Rice Genome-Scale Network Integration Reveals Transcriptional Regulators of Grass Cell Wall Synthesis

Kangmei Zhao[1†], Fan Lin[1†], Sandra P. Romero-Gamboa[2], Prasenjit Saha[1‡], Hyung-Jung Goh[3], Gynheung An[3], Ki-Hong Jung[3], Samuel P. Hazen[2] and Laura E. Bartley[1*]

[1] Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, United States, [2] Department of Biology, University of Massachusetts, Amherst, MA, United States, [3] Graduate School of Biotechnology & Crop Biotech Institute, Kyung Hee University, Yongin, South Korea

Grasses have evolved distinct cell wall composition and patterning relative to dicotyledonous plants. However, despite the importance of this plant family, transcriptional regulation of its cell wall biosynthesis is poorly understood. To identify grass cell wall-associated transcription factors, we constructed the Rice Combined mutual Ranked Network (RCRN). The RCRN covers >90% of annotated rice (*Oryza sativa*) genes, is high quality, and includes most grass-specific cell wall genes, such as mixed-linkage glucan synthases and hydroxycinnamoyl acyltransferases. Comparing the RCRN and an equivalent *Arabidopsis* network suggests that grass orthologs of most genetically verified eudicot cell wall regulators also control this process in grasses, but some transcription factors vary significantly in network connectivity between these divergent species. Reverse genetics, yeast-one-hybrid, and protoplast-based assays reveal that OsMYB61a activates a grass-specific acyltransferase promoter, which confirms network predictions and supports grass-specific cell wall synthesis genes being incorporated into conserved regulatory circuits. In addition, 10 of 15 tested transcription factors, including six novel Wall-Associated regulators (WAP1, WACH1, WAHL1, WADH1, OsMYB13a, and OsMYB13b), alter abundance of cell wall-related transcripts when transiently expressed. The results highlight the quality of the RCRN for examining rice biology, provide insight into the evolution of cell wall regulation, and identify network nodes and edges that are possible leads for improving cell wall composition.

Keywords: Network, cell wall, transcription factor, hydroxycinnamate, comparative analysis, regulatory evolution

INTRODUCTION

Cultivated grasses are the most abundant sustainable biomass source produced worldwide (Lal, 2005), and cell walls constitute the bulk of plant dry mass available for conversion to biofuels and other bioproducts. In vascular plants, primary walls surround growing cells; whereas, after cessation of growth, secondary walls form around cells such as tracheids, vessels, and fibers. Primary and secondary cell walls consist of both conserved components and those that vary across plant diversity (Liepman et al., 2010; Popper et al., 2011; Fangel et al., 2012). The major vascular plant cell wall

components are cellulose, lignin, and matrix polysaccharides, including hemicelluloses and pectins. Cellulose is synthesized by complexes of Cellulose Synthase A (CESA) proteins (Mutwil et al., 2008; Handakumbura et al., 2013; Schwerdt et al., 2015). Lignin is an aromatic polymer made from covalent crosslinking of phenylpropanoid monomers. Lignin is characteristic of secondary cell walls and forms a barrier for breakdown of cellulose and other wall polysaccharides, including during biofuel production (Bonawitz and Chapple, 2010; Vanholme et al., 2010). Most lignin biosynthesis enzymes function similarly in eudicots as in grasses, though evidence of pathway differences is starting to emerge (Shen et al., 2012; Takeda et al., 2018). Phylogenetic analyses have revealed orthologs of lignin biosynthesis genes and CESAs across eudicots and monocots (Hazen et al., 2002; Penning et al., 2009; Popper et al., 2011; Carpita, 2012).

In contrast to cellulose and lignin, grass cell wall hemicelluloses differ from those of eudicots in composition and relative abundance (Liepman et al., 2010; Pauly et al., 2013). While in eudicots the major hemicellulose classes are xyloglucans, xylans, and mannans, in grasses, xylans and ß-(1,3;1,4) mixed-linkage glucan (MLG) are most abundant. Proteins of the grass-expanded glycosyl transferase (GT) 61 clade, including OsXAX1, OsXAT1, and OsXYXT1, add arabinose and more complex branches to grass xylans that are absent in most eudicots (Anders et al., 2012; Chiniquy et al., 2012; Zhong et al., 2018). Cellulose synthase-like (CSL) proteins, including OsCSLF6, OsCSLF8 and OsCSLH1, incorporate MLG into grass primary and secondary cell walls (Burton et al., 2006; Vega-Sanchez et al., 2012; Kim et al., 2015). Phylogenetic reconstruction suggests that the MLG-synthesizing CSLs emerged in grasses from expansion of the relatively conserved cellulose synthase-like gene families (Hazen et al., 2002; Popper et al., 2011). In addition, commelinid monocot lignin and arabinoxylan are esterified with the hydroxycinnamic acids, ferulic acid and *p*-coumaric acid. Several so-called "BAHD" acyl-CoA acyltransferases (ATs) from grasses hydroxycinnamoylate cell wall precursors (Withers et al., 2012; Bartley et al., 2013; Petrik et al., 2014; Karlen et al., 2016; De Souza et al., 2018). The cell wall BAHD clade includes 0 to 2 members in eudicots, but 10 to 20 members in commelinids (Karlen et al., 2016; De Souza et al., 2018). Detailed phenotypic analysis of the closest Arabidopsis ortholog failed to reveal a cell wall phenotype (Rautengarten et al., 2012), suggesting that the monocot enzymes may have evolved this function subsequent to the divergence of eudicots and commelinids, which is consistent with occurrence patterns of feruloylation of arabinoxylan (Harris and Trethewey, 2010). To facilitate communication, we refer to genes encoding the GT61, CSLF/H, and AT clades mentioned above as "grass-specific" relative to Arabidopsis.

Along with delineating cell wall synthesis enzymes, researchers have made significant progress in understanding regulation of secondary cell wall synthesis, especially for eudicots such as Arabidopsis and *Medicago*. More than 30 eudicot secondary cell wall regulators have been confirmed through detailed forward and reverse genetic analyses [reviewed in: (Zhong et al., 2007; Zhong et al., 2010; Wang and Dixon, 2012)]. A few NAC (NAM, ATAF1, 2 and CUC2) proteins are top-level activators (Mitsuda et al., 2007; Zhong et al., 2010; Zhou et al., 2014). For example, Arabidopsis

NAC SECONDARY WALL THICKENINGS PROMOTING FACTOR1 (NST1), NST2, and Arabidopsis SECONDARY WALL-ASSOCIATED NAC PROTEIN1 (SND1, also known as NST3), function redundantly to activate overall secondary cell wall biosynthesis by enhancing expression of downstream transcription factors and cell wall biosynthesis genes (Appenzeller et al., 2004; Mitsuda et al., 2007; Wang and Dixon, 2012). Numerous R2R3 MYB family members function in secondary cell wall regulation (Zhao and Bartley, 2014). For example, AtMYB46 is a direct target of AtSND1 and can activate additional cell wall-associated transcription factors, *CESA*s, and lignin biosynthesis genes (Zhong et al., 2007; Ko et al., 2009; McCarthy et al., 2009; Zhong and Ye, 2014; Kim et al., 2014). Another example, AtMYB61, regulates plant resource allocation partly by activating cell wall synthesis genes and transcription factors (Newman et al., 2004; Romano et al., 2012). Recent large-scale promoter-transcription factor interaction experiments for Arabidopsis have expanded the likely cell wall-regulating transcription factor complement to over 200 proteins from multiple families and emphasized the feed-forward loop topology of the secondary cell wall biosynthesis regulatory network (Taylor-Teeples et al., 2015).

Fewer functional studies of cell wall regulation have been conducted in grasses [reviewed in: (Gray et al., 2012; Handakumbura and Hazen, 2012; Rao and Dixon, 2018)]. Our deepest understanding is arguably of a pair of negative regulators, ZmMYB31 and ZmMYB42, and their orthologs in rice, sorghum, and switchgrass, which repress secondary cell wall biosynthesis (Sonbol et al., 2009; Fornale et al., 2010; Rao et al., 2019) and bind to promoters of lignin biosynthesis gene *in vivo* (Agarwal et al., 2016). The Arabidopsis orthologs, AtMYB4 and AtMYB32, also repress lignin biosynthesis (Jin et al., 2000; Preston et al., 2004). Analysis of a battery of rice transgenics supports several orthologs of Arabidopsis secondary cell wall transcription factors as regulators of lignin biosynthesis (Hirano et al., 2013b). Similar results have been published recently with switchgrass (Rao et al., 2019). For example, OsMYB61a can bind to the promoters of and regulate secondary CESA expression (Hirano et al., 2013b; Huang et al., 2015). Likewise, AtSND1 orthologs in rice and other grasses (known as OsSWN1 in rice) activate secondary cell wall formation when expressed in Arabidopsis (Zhong et al., 2011) and when overexpressed in rice and switchgrass (Chai et al., 2015; Rao et al., 2019). From functional studies such as these, conservation of gene complement (Zhao and Bartley, 2014), and network analyses (Hirano et al., 2013b), secondary cell wall regulation appears to be conserved across angiosperms.

An outstanding gap that has not been systematically examined is the regulation of grass-specific cell wall genes. A recent analysis revealed that a trihelix family transcription factor, BdTHX1, binds to the promoter of Brachypodium *CSLF6* (Fan et al., 2018); however, the function of this gene has not been examined in eudicots. In general, grass-specific genes might be controlled either by novel or conserved regulators, or both.

To obtain an overview of grass cell wall regulation and distinguish between models of conservation and divergence in regulation of grass-specific genes, we turned to gene network analysis, which has been used successfully to decipher regulatory pathways and complex traits in many organisms (Movahedi et al.,

2011; Mutwil et al., 2011; Yeung et al., 2011; Hirano et al., 2013a; Hirano et al., 2013b; Sarkar et al., 2014; Obertello et al., 2015; Taylor-Teeples et al., 2015). Rice gene coexpression networks include the Rice Oligonucleotide Array Database (ROAD), Oryza Express, RiceArrayNet, Rice GeneNet Engine, and RiceFREND (Lee et al., 2011; Cao et al., 2012; Ficklin and Feltus, 2013; Hirano et al., 2013a). Among these, ROAD and PlaNet permit download of the whole network, facilitating large-scale comparisons and cross-validation. Other so-called "functional gene networks" combine co-expression data with protein-protein interactions and other functional and physical association evidence. In particular, RiceNet (version 1, v1) of Lee et al. (2011) combines various co-expression and protein-protein interaction data from rice, Arabidopsis, *C. elegans*, human, and yeast to provide a Bayesian likelihood score of a functional association. Analysis of high-scoring genes from RiceNet showed that 13 of 14 previously unstudied network neighbors were capable of protein–protein interactions with "bait" genes and reverse genetics revealed that at least three of five genes function in the predicted process (Lee et al., 2011). This validation rate is much higher than typical for coexpression networks. For example, the Arabidopsis component of PlaNet gave a validation rate of ~10% based on screening of T-DNA mutants for embryo lethality (Mutwil et al., 2011). This suggests that the approach of combining multiple types of data across multiple species provides a high-quality network with a reasonable ability to predict functional interactions. RiceNet v2 expands v1 by incorporating newer transcriptomics data and other updated genomic and molecular evidence (Lee et al., 2015). Similar co-expression and multi-data, Bayesian networks for Arabidopsis are also available (Obayashi et al., 2014; Lee et al., 2015).

Here, we combined publicly available, rice co-expression networks with a high-quality Bayesian network to create a novel, comprehensive, genome-scale network, the ice ombined mutual anked etwork (RCRN). Our goal was to study the regulation of grass cell wall biosynthesis relative to that of Arabidopsis. Our analysis suggests that orthologs of almost all Arabidopsis cell wall regulators are present in rice; however, some have different relative importance compared to those in Arabidopsis. Transient assays confirmed that four orthologs of Arabidopsis known cell wall transcription factor can activate cell wall biosynthesis genes. In addition, 6 out of 11 regulators that had not previously been examined for cell wall function control rice cell wall biosynthesis based on transient gene expression assays. Molecular genetics and direct binding assays show that OsMYB61a, a rice ortholog of the known cell wall regulator, AtMYB61, can bind to promoters of both CSL and AT grass-specific genes. This supports the model that grass-specific cell wall genes have been incorporated into regulatory cascades shared with eudicots.

METHODS

Generation of the Rice and Arabidopsis Combined Ranked Networks

We constructed the RCRN to be a comprehensive, high-quality rice genome scale network based on three publically available networks, namely, ROAD, PlaNet and RiceNet v2.

The goal of combining three networks is to expand the high quality network by recalibrating their association scores and covering more rice genes, which allows us to study grass-specific pathways. The three original rice networks, ROAD, PlaNet and RiceNet have three different score systems, Pearson Correlation Coefficient (PCC), Highest Reciprocal Rank (HRR) and Log Likelihood Score (LLS), respectively. For ROAD, we only included positive correlations with a score from 0.5 to 1 (Cao et al., 2012). PlaNet is a collection of different species networks and we only included the rice dataset into our study. PlatNet was built based on HRR and the score range is from 0 to 200 with increments of 1 (Mutwil et al., 2010; Mutwil et al., 2011). RiceNet v2 used log likelihood scores (LLS) to incorporate diverse proteomics, genomics and comparative genomics datasets likely related to rice biological process with scores ranging from 1 to 5 (Lee et al., 2010; Lee et al., 2011; Lee et al., 2015). To combine the three rice networks, we scaled different score systems using inverse mutual rank (MR) as follows: 1/MR = 1/sqrt (rank (A, B) × rank(B, A)) (Usadel et al., 2009). To apply the RiceNet scoring system to represent interactions between additional genes present in the other networks, we computed coefficients using a generalized linear (GLM) model in R based on 1,282 and 3,389 common edges among ROAD, PlaNet and RiceNet v2, respectively. This yielded the following equation:

$$\frac{1}{RiceNet.v2\_MR} = \frac{0.33}{ROAD\_MR} + \frac{0.025}{PlaNet\_MR}$$

(Equation I).

