

REGULATION OF AND BY THE PLANT CELL WALL

EDITED BY: Georgia Drakakaki, Laura Elizabeth Bartley, Charles T. Anderson
and Xiaolan Rao

PUBLISHED IN: Frontiers in Plant Science





frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88963-804-8

DOI 10.3389/978-2-88963-804-8

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

REGULATION OF AND BY THE PLANT CELL WALL

Topic Editors:

Georgia Drakakaki, University of California, Davis, United States

Laura Elizabeth Bartley, University of Oklahoma, United States

Charles T. Anderson, Pennsylvania State University (PSU), United States

Xiaolan Rao, University of North Texas, United States

Citation: Drakakaki, G., Bartley, L. E., Anderson, C. T., Rao, X., eds. (2020). Regulation of and by the Plant Cell Wall. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88963-804-8

Table of Contents

- 05 Editorial: Regulation of and by the Plant Cell Wall**
Xiaolan Rao, Laura E. Bartley, Georgia Drakakaki and Charles T. Anderson
- 08 Rice Genome-Scale Network Integration Reveals Transcriptional Regulators of Grass Cell Wall Synthesis**
Kangmei Zhao, Fan Lin, Sandra P. Romero-Gamboa, Prasenjit Saha, Hyung-Jung Goh, Gynheung An, Ki-Hong Jung, Samuel P. Hazen and Laura E. Bartley
- 24 Finding New Cell Wall Regulatory Genes in *Populus trichocarpa* Using Multiple Lines of Evidence**
Anna Furches, David Kainer, Deborah Weighill, Annabel Large, Piet Jones, Angelica M. Walker, Jonathon Romero, Joao Gabriel Felipe Machado Gazolla, Wayne Joubert, Manesh Shah, Jared Streich, Priya Ranjan, Jeremy Schmutz, Avinash Sreedasyam, David Macaya-Sanz, Nan Zhao, Madhavi Z. Martin, Xiaolan Rao, Richard A. Dixon, Stephen DiFazio, Timothy J. Tschaplinski, Jin-Gui Chen, Gerald A. Tuskan and Daniel Jacobson
- 41 Secondary Wall Regulating NACs Differentially Bind at the Promoter at a CELLULOSE SYNTHASE A4 Cis-eQTL**
Jennifer R. Olins, Li Lin, Scott J. Lee, Gina M. Trabucco, Kirk J.-M. MacKinnon and Samuel P. Hazen
- 53 Post-Golgi Trafficking and Transport of Cell Wall Components**
Rosalie Sinclair, Michel Ruiz Rosquete and Georgia Drakakaki
- 64 A New Calmodulin-Binding Protein Expresses in the Context of Secondary Cell Wall Biosynthesis and Impacts Biomass Properties in *Populus***
Raghuram Badmi, Raja S. Payyavula, Garima Bali, Hao-Bo Guo, Sara S. Jawdy, Lee E. Gunter, Xiaohan Yang, Kimberly A. Winkeler, Cassandra Collins, William H. Rottmann, Kelsey Yee, Miguel Rodriguez Jr., Robert W. Sykes, Stephen R. Decker, Mark F. Davis, Arthur J. Ragauskas, Gerald A. Tuskan and Udaya C. Kalluri
- 86 Progress and Opportunities in the Characterization of Cellulose – An Important Regulator of Cell Wall Growth and Mechanics**
Sintu Rongpipi, Dan Ye, Enrique D. Gomez and Esther W. Gomez
- 114 Recent Advances in the Transcriptional Regulation of Secondary Cell Wall Biosynthesis in the Woody Plants**
Jin Zhang, Meng Xie, Gerald A. Tuskan, Wellington Muchero and Jin-Gui Chen
- 128 Regulation of Lignin Biosynthesis and its Role in Growth-Defense Tradeoffs**
Meng Xie, Jin Zhang, Timothy J. Tschaplinski, Gerald A. Tuskan, Jin-Gui Chen and Wellington Muchero
- 137 Balancing Strength and Flexibility: How the Synthesis, Organization, and Modification of Guard Cell Walls Govern Stomatal Development and Dynamics**
Yue Rui, Yintong Chen, Baris Kandemir, Hojae Yi, James Z. Wang, Virendra M. Puri and Charles T. Anderson

