2017

Overexpression of *SbMyb60* in *Sorghum bicolor* impacts both primary and secondary metabolism

Erin D. Scully
USDA-ARS, erin.scully@ars.usda.gov

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

Nathan A. Palmer
USDA-ARS, nathan.palmer@ars.usda.gov

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

Deanna L. Funnell-Harris
USDA-ARS, Deanna.Funnell-Harris@ars.usda.gov

See next page for additional authors

New Phytologist

Research

Overexpression of *SbMyb60* in *Sorghum bicolor* impacts both primary and secondary metabolism

Erin D. Scully[1,2,3], Tammy Gries[1], Nathan A. Palmer[1], Gautam Sarath[1,2], Deanna L. Funnell-Harris[1,4], Lisa Baird[5], Paul Twigg[6], Javier Seravalli[7], Thomas E. Clemente[2,8] and Scott E. Sattler[1,2]

[1]Wheat, Sorghum, and Forage Research Unit, USDA-ARS, Lincoln, NE 68583, USA; [2]Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583, USA; [3]Stored Product Insect and Engineering Research Unit, USDA-ARS, Manhattan, KS 66502, USA; [4]Department of Plant Pathology, University of Nebraska-Lincoln, Lincoln, NE 68583, USA; [5]Department of Biology, Shiley Center for Science and Technology, University of San Diego, San Diego, CA 92110, USA; [6]Biology Department, University of Nebraska-Kearney, Kearney, NE 68849, USA; [7]Redox Biology Center and Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE 68583, USA; [8]Center for Plant Science Innovation, University of Nebraska, Lincoln, NE 68588, USA

Author for correspondence:
Scott E. Sattler
Tel: +1 402 472 5987
Email: scott.sattler@ars.usda.gov

Received: *29 June 2017*
Accepted: *23 August 2017*

New Phytologist (2017)
doi: 10.1111/nph.14815

Key words: aromatic amino acid biosynthesis, bioenergy, coexpression analysis, cofactors, lignocellulose, monolignol biosynthesis, phenylpropanoid metabolism, transcriptional activator.

Summary

- Few transcription factors have been identified in C_4 grasses that either positively or negatively regulate monolignol biosynthesis.
- Previously, the overexpression of *SbMyb60* in sorghum (*Sorghum bicolor*) has been shown to induce monolignol biosynthesis, which leads to elevated lignin deposition and altered cell wall composition. To determine how *SbMyb60* overexpression impacts other metabolic pathways, RNA-Seq and metabolite profiling were performed on stalks and leaves.
- *35S::SbMyb60* was associated with the transcriptional activation of genes involved in aromatic amino acid, *S*-adenosyl methionine (SAM) and folate biosynthetic pathways. The high coexpression values between *SbMyb60* and genes assigned to these pathways indicate that *SbMyb60* may directly induce their expression. In addition, *35S::SbMyb60* altered the expression of genes involved in nitrogen (N) assimilation and carbon (C) metabolism, which may redirect C and N towards monolignol biosynthesis. Genes linked to UDP-sugar biosynthesis and cellulose synthesis were also induced, which is consistent with the observed increase in cellulose deposition in the internodes of *35S::SbMyb60* plants. However, *SbMyb60* showed low coexpression values with these genes and is not likely to be a direct regulator of cell wall polysaccharide biosynthesis.
- These findings indicate that *SbMyb60* can activate pathways beyond monolignol biosynthesis, including those that synthesize the substrates and cofactors required for lignin biosynthesis.

Introduction

Worldwide, considerable efforts are being focused on the development of renewable resources for the production of energy, biofuels and bioproducts, which would reduce the reliance on fossil fuels and lead to a decrease in greenhouse gas emissions. Lignocellulosic biomass is one of the largest forms of renewable carbon (C) that could be exploited for these purposes; however, the quality of lignocellulosic biomass and its utility for downstream processes are largely determined by the abundance of lignin (Boerjan *et al.*, 2003). Biomass with decreased lignin content is preferable for biochemical conversion into fuels (Sun & Cheng, 2002), whereas biomass with higher levels of lignin and/or other cell wall phenolic compounds is preferred for thermochemical and green chemistry processes (Tuck *et al.*, 2012; Luterbacher *et al.*, 2014).

The subunits of the lignin polymer are derived from the amino acid L-phenylalanine through the monolignol biosynthesis pathway, which produces sinapyl, coniferyl and *p*-coumaryl alcohols through a series of 10 (monocots) or more (dicots) enzymatic reactions (Vanholme *et al.*, 2013). These monolignols are ultimately polymerized into syringyl (S-lignin), guaiacyl (G-lignin) and *p*-hydroxyphenyl (H-lignin) units, respectively, through oxidative processes catalyzed by both laccases and peroxidases (Boerjan *et al.*, 2003). The regulators of the monolignol biosynthesis pathway have been well characterized in Arabidopsis and other dicots, which largely include members of the MYB, WRKY and NAC transcription factor (TF) families (Wang *et al.*, 2010; Zhao & Dixon, 2011). In these plants, the entire monolignol biosynthesis pathway, with the exception of *ferulate 5-hydroxylase* (*F5H*), is positively regulated by two TFs, AtMyb58 and AtMyb63 (Zhao *et al.*, 2010). In turn, *AtMyb58* and *AtMyb63* are directly regulated by AtMyb46 and AtMyb83 (McCarthy *et al.*, 2009), which regulate the entire secondary cell wall biosynthesis program, including the genes required for the

biosynthesis of cellulose and hemicellulose. *SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1* (*SND1*) (Zhong *et al.*, 2006), a NAC TF, regulates *AtMyb46* and *AtMyb83* (McCarthy *et al.*, 2009), including *F5H* (Zhao *et al.*, 2010), and functions as one of the top tier positive transcriptional activators of secondary wall formation pathways. Other global regulators of secondary cell wall biosynthesis in dicots include *NST1*, *NST2*, *NST3*, *VND6* and *VND7* NAC TFs (Kubo *et al.*, 2005; Mitsuda *et al.*, 2005, 2007). In addition, negative correlations between the expression of monolignol biosynthesis genes and *cellulose synthase A* (*CesA*) have been observed in *4-coumarate:coenzyme A ligase* (*4CL*)-downregulated *Populus tremuloides*, which suggests that negative cross-talk between cellulose and lignin biosynthesis occurs in some plant species (Hu *et al.*, 1999). Although master regulators of secondary cell wall development have not yet been identified in monocots (Zhong & Ye, 2009), several positive and negative regulators of various branches of the secondary cell wall biosynthesis pathways have been identified and characterized (Sonbol *et al.*, 2009; Zhong & Ye, 2009; Fornalé *et al.*, 2012; Ibraheem *et al.*, 2015). Notably, monocot TFs can have different targets and different impacts on secondary cell wall biosynthesis relative to their dicot orthologs. For example, in *Oryza sativa* (rice), overexpression of *OsMyb58* and *OsMyb63* directly activates *CesA*, unlike in Arabidopsis (Noda *et al.*, 2015). Other variations in the regulation of secondary wall biosynthesis are likely to exist between monocots and dicots because of the extensive differences in secondary cell wall composition and plant anatomy.

Although comparatively less is known about the regulation of cell wall biosynthesis in monocots, grasses are commonly grown for forage and are being developed as lignocellulosic bioenergy feedstocks in the USA and throughout the world. Sorghum (*Sorghum bicolor* (L.) Moench) is a climate-resilient C_4 grass that can be grown on marginal lands and requires lower inputs of nitrogen (N) fertilizer than other crops (Rooney, 2004). Furthermore, biomass yields remain stable across a range of environments, including periods of water deficiency (Miller & McBee, 1993). In sorghum, mutations in *brown midrib* (*bmr*) loci are associated with reduced lignin content and enhanced yields from various conversion processes (Aydin *et al.*, 1999; Oliver *et al.*, 2004; Dien *et al.*, 2009; Sattler *et al.*, 2010; Xie *et al.*, 2015). These *Bmr* genes (*Bmr6*, *Bmr12* and *Bmr2*) encode three enzymes in the monolignol biosynthesis pathway: cinnamyl alcohol dehydrogenase (CAD), caffeic acid *O*-methyltransferase (COMT) and 4CL, respectively (Bout & Vermerris, 2003; Sattler *et al.*, 2009; Saballos *et al.*, 2012). Recently, *SbMyb60* has been identified as a positive regulator of the monolignol biosynthesis pathway (Scully *et al.*, 2016). The overexpression of this gene *in planta* activates the entire monolignol biosynthesis pathway, including *F5H*, which results in ectopic lignin deposition and altered lignin and secondary cell wall composition in comparison with wild-type (WT) plants (Scully *et al.*, 2016).

Despite the potential utility of *SbMyb60* in the manipulation of lignin composition in sorghum, little is known about the identity of other positive regulators of monolignol and phenylpropanoid biosynthesis in sorghum or other C_4 bioenergy grasses, and their impacts on cell wall composition. In addition, the impacts of *SbMyb60* and its grass orthologs on other primary and secondary metabolic pathways are not known in sorghum and most other monocots. For example, *SbMyb60* overexpression has the potential to impact primary metabolic pathways, because monolignol biosynthesis requires aromatic amino acids and three cofactors (*S*-adenosyl methionine (SAM), nicotinamide adenine dinucleotide phosphate (NADPH) and coenzyme A (CoA)). The alteration of the expression of genes associated with the monolignol biosynthesis pathway may also affect other metabolic pathways that share these substrates. In maize (*Zea mays*), *Brown midrib* (*Bm*) *2* and *4* loci encode gene products involved in SAM metabolism (Tang *et al.*, 2014; Li *et al.*, 2015). Both *bm2* and *bm4* plants showed modest reductions in lignin (Tang *et al.*, 2014; Li *et al.*, 2015), which demonstrates the importance of this cofactor in lignin biosynthesis. In addition, increased levels of aromatic amino acids have been observed during secondary cell wall formation (Ohtani *et al.*, 2016), suggesting that aromatic amino acid biosynthesis pathways might be coregulated with lignin biosynthesis. To understand how *SbMyb60* overexpression translates into changes in cell wall composition, metabolism and the plant ideotype, RNA-Seq and metabolite profiling were performed on sorghum stalks and leaves from *SbMyb60*-overexpressing plants. The goal was to discover potential developmental and metabolic pathways influenced by differential levels of *SbMyb60* expression.

Materials and Methods

Global gene expression analysis using RNA-Seq

Ten independent overexpression lines were generated by transgenetic insertion of a construct containing *SbMyb60* under the control of the E35S CaMV promoter, as described previously (Scully *et al.*, 2016). Three transformation events with the highest levels of *SbMyb60* overexpression, designated Myb10a, Myb15a and Myb2a, were selected for further characterization. RNA was extracted and purified from leaf and stalk tissue from 6-wk-old glasshouse-grown plants, as described previously ($n = 3$ biological replicates per line) (Scully *et al.*, 2016). PolyA RNA-Seq libraries were prepared and differential expression analysis was performed using edgeR (Supporting Information Methods S1). Weighted gene coexpression network analysis (WGCNA), gene ontology (GO) enrichment and KEGG pathway analysis were also performed (Methods S1). Raw Illumina reads are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive under BioProject PRJNA327395.

Analysis of wall-bound phenolics using GC/MS

Wall-bound phenolics were isolated from whole plants as described previously (Palmer *et al.*, 2008) and analyzed by integrated GC/MS equipped with a 7890B series gas chromatograph and a 5977A network mass spectrometer (Methods S1).

Metabolite analysis using LC/MS

Targeted analysis of polar metabolites of stalk and leaf samples ($n = 3$ biological replicates) from all four lines was performed by multiple reaction monitoring (LC-MRM-MS) analysis. Samples were extracted as described previously (Palmer *et al.*, 2014) from 70–90 mg of ground leaf or stalk tissue in 300 µl of 80% methanol spiked with 2.13 µM $^{13}C_5$-^{15}N-proline as an internal standard. LC/MS analysis was performed on a 4000 QTRAP hybrid triple quadrupole ion trap mass spectrometer (Sciex, Framingham, MA, USA), operating in either positive or negative ionization mode (Methods S1).

Light microscopy and scanning electron microscopy

Tissue from sorghum stalks was collected from the first internode below the peduncle from WT and *35S::SbMyb60* plants at the boot stage *c.* 8 wk post-germination. Internodes were treated with phloroglucinol to stain lignin (red) and FASGA (Spanish abbreviation for fuschin, alcian blue, safranin, glycerin and water) to stain lignified (red) and non-lignified (blue) cell walls for light microscopy, and dehydrated in an ethanol series for scanning electron microscopy (SEM) (Methods S1).

Statistical analysis

Metabolite concentrations were analyzed using sparse partial least-squares discriminant analysis (sPLS-DA) (Lê Cao *et al.*, 2011) employing MetaboAnalyst 3.0 (Xia *et al.*, 2015). Metabolites with high loading values were statistically analyzed using ANOVA followed by *F*-protected least-significant difference (LSD) *post hoc* comparisons. Levels of wall-bound phenolics and internode measurements were analyzed using SAS PROC MIXED v.9.2 (SAS, 2002–2008). Heterogeneous variances were addressed in PROC MIXED, and an *F*-protected LSD with an estimated experiment-wise error rate of $\alpha = 0.05$ was performed for *post hoc* comparisons.

Results

SbMyb60 overexpression impacts global gene expression and metabolite profiles

Overall, no major differences in read quality or the number of mapped reads between tissues or between the four lines were observed. Non-metric multidimensional scaling (NMDS) analysis clearly differentiated the expression profiles of *35S::SbMyb60* leaves and stalks from their respective WT tissues (Fig. 1a,b). Further, it was also apparent that the expression profiles of the three overexpression lines differed from one another in both tissues (Fig. 1a,b), which may be a result of the effects of different levels of *35S::SbMyb60* expression on global gene expression. In addition, the metabolite profiles from *35S::SbMyb60* leaves and stems were also clearly separated from their WT counterparts via sPLS-DA (Fig. 1c,d). In general, more variability among the metabolite profiles was observed between biological replicates of leaves (Fig. 1c) relative to stalks (Fig. 1d). Despite this variability, over 30 metabolites were differentially abundant in *35S::SbMyb60* leaves relative to WT, whereas over 100 metabolites were differentially abundant in *35S::SbMyb60* stalks relative to WT. Major metabolites impacted similarly in both tissues included the glycolysis intermediate glucose-1-phosphate, which was elevated in both tissues relative to WT, and *trans-trans*-farnesyl diphosphate and 2-ketoisoglutarate, which were lower in both tissues relative to WT. Although *35S::SbMyb60* overexpression affected some metabolites similarly in both tissues, others displayed strong differences between tissues. For example, metabolites that distinguished *35S::SbMyb60* leaves from WT leaves in sPLS-DA included phosphoenolpyruvate (higher in *35S::SbMyb60*), indole (higher in *35S::SbMyb60*), allantoin (lower in *35S::SbMyb60*), D-glucarate (higher in *35S::SbMyb60*) and 5,10-methylenetetrahydrofolate (lower in *35S::SbMyb60*), whereas glutamate (higher in *35S::SbMyb60*), glutathione (higher in *35S::SbMyb60*), ornithine (highest in Myb15a) and sedoheptulose-1,7-bisphosphate (lower in *35S::SbMyb60*) distinguished *35S::SbMyb60* stalks from WT stalks.

Overexpression of *SbMyb60* has distinct impacts on the transcriptomes of stalks and leaves

Compared with WT, combined totals of 1820 and 1922 genes were upregulated and downregulated, respectively, in leaves from the three *SbMyb60* overexpression lines, whereas, in stalks, combined totals of 6523 and 5162 genes were upregulated and downregulated, respectively. More differentially expressed genes (DEGs) were identified in stalk tissues relative to leaves (Fig. 1e). In general, the number of DEGs in leaves and stalks correlated with the expression level of *SbMyb60*, with Myb10a having the highest overall number of DEGs and the strongest differential expression levels (highest/lowest expression levels) of the three overexpression lines. In leaves, 312 up- and 130 downregulated genes were common to all three transgenic lines, whereas, in stalks, 2735 up- and 899 downregulated genes were common to all three events (Fig. 1e). In addition, *35S::SbMyb60* appeared to have different impacts on gene expression profiles in stalks and leaves. The majority of the DEGs identified in this study were differentially expressed in only one tissue, whereas very few DEGs were differentially expressed in both stalk and leaf tissues (< 10%) (Fig. 1f).