We then rearranged equation I to calculate the RCRN, as follows:

$$RCRN = \frac{1}{RiceNet.v2\_MR} + \frac{0.33}{ROAD\_MR} + \frac{0.025}{PlaNet\_MR}$$

(Equation II).

To facilitate examination of the hub genes in rice and Arabidopsis cell wall networks, we also created the combined Arabidopsis network by integrating the functional network, AraNet v2, and co-expression network from ATTEDII. AraNet v2 used log likelihood score (LLS) to incorporate diverse proteomics, genomics and comparative genomics datasets likely related to rice biological process with scores ranging from 1 to 5 (Lee et al., 2010; Lee et al., 2015). For ATTED II, we only included positive correlations with the score from 0.5 to 1 (Aoki et al., 2015). A GLM yielded the Arabidopsis combined ranked network by calibrating ATTED II network edges based on AraNet using Equation III:

$$ACRN = \frac{1}{AraNet.v2\_MR} + \frac{0.27}{ATTED\_MR}$$

These networks are available for download at: https://doi.org/10.5061/dryad.zgmsbcc69

Network Performance Assessment Based on Gene Ontology

We evaluated network quality based on Gene Ontology (GO) terms annotated by the Biofuel Feedstock Genomics Resources. In all, 40% of rice genes have been assigned the GO-biological process (BP) terms. As with assessment of RiceNet v2, we excluded 10 general GO-BP terms to avoid bias towards these common terms (Lee et al., 2015). We defined true positives as the number of edges with matched GO-BP terms with scores higher than a particular cutoff. True negatives are defined as the number of edges with unmatched GO-BP terms with scores lower than the cutoff. False positives are the number of edges unmatched GO-BP terms with scores higher than the cutoff. False negatives are defined as the number of edges with matched GO-BP terms with scores lower than the cutoff. For each network in the analysis, we applied 40 different cutoffs to generate the ROC curves by plotting true positive rate vs. false positive rate (Lee et al., 2004; Mcgary et al., 2007).

For the precision-recall analysis, precision was calculated as the proportion of true positive edges among all predictions at particular edge score cutoff. Recall represents the proportion of true positive edges relative to total true positives. Then, we defined the total number of edges with matched GO-BP terms within each whole (or trimmed) network as the Total True. At each network edge cut-off the fraction of True Positives (TP) is the number of edges with matched GO-BP terms over Total True Positives. The number of predictions (N) is defined as the number of edges within each network with a particular edge cut-off. Precision = TP/N. Recall = TP/Total True (Mcgary et al., 2007). As a control for this analysis, we built a random network by randomly assigning edges between a pair of genes within the rice genome.

Network Comparisons

We constructed cell wall-only networks in Arabidopsis and rice by extracting interactions from the ACRN and RCRN without cutoffs. We used Inparanoid (Remm et al., 2001) to identify orthologs of Arabidopsis cell wall related genes in rice (http://inparanoid.sbc.su.se/cgi-bin/index.cgi) and phylogenetic reconstructions of the R2R3 MYB family (Zhao and Bartley, 2014). To compare the network connectivity between species, we isolated interactions between each rice transcription factor with its Arabidopsis ortholog and other cell wall-related genes and the total number of genes in RCRN and ACRN without edge score cut-off. If co-orthologs exist, we counted the union of interactions of both co-orthologs. Fisher's exact test was used to determine the statistically different network connectivity for each set of (co-)orthologs.

Transcription Factor Expression Patterns

Expression data for Arabidopsis cell wall-associated transcription factors were extracted from the Arabidopsis gene expression atlas (Schmid et al., 2005). For rice, data were extracted from the rice gene expression atlas (Wang et al., 2010). The gene expression heatmaps were plotted with the heat.map 2 function in R using default hierarchical clustering for row dendrograms.

Construction of the Rice Cell Wall Network

To identify putative novel transcription factors controlling cell wall biosynthesis in rice, we constructed a 1-step network with 125 seed genes with the sum of inversed mutual rank score ≥0.03. This network includes 1,790 nodes and 215 of them are transcription factors. To better select candidates controlling rice cell wall biosynthesis, we excluded transcription factors with fewer than five edges with cell wall seed genes. In all, we predicted 96 transcription factors from 19 protein families as putative novel regulators of cell wall biosynthesis, as summarized in **Supplementary Table 4**.

Characterization of *myb61a-1* Insertion Mutants

We characterized an insertion mutant line for *OsMYB61a*, *PFG_2D-10906*, called *myb61a-1*, which possesses the T-DNA insertion from *pGA2707* in *Oryza sativa japonica* cv. *Dongjin* (An et al., 2005; Jeong et al., 2006). The line *2D-10906-11* was found to be homozygous for the insertion and line *2D-10906-8* was identified and used as the negative segregant. Genotyping primers are listed into **Supplementary Table 6**.

We measured gene expression using the 5th leaf (numbered from the bottom) harvested from 2-month old plants, choosing morphologically and developmentally matched leaves for analysis, based on plant size, leaf length and expansion. RNA was extracted with a Zymo Quick RNA Extraction Kit. We used 1 µg RNA to synthesize cDNA with Promega MMLV reverse transcriptase kit. We ran quantitative PCR with BioRad SYBR Green Master Mix and BioRad CFX96 thermocycler. qPCR primers and locus IDs for genes measured in this study are listed in **Supplementary Table 7**. To analyze gene expression data, we first calculated the real-time primer efficiency with LinRegPCR (Ruijter et al., 2009). Gene expression data were normalized to two reference genes, *Cc55* and *Ubi5*, which show stable expression level during rice development (Jain, 2009). Student two-tailed t-tests were used to compare expression between wild-type and mutant plants. False positives were controlled using q-value to estimate the false discovery rate <0.05 (Lee et al., 2015).

Cell Wall Assays

Five biological replicates of developmentally matched leaf and stem samples from 3-month old wild-type and *myb61a-1* plants were used for all cell wall assays. Alcohol insoluble residue (AIR) was prepared by boiling in 95% ethanol (1:4, w/v) for 30 min followed by washing with 70% ethanol and drying. Destarching to generate dsAIR, lignin *via* acetylbromide solubility, and cellulose content measurements were as previously described (Bartley et al., 2013). Mixed-linkage glucan (MLG) was measured by an enzyme-based kit (Megazyme, K-BGLU) with 5 mg of stem dsAIR as per the manufacturer's directions. Cell wall-associated hydroxycinamic acids (e.g., FA and *p*CA) were

examined in *myb61a-1* mutants and negative segregant plants. To release hydroxycinnamic acids from AIR, we treated 2 mg leaf AIR samples with 2 N NaOH for 24 h at 25 °C and analyzed the results with high performance liquid chromotography as described in Bartley et al. (2013). Student two-tailed t-tests were used to compare cell wall composition between wild-type and mutant plants.

Transient Gene Expression Assay in Rice Protoplast

All transcription factors were cloned from *Kitaake* rice RNA into a pENTRY-D TOPO vector (Invitrogen) with primers summarized in **Supplementary Table 6**. *p2GW7* was used for overexpression (Karimi et al., 2007) in 2-week old, dark-grown *Kitaake* rice seedlings protoplast as previously described (Bart et al., 2006). Gene expression was measured through qPCR as described above with primers as listed in **Supplementary Table 7**.

A dexamethasone (DEX) inducible system was used to examine the downstream targets of OsMYB61a in rice protoplasts with and without the treatment of a protein inhibitor, cycloheximide (CHX). The *OsMYB61a:GR* sequence was cloned into the overexpression vector of *p2GW7* (Karimi et al., 2007) and transformed into protoplasts. DEX (10 μM) was used to induce translocation of OsMYB61a:GR from the cytoplasm to nucleus; control cells were treated with ethanol without DEX. Protein synthesis was blocked by treating protoplasts with 2 μm CHX 30 min prior and during DEX induction. Protoplast were cultured 8 h with DEX before RNA extraction. Four replicates were used in each assay.

Yeast-One-Hybrid Assays

Full-length *OsMYB61a* coding sequence was cloned in-frame into the *GAL4* activation domain in the *pDEST22* vector (Invitrogen). Promoter fragments of (∼ > 1700 bp upstream of the transcription start site) of *OsCESA4*, *At3G62160*, *OsAT4*, *OsAT5*, *OsCSLF6*, and *OsCSLH1* were introduced into Gateway-compatible *pLUC* (*pLacZi* with replacement of *lacZ* with *gluc*) via LR recombination (Invitrogen), linearized with, and transformed

into the *Saccharomyces cerevisiae* strain YM4271 (Deplancke et al., 2004; Deplancke et al., 2006) using 50% PEG-3350, 10× TE, and LiAc, as previously described (Walhout and Vidal, 2001). Transformations were plated in SD-U media at 30 °C for 2 days and colonies were grown in deep plates with 375 μl of SD-U liquid media in a shaking incubator. Before screening, baits were tested for self-activation of the *Renilla* luciferase reporter enzyme using native coelenterazine (nCTZ) substrate (Sigma-Aldrich). Cell culture (20 μl) were transferred to a 96-well flat bottom black plate (Greiner Bio-one) and luciferase activity was measured in a microplate reader (SpectraMax M5) upon addition of nCTZ substrate Mix (1X PBS, 5M NaCl, 1 mg/ml nCTZ solution). Luciferase activity in relative luciferase units (RLU) was normalized by optical density (600 nm). Non-self-active colonies were transformed with the prey, *OsMYB61a*, and empty vector control constructs and SD-TU medium was used for selection of bait-pray transformation. Data were calculated as average fold change relative to the empty expression vector for three biological and four technical replicates. A two-tailed t-test was used to identify statistically different means.

RESULTS

Development of a High Coverage and High-Quality Rice Gene Network

Our goal was to utilize rice genome-scale networks to understand grass cell wall biosynthesis and regulation especially related to grass-specific aspects of the process. However, we found that the publicly available rice networks, ROAD, PlaNet and RiceNet v2, lacked some of the grass-specific cell wall genes available to use as "bait" genes, [including the 20 BAHD acyltransferase; 3 MLG biosynthesis genes, *OsCSLF6*, *OsCSLF8* and *OsCSLH1*; and 2 arabinoxylan modifying genes, *OsXAX1* and *OsXAT1* (grass-specific in **Supplementary Table 1**)]. The Bayesian functional network, RiceNet v2 lacks approximately one quarter of these genes (six, **Table 1**). Thus, this high-quality functional network may be incomplete with respect to grass-diverged cell wall synthesis. On the other hand, the two publicly available co-expression networks, ROAD and PlaNet, are only missing four and one of the grass cell wall genes, respectively. However, these rice co-expression

TABLE 1 | Summary of network features.

Network (Citation)	Network Scoring System	Included grass-specific cell wall genes[a]	Total nodes	Total edges	R² of the power law[b]
ROAD (Cao et al., 2012)	Pearson correlation coefficient (PCC)	21	28,626	8,520,163	0.76
PlaNet (Mutwil et al., 2011)	Highest reciprocal rank (HRR)	24	30,428	3,310,397	0.54
RiceNetv2 (Lee et al., 2015)	Bayesian log likelihood probability	19	25,765	1,775,000	0.88
RCRN (this study)	Sum of inverse rank (SIR)	24	36,419	13,185,506	0.74
ACRN (this study)	Sum of inverse rank (SIR)	NA[c]	24,648	5,319,766	0.69

[a]*Included grass-specific cell wall genes indicates the number out of 25 genes in this category as described in the text and (listed in* **Supplementary Table 1**).
[b]*The power law distribution is P(k) ~ k^t, which represents the probability of a node with k edges with γ as a constant for a given biological network.*
[c]*Not applicable.*

networks have not been experimentally validated and may have lower predictive power, i.e., quality, compared with RiceNet v2.