- 148** *New Insights Into Wall Polysaccharide O-Acetylation*
Markus Pauly and Vicente Ramírez
- 160** *Functional Analysis of Cellulose Synthase Cesa4 and Cesa6 Genes in Switchgrass (Panicum virgatum) by Overexpression and RNAi-Mediated Gene Silencing*
Mitra Mazarei, Holly L. Baxter, Mi Li, Ajaya K. Biswal, Keonhee Kim, Xianzhi Meng, Yunqiao Pu, Wegi A. Wuddineh, Ji-Yi Zhang, Geoffrey B. Turner, Robert W. Sykes, Mark F. Davis, Michael K. Udvardi, Zeng-Yu Wang, Debra Mohnen, Arthur J. Ragauskas, Nicole Labbé and C. Neal Stewart Jr.
- 176** *A Profusion of Molecular Scissors for Pectins: Classification, Expression, and Functions of Plant Polygalacturonases*
Yang Yang, Youjian Yu, Ying Liang, Charles T. Anderson and Jiashu Cao
- 192** *Current Models for Transcriptional Regulation of Secondary Cell Wall Biosynthesis in Grasses*
Xiaolan Rao and Richard A. Dixon
- 203** *Ethylene-Related Gene Expression Networks in Wood Formation*
Carolyn Seyfferth, Bernard Wessels, Soile Jokipii-Lukkari, Björn Sundberg, Nicolas Delhomme, Judith Felten and Hannele Tuominen



Editorial: Regulation of and by the Plant Cell Wall

Xiaolan Rao^{1,2*}, Laura E. Bartley^{3,4}, Georgia Drakakaki⁵ and Charles T. Anderson⁶

¹ College of Life Sciences, Hubei University, Wuhan, China, ² Department of Biological Sciences, BioDiscovery Institute, University of North Texas, Denton, TX, United States, ³ Microbiology and Plant Biology Department, University of Oklahoma, Norman, OK, United States, ⁴ Research Institute for the Sustainable Humanosphere, Kyoto University, Kyoto, Japan, ⁵ Department of Plant Sciences, University of California, Davis, Davis, CA, United States, ⁶ Department of Biology, The Pennsylvania State University, University Park, PA, United States

Keywords: plant cell wall, transcriptional regulation, cell wall biosynthesis, cell wall modification, vesicle-mediated trafficking, polysaccharide transport

Editorial on the Research Topic

Regulation of and by the Plant Cell Wall

The cell wall encapsulates plant cells and fundamentally influences their properties. Wall components and interactions vary among the major plant clades, throughout plant development, among different cell types, and even in response to external and internal stimuli. In addition to their fundamental roles in plant development and physiology, plant cell walls represent the most abundant terrestrial carbon sink and, thus, a key alternative to fossil carbon utilization for energy and materials (Youngs and Somerville, 2012).

This Frontiers in Plant Science virtual issue on “Regulation of and by the Plant Cell Wall” consists of 14 publications, including 8 reviews and 6 original research articles, which fall into three general topics: cell wall composition, synthesis, and modification; transcript-level regulation of cell wall synthesis; and the cell biology of walls. Understanding the regulation of plant cell wall biosynthesis and modification is fundamental to understanding plant development and provides insights for biotechnological innovation for the bioenergy, bio-product, and forage industries. As summarized below, each paper ties into the theme of regulation, gathering evidence of either direct regulation of cell wall components, homeostasis of cell wall composition, compensation among cell wall properties, or feedback of cell wall properties to plant physiology and development. Several papers present methods related to probing cell wall properties or biology.

CELL WALL COMPOSITION, SYNTHESIS, AND MODIFICATION

The most abundant plant cell wall polymer, cellulose, is a partially crystalline polymer and is thought to be the main load-bearing component of walls. The structure and arrangement of cellulose contributes to the mechanical properties of the cell wall and to anisotropic cell growth. Rongpipi et al. provide a detailed review of methods for physical characterization of cellulose microfibrils at different scales, including structural parameters such as shape, degree of polymerization, crystallinity, and spatial organization. The authors’ descriptions of the practical consideration of techniques such as x-ray diffraction, x-ray scattering, spectroscopy, and microscopy will certainly be of use to researchers. In related research, Mazarei et al. use transgenic approaches to characterize two cellulose synthase genes of switchgrass, *PvCesA6*, a predicted primary wall synthase, and *PvCesA4*, a predicted secondary wall synthase. (Secondary cell walls are deposited inside primary walls in many cell types after growth cessation). Both down-regulation and overexpression lines lead to reduced cellulose content and crystallinity and reduced plant