WGCNA further illustrated the differential impacts of *35S::SbMyb60* overexpression on sorghum stalks and leaves. In total, 13 coexpression modules were identified (Fig. S1), which could be divided into three primary groupings. Module 1 (M1), M3 and M4 consisted of genes with altered expression (either up- or downregulation) in *35S::SbMyb60* stalks relative to WT, whereas M2 and M6 consisted of genes with altered expression levels in *35S::SbMyb60* leaves relative to WT. M10, M11 and M13 included genes that were induced in both stalks and leaves in an event-specific manner (Fig. 2a,b).

Pathways involved in primary and secondary metabolism are impacted by *SbMyb60* overexpression

GO enrichment analysis showed that the impacts of *SbMyb60* overexpression extended far beyond the monolignol biosynthesis

New Phytologist

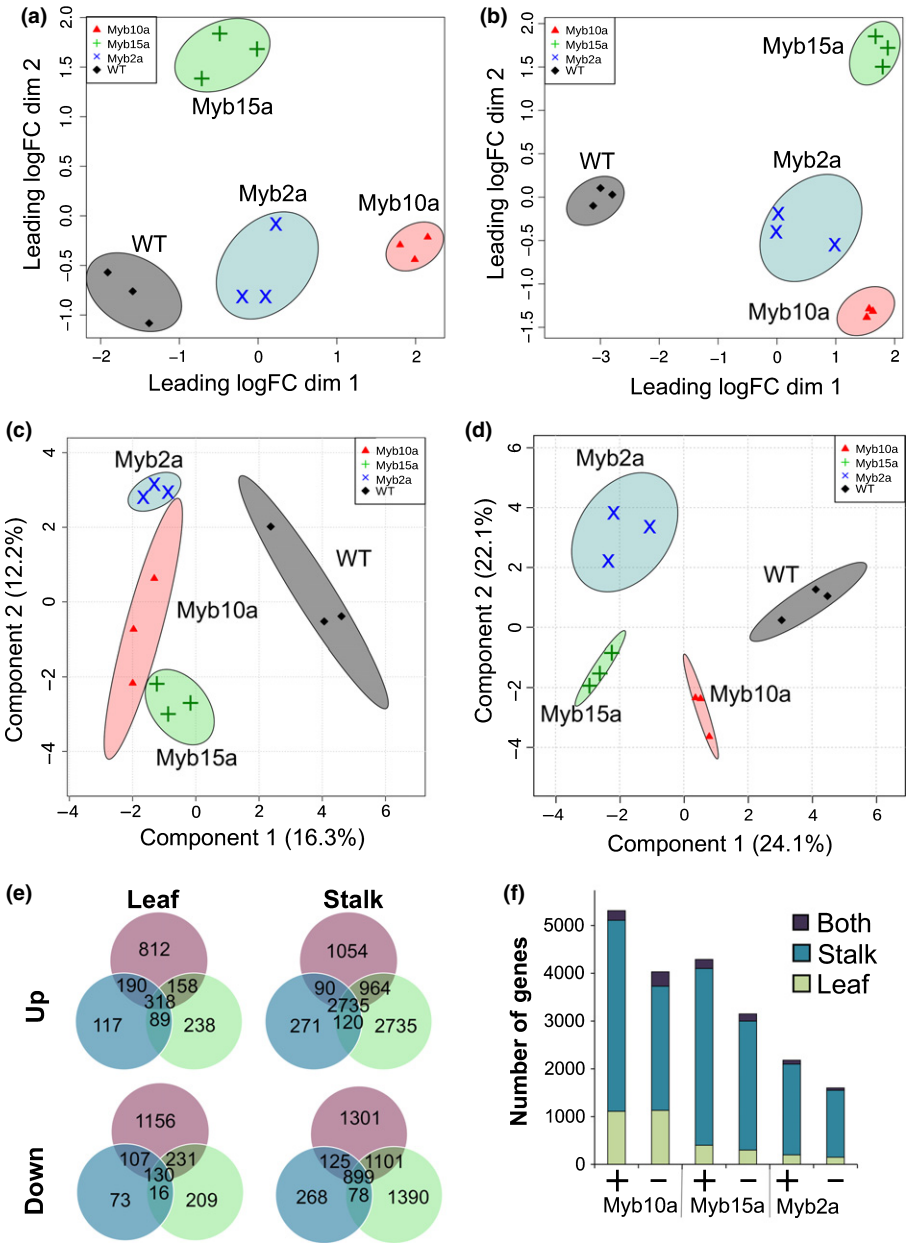
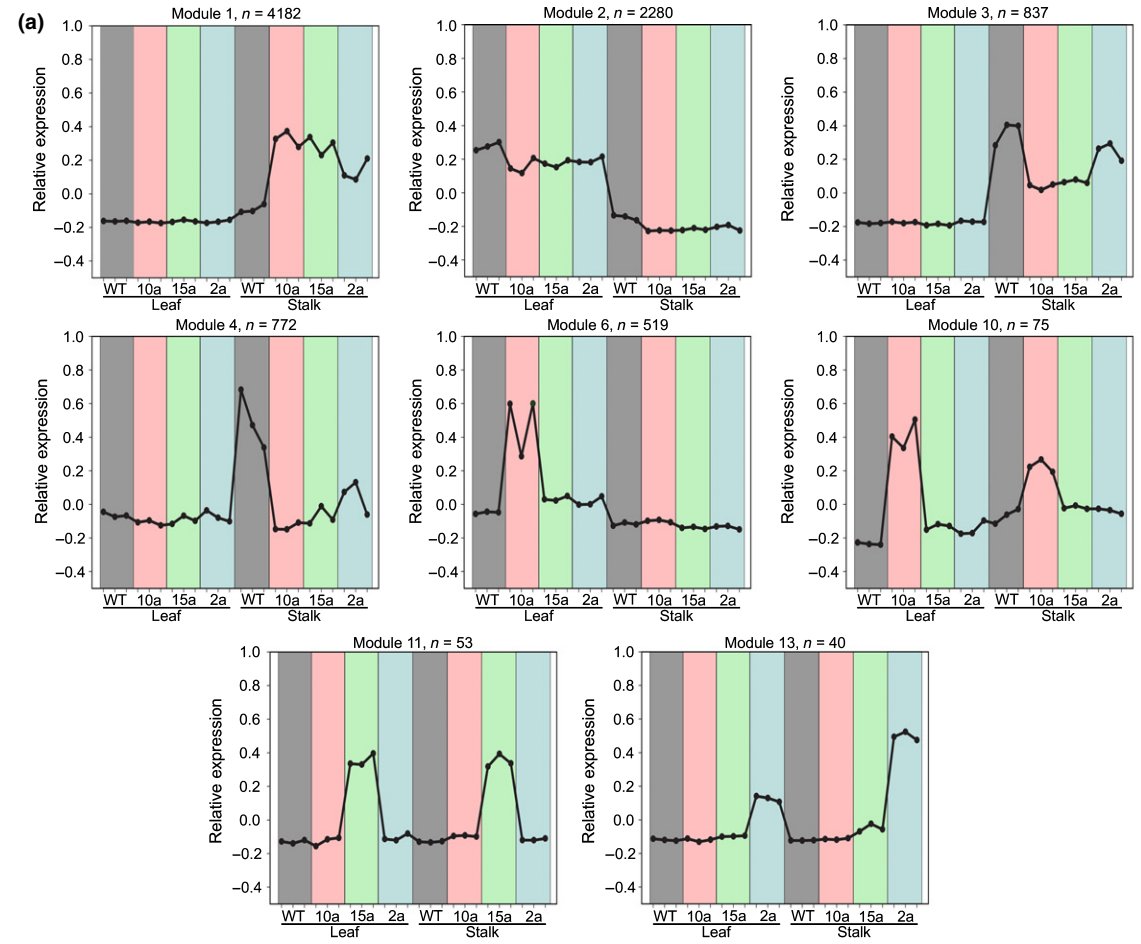
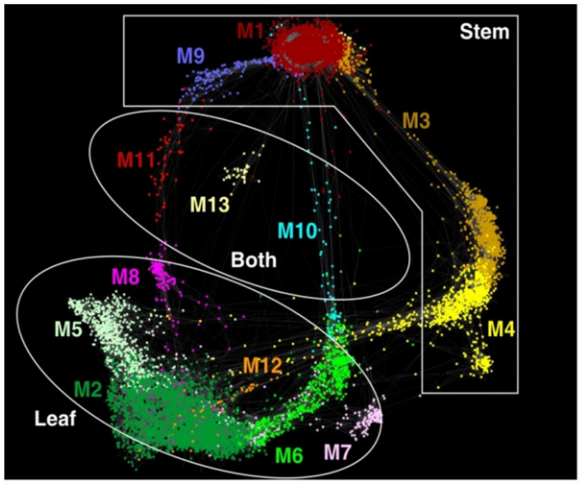


Fig. 1 RNA-Seq and metabolite profiles of sorghum leaves and stalks from wild-type (WT) and *35S::SbMyb60* plants. Non-metric multidimensional scaling (NMDS) analysis of RNA-Seq profiles from *35S::SbMyb60* and WT leaves (a) and stalks (b). RNA-Seq counts were normalized using the trimmed mean (TMM) method and genes with ≤1 count per million in at least three samples were removed from the analysis. NMDS plots for stalks and leaves were prepared separately using the 'plotMDS.DGEList' command in edgeR. Sparse partial least-squares discriminant analysis (sPLS-DA) of metabolite profiles from *35S::SbMyb60* and WT leaves (c) and stalks (d). Concentrations of metabolites detected during LC/MS positive and negative ionization modes were determined as described in the Materials and Methods section. Metabolite concentrations were normalized by sample median, log-transformed and scaled by mean centering. sPLS-DA plots were prepared for stalks and leaves separately using MetaboAnalyst 3.0. (e) Venn diagram of shared and unique upregulated (top) and downregulated (bottom) genes between leaves (left) and stalks (right) between each of the *35S::SbMyb60* overexpression lines and WT. Genes were defined as differentially expressed in each *35S::SbMyb60* event if the false discovery rate (FDR) was ≤0.05 and the absolute fold change (FC) was ≥2.0 relative to WT using a Fisher's exact test in edgeR. Purple, Myb10a; green, Myb15a; blue, Myb2a. (f) Stacked bar graph indicating the number of upregulated and downregulated genes unique to stalks and leaves and shared between stalks and leaves in each of the three *35S::SbMyb60* lines. Differentially expressed genes were identified as described in (e). Green, genes differentially expressed in the leaf only; blue, genes differentially expressed in the stalk only; purple, genes differentially expressed in both tissues. +, upregulated genes in each line; −, downregulated genes.

pathway and affected many primary and secondary metabolic pathways in both stems and leaves. For example, several categories associated with genes linked to secondary cell wall polysaccharide metabolism, DNA replication and translation were enriched in M1, whereas M6 was enriched in categories linked with genes involved in aromatic amino acid biosynthesis (Table 1). GO terms associated with M3 and M4 were largely linked to the oxidative stress response, whereas terms enriched in





(c)

	M1	M2	M3	M4	M6
α-Linolenic acid metabolism	0.3	1.2	1.2	3.2	0.0
Amino sugar and nucleotide sugar metabolism	3.4	0.3	3.5	2.1	1.3
Base excision repair	2.8	0.0	1.2	0.0	0.6
Biosynthesis of amino acids	4.5	3.4	3.5	2.1	10.8
Cell cycle	5.0	0.0	0.0	0.0	0.0
Circadian rhythm - plant	0.0	2.2	3.5	3.2	0.0
Cysteine and methionine metabolism	1.9	1.5	3.5	4.3	3.2
Diterpenoid biosynthesis	0.0	0.0	2.3	3.2	0.0
DNA replication	4.9	0.0	0.0	0.0	0.0
Glycine, serine and threonine metabolism	1.3	0.6	0.0	0.0	3.8
Glycolysis/Gluconeogenesis	2.7	3.1	1.2	4.3	1.3
Mismatch repair	3.0	0.0	0.0	0.0	0.0
Nitrogen metabolism	0.0	1.8	1.2	0.0	2.5
Nucleotide excision repair	2.5	0.0	0.0	0.0	0.0
One carbon pool by folate	0.2	0.3	0.0	0.0	2.5
Oxidative phosphorylation	1.1	2.5	0.0	0.0	3.2
Pentose phosphate pathway	1.1	1.8	1.2	1.1	3.2
Peroxisome	1.4	2.2	0.0	3.2	0.6
Phenylalanine metabolism	0.5	1.5	1.2	0.0	3.8
Phenylalanine, tyrosine and tryptophan biosynthesis	0.6	0.3	2.3	0.0	4.4
Phenylpropanoid biosynthesis	2.5	2.8	5.8	7.4	6.3
Plant hormone signal transduction	1.6	4.6	3.5	9.6	0.0
Plant-pathogen interaction	0.5	0.6	1.2	6.4	0.6
Purine metabolism	4.7	1.5	3.5	0.0	1.3
Pyrimidine metabolism	4.9	0.9	0.0	1.1	0.6
Ribosome biogenesis in eukaryotes	2.8	0.0	0.0	1.1	0.0
Starch and sucrose metabolism	4.4	2.5	5.8	1.1	1.3
Zeatin biosynthesis	0.3	0.3	3.5	0.0	0.0

Fig. 2 Gene network and KEGG pathway analysis of differentially expressed genes in sorghum *35S::SbMyb60* stalks and leaves. (a) Expression patterns of genes assigned to each coexpression module. The expression patterns of the selected coexpression modules were summarized using the module eigengene. Module 1 (M1) and M6 represent genes upregulated in *SbMyb60* stalks and leaves, respectively, whereas M3 and M4 are downregulated in *SbMyb60* stalks. M2 predominantly contains genes downregulated in *SbMyb60* leaves. WT, wild-type. (b) Network analysis based on coexpression analysis. Genes are color coded based on module membership. Genes with differential expression in stalks can be found in the upper and right portions of the chart (M1, M3, M4 and M9), whereas genes with differential expression in leaves are found primarily in the lower left quadrant (M2, M5, M6, M7, M8 and M12). Genes with differential expression in both leaves and stalks are found in the middle of the network (M10, M11 and M13). (c) Highly represented KEGG pathways in selected coexpression modules. Genes grouped to M1, M2, M3, M4 and M6 were assigned to KEGG pathways using the sorghum KEGG orthology (KO) assignments from Phytozome and the KEGG Pathway Mapper Tool. Values represent the relative abundance of the number of genes assigned to each KEGG pathway within each coexpression module. Blue, low relative abundance (low percentage of genes assigned to a particular KEGG term; 0–0.6%); yellow, moderate relative abundance (0.7–2.5%); orange/red, high relative abundance (> 2.6%).