To overcome the potential depth and quality limitations of existing networks, we developed a genome-scale integrated network suitable for mining grass-diverged traits. Our heuristic strategy was to use a generalized linear model (GLM) to recalibrate the edge scores between genes within ROAD and PlaNet to the scoring system of RiceNet v2. To scale the different scores to a similar range, we first calculated the inverse mutual rank for each network based on their original scores. Inverting the ranking makes greater scores reflect greater confidence. For ROAD, we used only positive correlations; whereas, positive scores for RiceNet and PlaNet include both positive and negative co-expression correlations. The result was the Rice Combined mutual Ranked Network (RCRN) (see Equations I and II in *Methods*).
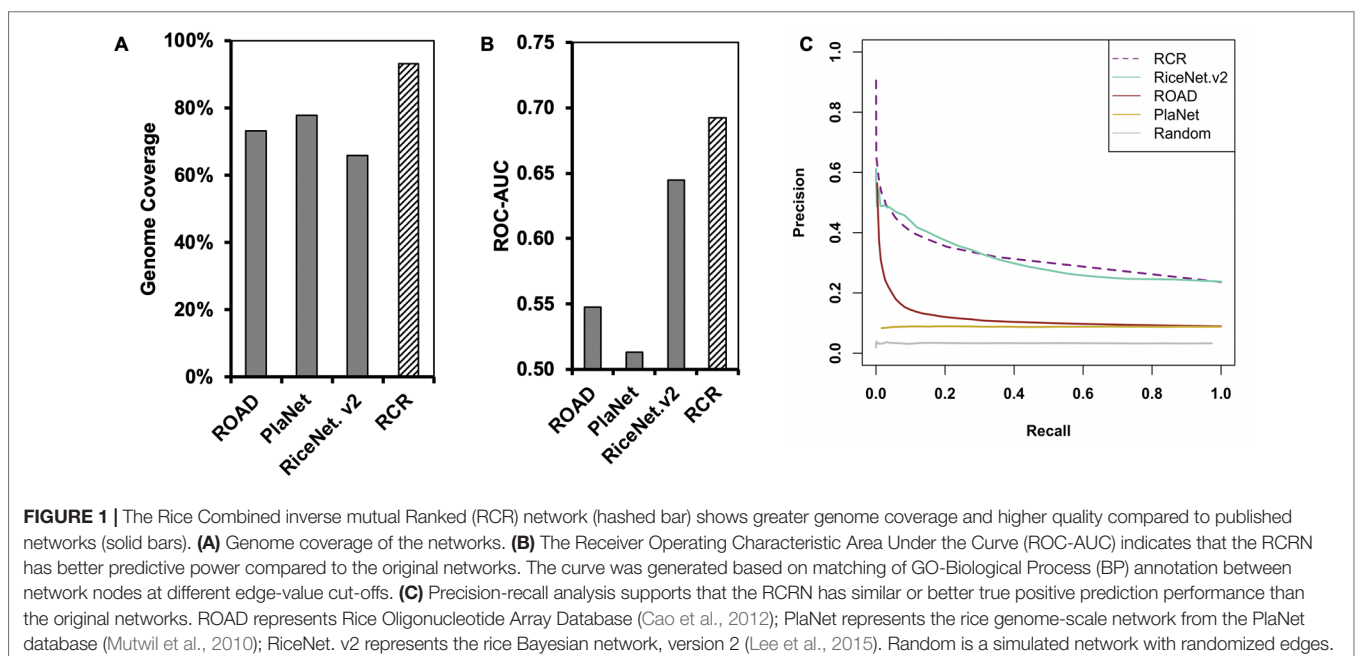
Compared to the original rice networks, the RCRN shows the highest genome coverage and maintains a scale-free topology. The RCRN covers 93% of rice genes (**Figure 1A**, **Table 1**) and misses only one (4%) of our list of grass-specific cell wall genes. This suggests that the RCRN is effective for study of specialized genes or traits of rice and other grasses. Moreover, we analyzed the topology of the networks by calculating fitness to the power law distribution, since biological networks have been found to be scale-free, with a few nodes having a very large number of edges (Barabási and Oltvai, 2004; Siegal et al., 2007; Baxter et al., 2015). All the rice networks fit the power law, though PlaNet fits least well (**Table 1**).

Besides improved genome coverage, the RCRN shows the highest predictive power compared to the three previous networks based on Gene Ontology (GO)-based evaluation. As genes involved in the same pathway tend to be co-expressed and co-regulated (Ashburner et al., 2000; Chang et al., 2013), we evaluated network quality based on Biological Process (BP) GO terms from the Biofuel Feedstock Genomics Resource (Childs et al., 2012).

Forty percent of rice genes have been assigned GO-BP terms. We excluded 10 common GO-BP terms to avoid bias from these high-level, generic terms (Lee et al., 2015). A Receiver-Operating Characteristic (ROC) curve measures the predictive power of each network at a series of edge score cut-offs. The ROC indicates the ratio of likely true positives with matched GO-BP terms compared the likely false positives with unmatched GO-BP terms. The area under the ROC curve (AUC) is higher for the RCRN (AUC = 0.69) than for the other networks (**Figure 1B** and **Supplementary Figure 1**). Precision-recall analysis, which focuses only on true positive predictions at different edge scores, also suggests that the RCRN exhibits a greater proportion of positive edges compared to the co-expression networks and a similar proportion to that of RiceNet v2 (**Figure 1C**).

Comparison of the Rice and Arabidopsis Cell Wall Regulatory Networks

To compare rice and Arabidopsis cell wall regulation, we first tested recall of known cell wall-related interactions in the RCRN by extracting edges between the cell wall "bait" (target) genes in three categories, 1) functionally characterized rice cell wall biosynthesis gene families including those of phenylpropanoid pathway genes, cellulose synthases, "Mitchell-Clade" BAHD acyltransferases, and xylan biosynthesis genes; 2) known grass cell wall-associated transcription factors; and 3) putative orthologs of known Arabidopsis cell wall-associated transcription factors (**Supplementary Table 1**). These 125 cell wall genes are highly interconnected in the RCRN, with their graph possessing 1177 edges when considered without edge-score cut-offs (**Supplementary Figure 2**). This recalls 92% (97 out of 105) of rice orthologs of known transcription factor-cell wall biosynthesis gene associations (**Supplementary Table 2**).

FIGURE 1 | The Rice Combined inverse mutual Ranked (RCR) network (hashed bar) shows greater genome coverage and higher quality compared to published networks (solid bars). **(A)** Genome coverage of the networks. **(B)** The Receiver Operating Characteristic Area Under the Curve (ROC-AUC) indicates that the RCRN has better predictive power compared to the original networks. The curve was generated based on matching of GO-Biological Process (BP) annotation between network nodes at different edge-value cut-offs. **(C)** Precision-recall analysis supports that the RCRN has similar or better true positive prediction performance than the original networks. ROAD represents Rice Oligonucleotide Array Database (Cao et al., 2012); PlaNet represents the rice genome-scale network from the PlaNet database (Mutwil et al., 2010); RiceNet. v2 represents the rice Bayesian network, version 2 (Lee et al., 2015). Random is a simulated network with randomized edges.
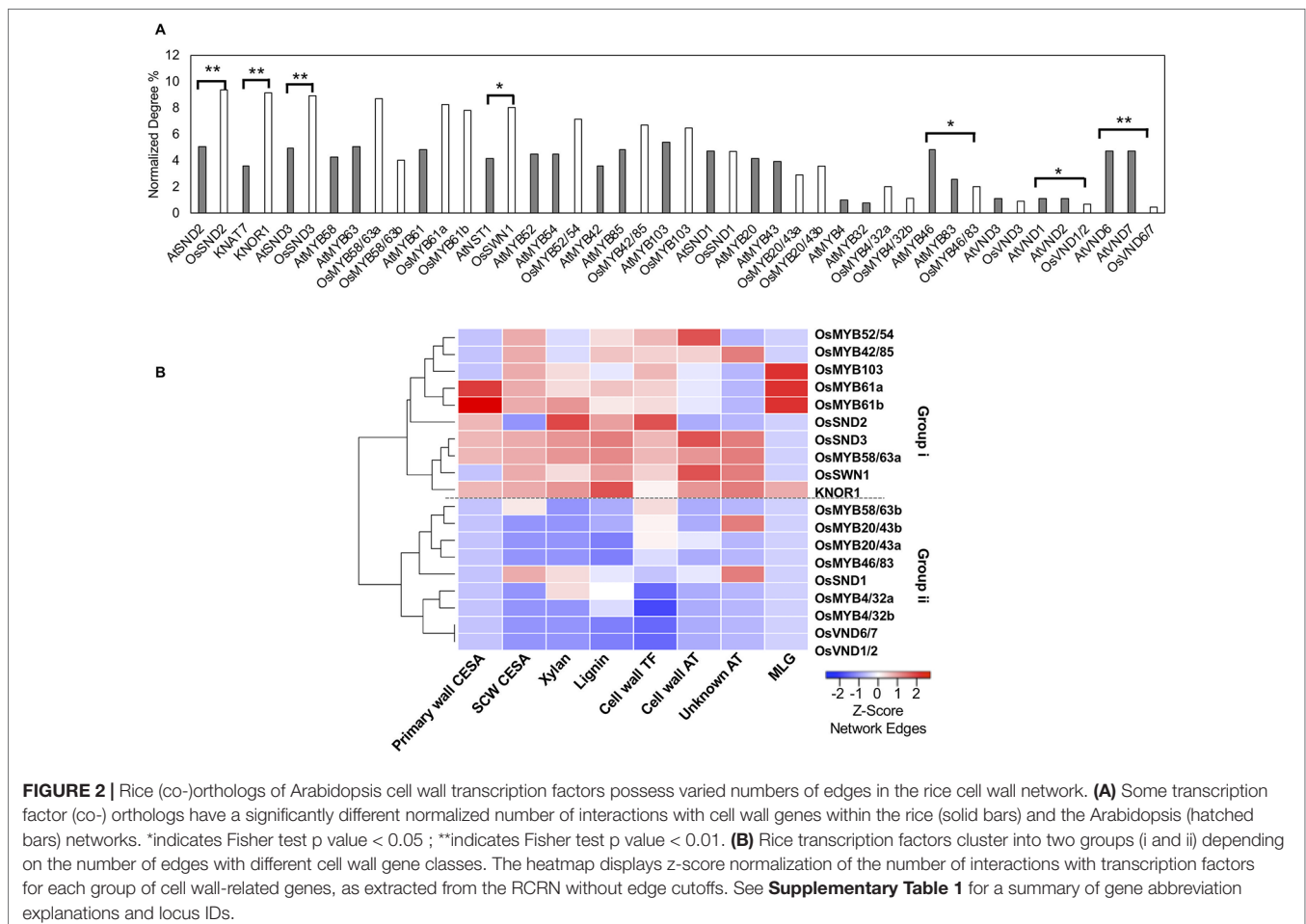
We then created the Arabidopsis Combined mutual Rank Network (ACRN) and extracted a similarly constructed cell wall network including genetically verified regulators, lignin, and cellulose synthesis genes. Like the RCRN, the ACRN combines the Arabidopsis Bayesian functional network, AraNet v2, and the co-expression network, ATTED II, through a GLM. Based on the number of edges with cell wall-related genes in the ACRN, many regulators are highly connected hubs, including AtSND1, AtSND2, AtSND3, AtNST1, AtVND6, and AtVND7, and their targets, including, AtMYB103, AtMYB63, and AtMYB46, among others (**Figure 2A**, **Supplementary Figures 2A**, **3**). That many genetically verified Arabidopsis cell wall regulators possess relatively high numbers of edges in the ACRN is consistent with the observation that "important" regulators are well connected within gene networks (Sorrells and Johnson, 2015). Additionally, many of these hub regulators are highly expressed in Arabidopsis stems compared to relatively less connected regulators in gene expression atlas data (**Supplementary Figure 4**).

We conducted a similar examination of rice orthologs of eudicot cell wall transcription factors in the cell wall network derived from the RCRN (**Supplementary Figure 2B**) and compared the results with those for Arabidopsis. To compare across species, we calculated the union of cell wall edges for co-orthologs (e.g., one gene in Arabidopsis vs. two genes in rice). The rice network has a more varied degree distribution, but still most orthologous gene sets possess similar numbers of edges between the rice and Arabidopsis networks (**Figure 2A**, **Supplementary Table 3**). For example, co-orthologs of AtMYB58 and AtMYB63, OsMYB58/63a and OsMYB58/63b are still highly connected, hub regulators. On the other hand, OsVND6/7, OsVND1/2 and OsMYB46/83 show significantly lower relative degree compared to their co-orthologs in Arabidopsis and relatively low gene expression (**Supplementary Figure 5**). In contrast, the rice ortholog of KNAT7 (named KNOTTED 1 of ice, KNOR1), OsSND2, OsSND3, and OsSWN1 possess significantly R more cell wall edges than their orthologs do in Arabidopsis (**Figure 2A**) and are among the more highly expressed putative rice cell wall transcription factors (**Supplementary Figure 5**). Beyond connections with just cell wall-related genes, we also investigated the connectivity using the total number of edges in the RCRN versus the ACRN and observed that most transcription factors show conserved connectivity, but a few are shifted (**Supplementary Figure 6**).

We further categorized the networks of rice cell wall-related genes based on the components that they synthesize (**Figure 2B**). Group i members have high degree with lignin and xylan biosynthesis genes and both secondary and primary cell wall CESAs. Group ii members show relatively lower degree with the

FIGURE 2 | Rice (co-)orthologs of Arabidopsis cell wall transcription factors possess varied numbers of edges in the rice cell wall network. **(A)** Some transcription factor (co-) orthologs have a significantly different normalized number of interactions with cell wall genes within the rice (solid bars) and the Arabidopsis (hatched bars) networks. *indicates Fisher test p value < 0.05 ; **indicates Fisher test p value < 0.01. **(B)** Rice transcription factors cluster into two groups (i and ii) depending on the number of edges with different cell wall gene classes. The heatmap displays z-score normalization of the number of interactions with transcription factors for each group of cell wall-related genes, as extracted from the RCRN without edge cutoffs. See **Supplementary Table 1** for a summary of gene abbreviation explanations and locus IDs.

classes of cell wall genes considered. The network connectivity and gene expression analysis suggest cases of both conserved and shifted importance in the cell wall biosynthesis regulatory networks between Arabidopsis and rice.