OPEN ACCESS

Edited and reviewed by:

David William McCurdy,
University of Newcastle, Australia

*Correspondence:

Xiaolan Rao
xiaolan.rao@unt.edu

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 18 March 2020

Accepted: 06 April 2020

Published: 29 April 2020

Citation:

Rao X, Bartley LE, Drakakaki G and
Anderson CT (2020) Editorial:
Regulation of and by the Plant Cell
Wall. *Front. Plant Sci.* 11:513.
doi: 10.3389/fpls.2020.00513

stature. Several lines have modified amounts of non-cellulosic cell wall polymers such as lignin and xylan, extending the evidence for functional compensation among cell wall polymers.

The polyphenolic, lignin, has been a subject of intensive research due to its role in preventing the efficient release and utilization of cell wall polysaccharides during the processing of plant biomass, and conversely as a valuable chemical precursor. Xie et al. review the transcriptional regulation, biosynthesis, and functions of lignin in growth and defense. While focusing mainly on well-studied pathways and networks from Arabidopsis, the authors also incorporate newer information about alternative pathways in other species. The review highlights the pleiotropic effects of lignin mutants, cataloging examples of growth defects that may connect lignin to different defense pathways.

Covalent modifications add further complexity to the composition and functions of cell wall polymers. This collection includes reviews of two polysaccharide-modifying enzyme groups, glycosyl O-acetyltransferases and polygalacturonases. Pauly and Ramírez provide a comprehensive update into the enzymes that O-acetylate the matrix polysaccharides xyloglucan, xylan, and pectin. The authors describe the protein families that conduct polysaccharide O-acetylation, compare similarities and differences in polysaccharide O-acetylation between plants and other organisms, and summarize evidence for roles in hormone signaling. Yang et al. highlight a major group of pectin-modifying enzymes, the polygalacturonases (PGs), which cleave pectic polysaccharides. Pectins are matrix polysaccharides in plant cell walls that are especially abundant in the eudicot primary walls. Yang et al. provide an overview of the functions of PGs in plant development, the classification and expression of PG family members, and their evolution across plant species, and describe potential regulatory functions of PGs in internal and external signaling.

TRANSCRIPT-LEVEL REGULATION OF CELL WALL SYNTHESIS

Hundreds of transcription factors have been implicated in regulating expression of genes related to cell wall biosynthesis and modification, either directly or indirectly. Two teams review recent advances in transcriptional and post-transcriptional regulation of enzymes involved in secondary cell wall synthesis in grass and woody plants. Rao and Dixon comprehensively discuss how the conservation and divergence of genes that regulate secondary wall deposition in grass and eudicot plants might be related to the distinct patterning and composition of the walls in these two plant groups. Similarly, Zhang et al. provide an update on understanding the regulation of secondary wall synthesis at the transcriptional level in woody species. They conclude that while many parallels exist, there appear to be added molecular complexities in the regulation of secondary wall deposition in trees relative to Arabidopsis.

This issue also reports new research on the transcriptional regulation of secondary wall synthesis in Arabidopsis, as well as crop and bioenergy species. In Arabidopsis, Olins et al. examine polymorphisms in the promoter of a secondary wall

cellulose synthase, *AtCesA4*, that form the basis of an expression quantitative trait locus in a Bay-o X Shahdara recombinant inbred population. A single nucleotide polymorphism in a NAC transcription factor-binding motif reduces *AtCesA4* gene expression by 2-fold. Interestingly, cellulose content appears to be unaltered in lines with this variant, providing an example of cell wall homeostasis.