Table 1 Gene ontology enrichment analysis of *35S::SbMyb60* sorghum stalks and leaves

Category	#DE	#Cat	Term	Ont	FDR
Module 1					
GO:0003735	119	304	Structural constituent of ribosome	MF	1.06E-35
GO:0005840	114	294	Ribosome	CC	5.75E-34
GO:0006412	114	296	Translation	BP	1.07E-33
GO:0003777	37	51	Microtubule motor activity	MF	1.40E-15
GO:0007018	37	51	Microtubule-based movement	BP	1.40E-15
GO:0004553	93	271	Hydrolase activity, hydrolyzing *O*-glycosyl compounds	MF	1.20E-14
GO:0005975	114	403	Carbohydrate metabolic process	BP	1.27E-11
GO:0006260	28	50	DNA replication	BP	9.65E-09
GO:0003677	199	1076	DNA binding	MF	0.000829
GO:0003887	13	24	DNA-directed DNA polymerase activity	MF	0.00197
GO:0005507	32	121	Copper ion binding	MF	0.00233
GO:0016818	12	23	Hydrolase activity, acting on phosphorus anhydrides	MF	0.00467
GO:0005618	25	82	Cell wall	CC	0.00470
GO:0015995	8	12	Chlorophyll biosynthetic process	BP	0.00588
GO:0005085	8	12	Guanyl-nucleotide exchange factor activity	MF	0.00592
GO:0016851	8	12	Magnesium chelatase activity	MF	0.00635
GO:0005088	7	11	Ras guanyl-nucleotide exchange factor activity	MF	0.0204
GO:0005089	7	11	Rho guanyl-nucleotide exchange factor activity	MF	0.0204
GO:0004003	6	8	ATP-dependent DNA helicase activity	MF	0.0204
GO:0006073	12	32	Cellular glucan metabolic process	BP	0.0204
GO:0016762	12	32	Xyloglucan: xyloglucosyl transferase activity	MF	0.0204
GO:0009416	12	28	Response to light stimulus	BP	0.0265
GO:0051258	8	18	Protein polymerization	BP	0.0500
GO:0006298	7	12	Mismatch repair	BP	0.0500
GO:0030983	7	12	Mismatched DNA binding	MF	0.0500
Module 2					
GO:0055085	96	592	Transmembrane transport	BP	2.36E-06
GO:0006855	17	54	Drug transmembrane transport	BP	0.000741
GO:0015238	17	54	Drug transmembrane transporter activity	MF	0.000741
GO:0015297	17	54	Antiporter activity	MF	0.000741
GO:0007585	5	6	Respiratory gaseous exchange	BP	0.00718
GO:0016758	38	229	Transferase activity, transferring hexosyl groups	MF	0.0162
GO:0050662	31	182	Coenzyme binding	MF	0.0216
GO:0043565	53	376	Sequence-specific DNA binding	MF	0.0297
GO:0042578	10	33	Phosphoric ester hydrolase activity	MF	0.0321
Module 3					
GO:0020037	36	534	Heme binding	MF	0.0118
GO:0016760	8	35	Cellulose synthase (UDP-forming) activity	MF	0.0248
GO:0004601	15	165	Peroxidase activity	MF	0.0252
GO:0030244	8	39	Cellulose biosynthetic process	BP	0.0256
GO:0006979	14	163	Response to oxidative stress	BP	0.0367
Module 4					
GO:0043531	29	326	ADP binding	MF	0.0147
GO:0016706	13	122	Oxidoreductase activity, acting on paired donors, with 2-oxoglutarate as one donor	MF	0.0280
GO:0020037	34	534	Heme binding	MF	0.028000458
GO:0004672	70	1258	Protein kinase activity	MF	0.0321
GO:0006468	70	1262	Protein phosphorylation	BP	0.0321
Module 6					
GO:0016746	17	227	Transferase activity, transferring acyl groups	MF	0.00657
GO:0055114	54	1465	Oxidation–reduction process	BP	0.00658
GO:0016841	4	8	Ammonia-lyase activity	MF	0.00830
GO:0016491	33	781	Oxidoreductase activity	MF	0.0145
GO:0016747	14	209	Transferase activity, transferring acyl groups other than amino-acyl groups	MF	0.0280
GO:0009073	3	6	Aromatic amino acid family biosynthetic process	BP	0.0453
GO:0009055	25	574	Electron carrier activity	MF	0.0456

Table 1 (Continued)

Category	#DE	#Cat	Term	Ont	FDR
Module 11					
GO:0006471	2	4	Protein ADP-ribosylation	BP	0.00795
GO:0003950	2	8	NAD^+ ADP-ribosyltransferase activity	MF	0.0174

Weighted gene coexpression network analysis (WGCNA) was used to generate modules of genes with similar expression patterns across the dataset. Gene ontology enrichment analysis was performed on several coexpression modules of interest using the entire set of genes with detectable expression levels to determine enrichment. Categories with false discovery rate (FDR)-corrected *P*-values of ≤ 0.05 were considered to be enriched in each module (M). No enriched terms were identified in M10 or M13. M1, highly expressed in *35S::SbMyb60* stalks; M2, expressed at the lowest levels in *35S::SbMyb60* leaves; M3, expressed at the lowest levels in *35S::SbMyb60* stalks; M4, expressed at the lowest levels in *35S::SbMyb60* stalks; M6, expressed at the highest levels in *35S::SbMyb60* stalks; M10, expressed at the highest levels in Myb10a stalks and leaves; M11, expressed at the highest levels in Myb15a stalks and leaves; M13, expressed at the highest levels in Myb2a stalks and leaves. #DE, number of differentially expressed genes (DEGs) assigned to each term; #Cat, total number of genes assigned to each category; MF, molecular function; CC, cellular component; BP, biological process.

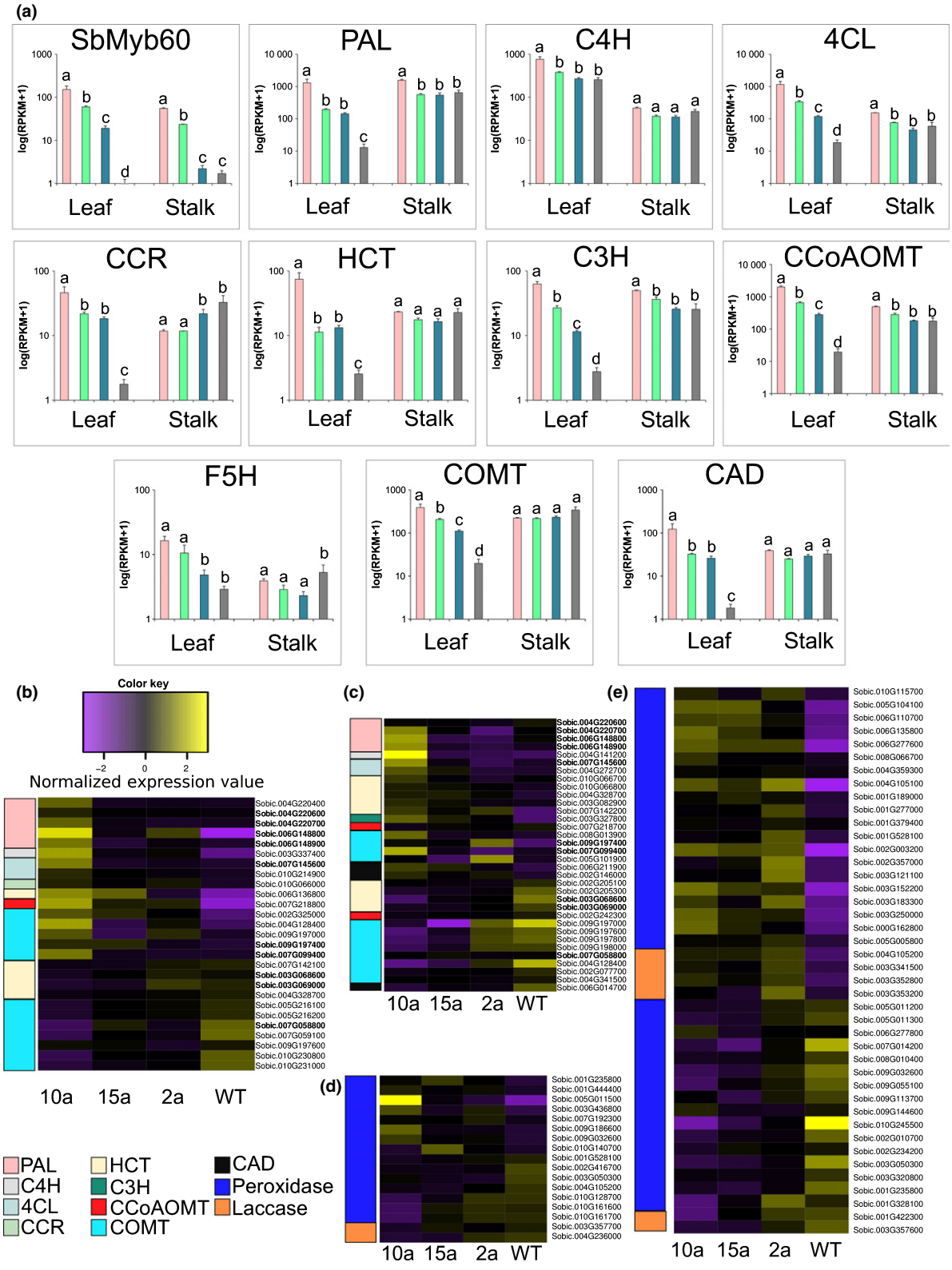
M2 were predominantly related to coenzyme binding, transmembrane transport and transfer of hexosyl groups. A comprehensive analysis of the specific metabolic pathways affected in the overexpression lines uncovered several additional pathways that were strongly impacted. In both leaves and stems, the most strongly impacted KEGG pathway in both up- and downregulated DEGs was the phenylpropanoid biosynthesis pathway (Fig. 2c). Meanwhile, in leaves, genes assigned to L-phenylalanine, tyrosine and tryptophan biosynthesis, the pentose phosphate pathway (PPP), and cysteine and methionine metabolism KEGG pathways tended to be upregulated (M6; Fig. 2c). In stalks, genes assigned to the ribosome, glutathione metabolism, amino and nucleotide sugar metabolism, DNA replication, starch and sucrose metabolism, and purine metabolism KEGG pathways were upregulated (M1), whereas genes assigned to the plant hormone signal transduction pathway tended to be downregulated (M3 and M4, Fig. 2c).

SbMyb60 overexpression impacts gene expression linked to monolignol and cell wall biosynthesis

Despite the differences in expression patterns noted between stalks and leaves overexpressing *SbMyb60*, several common pathways in both tissues were strongly impacted in the *35S::SbMyb60* lines. For example, genes associated with monolignol biosynthesis and other pathways linked to secondary cell wall biosynthesis were differentially expressed in both the stalks and leaves of *35S::SbMyb60* plants. In leaves, at least one gene for each of the 10 steps of monolignol biosynthesis was upregulated (Fig. 3a). In addition to these 10 genes, other homologs of pathway enzymes were also upregulated in *35S::SbMyb60* leaves (Fig. 3b). Genes coding for some monolignol biosynthesis pathway enzymes were activated in stalks, including four of the 10 candidate monolignol biosynthesis genes (*phenylalanine ammonia lyase* (*PAL*), *4CL*, *coumarate 3-hydroxylase* (*C3H*) and *caffeoyl-CoA O-methyltransferase* (*CCoAOMT*)) (Fig. 3a) and several additional pathway homologs (Fig. 3c). In contrast with leaves, *F5H* and *cinnamoyl CoA reductase* (*CCR*) were downregulated in stalks, as well as several other copies of *PAL* (three), *4CL* (one), *CCR* (one), *hydroxycinnamoyltransferase* (*HCT*) (five), *CCoAOMT* (two), *COMT* (nine) and *CAD* (one) (Fig. 3c). Few monolignol biosynthesis genes were differentially expressed in both tissues, indicating that *35S::SbMyb60* overexpression had divergent impacts on monolignol biosynthesis genes in stalk and leaf tissues. Related to monolignol biosynthesis, several secreted peroxidases and laccases were also differentially regulated in *35S::SbMyb60* leaves and stalks (Fig. 3d,e). Overall, the expression levels of more laccases and peroxidases, which could function in monolignol polymerization, were impacted in stalks relative to leaves, and most were upregulated. In addition, few peroxidases showed similar expression patterns in both tissues. For example, only Sobic.002G416700 and Sobic.003G050300 were downregulated in both tissues, whereas no laccases or peroxidases were commonly upregulated in both tissues. The expression levels of several BAHD acyltransferases were also impacted in *35S::SbMyb60* leaves and stems. This large family of proteins has been shown to esterify arabinoxylans with ferulic and *p*-coumaric acids, and is involved in the cross-linking of grass cell walls (Withers *et al.*, 2012; Hatfield *et al.*, 2017). In leaves, two BAHD acyltransferases were upregulated in each of the three lines, with one common BAHD (Sobic.003G043600) upregulated in all three. The expression levels of more BAHD genes were upregulated in *35S::SbMyb60* stalks relative to leaves, with Myb10a and Myb2a having the highest and lowest number of upregulated genes relative to WT, respectively.

Other components of the secondary cell wall biosynthesis program were also impacted in the *35S::SbMyb60* lines, including genes linked to the formation of UDP-sugars and cell wall polysaccharides, such as glucuronoarabinoxylan (GAX), cellulose and related polysaccharides. The strongest impacts on the biosynthesis of nucleotide sugars were observed in *35S::SbMyb60* stalks, where genes coding for the interconversion of various UDP-sugars were upregulated (Fig. 4a). In leaves, only *sucrose synthase* and *UDP-glucuronate 4-epimerase* were upregulated, and no genes were downregulated. Metabolite analysis showed slightly elevated levels of UDP-D-glucose in Myb10a relative to WT stalks, supporting the altered expression of genes linked to UDP-sugar biosynthesis in *35S::SbMyb60* stalks (Fig. 4b). One gene coding for cellulose synthase A2 (Sobic.002G094600), which could incorporate UDP-sugars into cellulose, was also upregulated in stalks (Fig. 4c). No *CesA* genes were downregulated in this tissue. In addition, three genes coding for CesA were upregulated in leaves, including Sobic.00G094600 (*CesA2*; also upregulated in stalks), Sobic.001G224300 (*CesA4*) and Sobic.002G118700 (*CesA2*), whereas none was downregulated (Fig. 4c).

New Phytologist



The expression levels of genes coding for pathways that lead to the production of ADP-sugars and starch were also impacted exclusively in *35S::SbMyb60* stalks (Fig. 4a). Supporting these expression changes, levels of sucrose/trehalose-6-phosphate and trehalose-sucrose were lower in Myb10a and Myb15a stalks relative to WT (Fig. 4d). Although the expression of these pathways was not altered in *35S::SbMyb60* leaves, levels of sucrose/trehalose-6-phosphate were higher in Myb10a and Myb15a leaves

Fig. 3 Impacts of *35S::SbMyb60* on monolignol biosynthesis pathway genes in sorghum. (a) Expression profiles of genes related to monolignol biosynthesis in leaves and stalks. Trimmed mean (TMM)-normalized RPKM (reads per kilobase per million mapped reads) counts from leaves (left) and stalks (right) were computed using the 'rpkm' command in edgeR. Bar graphs represent $\log_{10}$-transformed mean RPKM values and error bars represent + SE. Letters indicate statistical differences at $P \leq 0.05$ using a Fisher's exact test in edgeR. Pink bars, Myb10a; green bars, Myb15a; blue bars, Myb2a; gray bars, wild-type (WT). PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; CCR, cinnamoyl CoA reductase; HCT, hydroxycinnamoyl transferase; C3H, *p*-coumarate 3-hydroxylase; CCoAOMT, caffeoyl-CoA *O*-methyltransferase; F5H, ferulate 5-hydroxylase; COMT, caffeic acid *O*-methyltransferase; CAD, cinnamyl alcohol dehydrogenase. (b) Heatmap analysis of differentially expressed genes associated with the monolignol biosynthesis pathway in leaves. Raw counts of genes that were differentially expressed in at least one *35S::SbMyb60* line relative to WT were log-transformed and Z-score standardized. A heatmap was prepared using the heatmap.2 command from the GPLOTS package in R. Purple, low expression level; yellow, high expression level. Sidebar colors correspond to gene annotations and bold gene names indicate genes whose expression patterns were similarly impacted in both stalks and leaves. (c) Heatmap analysis of differentially expressed genes associated with the monolignol biosynthesis pathway in stalks. Heatmaps for differentially expressed genes in stalks were prepared as in (b). (d) Heatmap analysis of differentially expressed secreted peroxidases and laccases in *SbMyb60* leaves. Heatmaps were prepared for differentially expressed peroxidases and laccases in leaves that contained an identifiable signal peptide using SignalP. Heatmaps were prepared as in (b). (e) Heatmap analysis of differentially expressed secreted peroxidases and laccases in *SbMyb60* stalks. Signal peptides were identified as described in (d) and heatmaps were prepared as described in (b).

relative to WT (Fig. 4d). By contrast, the expression levels of genes encoding enzymes that lead to the formation of D-GDP-mannose and GDP-mannose polysaccharides were not impacted in either tissue, with the exception of *GDP mannose-4,6-dehydratase*, which was upregulated in stalks. Several other gene families previously linked to primary and secondary cell wall biosynthesis (McKinley *et al.*, 2016) were also differentially regulated in *35S::SbMyb60* leaves and stalks, which could impact the biosynthesis of other secondary wall polysaccharides comprising UDP sugars (Fig. 4e,f).

The impact of these transcriptional changes included elevated levels of wall-bound phenolic compounds. In Myb2a stalks, levels of sinapic, ferulic, caffeic and 4-hydroxybenzeneacetic acids were elevated by 96%, 22%, 36% and 43%, respectively, relative to WT (Fig. 5a). In Myb10a stalks, syringic acid, syringaldehyde, vanillic acid and 4-hydroxybenzoic acid levels were *c.* 70%, 83%, 37% and 30% higher, respectively, than those observed in WT (Fig. 5a). Overall, the results of the wall-bound analysis strongly mirrored the results of a previous soluble phenolics analysis conducted on *35S::SbMyb60* plant tissue (Scully *et al.*, 2016), suggesting that these monolignol biosynthesis pathway intermediates can readily be incorporated into the cell wall through esterification.