Identification of Additional Cell Wall-Associated Transcription Factors

To systematically identify transcription factors that may control rice cell wall biosynthesis, we examined the higher confidence edges of the Rice Cell Wall Network. This network extends from the 125 cell wall "bait genes" to include nodes from the RCRN with a sum of inverse rank (SIR) edge-score ≥0.03, i.e., the top 30 mutual rank interactions for each bait for a total of 1,790 non-bait nodes and 3,139 edges (**Supplementary Figure 7**). Of these, 215 are annotated as transcription factors and 96 connect with at least five cell wall bait genes. These 96 highly connected transcription factors are from 19 protein families, including multiple members of the MYB, NAC, TALE, AP2/ERF, HD-ZIP, bHLH, WRKY, DBB, C2H2, GATA, ARF, and MICK families along with seven others (**Supplementary Tables 4** and **5**). Twenty-one of the 96 high-degree transcription factors in the RCRN overlap with the novel transcription factors that are highly co-expressed (mutual rank >55) in a rice secondary cell wall network (Hirano et al., 2013a) (**Supplementary Table 5**). Furthermore, 79 have orthologs in Arabidopsis and 58 of those (73%) are part of the cell wall network in the ACRN, 16 (20%) of which have high degree with Arabidopsis secondary cell wall genes (**Supplementary Table 5**), consistent with conservation of cell wall association.

Based on their connection patterns with cell wall biosynthesis genes, the 96 putative uncharacterized wall-associated transcription factors can be divided into three groups (**Supplementary Figure 8**). Group i member share edges with most categories of cell wall genes, except primary cell wall CESAs and MLG biosynthesis genes. Group ii members are relatively less connected; however, a few show specific connections with primary cell wall CESAs and MLG biosynthesis genes. Group iii members connect mostly with cell wall transcription factors (**Supplementary Figure 8**).

Recruitment of Grass-Specific Cell Wall Genes to a Conserved Regulatory Network
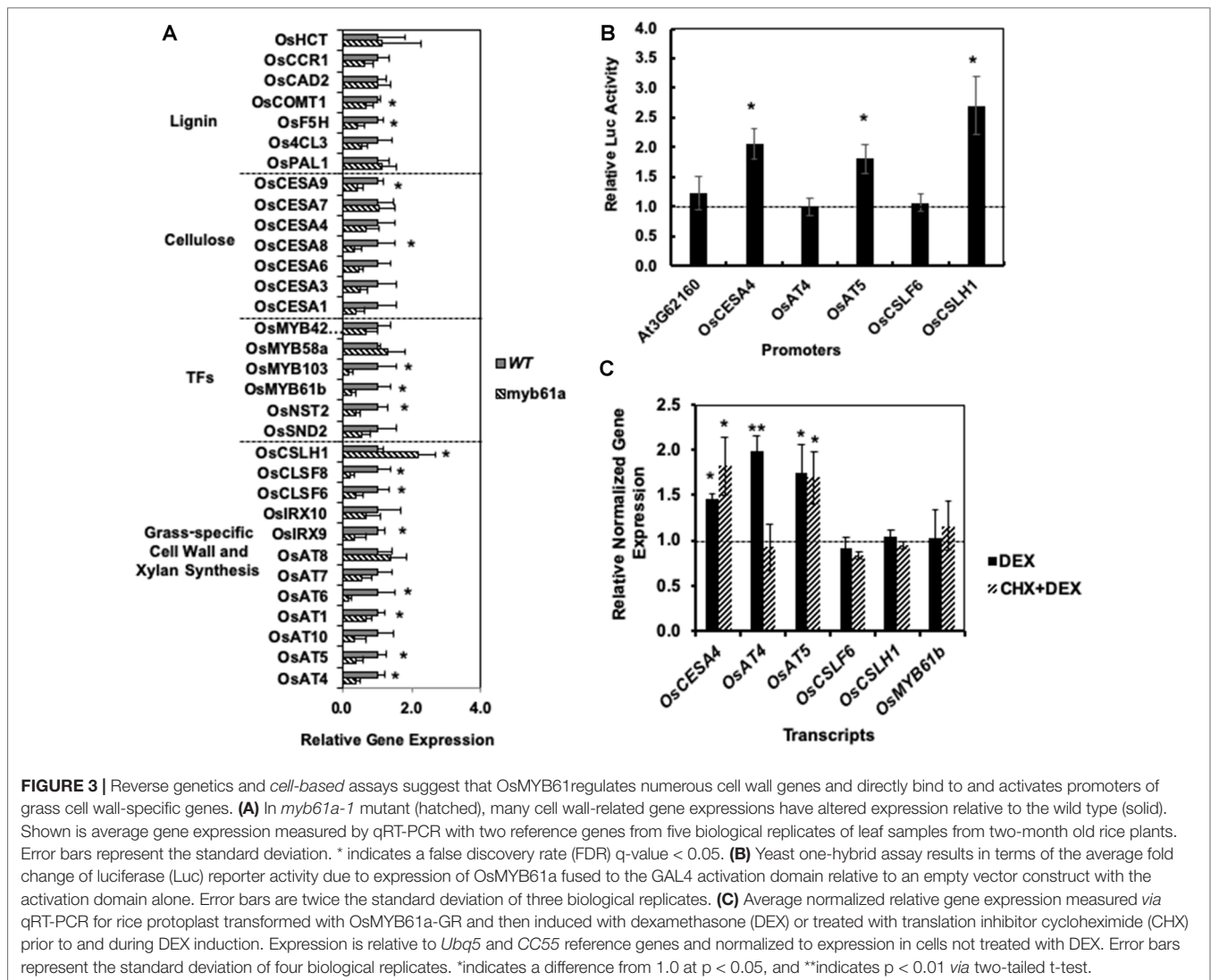
We next conducted functional analysis to validate the RCRN and explore regulation of grass-specific cell wall genes. OsMYB61a is one of two grass co-orthologs of AtMYB61, an activator of cell wall synthesis and other carbon-sink physiology (Romano et al., 2012). The RCRN suggests that OsMYB61a regulates CESA and lignin biosynthesis genes, as previously observed (Hirano et al., 2013b; Huang et al., 2015), and further, that OsMYB61a may control grass-specific cell wall genes (**Figure 2B**). To test this, we characterized a mutant line, *myb61a-1*, which has a T-DNA insertion in the third exon (**Supplementary Figure 9A**). Quantitative reverse transcription PCR (qRT-PCR) indicated that expression of *OsMYB61a* decreases at least five-fold in mature leaves of the mutant compared to those of negative segregant, wild-type plants (**Supplementary Figure 9B**).

Guided by potential regulatory interactions inferred from edges in the RCRN, we tested 32 cell wall-related genes for alterations in expression in *myb61a-1* mutant plants with qRT-PCR (**Figure 3A**). Fourteen genes show a change in gene expression with an average fold-change of 3 (± 1)-fold (q-value <0.05; **Figure 3A**). Expression of lignin biosynthesis genes, *OsCOMT1* and *OsF5H1*, and the secondary cell wall cellulose biosynthesis gene, *OsCESA9*, are modestly, but significantly reduced relative to wild-type plants. In addition, expression of all grass-specific cell wall genes connected with *OsMYB61a* in the RCRN, except for *OsCSLH1*, was significantly reduced in *myb61a-1* compared to in the wild-type (**Figure 3A**). Surprisingly, *OsCSLH1* showed increased expression in *myb61a-1* (2.2-fold, q = 0.04). Though lacking a connection with OsMYB61a in the RCR, two additional BAHD AT-encoding genes, *OsAT1* and *OsAT6*, also showed reduced expression in *myb61a* (**Figure 3A**), though *IRX10*, *OsAT7*, *OsAT8*, and *OsAT10* did not.

To examine whether OsMYB61a controls a regulatory cascade in rice, we measured the expression of six orthologs of Arabidopsis secondary cell wall-associated transcription factors that both share an edge with OsMYB61a and other cell wall synthesis genes in the RCRN and display relatively high expression in rice vegetative development (**Supplementary Figure 5**). *OsMYB61b*, *OsNST2*, and *OsMYB103* all show reduced expression in *myb61a-1* relative to wild type (**Figure 3A**), with *OsMYB103* showing the greatest reduction in expression of any gene assayed at sixfold, consistent with being downstream of OsMYB61a in the rice cell wall transcriptional network.

As expected from the reduction of expression of several cell wall synthesis genes, we found that *myb61a-1* mutant rice leaves and stems exhibit various cell wall phenotypes. Relative to the wild type, acetyl bromide soluble lignin (ABSL) and cellulose content of *myb61a-1* were reduced by 18% (p < 0.05) and 20% (p < 0.01), respectively (**Figure 4**). Furthermore, a lichenase assay showed that the grass-specific polymer, MLG, is reduced by 31% (p < 0.01) in mature *myb61a-1* stems (**Figure 4**). Finally, saponification of cell wall alcohol-insoluble residue (AIR) of leaf samples revealed a trend in reduction of FA and *p*CA of 17% and 11%, respectively (**Figure 4**), though these changes are not statistically significant (p = 0.2 and 0.3, respectively). Consistent with a defect in cell wall structural strength, *myb61a-1* plants also show a dwarf phenotype relative to the wild type (36% decrease, p < 10^5), with each internode of *myb61a-1* being smaller than those of the wild type (**Supplementary Figures 9C, D**).

Next, we assessed whether OsMYB61a can directly bind promoters of grass cell wall biosynthesis genes in two assays. Yeast one-hybrid assays show that OsMYB61a directly binds to the ~1.7-kb promoters of *OsCESA4*, *OsAT5*, and *OsCSLH1* (**Figure 3B**). As a negative control for this experiment, we tested the interaction between OsMYB61a and the promoter of *AT3G62160*, the Arabidopsis homolog of the rice "Mitchell clade" BAHD-acyltransferases, a knockout of which lacks a cell wall hydroxycinnamate phenotype (Rautengarten et al., 2012). We also analyzed the ability of OsMYB61a to directly alter transcription of grass-specific genes in rice seedling leaf-derived protoplasts when regulated by dexamethasone (DEX) with and without treatment with the protein biosynthesis inhibitor, cycloheximide (CHX). We observed that upon DEX-induction,

FIGURE 3 | Reverse genetics and *cell-based* assays suggest that OsMYB61regulates numerous cell wall genes and directly bind to and activates promoters of grass cell wall-specific genes. **(A)** In *myb61a-1* mutant (hatched), many cell wall-related gene expressions have altered expression relative to the wild type (solid). Shown is average gene expression measured by qRT-PCR with two reference genes from five biological replicates of leaf samples from two-month old rice plants. Error bars represent the standard deviation. * indicates a false discovery rate (FDR) q-value < 0.05. **(B)** Yeast one-hybrid assay results in terms of the average fold change of luciferase (Luc) reporter activity due to expression of OsMYB61a fused to the GAL4 activation domain relative to an empty vector construct with the activation domain alone. Error bars are twice the standard deviation of three biological replicates. **(C)** Average normalized relative gene expression measured *via* qRT-PCR for rice protoplast transformed with OsMYB61a-GR and then induced with dexamethasone (DEX) or treated with translation inhibitor cycloheximide (CHX) prior to and during DEX induction. Expression is relative to *Ubq5* and *CC55* reference genes and normalized to expression in cells not treated with DEX. Error bars represent the standard deviation of four biological replicates. *indicates a difference from 1.0 at p < 0.05, and **indicates p < 0.01 *via* two-tailed t-test.
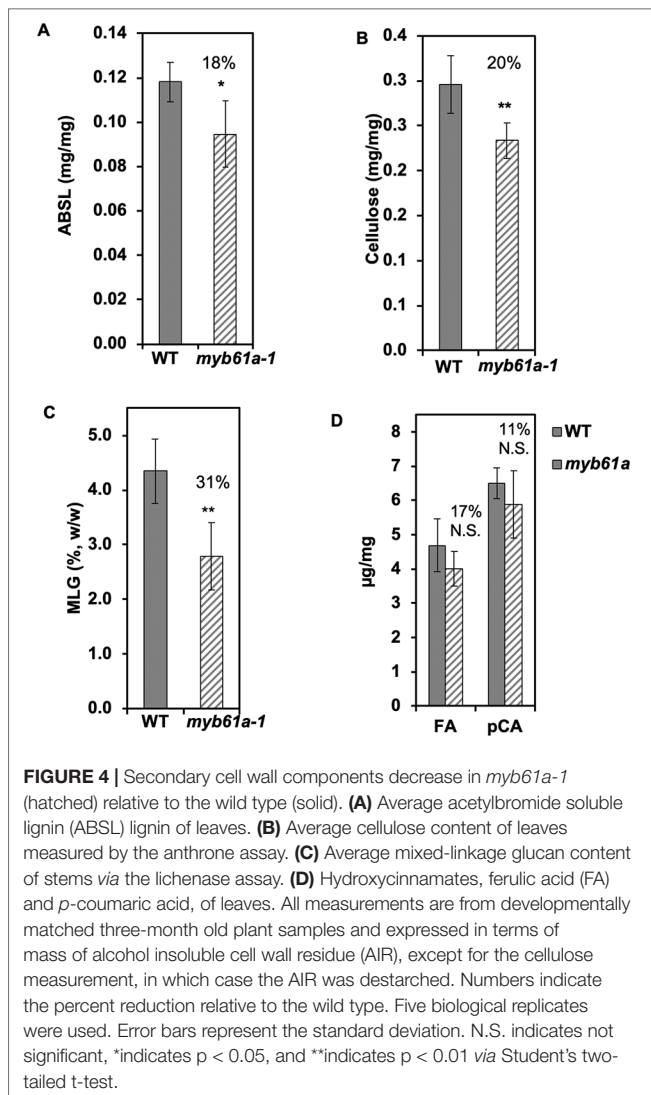
an OsMYB61a-glucocorticoid receptor ligand-binding domain (GR) fusion protein activated expression of *OsCESA4*, *OsAT4*, and *OsAT5*. However, only *OsCESA4* and *OsAT5* were still induced after treatment with CHX, suggesting that OsMYB61a binds directly to these promoters. In contrast *OsAT4* expression activation may rely on interactions with another transcription factor client induced by OsMYB61a (**Figure 3C**). Thus, OsMYB61a is able to directly regulate expression of some grass-specific cell wall genes that eudicots lack.