An important approach for dealing with the complexity of cell wall biology, three original research articles provide successful examples of exploring biological information derived from -omics data to identify secondary wall regulators. Furches et al. develop a multi-layered, network-based pipeline to search for novel secondary wall-related genes in poplar, combining datasets on gene co-expression, gene co-methylation, SNP correlations, and genome-wide association studies. Additional bioinformatics analysis supports a role in cell wall control for the transcription factor, PtGFR9, the Arabidopsis ortholog of which functions in drought-mediated growth inhibition. In another primarily bioinformatic study of woody species, Seyfferth et al. use an aspen transcriptome database to generate a co-expression network of genes involved in ethylene signaling, pursuing the role of ethylene in cambial growth. The aspen gene, *EIN3D*, is experimentally confirmed to function in ethylene signaling in Arabidopsis. Lastly, Zhao et al. develop a novel functional gene network in rice and mine it for cell wall regulators. They provide experimental evidence that a previously studied cell wall regulator, OsMYB61a, binds to the promoter of a grass-specific wall synthesis gene and that six out of 11 tested transcription factors function as novel regulators of secondary wall gene expression. The gene annotations generated by these studies provide additional wall regulatory candidates, and more generally, the analysis methods are likely to be useful for revealing crosstalk between biological pathways.

CELL WALL CELL BIOLOGY

Post-translational cellular processes that mediate wall formation and rearrangement are another critical aspect of cell wall regulation. Cellulose synthases transit through the Golgi apparatus on the way to the plasma membrane, and many other glycosyltransferases reside in the Golgi. The highly dynamic nature of the endomembrane system makes it challenging to assign unequivocal roles to specific vesicle populations in the synthesis and assembly of the cell wall. Sinclair et al. summarize current research on the transport and deposition of cell wall components by the endomembrane system during cell division and growth. The authors describe the coordinated trafficking of cell wall polysaccharides and proteins, and wall biosynthetic and modifying enzymes, and discuss promising avenues to gain insights into the trafficking of structural polysaccharides to the apoplast. Related to cellular control of cell wall biophysics, Rui et al. spotlight the heterogeneous and dynamic three-dimensional arrangement of the cell walls of stomatal guard cells. Mutant studies coupled with microscopy reveal the role of cell wall dynamics in the opening and closing of guard

cells for controlling gas diffusion at the plant surface. Badmi et al. report a new function for a putative calmodulin binding protein in poplar, PdIQD10, in cell wall biology, consistent with the multi-omic network analysis of Furches et al. Knocking down *PdIQD10* leads to larger poplar plants with increased cellulose content. Although the molecular mechanism by which PdIQD10 performs its cellular function remains elusive, the results suggest linkages between calcium signaling and secondary wall development.

We are deeply grateful to all of the authors, reviewers, and *ad hoc* editors for their contributions to the success of this Research Topic. We trust that the research community will benefit from this collection of knowledge at the frontiers of cell wall biosynthesis, its regulation, and

the impacts of the plant cell wall on plant development and physiology.

AUTHOR CONTRIBUTIONS

XR and LB drafted the manuscript. All authors revised and approved the final version.

FUNDING

LB was supported in part by RISH Kyoto University Mission linked Research Funding FY2019. GD was supported in part by NSF-MCB 1818219. CA was supported by the National Science Foundation under grant MCB-1616316.

REFERENCES

Youngs, H., and Somerville, C. (2012). Development of feedstocks for cellulosic biofuels. *F1000 Biol. Rep.* 4:10. doi: 10.3410/B4-10

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Rao, Bartley, Drakakaki and Anderson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



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

Kangmei Zhao^{1†}, Fan Lin^{1†}, Sandra P. Romero-Gamboa², Prasenjit Saha^{1‡}, Hyung-Jung Goh³, Gynheung An³, Ki-Hong Jung³, Samuel P. Hazen² and Laura E. Bartley^{1*}

OPEN ACCESS

Edited by:

Olga A. Zabotina,
Iowa State University,
United States

Reviewed by:

Ling Li,
Mississippi State University,
United States
Nobutaka Mitsuda,
National Institute of Advanced
Industrial Science and Technology
(AIST), Japan

*Correspondence:

Laura E. Bartley
lbartley@ou.edu

[†]Present address:

Department of Plant Biology,
Carnegie Institution for Science,
Palo Alto, CA, United States

[‡]Present address:

Calyx, Minneapolis, MN,
United States

Specialty section:

This article was submitted to
Plant Physiology,
a section of the journal
Frontiers in Plant Science

Received: 22 March 2019

Accepted: 12 September 2019

Published: 18 October 2019

Citation:

Zhao K, Lin F, Romero-Gamboa SP,
Saha P, Goh H-J, An G, Jung K-H,
Hazen SP and Bartley LE (2019) Rice
Genome-Scale Network Integration
Reveals Transcriptional Regulators of
Grass Cell Wall Synthesis.
Front. Plant Sci. 10:1275.
doi: 10.3389/fpls.2019.01275

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 β -(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, CESAs, 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 RiceFRIEND (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 rice combined mutual ranked network (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. PlaNet 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: $1/MR = 1/\sqrt{\text{rank}(A, B) \times \text{rank}(B, A)}$ (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:

$$\frac{1}{\text{RiceNet.v2_MR}} = \frac{0.33}{\text{ROAD_MR}} + \frac{0.025}{\text{PlaNet_MR}} \quad (\text{Equation I}).$$

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

$$\text{RCRN} = \frac{1}{\text{RiceNet.v2_MR}} + \frac{0.33}{\text{ROAD_MR}} + \frac{0.025}{\text{PlaNet_MR}} \quad (\text{Equation II}).$$

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:

$$\text{ACRN} = \frac{1}{\text{AraNet.v2_MR}} + \frac{0.27}{\text{ATTED_MR}}$$

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 μ g 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 hydroxycinnamic acids (e.g., FA and pCA) 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 chromatography 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 μ M) 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 μ M 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 (\sim > 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 \times 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 μ l 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 μ l) 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	R ² 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

^aIncluded 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**).

^bThe power law distribution is $P(k) \sim k^{-\gamma}$, which represents the probability of a node with k edges with γ as a constant for a given biological network.

^cNot 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*).

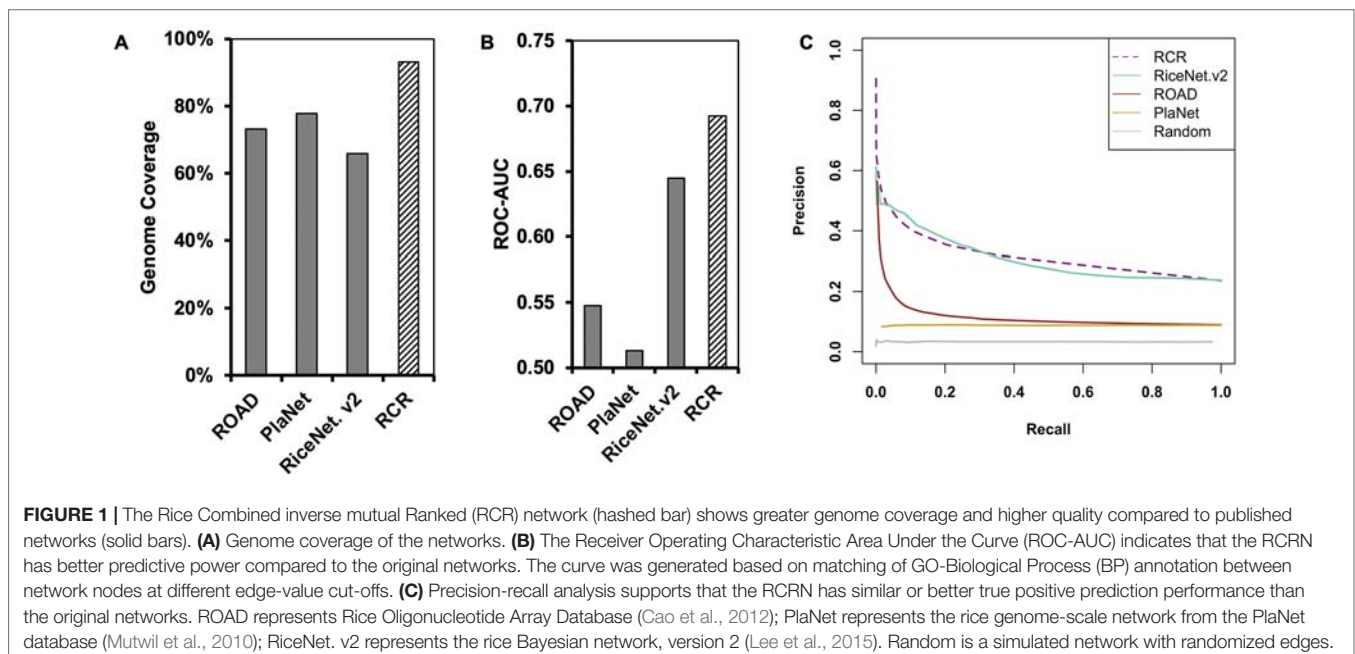
Compared to the original rice networks, the RCRN shows the highest genome coverage and maintains a scale-free topology. The RCRN covers 93% of rice genes (**Figure 1A, Table 1**) and misses only one (4%) of our list of grass-specific cell wall genes. This suggests that the RCRN is effective for study 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**).