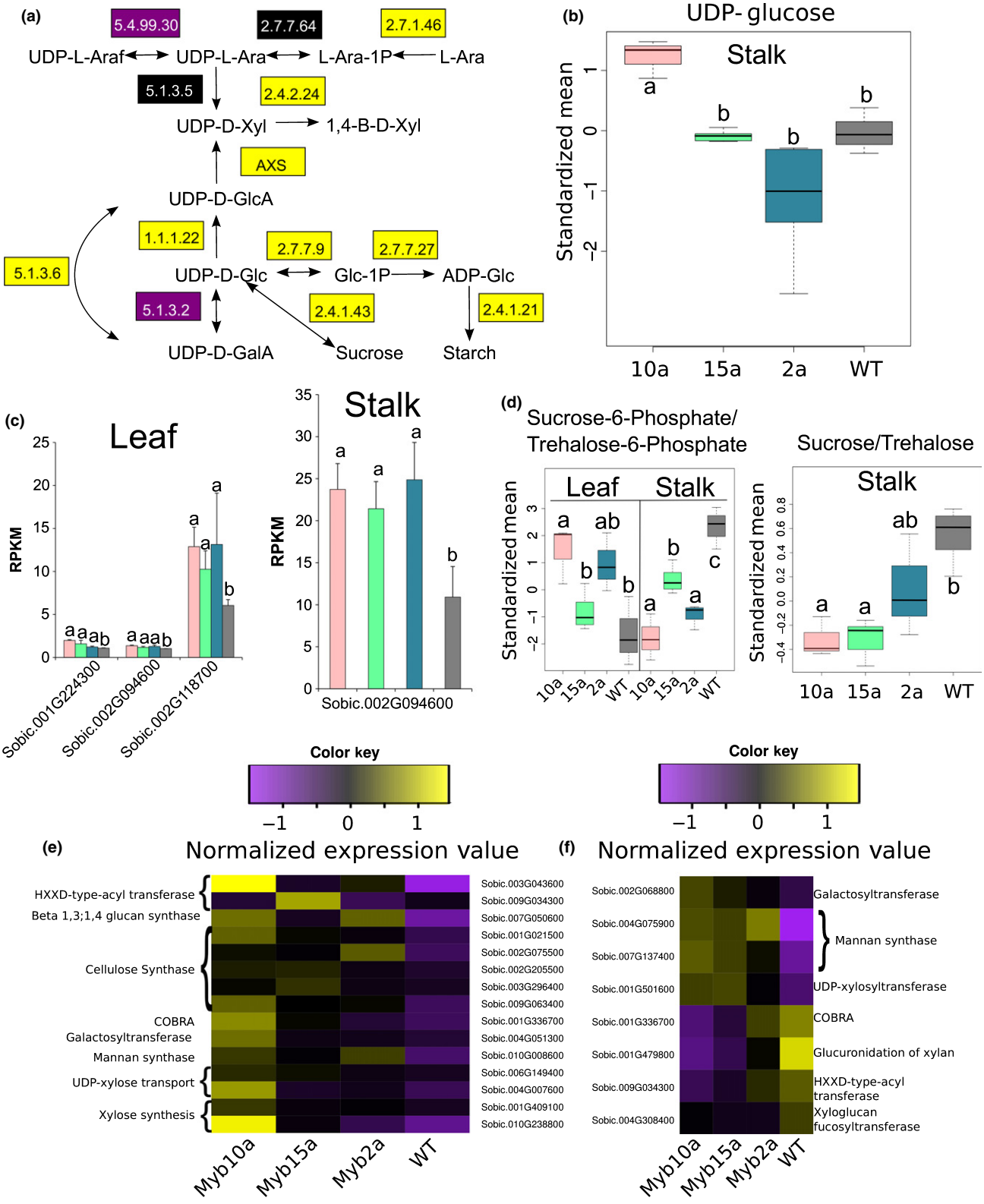
Stalk cross-sections were stained with either phloroglucinol or FASGA to observe differences in cell wall composition between *35S::SbMyb60* and WT stalks. These results indicated that *SbMyb60* overexpression impacted both lignin and cellulose deposition (Fig. 5b,c). In phloroglucinol-stained sorghum stalks, increased lignin deposition (pink) was observed in sclerenchyma cell walls of vascular bundles in Myb10a and Myb15a, and lignification extended further into the parenchymal cells than in WT (Fig. 5b). In FASGA-stained stalk cross-sections, parenchyma cell walls stained purple, indicative of both lignin (red) and cellulose (blue), and extended further from the cortical region into the interior regions of Myb10a and Myb15a stalks in comparison with WT (Fig. 5c). The sclerenchyma layer surrounding the vascular bundles appeared to be enlarged, whereas the cortical region appeared to expand towards the center of the Myb10a and Myb15a stalks relative to WT (Fig. 5c). Overall, the Myb2a stalk sections contained fewer vascular bundles relative to WT, and the other two Myb lines and their cell walls appeared blue after FASGA staining, indicative of less lignin deposition in secondary cell walls (Fig. 5b,c). Although the vascular bundles of Myb2a stalks were smaller than those of WT, Myb10a or Myb15a, the distribution of these structures in Myb2a was similar to that in WT (Fig. 5c).

Pathways for the biosynthesis of L-phenylalanine and the cofactors required for the production of monolignols

The monolignol biosynthesis pathway requires L-phenylalanine as a substrate, and SAM, CoA and NADPH as cofactors; the genes assigned to the pathways responsible for the synthesis of these compounds were impacted in both leaves and stalks. The genes assigned to the majority of the steps in the PPP, whose products can be used for aromatic amino acid biosynthesis via the shikimate pathway, were upregulated in both stalks and leaves (Fig. 6a). Moreover, the PPP also regenerates NADPH, a cofactor for the monolignol biosynthesis pathway. In addition to the PPP, genes encoding the entire portions of the shikimate pathway, leading to the production of L-phenylalanine and tyrosine, were upregulated in both leaves and stalks, whereas only two genes leading to the biosynthesis of tryptophan were upregulated in *35S::SbMyb60* in both tissues (Fig. 6b). Although no genes associated with aromatic amino acid catabolism were upregulated in *35S::SbMyb60* leaves, four genes involved in the catabolism of L-phenylalanine and tyrosine were upregulated in *35S::SbMyb60* stalks.

The concentrations of both PPP and shikimate pathway metabolites were altered (Fig. 6c,d), indicative of increased aromatic amino acid biosynthesis. L-Phenylalanine levels were elevated, particularly in Myb10a leaves, whereas an intermediate in the biosynthesis of L-phenylalanine, shikimate-3-phosphate, was lower in Myb10a and Myb15a leaves (Fig. 6d). In addition, phosphoenolpyruvate was significantly lower in Myb10a leaves, which is a substrate for chorismate biosynthesis (Fig. 5c). This finding supports the utilization of PPP products for the biosynthesis of the aromatic monolignol precursor L-phenylalanine. Although the levels of genes involved in the biosynthesis of tryptophan were not differentially expressed, tryptophan was elevated

New Phytologist



in Myb10a leaves (Fig. 6d). In stalks, both shikimate and shikimate-3-phosphate were present in lower concentrations in Myb10a and Myb15a relative to WT, but L-phenylalanine was not differentially abundant in this tissue (Fig. 6d). In addition, the concentrations of five metabolites from the PPP were greatly impacted (Fig. 6c). In most cases, concentrations were lowest in Myb10a stalks; however, the concentrations of glycerate and ribose-phosphate were lowest in Myb2a stalks.

Fig. 4 Impacts of *35S::SbMyb60* on secondary cell wall biosynthesis genes in sorghum. (a) Impacts of *35S::SbMyb60* on UDP- and ATP-sugar biosynthesis in stalks. Genes linked to the biosynthesis and interconversion of UDP- and ATP-sugars were identified using the KEGG Pathway Mapper Tool. Purple indicates genes that were downregulated in *35S::SbMyb60* stalks, black indicates genes that were not differentially expressed, and yellow indicates genes that were upregulated in *35S::SbMyb60* stalks relative to wild-type (WT). UDP-L-Araf, UDP-L-arabinofuranose; UDP-L-Ara, UDP-L-arabinose; L-Ara-1P, β-L-arabinose-1-phosphate; L-Ara, L-arabinose; UDP-D-Xyl, UDP-D-xylose; 1,4-β-D-Xyl, 1,4-β-D-xylan; UDP-D-GlcA, UDP-D-glucuronate; UDP-D-Glc, UDP-α-D-glucose; UDP-D-GalA, UDP-α-D-galactose; Glc-1P, glucose-1-phosphate; ADP-Glc, ADP-α-D-glucose. (b) Relative abundances of metabolites associated with UDP-glucose in *35S::SbMyb60* stalks. Concentrations of UDP-glucose in stalks were measured using LC/MS. Letters indicate statistical differences at $P \leq 0.05$ using ANOVA followed by a Fisher's least-significant difference (LSD) test for *post hoc* comparisons. Pink bars, Myb10a; green bars, Myb15a; blue bars, Myb2a; gray bars, WT. Error bars represent ± SE and letters represent significant differences at $P \leq 0.05$. (c) Relative abundances of metabolites associated with ADP-sugar biosynthesis in *35S::SbMyb60* stalks and leaves. Metabolite concentrations were measured by LC/MS as described in (b). Error bars represent + SE and letters represent significant differences at $P \leq 0.05$. (d) Expression levels of *CesA* genes differentially expressed in *35S::SbMyb60* leaves and stalks. RPKM (reads per kilobase per million mapped reads) values were computed using edgeR. Bars represent average RPKM values, error bars represent ± SE and letters represent significant differences at a false discovery rate (FDR) ≤ 0.05. Impacts of *35S::SbMyb60* on other genes previously linked to secondary cell wall biosynthesis in leaves (e) and stalks (f). RPKM values of genes linked to secondary cell wall biosynthesis in a previous study were determined (McKinley *et al.*, 2016). Heatmaps were prepared as described for Fig. 3.

Two *serine hydroxymethyltransferase* (*SHMT*) genes and one *methylenetetrahydrofolate reductase* (*MTHFR*) gene associated with SAM metabolism were upregulated in *35S::SbMyb60* leaves and stalks (Fig. 7a). MTHFR is the rate-limiting enzyme for the SAM cycle and serves as a primary flux control point between folate and SAM biosynthesis. Genes encoding two major enzymes leading to the production of SAM from homocysteine, homocysteine *S*-methyltransferase and SAM synthetase, were also upregulated in *35S::SbMyb60* leaves and stalks (Fig. 7a). The serine and glycine pools serve as important inputs for one-C metabolism, the production of folate and the production of methionine, all of which are critical for SAM biosynthesis (Fig. 7b). Genes coding for the entire pathway leading to the biosynthesis of serine from 3P-D-glycerate (Fig. 7b) and the entire glycine cleavage system, which can generate glycine in the reverse direction, were also upregulated in *35S::SbMyb60* leaves (Fig. 7b). A gene encoding a cysteine synthase A, which converts serine to cysteine, was strongly upregulated in *35S::SbMyb60* leaves and stalks. In stalks, genes encoding the enzymes required for the biosynthesis of serine were not impacted; however, the majority of the genes in the glycine cleavage system were upregulated, suggesting that glycine and serine could also be regenerated in this tissue (Fig. 7b).

The expression levels of genes coding for enzymes linked to CoA metabolism were affected in *35S::SbMyb60* stalks and leaves (Fig. 7c). In *35S::SbMyb60* leaves, two genes encoding enzymes required for the biosynthesis of pantothenate, the major substrate for CoA biosynthesis, and one gene leading to the biosynthesis of CoA from pantothenate were upregulated (Fig. 7c). *Acetolactate synthase* and *3-methyl-2-oxobutanoate* were also upregulated in *35S::SbMyb60* stalks, as well as a gene encoding dihydroxy-acid dehydratase, required for the biosynthesis of pantothenate (Fig. 7c). The biosynthesis of pantothenate via β-alanine requires pantoate-β-alanine ligase. Although the gene encoding this enzyme was not differentially expressed in either stalks and leaves, two genes (*spermine oxidase* and *polyamine oxidase*) and three genes (*spermine oxidase*, *polyamine oxidase* and *primary-amino oxidase*) encoding enzymes responsible for the biosynthesis of β-alanine from spermine and L-aspartate were upregulated in leaves and stalks, respectively. These transcriptional changes could provide more β-alanine for diversion to the pantothenate/CoA biosynthesis pathway. No genes assigned to the pantothenate and CoA biosynthesis pathway were downregulated in either *35S::SbMyb60* leaves or stalks. Although NADPH is required as a cofactor by several enzymes in monolignol biosynthesis, only one NAD kinase (Sobic.009G123700) was upregulated in *35S::SbMyb60* stalks. However, the expression levels of *glucose-6-phosphate dehydrogenase* from the PPP were upregulated in both *35S::SbMyb60* leaves and stalks, which may maintain $NADP^+$/NADPH.

Metabolite analysis supported the increased utilization of the cofactors CoA and NAD^+ in both *35S::SbMyb60* leaves and stalks. For example, 5-methyltetrahydrofolate (folate metabolism) was present at a lower abundance in *35S::SbMyb60* leaves, whereas the levels of glycine and serine associated with the glycine cleavage system were generally higher in *35S::SbMyb60* stalks relative to WT, and guanidinoacetic acid levels were lower in Myb2a stalks relative to WT (Fig. 7d). Metabolites associated with SAM metabolism, including cystine, which is a metabolite of cysteine, homocysteine, *O*-acetyl-L-serine and *S*-methyl-5-thioadenosine, which is a metabolite involved in methionine salvage, were higher in *35S::SbMyb60* stalks (Fig. 7d). In addition, 2-oxo-4-methylthiobutanoate levels were lower in Myb10a and Myb15a stalks relative to WT, whereas levels of *S*-adenosyl-L-methioninamine were higher in Myb15a and Myb2a relative to WT stalks (Fig. 7d). Further, levels of NAD^+ were generally higher in *35S::SbMyb60* leaves relative to WT, and lower in Myb15a stalks relative to WT (Fig. 8). In stalks, several other metabolites associated with nicotinate metabolism were impacted, including $NADP^+$ (elevated in Myb10a relative to WT), nicotinamide (higher in Myb10a and Myb15a relative to WT), nicotinate (lower in Myb2a relative to WT) and quinolate (lower in Myb10a and Myb15a relative to WT) (Fig. 8).

Nitrogen flux and lysine catabolism are impacted in *35S::SbMyb60* plants

Nitrogen is a central element for many of the substrates and cofactors used for monolignol biosynthesis, and thus it is not unexpected that the metabolites and pathways linked to N flux were impacted in both *35S::SbMyb60* stalks and leaves. Most notably, genes coding for glutamine synthetase and glutamate