Functional Validation of Orthologs of Arabidopsis Cell Wall Transcription Factors

To accelerate functional exploration of the RCRN, we tested the ability of four orthologs of known cell wall regulators to alter cell wall-related gene expression in rice seedling-derived protoplasts. Transient transcription factor overexpression was driven by the cauliflower mosaic virus *35S* promoter, which is moderately strong in grass cells (Terada and Shimamoto, 1990).

To test the transient protoplast assay sensitivity and accuracy, we overexpressed *OsMYB61a* and were able to recapitulate many gene expression changes expected from the whole plant studies. We observed that four of nine genes that were decreased in leaves of *myb61a-1* knockout plants increased (1.5- to 2-fold, P < 0.05, **Figure 3A**) in protoplasts over expressing *OsMYB61a*, including, *OsF5H1*, *OsCESA9*, *OsCSLF6* and *OsAT4* (**Table 2**). From this, we conclude that this assay may be less sensitive than whole plant genetic manipulation, but nonetheless, the results support the conclusion that OsMYB61a activates grass-specific cell wall genes.

Next, we examined the effect on cell wall gene expression of overexpression of orthologs of three other characterized Arabidopsis genes, *OsMYB61b*, *OsMYB58/63a* and *OsSND2*, which may also act as hub regulators of rice cell wall synthesis based on network connectivity (**Figure 2**). We found that OsMYB61b, a paralog of OsMYB61a, also activates lignin and cellulose biosynthesis gene expression (Koshiba et al., 2017), and the grass-specific cell wall genes, *OsAT4* and *OsAT5* (**Table 2**). A co-ortholog of the Arabidopsis lignin biosynthesis transcriptional activator

FIGURE 4 | Secondary cell wall components decrease in *myb61a-1* (hatched) relative to the wild type (solid). **(A)** Average acetylbromide soluble lignin (ABSL) lignin of leaves. **(B)** Average cellulose content of leaves measured by the anthrone assay. **(C)** Average mixed-linkage glucan content of stems *via* the lichenase assay. **(D)** Hydroxycinnamates, ferulic acid (FA) and *p*-coumaric acid, of leaves. All measurements are from developmentally matched three-month old plant samples and expressed in terms of mass of alcohol insoluble cell wall residue (AIR), except for the cellulose measurement, in which case the AIR was destarched. Numbers indicate the percent reduction relative to the wild type. Five biological replicates were used. Error bars represent the standard deviation. N.S. indicates not significant, *indicates p < 0.05, and **indicates p < 0.01 *via* Student's two-tailed t-test.

(Zhou et al., 2009), OsMYB58/63a activates four out of five tested lignin biosynthesis genes in protoplasts (**Table 2**, **Figure 5**), consistent with the rice protein having a conserved function with AtMYB58/63. Unexpectedly, OsSND2 may repress cell wall synthesis gene expression in rice, as transient overexpression of *OsSND2* reduced expression of *OsAT5* and *OsCESA9* by approximately 3-fold (**Table 2**, **Figure 5**). The literature reveals some ambiguity in the Arabidopsis ortholog's role as a positive or negative regulator (Zhong et al., 2008; Hussey et al., 2011).

Functional Validation of Novel Putative Rice Cell Wall Regulators

To extend our understanding of secondary cell wall regulation, we selected eleven unstudied, putative cell wall transcription factors from the 96 high degree rice transcription factors in the RCRN (**Figure 5**, **Supplementary Figure 8**). For each transcription factor, we tested its ability to alter expression of cell wall genes with edges in the RCRN. Transient

overexpression in rice protoplast showed statistically significant and repeatable alterations in expression of cell wall genes for 55% (6 out of 11) of the uncharacterized transcription factors consistent with these proteins regulating cell wall biosynthesis (**Table 3**, **Figure 5**).

Among the validated uncharacterized transcription factors, five out of six are activators. The overexpression of Wall-Associated AP2/ERF family protein, WAP1, encoded by *LOC_Os03g08470*, significantly activated *OsF5H1* (**Table 3**, **Figure 5**). To our knowledge, the only AP2/ERF protein previously experimentally demonstrated to function in cell wall regulation is SHINE2/WAX INDUCER (SHN2/WIN) (Ambavaram et al., 2011). WAP1 also has relatively high rank in the rice cell wall network of Hirano et al. (2013a). In addition to WAP1, the Wall-Associated basic Helix-Loop-helix family protein, WAHL1, encoded by *LOC_Os01g11910*, also significantly activated *OsF5H1* (**Table 3**, **Figure 5**).

Expression of the Wall-Associated Homeoomain protein, WAHD1, encoded by *LOC_Os12g43950*, significantly activated *OsCAD2* expression (**Table 3**, **Figure 5**). WADH1 is in the clade neighboring OsBLH6 (Jain et al., 2008), which is another bell-type homeodomain protein in the list of potential cell wall regulators (**Supplementary Table 5**), and which has been shown to activate lignin biosynthesis (Hirano et al., 2013b).

OsMYB13a, encoded by *LOC_Os02g41510*, and OsMYB13b, encoded by *LOC_Os04g43680*, both activated *Os4CL3* transcription. OsMYB13a also activated *OsCOMT1,* whereas OsMYB13b activated *OsCAD2* (**Table 3**, **Figure 5**). We named the two wall-associated R2R3 MYB proteins based on their ortholog in Arabidopsis, AtMYB13, which has not been associated with cell wall regulation to our knowledge.

We observed one repressor, Wall Associated C2H2, WACH1, encoded by *LOC_Os04g08060* (**Table 3**, **Figure 5**) in the protoplast assay. WACH1 repressed *Os4CL3* and the secondary cell wall-associated *OsCESA4*. The Arabidopsis ortholog is involved in stress responses (Ciftci-Yilmaz and Mittler, 2008). This protein also has relatively high rank in both the rice and Arabidopsis cell wall networks and in Hirano et al. (2013a).

DISCUSSION

The altered patterning and composition of grass cell walls compared with eudicots presents the need for regulatory innovation over the course of evolution. This work expands the systematic identification and experimental validation of transcription factors involved in grass cell wall synthesis regulation.

RCRN Promotes Understanding of Rice Molecular Pathways

The RCRN shows greater genome coverage and quality compared to previous, publicly available rice gene networks. The heuristic approach for constructing the RCRN applied inverse mutual rank to the three original networks and then used a general linearized model to calibrate the co-expression network edges relative to the high-quality Bayesian comparative network, RiceNet v2.

TABLE 2 | Expression of cell wall biosynthesis genes in rice protoplasts overexpressing rice (co-)orthologs of Arabidopsis secondary cell wall transcription factors.

	LOC_Os Name	*01g18240* OsMYB61a	*05g04820* OsMYB61b	*02g46780* OsMYB58/63a	*01g48130* OsSND2
Transcription Factors		*190 ± 30[a]*	*320 ± 10[b]*	*590 ± 20[b]*	*270 ± 10[b]*
Lignin	*Os4CL3*	1.0 ± 0.2	*1.4 ± 0.2[a]*	*4.7 ± 0.3[a]*	ND[c]
	OsCCR1	1.2 ± 0.2	1.3 ± 0.2	1.2 ± 0.2	1.0 ± 0.1
	OsCOMT1	1.3 ± 0.3	*1.6 ± 0.1[a]*	*5.1 ± 0.7[a]*	0.8 ± 0.1
	OsF5H1	*1.6 ± 0.2[a]*	*1.6 ± 0.2[a]*	*1.8 ± 0.1[a]*	0.9 ± 0.1
	OsCAD2	ND[c]	ND[c]	*2.5 ± 0.2[a]*	1.3 ± 0.2
Secondary wall CESA	*OsCESA4*	ND[c]	ND[c]	0.9 ± 0.1	0.4 ± 0.3
	OsCESA9	*1.9 ± 0.2[a]*	*2.1 ± 0.3[a]*	1.0 ± 0.3	*0.3 ± 0.1[a]*
Mixed-linkage glucan	*OsCSLF6*	*1.6 ± 0.1[a]*	1.3 ± 0.1	ND[c]	ND[c]
	OsCSLH1	1.1 ± 0.1	1.1 ± 0.2	ND[c]	ND[c]
Hydroxycinnamic acid	*OsAT4*	1.8 ± 0.1[a]	1.7 ± 0.1[a]	1.8 ± 0.2[a]	0.3 ± 0.2[a]
	OsAT5	1.3 ± 0.2	2.0 ± 0.3[a]	1.1 ± 0.1	*0.3 ± 0.1[a]*

Data represent average fold change and standard deviation of the normalized expression (based on reference genes, Ubi5 and CC55) in three biological replicates upon expression of the regulators under control of the 35S promoter relative to empty vector controls. Data are from a single representative experiment. All experiments were repeated independently two to three times and bold and italic font demarcates repeatable significant differences.

[a]*Two-tailed t-test p-value < 0.05.*
[b]*Two-tailed t-test p-value < 0.01.*
[c]*ND indicates the interactions were not determined in this assay, which examined interactions based on RCR network predictions.*
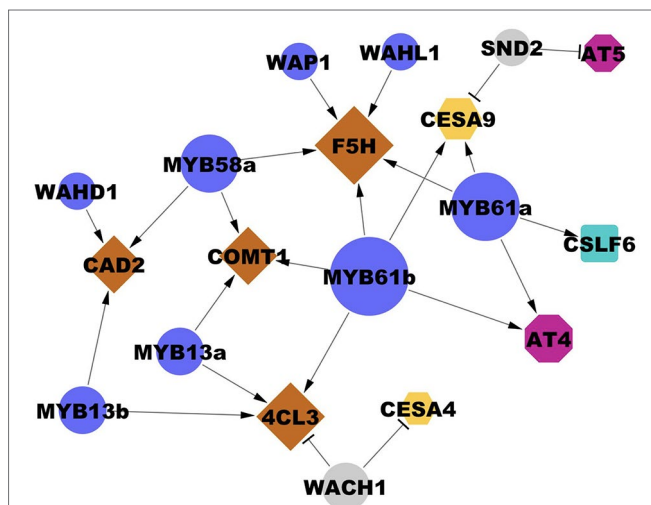

FIGURE 5 | Interactions between transcription factors and cell wall biosynthesis genes validated in the transient assay. Edges ending in a triangle (i.e., arrow head) indicate activation; edges ending in a bar indicate repression. Gene node size is proportional to degree (number of regulatory connections). Blue and grey circles represent cell wall-associated activators and repressors, respectively. Brown diamonds represent lignin biosynthesis genes. Yellow hexagons represent CESAs. Magenta octagons represent cell wall-associated acyltransferases (AT). The blue square represents a mixed-linkage glucan biosynthesis gene. Locus IDs of transcription factors are listed in **Table 2** and **3** and locus IDs of all genes are in **Supplementary Table 1**. Three biological replicates were used in each experiment and each result was observed in at least two independently replicated experiments.

The slightly superior quality of the RCRN over even RiceNet v2 based on gene ontology similarity of connected nodes (**Figure 1**, **Supplementary Figure 1**) may be due to the observation that mutual rank improves reproducibility and overall performance of gene networks (Obayashi et al., 2014).

The RCRN also shows a lower false negative rate than RiceNet v2 based on the experimental gene expression network derived from characterizing the *myb61a-1* mutant line. RiceNet v2 and the RCRN predict 15 and 36 interactions between OsMYB61a and cell wall genes, respectively (**Supplementary Figure 10**). Compared to the gene expression measurements, RiceNet v2 and the RCRN have similar true positive rates of 40% (4 out of 9 validated interactions) and 39% (9 out 23 validated interactions), respectively. On the other hand, the networks differ in their relative false negative rate, which represents validated interactions in the gene expression analysis not predicted by the networks. The false negative percentage for RiceNet v2 is 53%, which is much higher than that of the RCRN, at 8.3%. In particular, RiceNet v2 misses interactions with BAHD-ATs and rice xylan biosynthesis genes (**Supplementary Figure 10**).

Conservation and Divergence of Known Cell Wall Regulators in Angiosperms

The RCRN and the experimental evidence we report add to the literature to suggest that most orthologs of genetically verified secondary cell wall regulators maintain general functional conservation between grasses and eudicots but may differ in mechanistic details. In the RCRN, all rice orthologs of genetically tested Arabidopsis secondary cell wall regulators connect with cell wall-related genes (**Figure 2B**, **Supplementary Table 2**). Further, transient expression experiments with OsMYB58/63a indicate general conservation of function of this protein between rice and Arabidopsis (**Table 2**), in line with recent molecular genetic analysis in sorghum and switchgrass (Scully et al., 2016; Rao et al., 2019). Likewise, stable genetic and transient analyses in this study (**Figure 3**, **Supplementary Figure 9**, **Supplementary Table 2**) and the literature (Hirano et al., 2013b; Huang et al., 2015) suggest that the function of MYB61 in cell wall regulation is also broadly conserved. Taken together the data presented here add to a model that most secondary cell wall-associated transcription factors originated before the divergence of eudicots and monocots and have maintained similar functions during evolution (Rao et al., 2019).