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 $3 (\pm 1)$ -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-1* (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 pCA 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,

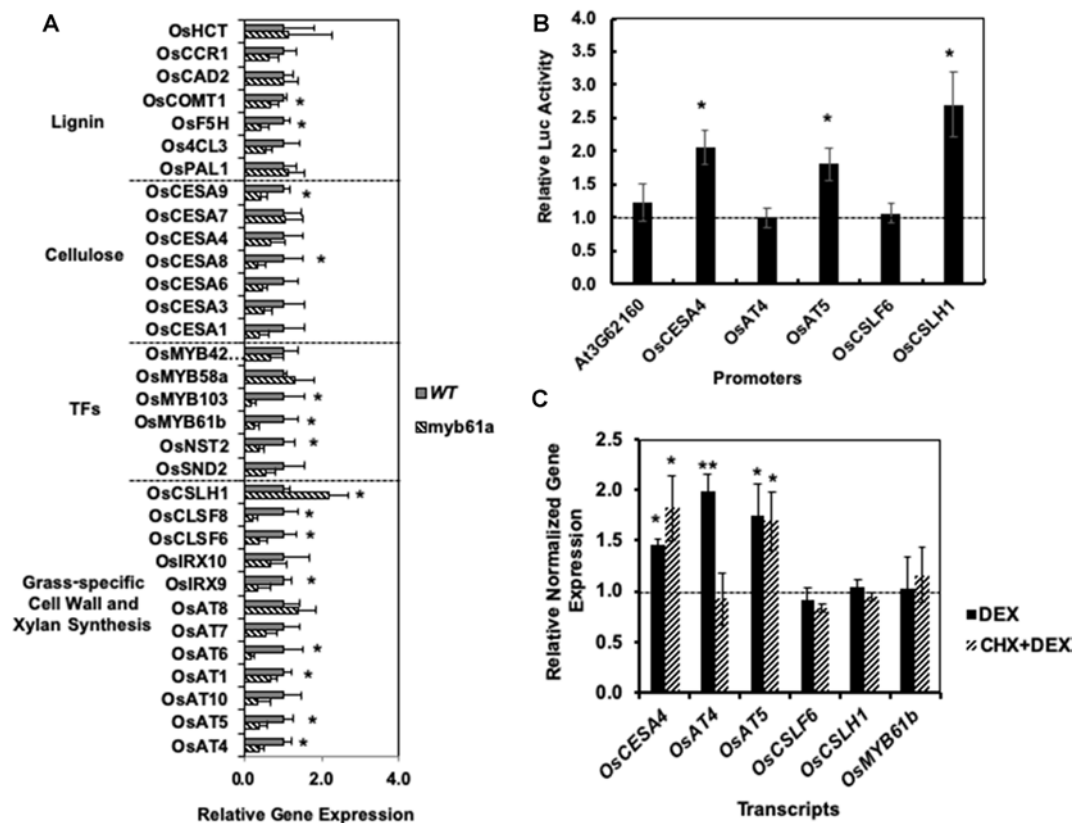


FIGURE 3 | Reverse genetics and cell-based assays suggest that OsMYB61a regulates numerous cell wall genes and directly binds to and activates promoters of grass cell wall-specific genes. **(A)** In *myb61a-1* mutant (hatched), many cell wall-related gene expressions have altered expression relative 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 *Ubg5* 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.