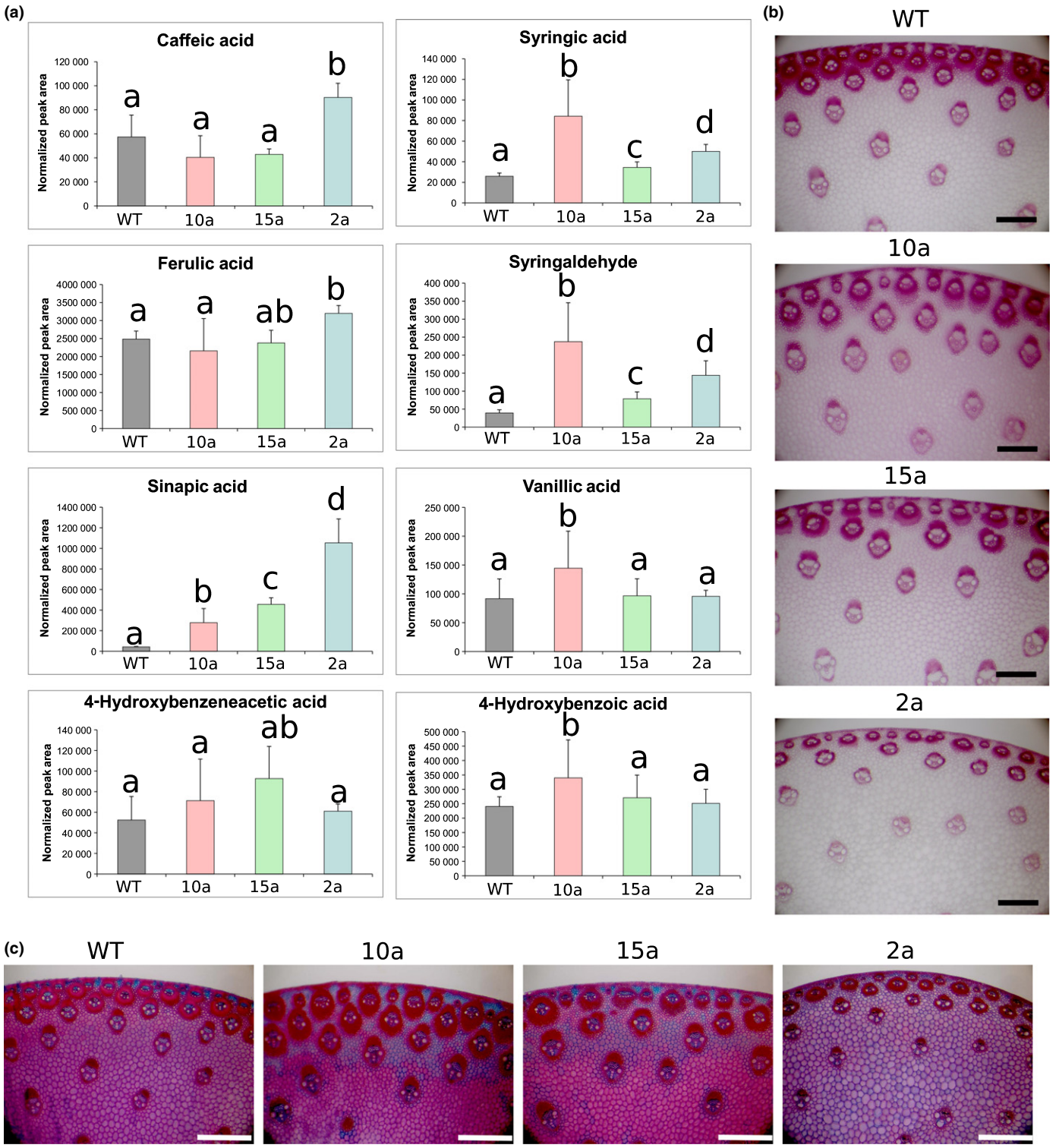


Fig. 5 Phenotypic impacts of *35S::SbMyb60* on sorghum cell wall composition. (a) Relative abundances of wall-bound phenolics from stover. Wall-bound phenolics were analyzed from stover using GC/MS. Statistical comparisons were performed using PROC Mixed with a *post hoc* Fisher's least-significant difference (LSD) test. The experiment-wise error rate was controlled at 0.05, error bars represent + SE and letters represent mean separation at $P \leq 0.05$. Pink bars, Myb10a; green bars, Myb15a; blue bars, Myb2a; gray bars, wild-type (WT). Impacts of *35S::SbMyb60* overexpression on lignin (b) and cellulose (c) content. Phloroglucinol (right) and FASGA staining (bottom) of 8-wk-old stalks collected from WT and each *35S::SbMyb60* line. Bars, 500 μm.

synthase were upregulated in both *35S::SbMyb60* leaves and stalks, which probably serves as a mechanism to recycle ammonia released by PAL or other enzymes for anabolic purposes. In addition, several metabolites with relevance to N cycling were differentially abundant in both tissues (Fig. S2), including three metabolites linked to lysine degradation.

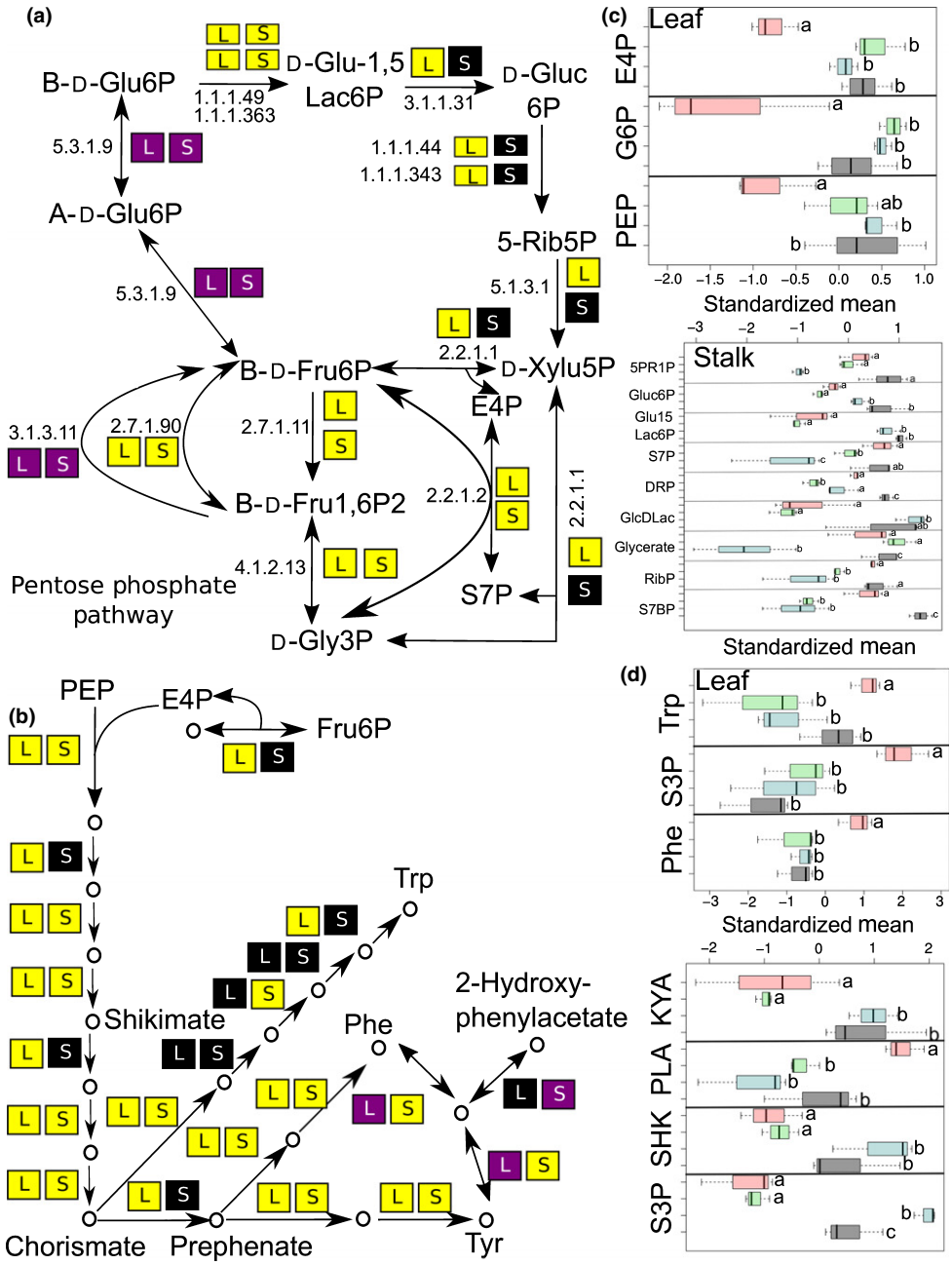


Fig. 6 Impacts of *35S::SbMyb60* on the pentose phosphate (PPP) and shikimate pathways in sorghum. (a) Impacts of *35S::SbMyb60* on PPP in leaves and stalks. PPP pathway genes were identified by KEGG orthology (KO) assignments. Purple, genes downregulated in *35S::SbMyb60* leaves (L) and stalks (S) relative to wild-type (WT); black, genes that were not differentially expressed; yellow, genes upregulated in *35S::SbMyb60*. β-D-Glu6P, β-D-glucose-6-phosphate; D-Glu-1,5Lac6P, D-glucono-1,5-lactone-6-phosphate; D-Gluc6P, 6-phospho-D-gluconate; 5-Rib5P, D-ribulose-5-phosphate; D-Xylu5P, D-xylulose-5-phosphate; S7P, D-sedo-heptulose-7-phosphate; E4P, D-erythrose-4-phosphate; β-D-Fru6P, β-D-fructose-6-phosphate; β-D-Fru1,6P2, β-D-fructose-1,6-bisphosphate; D-Gly3P, D-glyceraldehyde-3-phosphate. (b) Impacts of *35S::SbMyb60* on the shikimate pathway in leaves and stalks. Shikimate pathway genes were identified and color coded as described in (a). Fru6P, fructose-6-phosphate; E4P, erythrose-4-phosphate; PEP, phosphoenolpyruvate; Trp, tryptophan; Tyr, tyrosine; Phe, L-phenylalanine. (c) Concentrations of PPP metabolites in leaves (upper) and stalks (lower). Concentrations of metabolites derived from the PPP were identified using LC/MS. The peak area of each metabolite was used to determine the abundance, and samples were normalized by sample weight and the abundance of an internal standard. Samples were log-transformed and means were standardized. Line represents mean, box represents interquartile range (IQR) and bars represent highest and lowest values. Samples were statistically analyzed using ANOVA followed by Fisher's least-significant difference (LSD) test for *post hoc* comparisons; letters indicate mean separation at $P \leq 0.05$. Pink, Myb10a; green, Myb15a; blue, Myb2a; gray, WT. E4P, erythrose-4-phosphate; G6P, glucose-6-phosphate; PEP, phosphoenolpyruvate; 5PR1P, 5-phosphoribosyl-1-pyrophosphate; Gluc6P, 6-phospho-D-gluconate; Glu15Lac6P, 6-phospho-D-glucono-1,5-lactone; S7P, D-sedoheptulose-1,7-phosphate; DRP, deoxyribose phosphate; GlcDLac, glucono-D-lactone; RibP, ribose phosphate; S7BP, sedoheptulose-1,7-bisphosphate. (d) Concentrations of shikimate pathway metabolites in leaves (upper) and stalks (lower). Metabolite concentrations were determined as described in (c). Trp, tryptophan; S3P, shikimate-3-phosphate; Phe, L-phenylalanine; SHK, shikimate; PLA, phenyllactic acid; KYA, kynurenic acid.

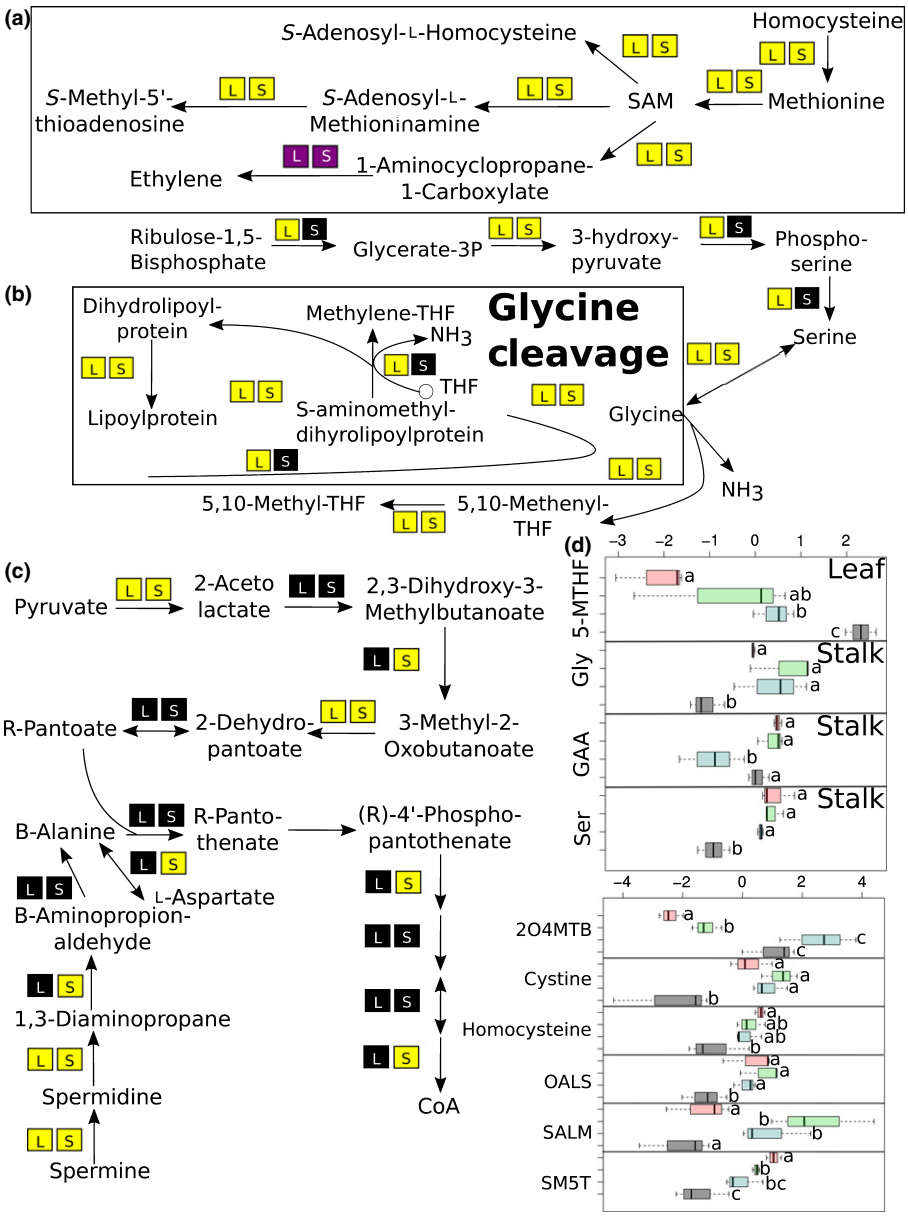


Fig. 7 Impacts of *35S::SbMyb60* on the pathways for the biosynthesis of the cofactors required for monolignol biosynthesis in sorghum. (a) Impacts of *35S::SbMyb60* on genes assigned to the *S*-adenosyl methionine (SAM) biosynthesis pathway. Genes coding for SAM biosynthesis enzymes were identified using KEGG orthology (KO) terms generated by Phytozome and the KEGG Pathway Mapper Tool. Purple, genes expressed at lower levels in *35S::SbMyb60* leaves (L) and stalks (S); yellow, genes expressed at higher levels in these two tissues in *35S::SbMyb60* plants. (b) Impacts of *35S::SbMyb60* on genes assigned to the serine, glycine and folate biosynthesis pathways. Genes coding for enzymes assigned to these pathways were identified and color coded as described in (a). Black, genes that were not differentially expressed. Glycerate-3P, glycerate-3-phosphate; THF, tetrahydrofolate. (c) Impacts of *35S::SbMyb60* on pantothenate, β-alanine and coenzyme A (CoA) biosynthesis. Genes coding for enzymes assigned to these pathways were identified and color coded as described in (a). (d) Concentrations of metabolites associated with serine, glycine and folate biosynthesis (upper) and SAM metabolism (lower). The concentrations of metabolites derived from these pathways were identified using LC/MS. The peak area of each metabolite was used to determine the abundance, and samples were normalized by sample weight and the abundance of an internal standard. Samples were log-transformed and means were standardized. Line represents mean, box represents interquartile range (IQR) and bars represent highest and lowest values. Samples were statistically analyzed using ANOVA followed by Fisher's least-significant difference (LSD) test for *post hoc* comparisons; letters indicate mean separation at $P \leq 0.05$. Pink, Myb10a; green, Myb15a; blue, Myb2a; gray, wild-type (WT). 5-MTHF, 5-methyltetrahydrofolate; Gly, glycine; GAA, guanidoacetic acid; Ser, serine; 2O4MTB, 2-oxo-4-methylthiobutanoate; OALS, *O*-acetyl-L-serine; SALM, *S*-adenosyl-L-methioninamine; SM5T, *S*-methyl-5-thioadenosine.