TABLE 3 | Expression of cell wall biosynthesis genes in rice protoplasts overexpressing putative cell wall transcription factors. Data represent average fold change and standard deviation of three biological replicates of normalized expression (based on expression of references genes, *Ubi5* and *CC55*) upon expression of the regulators under control of the 35S promoter relative to empty vector controls. Data are from a single representative experiment. All experiments were repeated independently two to three times and bold and italic font demarcates repeatable significant differences.

LOC_Os Gene/Family	02g41510 MYB13a	04g43680 MYB13b	03g08470 WAP1	01g11910 WAHL1	12g43950 WAHD1	04g08060 WACH1	10g39030 bHLH	01g39330 bHLH	06g43860 KNOX	06g46270 NAC	07g48550 NAC
Transcription factors											
Lignin											
Os4CL3	***162 ± 12a***	***209 ± 31b***	***95 ± 16b***	***365 ± 34b***	***164 ± 12b***	***736 ± 49b***	***164 ± 20b***	***216 ± 19b***	***289 ± 37b***	***318 ± 19b***	***371 ± 26b***
OsCCR1	***1.4 ± 0.1a***	***1.7 ± 0.1a***	1.1 ± 0.1	1.3 ± 0.1	1.0 ± 0.1	0.4 ± 0.1a	1.2 ± 0.1	1.3 ± 0.2	1.0 ± 0.1	1.2 ± 0.1	1.0 ± 0.1
OsCOMT1	1.6 ± 0.2	1.4 ± 0.2	1.0 ± 0.4	1.2 ± 0.1a	1.0 ± 0.1	1.0 ± 0.2	1.2 ± 0.1	1.1 ± 0.1	1.1 ± 0.1	1.2 ± 0.1	1.0 ± 0.1
OsF5H1	***1.6 ± 0.1a***	1.1 ± 0.1	***2.0 ± 0.2a***	1.1 ± 0.1	0.9 ± 0.1	0.6 ± 0.1a	0.9 ± 0.1	0.9 ± 0.2	1.0 ± 0.1	1.0 ± 0.1	1.1 ± 0.1
OsCAD2	2.4 ± 0.5a	1.3 ± 0.1	0.9 ± 0.1	***1.4 ± 0.1b***	NDc	2.3 ± 0.3a	0.8 ± 0.1	1.2 ± 0.1a	NDc	1.2 ± 0.2	1.1 ± 0.1
Secondary wall											
CESA											
OsCESA4	0.8 ± 0.1	***1.4 ± 0.1a***	1.2 ± 0.2	NDc	***1.8 ± 0.2a***	***0.2 ± 0.1a***	0.9 ± 0.2	1.3 ± 0.1	0.9 ± 0.1	NDc	NDc
OsCESA9	NDc	NDc	1.2 ± 0.3	1.3 ± 0.1	1.5 ± 0.3	1.6 ± 0.4	1.2 ± 0.01	1.1 ± 0.1	0.6 ± 0.2	1.0 ± 0.3	1.0 ± 0.1
MLG											
OsCSLF6	NDc	NDc	1.1 ± 0.2	NDc	1.0 ± 0.1	0.9 ± 0.4	NDc	NDc	NDc	1.6 ± 0.2	1.1 ± 0.1
OsCSLH1	NDc	NDc	1.4 ± 0.2	NDc	1.1 ± 0.1	1.1 ± 0.3	NDc	NDc	NDc	NDc	NDc
Hydroxycinnamic acid											
OsAT4	1.0 ± 0.1	1.0 ± 0.1	1.0 ± 0.1	1.2 ± 0.1	1.2 ± 0.4	1.8 ± 0.1a	0.8 ± 0.2	1.0 ± 0.2	1.1 ± 0.1	1.2 ± 0.3	NDc
OsAT5	1.0 ± 0.2	0.7 ± 0.2	NDc	0.9 ± 0.2	NDc	1.2 ± 0.2	0.8 ± 0.1	0.9 ± 0.2	0.9 ± 0.3	0.9 ± 0.2	0.7 ± 0.3

[a]Two-tailed t-test p value < 0.05.
[b]Two-tailed t-test p value < 0.01.
[c]ND indicates the interactions were not determined in this assay, which examined relationships based on RCR network predictions.

Despite general conservation, our analyses support the notion that at least some of the molecular details of secondary cell wall regulation vary between rice and Arabidopsis (**Figure 2A**, **Supplementary Figure 2**). As network connectivity (node degree) reflects essentiality (Batada et al., 2006; He and Zhang, 2006; Yang et al., 2014), this metric suggests that several rice orthologs of known Arabidopsis cell wall regulators, especially OsVND6/7, may have altered importance relative to their roles in eudicots. This absence was observed previously, leading to the hypothesis that OsVND6/7 might have specialized to regulate other aspects of xylem differentiation in grasses (Hirano et al., 2013a). An alternative hypothesis, consistent with the activity of *Physcomitrella patens* VND7 homologs regulating secondary cell wall gene expression in Arabidopsis (Xu et al., 2014), is that NAC activity in rice is modulated by interactions with other proteins (e.g., Yamaguchi et al., 2010). The relatively low degree of the rice AtMYB46/83 ortholog was also surprising, given this protein's orthologs' important and conserved function in activating cell wall biosynthesis (Zhong et al., 2015). To our knowledge, this protein's function has not been tested genetically in grasses, but we would predict that though its function in controlling secondary cell wall biosynthesis gene expression is retained, its role is diminished relative to the large number of other regulators that grasses utilize (Hirano et al., 2013a; Yan et al., 2013; Rao et al., 2019). While much remains to be elucidated, we speculate that differences in specific molecular interactions within the regulatory networks between grasses and eudicots may lead to variation in stem anatomy and secondary cell wall patterning between these lineages. Indeed, the literature that compares regulatory networks across species suggests that general conservation but subtle divergence across evolution might be more the rule than the exception. In another example, orthologs of the Arabidopsis stomatal initiation regulators also control stomatal development in *Brachypodium*, but the function of individual regulators and the relationships among them appear to differ (Raissig et al., 2016). Similarly, HOX genes regulate body-plan development in animals, but have evolved to also control abdomen pigmentation in some *Drosophila* species (Jeong et al., 2006; Rebeiz et al., 2009). Even within the grasses, ZmMYB31 and ZmMYB42 orthologs in rice and sorghum show distinct promoter occupancies and gene expression correlations *in vivo* (Agarwal et al., 2016).

Incorporation of Grass-Expanded Genes Into Cell Wall Regulatory Networks

In contrast to cell pattern alterations, compositional differences between grasses and eudicots are better understood. Grass-specific cell wall synthesis enzymes fall into two classes, those with close homology to cell wall synthesis enzymes in eudicots (i.e., MLG synthesis) and those the close homologs of which appear to have other functions in eudicots (i.e., Mitchel clade BAHDs). We considered two models for evolution of regulation of these grass-specific cell wall synthesis genes: 1) that the grass-specific genes have been incorporated into conserved regulatory networks and 2) that grass-specific genes are regulated by novel regulators, not involved in cell wall synthesis regulation in other lineages. Our analysis supports the model that orthologs of known cell wall-associated transcription factors (i.e.,

OsMYB61a, OsMYB61b, and OsSND2) regulate grass-specific cell wall biosynthesis genes (**Table 2**, **Figures 4** and **5**).

Unsurprisingly, the differences among assays probing the function of OsMYB61a imply that there are additional regulators of grass-specific cell wall biosynthesis genes. Gene expression analyses in the *myb61a-1* mutant (**Figure 3**) and protoplast-based assays (**Table 2**) are consistent with OsMYB61a broadly regulating multiple classes of cell wall-related genes, including other regulators. This builds on previous results showing that OsMYB61a directly activates the promoters of rice *CESA4*, *CESA7*, and *CESA9* (Huang et al., 2015). However, even when OsMYB61a is capable of binding to a particular promoter, additional regulation is implicated. For example, the absence of expression changes of *OsCSLH1* with increased expression *OsMYB61a* and with DEX-induction of OsMYB61a-GR (**Figure 3C**) suggests that transcriptional repression of *OsCSLH1* might depend on other proteins absent in seedling-derived protoplasts, despite OsMYB61a-*OsCSLH1* promoter interaction capability (**Figure 3B**). Indeed, the modest cell wall compositional alterations (**Figure 4**), despite numerous gene expression changes observed in the *myb61a-1* mutant (**Figure 3**), are consistent with OsMYB61a controlling cell wall synthesis in concert with other regulators. This is consistent with the general architecture of cell wall regulation with many regulators binding both to the promoters of other transcription factors and directly to the promoters of cell wall enzymes (Taylor-Teeples et al., 2015).

Cell Wall-Associated Transcription Factors

We provide experimental evidence for six previously unexamined transcription factors participating in cell wall regulation. Specifically, ectopic expression in protoplasts suggests that OsMYB13a, OsMYB13b, WAP1, WAHD1 and WAHL1 may activate lignin biosynthesis genes and that WACH1 may repress CESA and lignin biosynthesis genes (**Table 3**, **Figure 5**). Inparanoid analysis suggests that three out of six of the new wall-associated regulators have (co-)orthologs in Arabidopsis (**Supplementary Table 5**). Of these three, orthologs of MYB13a and WAHD1 are connected with known cell wall genes in the ACRN, suggesting that they are also likely to be wall-associated regulators in eudicots (**Supplementary Table 5**).

Cell Wall-Associated Repressors and Applications

Besides identifying cell wall biosynthesis activators, this study uncovered two possible transcriptional repressors, OsSND2 and WACH1, which present both the opportunity to better understand the biology of cell wall patterning and to apply this to biotechnological biomass improvement. From a biological perspective, the role of negative regulators remains unclear. Their expression patterns tend to be largely similar to those of activators (**Supplementary Figures 5** and **8**) (Fornalé et al., 2006; Hussey et al., 2011; Shen et al., 2012), though negative correlations are apparent in some species (Agarwal et al., 2016). These proteins may function to repress expression in cells adjacent to those undergoing secondary cell wall synthesis, leading to tissue level

patterning, or to moderate cell wall synthesis, halting the feed-forward loop that characterizes cell wall synthesis and other developmental events (Taylor-Teeples et al., 2015).

Especially as we learn more about the cellular mechanism for wall accumulation of components with roles in recalcitrance or as desirable bioproducts, regulation of transcriptional modules in a cell-type dependent fashion, as opposed to altering expression of single biosynthesis genes, may be more effective. At this point, WACH1, which is also present in Arabidopsis, is an attractive target for up-regulation to inhibit components of secondary cell wall synthesis, as has been demonstrated for switchgrass PvMYB4 (Shen et al., 2012; Baxter et al., 2015). SND2 may also be amenable for use as a negative modulator of secondary cell wall gene expression, though achieving this may require fine tuning. Indeed, SND2 was originally identified in Arabidopsis as a downstream target of SND1 and shown to be capable of activating transcription in yeast (Zhong et al., 2008). Zhong et al. (2008) also found that over expression of a dominant negative fusion protein showed thinner interfasicular fiber cell walls. On the other hand, over expression in Arabidopsis with a double *35S* promoter also decreased fiber cell wall thickness (Hussey et al., 2011), and *PvUbiquitin* promoter-controlled SND2-RNA interference in switchgrass resulted in marginal to no effects (Rao et al., 2019). Thus, SND2 activity may be sensitive to dosage and cellular context that alter its molecular partners as has been observed for other cell wall network regulators (Taylor-Teeples et al., 2015).

Finally, just as we have used the RCRN and ACRN to interrogate cell wall synthesis regulation, these high-quality networks should be useful for delineating other molecular pathways and their divergence between rice and Arabidopsis.

ACCESSION NUMBERS

Rice and Arabidopsis loci and nomenclature used in this work are listed in **Supplementary Tables 1** and **5**.

DATA AVAILABILITY STATEMENT

The RCRN and ACRN are available at: https://doi.org/10.5061/dryad.zgmsbcc69. Other datasets generated for this study are available provided with this manuscript or on request to the corresponding author.

AUTHOR CONTRIBUTIONS

KZ, LB, FL, SH, and K-HJ designed this study. KZ, FL, SR-G, H-JG, and PS performed the experiments. GA provided novel reagents. KZ, SR-G, and FL analyzed the data. KZ SR-G, SH, and LB wrote the manuscript and all authors approved the manuscript.

