs/abstracts/53/3/853 [Accessed October 5, 2015].

Whetten, R. and Sederoff, R. (1995) Lignin biosynthesis. *Plant Cell*, **7**, 1001–1013. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=160901&tool=pmcentrez&rendertype=abstract [Accessed October 22, 2014].

Xie, S., Qin, X., Cheng, Y. *et al.* (2015) Simultaneous conversion of all cell wall components by an oleaginous fungus without chemi-physical pretreatment. *Green Chem.* **17**, 1657–1667. Available at: http://pubs.rsc.org/en/content/articlehtml/2015/gc/c4gc01529k [Accessed May 21, 2015].

Yaman, S. (2004) Pyrolysis of biomass to produce fuels and chemical feedstocks. *Energy Convers. Manage.* **45**, 651–671. Available at: http://www.sciencedirect.com/science/article/pii/S0196890403001778 [Accessed July 9, 2014].

Yang, A., Dai, X. and Zhang, W.-H. (2012) A R2R3-type MYB gene, *OsMYB2*, is involved in salt, cold, and dehydration tolerance in rice. *J. Exp. Bot.* **63**, 2541–2556. Available at: http://jxb.oxfordjournals.org/content/early/2012/02/02/jxb.err431.abstract [Accessed May 24, 2015].

Yilmaz, A., Nishiyama, M.Y., Fuentes, B.G., Souza, G.M., Janies, D., Gray, J. and Grotewold, E. (2009) GRASSIUS: a platform for comparative regulatory genomics across the grasses. *Plant Physiol.* **149**, 171–180. Available

at: http://www.plantphysiol.org/content/149/1/171.short [Accessed April 23, 2015].

Zhong, R. and Ye, Z.-H. (2009) Transcriptional regulation of lignin biosynthesis. *Plant Signal. Behav.* **4**, 1028–1034. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2819510&tool=pmcentrez&rendertype=abstract [Accessed November 20, 2014].

Zhou, J., Lee, C., Zhong, R. and Ye, Z.-H. (2009) *MYB58* and *MYB63* are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell*, **21**, 248–266. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2648072&tool=pmcentrez&rendertype=abstract [Accessed January 5, 2015].

Zwickl, D.J. (2006) *Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets Under the Maximum Likelihood Criterion*. Austin, TX: The University of Texas at Austin.