35S::SbMyb60 affects the expression of the tricarboxylic acid (TCA) cycle and glycolysis-related genes and metabolites

Genes encoding enzymes in glycolysis and the TCA cycle were upregulated in *35S::SbMyb60*, with relatively more glycolysis and TCA pathway enzymes upregulated in *35S::SbMyb60* leaves relative to *35S::SbMyb60* stalks (Fig. 9a). Specifically, in leaves, genes encoding TCA enzymes that convert oxaloacetate to 2-oxoglutarate were strongly upregulated (Fig. 9a). Four genes encoding enzymes responsible for the conversion of pyruvate to acetyl-CoA were also upregulated in

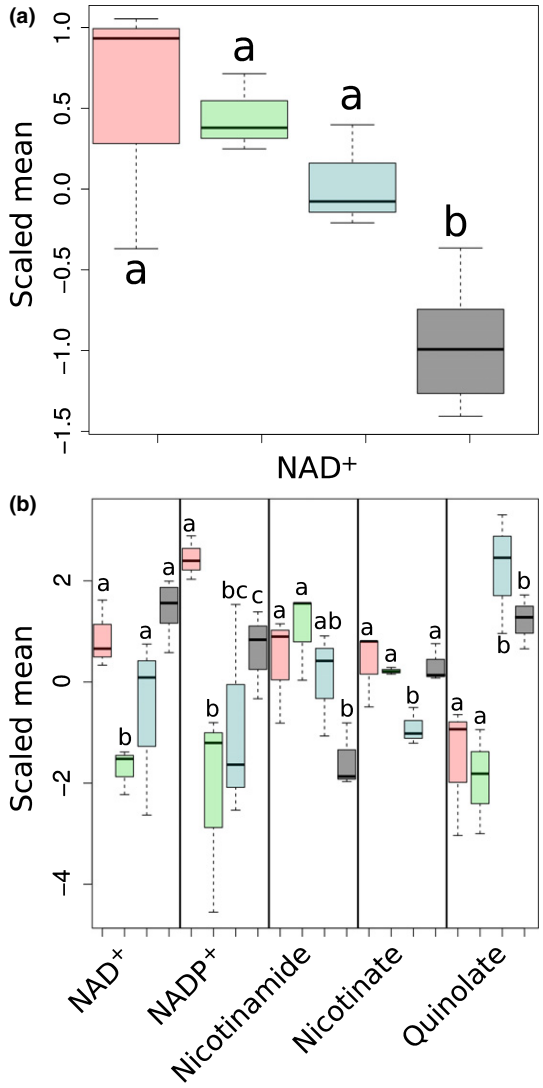


Fig. 8 Levels of NAD$^+$ and NAD$^+$-related metabolites in *35S::SbMyb60* sorghum leaves and stalks. (a) Levels of NAD$^+$ in leaves. Concentrations of NAD$^+$ were identified using LC/MS. The peak area of each metabolite was used to determine the abundance, and samples were normalized by sample weight and the abundance of an internal standard. Samples were log-transformed and means were standardized. Line represents mean, box represents interquartile range (IQR) and bars represent highest and lowest values. Samples were statistically analyzed using ANOVA followed by Fisher's least-significant difference (LSD) test for *post hoc* comparisons; letters indicate mean separation at $P \leq 0.05$. Pink, Myb10a; green, Myb15a; blue, Myb2a; gray, wild-type (WT). (b) Levels of NAD$^+$ and other nicotinamide metabolites in stalks. Metabolite concentrations were determined and analyzed as described in (a).

35S::SbMyb60 leaves. By contrast, in stalks, genes encoding TCA enzymes capable of converting oxaloacetate to succinyl-CoA were upregulated, whereas only two genes encoding enzymes responsible for acetyl-CoA biosynthesis were upregulated in this tissue (Fig. 9a). In *35S::SbMyb60* leaves, five genes encoding glycolysis pathway enzymes were upregulated, whereas four glycolysis genes were upregulated in *35S::SbMyb60* stalks (Fig. 9b). Few TCA cycle or glycolysis genes were downregulated in either tissue.

The impacts of the transcriptional changes observed on the glycolysis and TCA pathways were also apparent in metabolite concentrations in stalk and leaf tissues (Fig. 9c,d). In stalks, the TCA cycle intermediate aconitate accumulated to significantly higher levels in *35S::SbMyb60* (Fig. 9c). In addition, citrate was lower in Myb2a stalks relative to WT, whereas isocitrate was lower in Myb15a stalks relative to WT (Fig. 9c). No TCA intermediates were differentially abundant in *35S::SbMyb60* leaves. Intermediates of glycolysis that were altered in *35S::SbMyb60* leaves included fructose-1,6-phosphate, which was lower in Myb15a relative to WT, and glucose-1-phosphate, which was higher in Myb10a relative to WT (Fig. 9d). In stalks, levels of 3-phosphoglycerate, D-glyceraldehyde-3-phosphate, dihydroxyacetone phosphate fructose-6-phosphate, glucose-1-phosphate and glucose-6-phosphate were elevated in all three Myb lines relative to WT (Fig. 9d). Genes encoding the major mitochondrial complexes (1–5) involved in oxidative phosphorylation were also strongly upregulated in *35S::SbMyb60* stalks (Fig. 9e). These observations suggest that TCA products in this tissue may be used directly for ATP and NADH production (Fig. 9e). By contrast, only genes encoding cytochrome C oxidase (complex 4) and NADH dehydrogenase (complex 1) from the oxidative phosphorylation pathway were upregulated in *35S::SbMyb60* leaves.

Genes related to hormone biosynthesis were altered in *35S::SbMyb60* stalks

Genes encoding enzymes for the majority of the methylerythritol pathway responsible for the conversion of *trans*-geranylgeranyl diphosphate to gibberellin were downregulated in *35S::SbMyb60* stalks in both Myb10a and Myb15a plants (Fig. 10a). Furthermore, five genes linked to the biosynthesis of jasmonic acid (JA) from its precursor α-linolenic acid were also downregulated in stalks (Fig. 10b). Supporting the possible disruption of hormone signaling in *35S::SbMyb60* stalks, levels of indole-3-carboxylic acid were lower in *35S::SbMyb60*, which suggests that flux through the indole-3-acetic acid (IAA) biosynthesis pathway was altered. Despite the major transcriptional changes to the hormone signaling pathways observed in *35S::SbMyb60* stalks, fewer transcriptional impacts to these pathways were observed in *35S::SbMyb60* leaves.

Cell elongation and internode development were impacted in *35S::SbMyb60* lines

Internode development and cell expansion in *35S::SbMyb60* sorghum stalks were differentially impacted, potentially as a consequence of changes in hormone signaling. For example, internodes in Myb10a and Myb15a plants were significantly shorter in length compared with the internodes of WT stalks, whereas the internodes in Myb2a were similar to those of WT in length (Figs 10c, S3a). In addition, Myb10a had significantly more internodes than WT (Fig. S3b), but the mean diameter of the internodes did not differ (Fig. S3c), which may reflect the prolonged vegetative growth period observed previously in this line (Scully *et al.*, 2016). In SEM images of the longitudinal sections

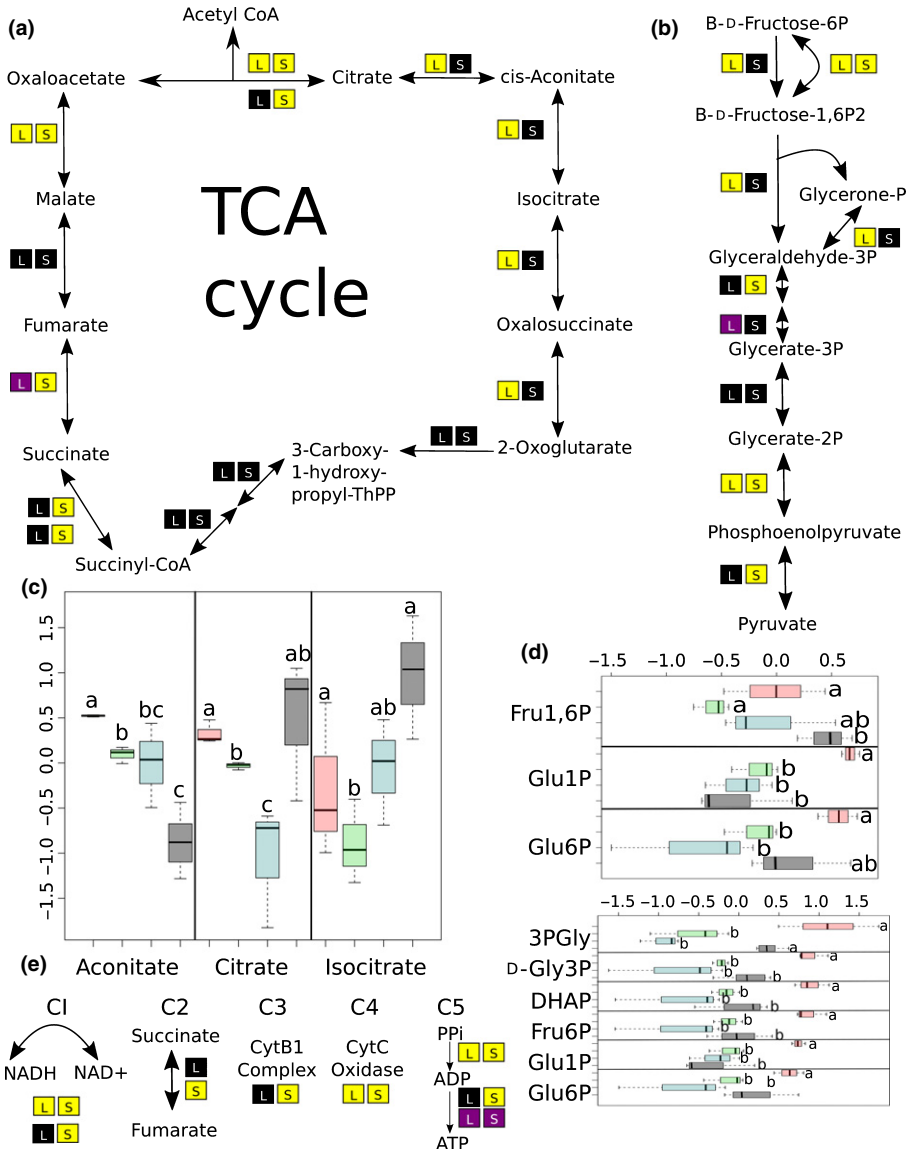


Fig. 9 Impacts of *35S::SbMyb60* on the expression of the tricarboxylic acid (TCA) cycle and glycolysis genes in sorghum. (a) Impacts of *35S::SbMyb60* on the expression of genes assigned to the TCA cycle. TCA cycle genes were identified using KEGG orthology (KO) annotations and the KEGG Pathway Mapper Tool. Purple, genes downregulated in *35S::SbMyb60* leaves (L) and stalks (S) relative to wild-type (WT); black, genes that were not differentially expressed; yellow, genes upregulated in *35S::SbMyb60* leaves or stalks. 3-Carboxy-1-hydroxy-propyl-ThPP, 3-carboxy-1-hydroxypropyl-thiamine diphosphate. (b) Impacts of *35S::SbMyb60* on the expression of genes assigned to glycolysis. Genes assigned to the glycolysis pathway were identified and color coded as described in (a). P, phosphate; P2, bisphosphate. (c) Concentrations of TCA-derived metabolites in stalks. Metabolite concentrations were determined via LC/MS. The peak area of each metabolite was used to determine the abundance, and samples were normalized by sample weight and the abundance of an internal standard. Samples were log-transformed and means were standardized. Line represents mean, box represents interquartile range (IQR) and bars represent highest and lowest values. Samples were statistically analyzed using ANOVA followed by Fisher's least-significant difference (LSD) test for *post hoc* comparisons; letters indicate mean separation at $P \leq 0.05$. Pink, Myb10a; green, Myb15a; blue, Myb2a; gray, wild-type (WT). The *y*-axis is the scaled mean. (d) Concentrations of glycolysis-derived metabolites in leaves (upper) and stalks (lower). Metabolite concentrations were determined and analyzed as described in (c). Fru1,6P, fructose-1,6-phosphate; Glu1P, glucose-1-phosphate; Glu6P, glucose-6-phosphate; 3PGly, 3-phosphoglycerate; D-Gly3P, D-glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone phosphate; Fru6P, fructose-6-phosphate. The *y*-axis is the scaled mean. (e) Impacts of *35S::SbMyb60* on genes from the oxidative phosphorylation pathway. Genes encoding enzymes involved in oxidative phosphorylation were identified and color coded as described in (a). C1–C5, complexes 1–5; Cyt, cytochrome.

of stalks, the structures of cells from WT stalks were more regular and consistent in shape and diameter in comparison with those of the three *35S::SbMyb60* overexpression lines (Fig. 10d). Overall, the cell shapes of Myb10a and Myb15a stalks were more distinct from WT in comparison with those of Myb2a.

Potential regulatory roles of *SbMyb60* outside the monolignol biosynthesis pathway

By using WGCNA, several genes that shared high coexpression values (Topological Overlap Matrix (TOM) values > 0.10) with

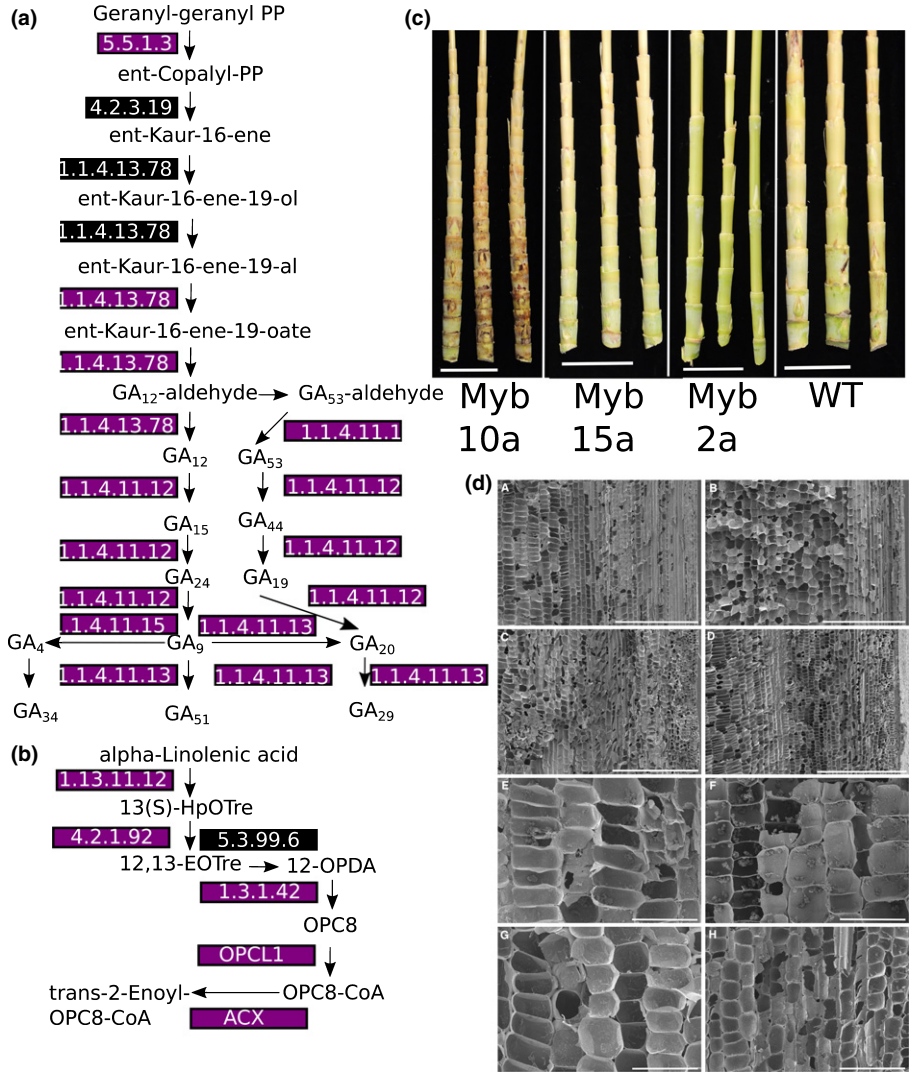
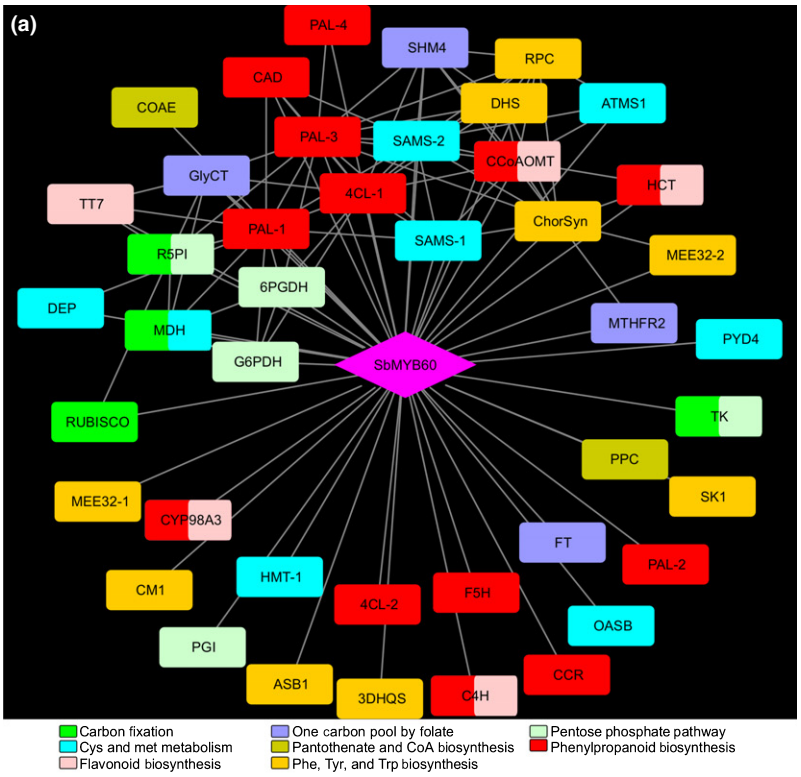


Fig. 10 Impacts of *35S::SbMyb60* on internode development and expression levels of genes linked to hormone biosynthesis in sorghum. (a) Impacts of *35S::SbMyb60* on expression of genes coding for gibberellin biosynthesis enzymes in stalks. Genes coding for diterpenoid biosynthesis pathway enzymes were identified using KEGG orthology (KO) annotations and the KEGG Pathway Mapper Tool. Purple, genes downregulated in *35S::SbMyb60* stalks relative to wild-type (WT); black, genes that were not differentially expressed. PP, bisphosphate; GA, gibberellin. (b) Impacts of *35S::SbMyb60* on the expression of genes coding for α-linolenic acid biosynthetic enzymes. Genes encoding enzymes linked to the production of α-linolenic acid were identified and color coded as in (a). 13(S)-HpOTre, (9*Z*,11*E*,13*S*,15*Z*)-13-hydroperoxyoctadeca-9,11,15-trienoate; 12,13-EOTre, (9*Z*,15*Z*)-(13*S*)-12,13-epoxyoctadeca-9,11,15-trienoate; 12-OPDA, (15*Z*)-12-oxophyto-10,15-dienoate; OPC8, 8-[(1*R*,2*R*)-3-oxo-2-{(*Z*)-pent-2-enyl}cyclopentyl]octanoate. (c) Impacts of *35S::SbMyb60* on internode development. Stalks were collected from 8-wk-old sorghum plants for internode measurements. Bars, 5 cm. (d) Impacts of *35S::SbMyb60* on cell elongation in stalks. Stalk tissue from 8-wk-old sorghum plants was subjected to scanning electron microscopy (SEM). Bars: (A–D) 500 μm; (E–H) 100 μm. A and E, wild-type (WT); B and F, Myb2a; C and G, Myb15a; D and H, Myb10a.