FUNDING

and by a grant from the Research Council of the University of Oklahoma Norman Campus to LB.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01275/full#supplementary-material

REFERENCES

Agarwal, T., Grotewold, E., Doseff, A. I., and Gray, J. (2016). MYB31/MYB42 syntelogs exhibit divergent regulation of phenylpropanoid genes in maize, sorghum and rice. *Sci. Rep.* 6, 28502. doi: 10.1038/srep28502

Ambavaram, M. M., Krishnan, A., Trijatmiko, K. R., and Pereira, A. (2011). Coordinated activation of cellulose and repression of lignin biosynthesis pathways in rice. *Plant Physiol.* 155, 916–931. doi: 10.1104/pp.110.168641

Anders, N., Wilkinson, M. D., Lovegrove, A., Freeman, J., Tryfona, T., Pellny, T. K., et al. (2012). Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proc. Natl. Acad. Sci. U. S. A.* 109, 989–993. doi: 10.1073/pnas.1115858109

An, G., Lee, S., Kim, S. H., and Kim, S. R. (2005). Molecular genetics using T-DNA in rice. *Plant Cell Physiol.* 46 (1), 14–22. doi: 10.1093/pcp/pci502

Aoki, Y., Okamura, Y., Tadaka, S., Kinoshita, K., and Obayashi, T. (2015). ATTED-II in 2016: a plant coexpression database towards lineage-specific coexpression. *Plant Cell Physiol.* 57 (1), e5–e5. doi: 10.1093/pcp/pcv165

Appenzeller, L., Doblin, M., Barreiro, R., Wang, H., Niu, X., Kollipara, K., et al. (2004). Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family. *Cellulose* 11, 287–299. doi: 10.1023/B:CELL.0000046417.84715.27

Ashburner, M., Ball, C., Blake, J., Botstein, D., Butler, H., Cherry, J., et al. (2000). Gene ontology: tool for the unification of biology. *Gene Ontol. Consort. Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Barabási, A.-L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi: 10.1038/nrg1272

Bart, R., Chern, M., Park, C.-J., Bartley, L., and Ronald, P. C. (2006). A novel system for gene silencing using siRNAs in rice leaf and stem-derived protoplasts. *Plant Methods* 2, 13. doi: 10.1186/1746-4811-2-13

Bartley, L. E., Peck, M. L., Kim, S. R., Ebert, B., Manisseri, C., Chiniquy, D. M., et al. (2013). Overexpression of a BAHD acyltransferase, OsAt10, alters rice cell wall hydroxycinnamic acid content and saccharification. *Plant Physiol.* 161, 1615–1633. doi: 10.1104/pp.112.208694

Batada, N. N., Reguly, T., Breitkreutz, A., Boucher, L., Breitkreutz, B.-J., Hurst, L. D., et al. (2006). Stratus not altocumulus: a new view of the yeast protein interaction network. *PLoS Biol.* 4, e317. doi: 10.1371/journal.pbio.0040317

Baxter, H. L., Poovaiah, C. R., Yee, K. L., Mazarei, M., Rodriguez, M., Thompson, O. A., et al. (2015). Field evaluation of transgenic switchgrass plants overexpressing PvMYB4 for reduced biomass recalcitrance. *BioEnergy Res.* 8, 910–921. doi: 10.1007/s12155-014-9570-1

Bonawitz, N. D., and Chapple, C. (2010). The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu. Rev. Genet.* 44, 337–363. doi: 10.1146/annurev-genet-102209-163508

Burton, R. A., Wilson, S. M., Hrmova, M., Harvey, A. J., Shirley, N. J., Medhurst, A., et al. (2006). Cellulose synthase-like CslF genes mediate the synthesis of cell wall 3. (1,3;1,4)-beta-D-glucans. *Science*, 311(5769), 1940–1942. doi: 10.1126/science.1122975

Cao, P., Jung, K.-H., Choi, D., Hwang, D., Zhu, J., and Ronald, P. C. (2012). The rice oligonucleotide array database: an atlas of rice gene expression. *Rice* 5, 1–9. doi: 10.1186/1939-8433-5-17

Carpita, N. C. (2012). Progress in the biological synthesis of the plant cell wall: new ideas for improving biomass for bioenergy. *Curr. Opin. Biotechnol.* 23, 330–337. doi: 10.1016/j.copbio.2011.12.003

Chai, M. F., Bellizzi, M., Wan, C. X., Cui, Z. F., Li, Y. B., and Wang, G. L. (2015). The NAC transcription factor OsSWN1 regulates secondary cell wall development in *Oryza sativa*. *J. Plant Biol.* 58, 44–51. doi: 10.1007/s12374-014-0400-y

Chang, B., Kustra, R., and Tian, W. (2013). Functional-network-based gene set analysis using gene-ontology. *PLoS One* 8, e55635. doi: 10.1371/journal.pone.0055635

Chiniquy, D., Sharma, V., Schultink, A., Baidoo, E. E., Rautengarten, C., Cheng, K., et al. (2012). XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan. *Proc. Natl. Acad. Sci. U. S. A.* 109, 17117–17122. doi: 10.1073/pnas.1202079109

Childs, K. L., Konganti, K., and Buell, C. R. (2012). The Biofuel Feedstock Genomics Resource: a web-based portal and database to enable functional genomics of plant biofuel feedstock species. Database 2012. doi: 10.1093/database/bar061

Ciftci-Yilmaz, S., and Mittler, R. (2008). The zinc finger network of plants. *Cell. Mol. Life Sci.* 65, 1150–1160. doi: 10.1007/s00018-007-7473-4

De Souza, W. R., Martins, P. K., Freeman, J., Pellny, T. K., Michaelson, L. V., Sampaio, B. L., et al. (2018). Suppression of a single BAHD gene in *Setaria viridis* causes large, stable decreases in cell wall feruloylation and increases biomass digestibility. *New Phytol.* 218, 81–93. doi: 10.1111/nph.14970

Deplancke, B., Dupuy, D., Vidal, M., and Walhout, A. J. M. (2004). A gateway-compatible yeast one-hybrid system. *Genome Res.* 14, 2093–2101. doi: 10.1101/gr.2445504

Deplancke, B., Vermeirssen, V., Arda, H. E., Martinez, N. J., and Walhout, A. J. M., (2006). Gateway-compatible yeast one-hybrid screens. *Cold Spring Harb. Protoc.* 2006, pdb.prot4590. doi: 10.1101/pdb.prot4590

Fan, M., Herburger, K., Jensen, J. K., Zemelis-Durfee, S., Brandizzi, F, Fry, S. C., et al. (2018). A trihelix family transcription factor is associated with key genes in mixed-linkage glucan accumulation. *Plant Physiol.* 178(3), 1207–1221. doi: 10.1104/pp.18.00978

Fangel, J. U., Ulvskov, P., Knox, J. P., Mikkelsen, M. D., Harholt, J., Popper, Z. A., et al. (2012). Cell wall evolution and diversity. *Front. Plant Sci.* 3, 152. doi: 10.3389/fpls.2012.00152

Ficklin, S. P., and Feltus, F. A. (2013). A systems-genetics approach and data mining tool to assist in the discovery of genes underlying complex traits in *Oryza sativa*. *PLoS One* 8, e68551. doi: 10.1371/journal.pone.0068551

Fornale, S., Shi, X., Chai, C., Encina, A., Irar, S., Capellades, M., et al. (2010). ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant J.* 64, 633–644. doi: 10.1111/j.1365-313X.2010.04363.x

Fornalé, S., Sonbol, F.-M., Maes, T., Capellades, M., Puigdomènech, P., Rigau, J., et al. (2006). Down-regulation of the maize and *Arabidopsis thaliana* caffeic acid O-methyl-transferase genes by two new maize R2R3-MYB transcription factors. *Plant Mol. Biol.* 62, 809–823. doi: 10.1007/s11103-006-9058-2

Gray, J., Caparros-Ruiz, D., and Grotewold, E. (2012). Grass phenylpropanoids: regulate before using! *Plant Sci.* 184, 112–120. doi: 10.1016/j.plantsci.2011.12.008

Handakumbura, P. P., and Hazen, S. P. (2012). Transcriptional regulation of grass secondary cell wall biosynthesis: playing catch-up with *Arabidopsis thaliana*. *Front. Plant Sci.* 3, 74. doi: 10.3389/fpls.2012.00074

Handakumbura, P. P., Matos, D. A., Osmont, K. S., Harrington, M. J., Heo, K., Kafle, K., et al. (2013). Perturbation of *Brachypodium distachyon* CELLULOSE SYNTHASE A4or7results in abnormal cell walls. *BMC Plant Biol.* 13, 131. doi: 10.1186/1471-2229-13-131

Harris, P. J., and Trethewey, J.a.K. (2010). The distribution of ester-linked ferulic acid in the cell walls of angiosperms. *Phytochem. Rev.* 9, 19–33. doi: 10.1007/s11101-009-9146-4

Hazen, S. P., Scott-Craig, J. S., and Walton, J. D. (2002). Cellulose synthase-like genes of rice. *Plant Physiol.* 128, 336–340. doi: 10.1104/pp.010875

He, X., and Zhang, J. (2006). Why do hubs tend to be essential in protein networks? *PLoS Genet.* 2, e88. doi: 10.1371/journal.pgen.0020088

Hirano, K., Aya, K., Morinaka, Y., Nagamatsu, S., Sato, Y., Antonio, B. A., et al. (2013a). Survey of genes involved in rice secondary cell wall formation through a co-expression network. *Plant Cell Physiol.* 54, 1803–1821. doi: 10.1093/pcp/pct121

Hirano, K., Kondo, M., Aya, K., Miyao, A., Sato, Y., Antonio, B. A., et al. (2013b). Identification of transcription factors involved in rice secondary cell wall formation. *Plant Cell Physiol.* 54, 1791–1802. doi: 10.1093/pcp/pct122

Huang, D., Wang, S., Zhang, B., Shang-Guan, K., Shi, Y., Zhang, D., et al. (2015). A gibberellin-mediated DELLA-NAC signaling cascade regulates cellulose synthesis in rice. *Plant Cell* 27, 1681–1696. doi: 10.1105/tpc.15.00015

Hussey, S. G., Mizrachi, E., Spokevicius, A. V., Bossinger, G., Berger, D. K., and Myburg, A. A. (2011). SND2, a NAC transcription factor gene, regulates genes involved in secondary cell wall development in Arabidopsis fibres and increases fibre cell area in Eucalyptus. *BMC Plant Biol.* 11, 173. doi: 10.1186/1471-2229-11-173

Jain, M. (2009). Genome-wide identification of novel internal control genes for normalization of gene expression during various stages of development in rice. *Plant Sci.* 176, 702–706. doi: 10.1016/j.plantsci.2009.02.001

Jain, M., Tyagi, A. K., and Khurana, J. P. (2008). Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. *FEBS J.* 275, 2845–2861. doi: 10.1111/j.1742-4658.2008.06424.x

Jeong, S., Rokas, A., and Carroll, S. B. (2006). Regulation of body pigmentation by the abdominal-B Hox protein and its gain and loss in Drosophila evolution. *Cell* 125, 1387–1399. doi: 10.1016/j.cell.2006.04.043

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., et al. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *EMBO J.* 19, 6150–6161. doi: 10.1093/emboj/19.22.6150

Karimi, M., Depicker, A., and Hilson, P. (2007). Recombinational cloning with plant gateway vectors. *Plant Physiol.* 145, 1144–1154. doi: 10.1104/pp.107.106989

Karlen, S. D., Zhang, C., Peck, M. L., Smith, R. A., Padmakshan, D., Helmich, K. E., et al. (2016). Monolignol ferulate conjugates are naturally incorporated into plant lignins. *Sci. Adv.* 2, e1600393. doi: 10.1126/sciadv.1600393

Kim, S. J., Zemelis, S., Keegstra, K., and Brandizzi, F. (2015). The cytoplasmic localization of the catalytic site of CSLF6 supports a channeling model for the biosynthesis of mixed-linkage glucan. *Plant J.* 81, 537–547. doi: 10.1111/tpj.12748

Kim, W. C., Kim, J. Y., Ko, J. H., Kang, H., and Han, K. H. (2014). Identification of direct targets of transcription factor MYB46 provides insights into the transcriptional regulation of secondary wall biosynthesis. *Plant Mol. Biol.* 85 (6), 589–599. doi: 10.1007/s11103-014-0205-x

Ko, J. H., Kim, W. C., and Han, K. H. (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. *Plant J.* 60 (4), 649–665. doi: 10.1111/j.1365-313X.2009.03989

Koshiba, T., Yamamoto, N., Tobimatsu, Y., Yamamura, M., Suzuki, S., Hattori, T., et al. (2017). MYB-mediated upregulation of lignin biosynthesis in *Oryza sativa* towards biomass refinery. *Plant Biotechnol.* 34, 7–15. doi: 10.5511/plantbiotechnology.16.1201a

Lal, R. (2005). World crop residues production and implications of its use as a biofuel. *Environ. Int.* 31, 575–584. doi: 10.1016/j.envint.2004.09.005

Lee, I., Ambaru, B., Thakkar, P., Marcotte, E. M., and Rhee, S. Y. (2010). Rational association of genes with traits using a genome-scale gene network for *Arabidopsis thaliana*. *Nat. Biotech.* 28, 149–156. doi: 10.1038/nbt.1603

Lee, I., Date, S. V., Adai, A. T., and Marcotte, E. M. (2004). A probabilistic functional network of yeast genes. *Science* 306, 1555–1558. doi: 10.1126/science.1099511

Lee, I., Seo, Y.-S., Coltrane, D., Hwang, S., Oh, T., Marcotte, E. M., et al. (2011). Genetic dissection of the biotic stress response using a genome-scale gene network for rice. *Proc. Natl. Acad. Sci.* 108, 18548–18553. doi: 10.1073/pnas.1110384108

Lee, T., Oh, T., Yang, S., Shin, J., Hwang, S., Kim, C. Y., et al. (2015). RiceNet v2: an improved network prioritization server for rice genes. *Nucleic Acids Res.* 43, W122–W127. doi: 10.1093/nar/gkv253

Liepman, A. H., Wightman, R., Geshi, N., Turner, S. R., and Scheller, H. V. (2010). Arabidopsis—a powerful model system for plant cell wall research. *Plant J.* 61, 1107–1121. doi: 10.1111/j.1365-313X.2010.04161.x

McCarthy, R. L., Zhong, R., and Ye, Z.-H. (2009). MYB83 is a direct target of snd1 and acts redundantly with myb46 in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell Physiol.* 50, 1950–1964. doi: 10.1093/pcp/pcp139

Mcgary, K. L., Lee, I., and Marcotte, E. M. (2007). Broad network-based predictability of *Saccharomyces cerevisiae* gene loss-of-function phenotypes. *Genome Biol.* 8, R258. doi: 10.1186/gb-2007-8-12-r258

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (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. doi: 10.1105/tpc.106.047043

Movahedi, S., Van De Peer, Y., and Vandepoele, K. (2011). Comparative network analysis reveals that tissue specificity and gene function are important factors influencing the mode of expression evolution in Arabidopsis and rice. *Plant Physiol.* 156, 1316–1330. doi: 10.1104/pp.111.177865

Mutwil, M., Klie, S., Tohge, T., Giorgi, F. M., Wilkins, O., Campbell, M. M., et al. (2011). PlaNet: combined sequence and expression comparisons across plant networks derived from seven species. *Plant Cell* 23, 895–910. doi: 10.1105/tpc.111.083667

Mutwil, M., Obro, J., Willats, W. G., and Persson, S. (2008). GeneCAT—novel webtools that combine BLAST and co-expression analyses. *Nucleic Acids Res.* 36, W320–W326. doi: 10.1093/nar/gkn292

Mutwil, M., Usadel, B., Schutte, M., Loraine, A., Ebenhoh, O., and Persson, S. (2010). Assembly of an interactive correlation network for the Arabidopsis genome using a novel heuristic clustering algorithm. *Plant Physiol.* 152, 29–43. doi: 10.1104/pp.109.145318

Newman, L. J., Perazza, D. E., Juda, L., and Campbell, M. M. (2004). Involvement of the R2R3-MYB, AtMYB61, in the ectopic lignification and dark-photomorphogenic components of the det3 mutant phenotype. *Plant J.* 37, 239–250. doi: 10.1046/j.1365-313X.2003.01953.x

Obayashi, T., Okamura, Y., Ito, S., Tadaka, S., Aoki, Y., Shirota, M., et al.. (2014). ATTED-II in 2014: Evaluation of gene coexpression in agriculturally important plants. *Plant Cell Physiol.* 55, e6. doi: 10.1093/pcp/pct178

Obertello, M., Shrivastava, S., Katari, M. S., and Coruzzi, G. M. (2015). Cross-species network analysis uncovers conserved nitrogen-regulated network modules in rice. *Plant Physiol.* 168, 1830–1843. doi: 10.1104/pp.114.255877

Pauly, M., Gille, S., Liu, L., Mansoori, N., De Souza, A., Schultink, A., et al.. (2013). Hemicellulose biosynthesis. *Planta* 238, 627–642. doi: 10.1007/s00425-013-1921-1

Penning, B. W., Hunter, C. T., Tayengwa, R., Eveland, A. L., Dugard, C. K., Olek, A. T., et al. (2009). Genetic resources for maize cell wall biology. *Plant Physiol.* 151, 1703–1728. doi: 10.1104/pp.109.136804

Petrik, D. L., Karlen, S. D., Cass, C. L., Padmakshan, D., Lu, F., Liu, S., et al.. (2014). p-Coumaroyl-CoA: monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in *Brachypodium distachyon*. *Plant J.* 77, 713–726. doi: 10.1111/tpj.12420

Popper, Z. A., Michel, G., Hervé, C., Domozych, D. S., Willats, W. G. T., Tuohy, M. G., et al. (2011). Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu. Rev. Plant Biol.* 62, 567–590. doi: 10.1146/annurev-arplant-042110-103809

Preston, J., Wheeler, J., Heazlewood, J., Li, S. F., and Parish, R. W. (2004). AtMYB32 is required for normal pollen development in *Arabidopsis thaliana*. *Plant J.* 40, 979–995. doi: 10.1111/j.1365-313X.2004.02280.x

Raissig, M. T., Abrash, E., Bettadapur, A., Vogel, J. P., and Bergmann, D. C. (2016). Grasses use an alternatively wired bHLH transcription factor network to establish stomatal identity. *Proc. Natl. Acad. Sci.* 113, 8326–8331. doi: 10.1073/pnas.1606728113

Rao, X., Chen, X., Shen, H., Ma, Q., Li, G., Tang, Y., et al. (2019). Gene regulatory networks for lignin biosynthesis in switchgrass (*Panicum virgatum*). *Plant Biotechnol. J.* 17, 580–593. doi: 10.1111/pbi.13000

Rao, X., and Dixon, R. A. (2018). Current Models for transcriptional regulation of secondary cell wall biosynthesis in grasses. *Front. Plant Sci.* 9, 399. doi: 10.3389/fpls.2018.00399

Rautengarten, C., Ebert, B., Ouellet, M., Nafisi, M., Baidoo, E. E., Benke, P., et al. (2012). Arabidopsis deficient in cutin ferulate encodes a transferase required for feruloylation of omega-hydroxy fatty acids in cutin polyester. *Plant Physiol.* 158, 654–665. doi: 10.1104/pp.111.187187

Rebeiz, M., Pool, J. E., Kassner, V. A., Aquadro, C. F., and Carroll, S. B. (2009). Stepwise modification of a modular enhancer underlies adaptation in a Drosophila population. *Science (New York, N.y.)* 326, 1663–1667. doi: 10.1126/science.1178357

Remm, M., Storm, C. E., and Sonnhammer, E. L. (2001). Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* 314, 1041–1052. doi: 10.1006/jmbi.2000.5197

Romano, J. M., Dubos, C., Prouse, M. B., Wilkins, O., Hong, H., Poole, M., et al. (2012). AtMYB61, an R2R3-MYB transcription factor, functions as a pleiotropic regulator via a small gene network. *New Phytol.* 195, 774–786. doi: 10.1111/j.1469-8137.2012.04201.x

Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., Van Den Hoff, M. J. B., et al. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* 37, e45–e45. doi: 10.1093/nar/gkp045

Sarkar, N., Kim, Y.-K., and Grover, A. (2014). Coexpression network analysis associated with call of rice seedlings for encountering heat stress. *Plant Mol. Biol.* 84, 125–143. doi: 10.1007/s11103-013-0123-3

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., et al. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* 37, 501–506. doi: 10.1038/ng1543

Schwerdt, J. G., Mackenzie, K., Wright, F., Oehme, D., Wagner, J. M., Harvey, A. J., et al. (2015). Evolutionary dynamics of the cellulose synthase gene superfamily in grasses. *Plant Physiol.* 168, 968–983. doi: 10.1104/pp.15.00140

Scully, E. D., Gries, T., Sarath, G., Palmer, N. A., Baird, L., Serapiglia, M. J., et al. (2016). Overexpression of SbMyb60 impacts phenylpropanoid biosynthesis and alters secondary cell wall composition in *Sorghum bicolor*. *Plant J.* 85, 378–395. doi: 10.1111/tpj.13112

Shen, H., He, X., Poovaiah, C. R., Wuddineh, W. A., Ma, J., Mann, D. G. J., et al. (2012). Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytol.* 193, 121–136. doi: 10.1111/j.1469-8137.2011.03922.x

Siegal, M. L., Promislow, D. E. L., and Bergman, A. (2007). Functional and evolutionary inference in gene networks: does topology matter? *Genetica* 129, 83–103. doi: 10.1007/s10709-006-0035-0

Sonbol, F. M., Fornale, S., Capellades, M., Encina, A., Tourino, S., Torres, J. L., et al. (2009). The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in *Arabidopsis thaliana*. *Plant Mol. Biol.* 70, 283–296. doi: 10.1007/s11103-009-9473-2

Sorrells, T., ,. R., and Johnson, A. D. (2015). Making sense of transcription networks. *Cell* 161, 714–723. doi: 10.1016/j.cell.2015.04.014

Takeda, Y., Tobimatsu, Y., Karlen, S. D., Koshiba, T., Suzuki, S., Yamamura, M., et al. (2018). Downregulation of p-COUMAROYL ESTER 3-HYDROXYLASE in rice leads to altered cell wall structures and improves biomass saccharification. *Plant J.* 95(5), 796–811. doi: 10.1111/tpj.13988

Taylor-Teeples, M., Lin, L., De Lucas, M., Turco, G., Toal, T. W., Gaudinier, A., et al. (2015). An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature* 517, 571–575. doi: 10.1038/nature14099

Terada, R., and Shimamoto, K. (1990). Expression of CaMV35S-GUS gene in transgenic rice plants. *Mol. Gen. Genetics MGG* 220, 389–392. doi: 10.1007/BF00391743

Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F. M., Bassel, G. W., Tanimoto, M., et al. (2009). Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. *Plant, Cell Environ.* 32, 1633–1651. doi: 10.1111/j.1365-3040.2009.02040.x

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiol.* 153, 895–905. doi: 10.1104/pp.110.155119

Vega-Sanchez, M. E., Verhertbruggen, Y., Christensen, U., Chen, X., Sharma, V., Varanasi, P., et al. (2012). Loss of cellulose synthase-like F6 function affects mixed-linkage glucan deposition, cell wall mechanical properties, and defense responses in vegetative tissues of rice. *Plant Physiol.* 159, 56–69. doi: 10.1104/pp.112.195495

Walhout, A. J., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. *Methods* 24, 297–306. doi: 10.1006/meth.2001.1190

Wang, H. Z., and Dixon, R. A. (2012). On-off switches for secondary cell wall biosynthesis. *Mol. Plant.* 5, 297–303. doi: 10.1093/mp/ssr098

Wang, L., Xie, W., Chen, Y., Tang, W., Yang, J., Ye, R., et al.. (2010). A dynamic gene expression atlas covering the entire life cycle of rice. *Plant J.* 61, 752–766. doi: 10.1111/j.1365-313X.2009.04100.x

Withers, S., Lu, F., Kim, H., Zhu, Y., Ralph, J., and Wilkerson, C. G. (2012). Identification of grass-specific enzyme that acylates monolignols with p-coumarate. *J. Biol. Chem.* 287, 8347–8355. doi: 10.1074/jbc.M111.284497

Xu, B., Ohtani, M., Yamaguchi, M., Toyooka, K., Wakazaki, M., Sato, M., et al. (2014). Contribution of NAC transcription factors to plant adaptation to land. *Science* 343, 1505–1508. doi: 10.1126/science.1248417

Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., et al. (2010). VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in Arabidopsis. *Plant Cell* 22, 1249–1263. doi: 10.1105/tpc.108.064048

Yan, L., Xu, C., Kang, Y., Gu, T., Wang, D., Zhao, S., et al.. (2013). The heterologous expression in *Arabidopsis thaliana* of sorghum transcription factor SbbHLH1 downregulates lignin synthesis. *J. Exp. Botany* 64, 3021–3032. doi: 10.1093/jxb/ert150

Yang, Y., Han, L., Yuan, Y., Li, J., Hei, N., and Liang, H. (2014). Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. *Nat. Commun.* 5, 3231. doi: 10.1038/ncomms4231

Yeung, K. Y., Dombek, K. M., Lo, K., Mittler, J. E., Zhu, J., and Schadt, E. E. (2011). Construction of regulatory networks using expression time-series data of a genotyped population. *Proc. Natl. Acad. Sci. U. S. A.* 108 (48), 19436–19441.. doi: 10.1073/pnas.1116442108

Zhao, K., and Bartley, L. (2014). Comparative genomic analysis of the R2R3 MYB secondary cell wall regulators of Arabidopsis, poplar, rice, maize, and switchgrass. *BMC Plant Biol.* 14, 135. doi: 10.1186/1471-2229-14-135

Zhao, Q., and Dixon, R. A. (2014). Altering the cell wall and its impact on plant disease: from forage to bioenergy. *Annu. Rev. Phytopathol.* 52, 69–91. doi: 10.1146/annurev-phyto-082712-102237

Zhong, R., Cui, D., Phillips, D. R., and Ye, Z.-H. (2018). A novel rice xylosyltransferase catalyzes the addition of 2-O-xylosyl side chains onto the xylan backbone. *Plant Cell Physiol.* 59, 554–565. doi: 10.1093/pcp/pcy003

Zhong, R., and Ye, Z. H. (2014). Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant Cell Physiol.* 56 (2), 195–214. doi: 10.1093/pcp/pcu140

Zhong, R., Lee, C., Mccarthy, R. L., Reeves, C. K., Jones, E. G., and Ye, Z. H. (2011). Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant Cell Physiol.* 52, 1856–1871. doi: 10.1093/pcp/pcr123

Zhong, R., Lee, C., and Ye, Z. H. (2010). Functional characterization of poplar wood-associated NAC domain transcription factors. *Plant Physiol.* 152, 1044–1055. doi: 10.1104/pp.109.148270

Zhong, R., Lee, C., Zhou, J., Mccarthy, R. L., and Ye, Z. H. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell* 20, 2763–2782. doi: 10.1105/tpc.108.061325

Zhong, R., Richardson, E. A., and Ye, Z. H. (2007). The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. *Plant Cell* 19, 2776–2792. doi: 10.1105/tpc.107.053678

Zhong, R., Yuan, Y., Spiekerman, J. J., Guley, J. T., Egbosiuba, J. C., and Ye, Z.-H. (2015). Functional characterization of NAC and MYB transcription factors involved in regulation of biomass production in switchgrass (*Panicum virgatum*). *PLoS One* 10, e0134611. doi: 10.1371/journal.pone.0134611

Zhou, J., Lee, C., Zhong, R., and Ye, Z. H. (2009). MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *Plant Cell* 21, 248–266. doi: 10.1105/tpc.108.063321

Zhou, J., Zhong, R., and Ye, Z.-H. (2014). Arabidopsis NAC domain proteins, VND1 to VND5, are transcriptional regulators of secondary wall biosynthesis in vessels. *PLoS One* 9, e105726. doi: 10.1371/journal.pone.0105726