SbMyb60 were identified that could potentially serve as direct targets for this TF. Although *SbMyb60* shared high coexpression levels with genes involved in monolignol biosynthesis, it also shared similar coexpression values with genes involved in primary metabolism, including genes coding for the aromatic amino acid biosynthesis enzymes arogenate dehydratase, prephenate dehydratase and chorismate synthetase, and the SAM metabolic gene *SAM synthetase* (Fig. 11a). Although this finding suggests that *SbMyb60* directly regulates enzymes in the monolignol biosynthesis pathway, it also suggests that it may activate genes outside of this pathway which encode enzymes that synthesize substrates required by the pathway. In addition, *SbMyb60* probably triggers the activation of other TFs that may serve roles in both primary and secondary metabolism (Fig. 11b). For example, TFs with high coexpression values with *SbMyb60* included a myb (Sobic.005G062000), a basic helix–loop–helix (bHLH) (Sobic.008G186100) and a Broad complex, Tramtrack, Bric-à-brac (BTB) (Sobic.002G204300), which all shared high TOM values with genes linked to monolignol biosynthesis, aromatic amino acid biosynthesis, one-C metabolism and N metabolism (Table 2).

In addition, *SbMyb60* did not share high coexpression values with genes associated with UDP-sugar or hormone biosynthesis, indicating that it was not likely to be responsible for the



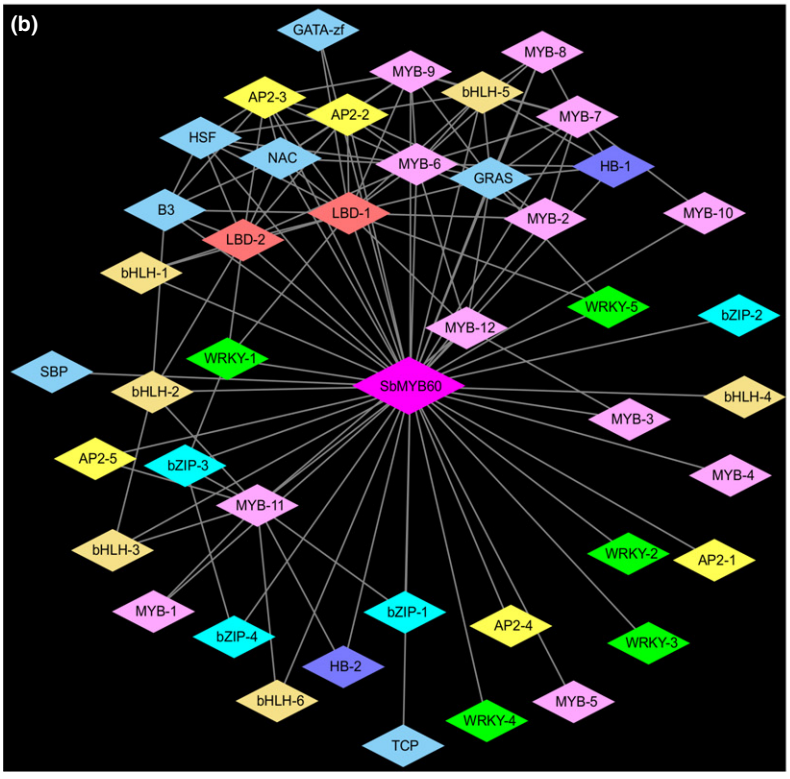


Fig. 11 Sorghum *SbMyb60* coexpression analysis. (a) Genes linked to aromatic amino acid and cofactor biosynthesis with high coexpression values with *SbMyb60*. Several genes with high coexpression values (Topological Overlap Matrix (TOM) ≥ 0.10) with *SbMyb60* were assigned to pathways outside of monolignol biosynthesis, including the pentose phosphate pathway (PPP), aromatic amino acid biosynthesis, *S*-adenosyl methionine (SAM) metabolism and coenzyme A (CoA) biosynthesis. Genes are color coded by KEGG pathway assignment, gene names are defined in Supporting Information Table S1 and lines indicate TOM ≥ 0.10. (b) Transcription factors (TFs) with high coexpression values with *SbMyb60*. TFs with TOM values ≥ 0.10 with *SbMyb60* are shown. TFs were assigned to families based on PFAM domain assignments and named according to their highest scoring BLASTP match from Arabidopsis. TFs are color coded by families, with light blue representing families containing only a single gene (NAC, SPB, B3, HSF, GATA-zf and TCP). Sorghum locus IDs are provided in Table S2.

regulation of these pathways. TFs with high coexpression levels with hormone biosynthesis genes included Sobic.009G024600 (AP2-like TF) and Sobic.002G153100 (bZIP), which had high coexpression values with all genes associated with the hormone signaling pathway, and Sobic.002G031000 (bHLH) and Sobic.001G078900 (MYB), which had high coexpression values with all but two of the hormone signaling genes. Transgenic down-regulation of the Arabidopsis ortholog of Sobic.002G024600

Table 2 Number of sorghum genes assigned to various KEGG pathways with high coexpression values with transcription factors (TFs) induced by *SbMyb60*

KEGG path	*SbMyb60*	Myb	bHLH	BTB	Shared
Biosynthesis of amino acids	22	23	22	23	22
Carbon metabolism	16	21	16	17	16
Phenylpropanoid biosynthesis	13	13	12	13	12
Glyoxylate and dicarboxylate metabolism	10	11	10	10	10
L-Phenylalanine, tyrosine and tryptophan biosynthesis	9	9	9	9	9
L-Phenylalanine metabolism	8	9	8	8	8
Cysteine and methionine metabolism	8	8	8	8	8
Glycine, serine and threonine metabolism	7	7	7	7	7
Nitrogen metabolism	6	7	7	6	6
Pentose phosphate pathway	5	7	5	6	5
Oxidative phosphorylation	5	7	7	6	5
Carbon fixation in photosynthetic organisms	4	7	4	5	4
Stilbenoid, diarylheptanoid and gingerol biosynthesis	6	5	5	6	5
Glycolysis/gluconeogenesis	4	6	4	5	4
Pyruvate metabolism	4	6	4	4	4
Alanine, aspartate and glutamate metabolism	5	5	5	5	5
Flavonoid biosynthesis	5	4	4	5	4
Photosynthesis	4	4	4	4	4
One carbon pool by folate	4	4	4	4	4
Ubiquinone and other terpenoid-quinone biosynthesis	3	4	3	3	3
Arginine biosynthesis	3	3	3	3	3
Purine metabolism	3	3	3	3	3
Fatty acid biosynthesis	2	3	2	2	2
Glutathione metabolism	2	3	2	2	2
Glycerolipid metabolism	3	3	3	2	2
Fatty acid metabolism	2	3	2	2	2
Peroxisome	2	3	2	2	2
Fructose and mannose metabolism	1	3	1	2	1

SbMyb60 shared high Topological Overlap Matrix (TOM) values with three TFs: Sobic.005G062000 (Myb), Sobic.008G186100 (bHLH) and Sobic.002G204300 (BTB). In turn, these transcription factors shared high coexpression levels with genes linked to several primary metabolic pathways. The numbers of genes in each pathway with TOM ≥ 0.10 for each TF are listed and the numbers of genes with high coexpression levels with all four TFs are presented in the shared column.

(At2g28550) delayed floral development, whereas the ortholog for Sobic.002G153100 (At5g06839) was required for anther development in Arabidopsis, supporting their roles in hormone signaling and development in sorghum. In concert with phytohormone signaling pathways, several genes assigned to the circadian rhythm pathway also shared high coexpression values with these four TFs. In terms of the methylerythritol and α-linolenic acid metabolism pathways, *SbMyb60* did not share high coexpression values with any of the genes assigned to these pathways, indicating that it is unlikely to serve as a direct negative regulator of these genes. Instead, 14 other TFs shared high coexpression values with 11 of the 12 genes assigned to both pathways and are candidates for the regulation of these pathways (Table 3). *SbMyb60* did not share high coexpression levels with any of these TFs, indicating that other transcriptional changes occurring in the stalk are probably responsible for the control of these TFs, and that *SbMyb60* is indirectly responsible for the transcriptional changes linked to the methylerythritol and α-linolenic acid biosynthesis pathways.

Discussion

Lignin biosynthesis requires phenylalanine, NADPH, SAM and CoA; therefore, the manipulation of monolignol biosynthesis to optimize cell wall composition for bioenergy conversion has the potential to either directly or indirectly impact other primary and secondary metabolic pathways. Despite the large number of studies devoted to the determination of the roles of monolignol biosynthesis enzymes in cell wall composition (Boerjan *et al.*, 2003; Vanholme *et al.*, 2010), very few studies have examined the phenotypic impacts of the manipulation of this pathway on other metabolic pathways. Several plants with altered monolignol biosynthetic capabilities have also been associated with growth and developmental defects (Li *et al.*, 2010; Zhao *et al.*, 2013; Bonawitz *et al.*, 2014); however, it is uncertain whether these phenotypes are directly caused by the manipulation of monolignol biosynthesis or are the result of indirect impacts to other metabolic pathways. In this study, global RNA-Seq and metabolite profiling of three sorghum transformants overexpressing *SbMyb60*, a positive regulator of monolignol biosynthesis, demonstrated that this TF also has the potential to directly regulate genes associated with aromatic amino acid, SAM and CoA biosynthesis, and may also influence NADPH biosynthesis and turnover. *35S::SbMyb60* indirectly impacts several other pathways, including UDP-sugar and various hormone biosynthesis pathways, and has different impacts on the global transcriptomic and metabolomic profiles of leaves and stalks.

SbMyb60 overexpression alters gene expression and metabolite levels in pathways that supply substrates and cofactors for the monolignol biosynthesis pathway, including PPP, L-phenylalanine, SAM and CoA biosynthesis. Genes linked to NADPH cycling and catabolism were also upregulated in both leaves and stalks, and metabolite levels were altered accordingly. The impacts to these pathways are anticipated, given that maize *bm2* and *bm4* mutants have an impact on SAM metabolism and result in modest reductions in Klason lignin relative to WT (Tang *et al.*, 2014; Li *et al.*, 2015), and given that aromatic amino acid biosynthesis pathways are commonly upregulated in lignifying tissues (Ehlting *et al.*, 2005; Dauwe *et al.*, 2007; Ohtani *et al.*, 2016). In the current study, genes linked to the majority of the PPP and aromatic amino acid biosynthesis pathways also shared high TOM scores with *SbMyb60*, which indicates that *SbMyb60* may also regulate these genes in order to increase the concentrations of substrates and cofactors required for monolignol biosynthesis. Genes related to C and N metabolism were also altered in *35S::SbMyb60* stalks and leaves. L-Phenylalanine is the major substrate for PAL, and its byproduct, NH_3, can be recycled through glutamine synthetase/

New Phytologist

Table 3 Transcription factors (TFs) with high coexpression values with genes assigned to the diterpenoid and α-linolenic acid biosynthesis pathways in *35S::SbMyb60* sorghum stalks and leaves

Gene ID	TF	AT homolog	*n*	Ave. TOM	Function of AT homolog
Sobic.004G037800	B3	AT5G20730	11	0.292	Auxin-regulated; response to ethylene
Sobic.006G107700	MADS	AT3G57230	11	0.278	N/A
Sobic.005G019800	ZF-HD	AT3G28917	11	0.263	N/A
Sobic.010G073600	B3	AT5G20730	11	0.26	Auxin-regulated; response to ethylene
Sobic.003G353700	MYB	AT3G46640	11	0.256	Circadian clock
Sobic.002G212000	AP2	AT1G53910	11	0.256	Ethylene response
Sobic.001G481600	AP2	AT1G72360	11	0.254	Ethylene response
Sobic.006G175500	bHLH	AT1G63650	11	0.252	Jasmonate-induced anthocyanin accumulation
Sobic.005G230000	GRAS	AT1G07530	11	0.251	Stress response
Sobic.003G034500	MYB	AT2G36890	11	0.24	Axillary meristem formation
Sobic.007G217700	WRKY	AT2G38470	11	0.234	Abiotic stress
Sobic.001G386600	HB	AT3G03660	11	0.229	Root organogenesis
Sobic.009G216900	MYB	AT5G26660	11	0.228	N/A
Sobic.010G085400	MADS	AT2G22540	11	0.22	Thermosensory pathway
Sobic.010G253300	B3	AT1G19220	10	0.262	Ethylene response
Sobic.005G224700	MYB	AT3G46130	10	0.25	N/A
Sobic.001G002500	AP2	AT3G23240	10	0.244	Ethylene response
Sobic.008G020700	ZF-HD	AT3G28917	10	0.241	N/A
Sobic.002G235500	bZIP	AT3G30530	10	0.241	N/A
Sobic.005G120600	SBP	AT5G50670	10	0.236	N/A
Sobic.003G052900	LBD	AT5G63090	10	0.235	Lateral organ development
Sobic.001G389000	WRKY	AT4G11070	10	0.235	N/A
Sobic.002G272900	HB	AT2G46680	10	0.227	ABA-dependent; drought response
Sobic.001G503000	LBD	AT2G42440	10	0.226	N/A
Sobic.004G267400	bHLH	AT4G34530	10	0.225	Floral initiation
Sobic.003G076600	MYB	AT5G58900	10	0.222	N/A
Sobic.002G242500	WRKY	AT5G56270	10	0.209	Embryo development
Sobic.003G085600	AP2	AT1G68550	10	0.205	N/A
Sobic.002G287400	Nin-like	AT1G76350	10	0.198	N/A

Several TF candidates linked to the transcriptional downregulation of methylerythritol and α-linolenic acid biosynthesis pathways were identified using coexpression analysis. Candidates were considered if they had high Topological Overlap Matrix (TOM) values (≥0.10) with at least 10 of the genes assigned to the two pathways. AT homolog, highest scoring Arabidopsis BLASTP match; Ave. TOM, average TOM score between the TF and all of the genes assigned to the two pathways; N/A, unknown.

glutamine oxoglutarate aminotransferase (GS/GOGAT) (Razal *et al.*, 1996), which explains the observed increased gene expression and metabolite levels linked to N recycling pathways in both *35S::SbMyb60* stalks and leaves. In addition to N metabolism, the expression levels of glycolysis-related genes were also altered in both tissues, which could produce more C substrates (PEP) for aromatic amino acid biosynthesis (Dauwe *et al.*, 2007). Finally, the levels of several metabolites of lysine were elevated in *SbMyb60* stalks relative to WT, which suggests that this amino acid was catabolized in this tissue in *35S::SbMyb60* plants. The levels of these same metabolites were elevated in developing protoxylem elements, which may provide N for L-phenylalanine biosynthesis (Ohtani *et al.*, 2016). These expression changes of genes involved in N recycling and TCA/glycolysis could increase the available N and C for the biosynthesis of L-phenylalanine precursors.

Despite the fact that monolignol, aromatic amino acid, SAM and CoA biosynthesis pathways were similarly impacted in both tissues, *35S::SbMyb60* had divergent effects on the global transcriptomic and metabolomic profiles in these two tissues. Overall, a significantly larger number of DEGs were observed in *35S::SbMyb60* stalks, with 4073 and 3240 more genes upregulated and downregulated in stalks relative to leaves, respectively. DEGs present in stalks were also associated with a larger number of KEGG pathways. The majority of genes whose expression levels were impacted in stalks were not directly related to monolignol biosynthesis, and instead included the biosynthesis of purine and pyrimidine nucleotides, protein biosynthesis and processing, DNA replication, and starch and sucrose metabolism. *SbMyb60* expression levels were significantly higher in leaves relative to stalks, which, in part, may explain some of the differences in the global transcript profiles between these two tissues. In addition, the interactions between *SbMyb60* and the different regulatory factors present in these two tissues probably led to the distinct transcriptomic profiles observed in leaf vs stalk tissue.

In addition to affecting lignin deposition, *35S::Myb60* also impacted other phenolic compounds in cell walls. The esterified cell wall-bound phenolic compounds sinapic, ferulic, caffeic and 4-hydroxybenzeneacetic acids were elevated in Myb2a relative to WT. Previously, elevated levels of soluble caffeic, ferulic and sinapic acids have been detected in biomass from this line (Scully *et al.*, 2016). In concert, the expression levels of BAHD acyltransferases, whose gene products acylate GAX with hydroxycinnamyl-CoA and lead to the formation of hydroxycinnamyl acids

esterified to hemicellulose (D'Auria, 2006), were elevated in the current study. In grass cell walls, significant levels of these esterified phenolic compounds are present outside of lignin polymers that allow for cross-linking (Hatfield *et al.*, 2017). Together, these findings suggest that perturbation of the monolignol biosynthesis pathway can also affect the prevalence of these phenolic groups within cell walls. Surprisingly, the levels of wall-bound phenolics in the current study and the levels of total energy measured in a previous study (Scully *et al.*, 2016) were inversely correlated with the *SbMyb60* expression level: Myb2a stover had the highest overall total energy level and the highest levels of most soluble and wall-bound phenolic compounds. In this line, genes coding for only eight of the 10 monolignol pathway genes showed increased expression relative to WT, and Myb2a had lower PAL and 4CL enzymatic activity relative to the other two overexpression lines (Scully *et al.*, 2016). Incomplete activation of this pathway may result in the accumulation of monolignol intermediates and their subsequent incorporation into the cell wall. The higher total energy level in Myb2a suggests that the energy levels are a result of increased phenolic esters and not lignin. Thus, increasing levels of phenolic esters could be a potential target to increase the total energy levels of biomass for thermoconversion.

In addition, *35S::Myb60* also had an indirect impact on the levels of cell wall polysaccharides. In stalks, *35S::SbMyb60* was associated with the activation of UDP-sugar and cellulose biosynthesis, as indicated by the upregulation of genes associated with these pathways, increased FASGA staining in stalk tissue and elevated UDP-glucose levels. Although UDP-glucose levels and FASGA staining intensity were correlated with *35S::SbMyb60* expression levels, *SbMyb60* expression did not share high coexpression values with genes associated with cellulose or UDP-sugar biosynthesis, and therefore it is probably not a direct regulator of these processes. This finding contrasts with previous studies in which inverse correlations were found between cellulose and lignin levels, which suggests that negative cross-talk can occur between these two pathways (Hu *et al.*, 1999). This finding is also inconsistent with our previous analysis, which showed lower levels of cellulose in stover from *35S::SbMyb60* plants; however, these plants were collected at different developmental stages: at maturity in the previous study and at 8-wk post-germination in the current study (Scully *et al.*, 2016). These observations suggest that the relationship between lignin and cellulose can vary over the course of development, which is supported by reductions in cellulose content post-anthesis (McKinley *et al.*, 2016).

In addition to its impacts on cell wall composition, *35S::SbMyb60* was also associated with stunted growth and delayed flowering relative to WT plants. RNA-Seq and metabolite analyses provided some insights into the potential causes of the delayed flowering and reduced stature of Myb10a and Myb15a plants. Excess lignin deposition may have prevented internode cell expansion and elongation. Internode lengths were significantly shorter in Myb10a and Myb15a stalks relative to WT or Myb2a. SEM indicated irregular cell shapes in Myb10a and Myb15a, which suggests that cell expansion may have been impeded. In addition, there is evidence that *SbMyb60* overexpression shunts C and N resources towards phenylpropanoid metabolism, which may, in part, divert C and N resources away from other metabolic pathways, leading to impaired growth in Myb10a and Myb15a. *SbMyb60* overexpression was also associated with reduced expression levels of genes linked to gibberellin biosynthesis, which may have interfered with internode elongation and delayed flowering in the Myb10a and Myb15a lines. Previously, gibberellins have been linked to growth and cell wall expansion, and gibberellin biosynthesis mutants typically display defects in stalk elongation and flowering time (Silverstone *et al.*, 1998; de Lucas *et al.*, 2008). Despite the lower expression of genes associated with the gibberellin biosynthesis pathway observed in the *35S::SbMyb60* lines, *SbMyb60* was not highly coexpressed with genes assigned to these pathways and is unlikely to be a direct negative regulator.

This study demonstrated that *SbMyb60* not only regulates lignin biosynthesis, but also appears to direct plant metabolism towards lignin biosynthesis. However, one question that remains unresolved is which biochemical steps and biological processes limit the amount of lignin that can be synthesized by sorghum overexpressing *SbMyb60*. As industries shift away from the manufacture of petroleum-based chemical precursors and move towards the use of renewable plant biomass for bioenergy and chemical manufacturing, lignin valorization will be critical for the development of advanced biofuels and bio-based products in sorghum and other C_4 bioenergy grasses. *SbMyb60* may represent a key tool for the modulation of lignin content and cell wall composition in sorghum and other bioenergy feedstocks.

Acknowledgements

We thank John Toy, Sarah Finegan, Nathan Peroutka-Bigus and Dylan Oates for technical assistance, Pat O'Neill for statistical support, and Drs Jim Eudy and Alok Dhar for sequencing assistance. Illumina HiSeq 2000 sequencing was performed at the University of Nebraska Medical Center Genomics Core Facility, Omaha, NE, USA. This research was supported by the US Department of Agriculture – National Institute of Food and Agriculture AFRI grant no. 2011-67009-30026 (S.E.S., T.E.C. and D.L.F-H.) and additional funding from the US Department of Agriculture, Agricultural Research Service (USDA-ARS) CRIS projects 3042-21220-032-00D (S.E.S. and D.L.F-H.) and 3042-21000-030-00D (G.S.). The USDA-ARS is an equal opportunity/affirmative action employer and all agency services are available without discrimination. The mention of commercial products and organizations in this article is solely to provide specific information. It does not constitute endorsement by USDA-ARS over other products and organizations not mentioned.

Author contributions

E.D.S., T.E.C. and S.E.S. designed the research; E.D.S., T.G., N.A.P., L.B., P.T. and J.S. performed the research; E.D.S., N.A.P., G.S., D.L.F-H., L.B., J.S., T.E.C. and S.E.S. analyzed

and interpreted the data; E.D.S., T.G., N.A.P., G.S., D.L.F-H., L.B., P.T., J.S., T.E.C. and S.E.S. wrote the manuscript.

References

Aydin G, Grant RJ, O'Rear J. 1999. *Brown midrib* sorghum in diets for lactating dairy cows. *Journal of Dairy Science* **82**: 2127–2135.

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

Bonawitz ND, Im Kim J, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M *et al.* 2014. Disruption of Mediator rescues the stunted growth of a lignin-deficient *Arabidopsis* mutant. *Nature* **509**: 376.

Bout S, Vermerris W. 2003. A candidate-gene approach to clone the sorghum *Brown midrib* gene encoding caffeic acid O-methyltransferase. *Molecular Genetics and Genomics* **269**: 205–214.

D'Auria JC. 2006. Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology* **9**: 331–340.

Dauwe R, Morreel K, Goeminne G, Gielen B, Rohde A, Van Beeumen J, Ralph J, Boudet A-M, Kopka J, Rochange SF *et al.* 2007. Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. *Plant Journal* **52**: 263–285.

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

Ehlting J, Mattheus N, Aeschliman DS, Li E, Hamberger B, Cullis IF, Zhuang J, Kaneda M, Mansfield SD, Samuels L *et al.* 2005. Global transcript profiling of primary stems from *Arabidopsis thaliana* identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. *Plant Journal* **42**: 618–640.

Fornalé S, Capellades M, Encina A, Wang K, Irar S, Lapierre C, Ruel K, Joseleau J-P, Berenguer J, Puigdomènech P *et al.* 2012. Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for *cinnamyl alcohol dehydrogenase*. *Molecular Plant* **5**: 817–830.

Hatfield RD, Rancour DM, Marita JM. 2017. Grass cell walls: a story of cross-linking. *Frontiers in Plant Science* **7**: 2056.

Hu WJ, Harding SA, Lung J, Popko JL, Ralph J, Stokke DD, Tsai CJ, Chiang VL. 1999. Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nature Biotechnology* **17**: 808–812.

Ibraheem F, Gaffoor I, Tan Q, Shyu C-R, Chopra S. 2015. A sorghum MYB transcription factor induces 3-deoxyanthocyanidins and enhances resistance against leaf blights in maize. *Molecules (Basel, Switzerland)* **20**: 2388–2404.

Kubo M, Udagawa M, Nishikubo N, Horiguchi G, Yamaguchi M, Ito J, Mimura T, Fukuda H, Demura T. 2005. Transcription switches for protoxylem and metaxylem vessel formation. *Genes & Development* **19**: 1855–1860.

Lê Cao K-A, Boitard S, Besse P. 2011. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* **12**: 253.

Li L, Hill-Skinner S, Liu S, Beuchle D, Tang HM, Yeh C-T, Nettleton D, Schnable PS. 2015. The maize *brown midrib4* (*bm4*) gene encodes a functional folylpolyglutamate synthase. *Plant Journal* **81**: 493–504.

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

de Lucas M, Daviere J-M, Rodriguez-Falcon M, Pontin M, Iglesias-Pedraz JM, Lorrain S, Fankhauser C, Blazquez MA, Titarenko E, Prat S. 2008. A molecular framework for light and gibberellin control of cell elongation. *Nature* **451**: 480–484.

Luterbacher JS, Martin Alonso D, Dumesic JA. 2014. Targeted chemical upgrading of lignocellulosic biomass to platform molecules. *Green Chemistry* **16**: 4816–4838.

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

McKinley B, Rooney W, Wilkerson C, Mullet J. 2016. Dynamics of biomass partitioning, stem gene expression, cell wall biosynthesis, and sucrose accumulation during development of *Sorghum bicolor*. *Plant Journal* **88**: 662–680.

Miller FR, McBee GG. 1993. Genetics and management of physiologic systems of sorghum for biomass production. *Biomass and Bioenergy* **5**: 41–49.

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

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

Noda S, Koshiba T, Hattori T, Yamaguchi M, Suzuki S, Umezawa T. 2015. The expression of a rice secondary wall-specific cellulose synthase gene, *OsCesA7*, is directly regulated by a rice transcription factor, OsMYB58/63. *Planta* **242**: 589–600.

Ohtani M, Morisaki K, Sawada Y, Sano R, Uy ALT, Yamamoto A, Kurata T, Nakano Y, Suzuki S, Matsuda M *et al.* 2016. Primary metabolism during biosynthesis of secondary wall polymers of protoxylem vessel elements. *Plant Physiology* **172**: 1612–1624.

Oliver AL, Grant RJ, Pedersen JF, 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. *Journal of Dairy Science* **87**: 637–644.

Palmer NA, Saathoff AJ, Tobias CM, Twigg P, Xia Y, Vogel KP, Madhavan S, Sattler SE, Sarath G. 2014. Contrasting metabolism in perenniating structures of upland and lowland switchgrass plants late in the growing season. *PLoS ONE* **9**: e105138.

Palmer NA, Sattler SE, Saathoff AJ, Funnell D, Pedersen JF, Sarath G. 2008. Genetic background impacts soluble and cell wall-bound aromatics in *brown midrib* mutants of sorghum. *Planta* **229**: 115–127.

Razal RA, Ellis S, Singh S, Lewis NG, Towers GHN. 1996. Nitrogen recycling in phenylpropanoid metabolism. *Phytochemistry* **41**: 31–35.

Rooney W. 2004. Sorghum improvement – integrating traditional and new technology to produce improved genotypes. *Advances in Agronomy* **83**: 37–109.

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

Sattler SE, Funnell-Harris DL, Pedersen JF. 2010. Efficacy of singular and stacked *brown midrib 6* and *12* in the modification of lignocellulose and grain chemistry. *Journal of Agricultural and Food Chemistry* **58**: 3611–3616.

Sattler SE, Saathoff AJ, Haas EJ, Palmer NA, Funnell-Harris DL, Sarath G, Pedersen JF. 2009. A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the sorghum *brown midrib6* phenotype. *Plant Physiology* **150**: 584–595.

Scully ED, Gries T, Sarath G, Palmer NA, Baird L, Serapiglia MJ, Dien BS, Boateng AA, Ge Z, Funnell-Harris DL *et al.* 2016. Overexpression of *SbMyb60* impacts phenylpropanoid biosynthesis and alters secondary cell wall composition in *Sorghum bicolor*. *Plant Journal* **85**: 378–395.

Silverstone AL, Ciampaglio CN, Sun T. 1998. The *Arabidopsis* RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. *Plant Cell* **10**: 155–169.

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

Sun Y, Cheng J. 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology* **83**: 1–11.

Tang HM, Liu S, Hill-Skinner S, Wu W, Reed D, Yeh C-T, Nettleton D, Schnable PS. 2014. The maize *brown midrib2* (*bm2*) gene encodes a methylenetetrahydrofolate reductase that contributes to lignin accumulation. *Plant Journal* **77**: 380–392.

Tuck CO, Pérez E, Horváth IT, Sheldon RA, Poliakoff M. 2012. Valorization of biomass: deriving more value from waste. *Science* **337**: 695–699.

Vanholme R, Cesarino I, Rataj K, Xiao Y, Sundin L, Goeminne G, Kim H, Cross J, Morreel K, Araujo P *et al.* 2013. Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* **341**: 1103–1106.

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

Wang H, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA. 2010. Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proceedings of the National Academy of Sciences, USA* **107**: 22338–22343.

Withers S, Lu F, Kim H, Zhu Y, Ralph J, Wilkerson CG. 2012. Identification of grass-specific enzyme that acylates monolignols with p-coumarate. *Journal of Biological Chemistry* **287**: 8347–8355.

Xia J, Sinelnikov IV, Han B, Wishart DS. 2015. MetaboAnalyst 3.0: making metabolomics more meaningful. *Nucleic Acids Research* **43**: W251–W257.

Xie S, Qin X, Cheng Y, Laskar D, Qiao W, Sun S, Reyes LH, Wang X, Dai SY, Sattler SE *et al.* 2015. Simultaneous conversion of all cell wall components by an oleaginous fungus without chemi-physical pretreatment. *Green Chemistry* **17**: 1657–1667.

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

Zhao Q, Tobimatsu Y, Zhou R, Pattathil S, Gallego-Giraldo L, Fu C, Jackson LA, Hahn MG, Kim H, Chen F *et al.* 2013. Loss of function of *cinnamyl alcohol dehydrogenase 1* leads to unconventional lignin and a temperature-sensitive growth defect in *Medicago truncatula*. *Proceedings of the National Academy of Sciences, USA* **110**: 13660–13665.

Zhao Q, Wang H, Yin Y, Xu Y, Chen F, Dixon RA. 2010. Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. *Proceedings of the National Academy of Sciences, USA* **107**: 14496–14501.

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

Zhong R, Ye Z-H. 2009. Transcriptional regulation of lignin biosynthesis. *Plant Signaling & Behavior* **4**: 1028–1034.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Weighted gene coexpression network analysis (WGCNA) expression profiles for all 13 coexpression modules identified from sorghum stalks and leaves.

Fig. S2 Impacts of *35S::SbMyb60* on nitrogen flux and amino acid metabolism in sorghum leaves and stalks.

Fig. S3 Impacts of *35S::SbMyb60* on sorghum internodes.

Table S1 Genes with high coexpression values with *SbMyb60*

Table S2 Transcription factors with high coexpression values with *SbMyb60*

Methods S1 Supplemental methods.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.