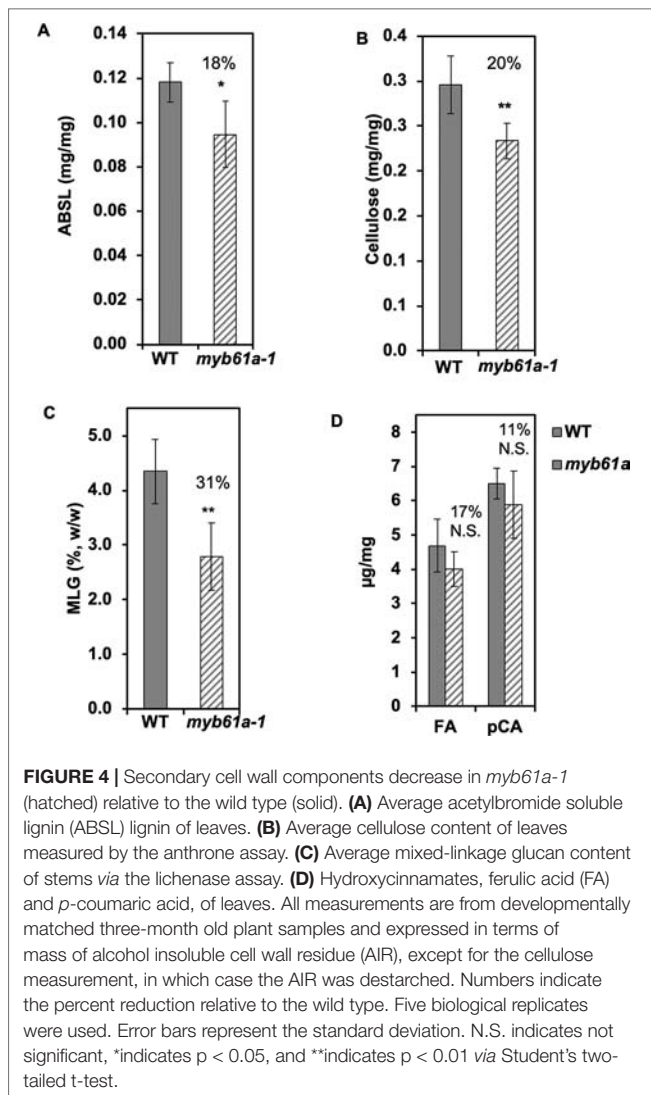
an OsMYB61a-glucocorticoid receptor ligand-binding domain (GR) fusion protein activated expression of *OsCESA4*, *OsAT4*, and *OsAT5*. However, only *OsCESA4* and *OsAT5* were still induced after treatment with CHX, suggesting that OsMYB61a binds directly to these promoters. In contrast *OsAT4* expression 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



(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 Homeomain 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					
Lignin	<i>Os4CL3</i>	190 ± 30 ^a	320 ± 10 ^b	590 ± 20 ^b	270 ± 10 ^b
	<i>OsCCR1</i>	1.0 ± 0.2	1.4 ± 0.2 ^a	4.7 ± 0.3 ^a	ND ^c
	<i>OsCOMT1</i>	1.2 ± 0.2	1.3 ± 0.2	1.2 ± 0.2	1.0 ± 0.1
	<i>OsF5H1</i>	1.3 ± 0.3	1.6 ± 0.1 ^a	5.1 ± 0.7 ^a	0.8 ± 0.1
	<i>OsF5H1</i>	1.6 ± 0.2 ^a	1.6 ± 0.2 ^a	1.8 ± 0.1 ^a	0.9 ± 0.1
Secondary wall CESA	<i>OsCAD2</i>	ND ^c	ND ^c	2.5 ± 0.2 ^a	1.3 ± 0.2
	<i>OsCESA4</i>	ND ^c	ND ^c	0.9 ± 0.1	0.4 ± 0.3
	<i>OsCESA9</i>	1.9 ± 0.2 ^a	2.1 ± 0.3 ^a	1.0 ± 0.3	0.3 ± 0.1 ^a
Mixed-linkage glucan	<i>OsCSLF6</i>	1.6 ± 0.1 ^a	1.3 ± 0.1	ND ^c	ND ^c
	<i>OsCSLF1</i>	1.1 ± 0.1	1.1 ± 0.2	ND ^c	ND ^c
Hydroxycinnamic acid	<i>OsAT4</i>	1.8 ± 0.1 ^a	1.7 ± 0.1 ^a	1.8 ± 0.2 ^a	0.3 ± 0.2 ^a
	<i>OsAT5</i>	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.

^aTwo-tailed t-test p-value < 0.05.

^bTwo-tailed t-test p-value < 0.01.

^cND indicates the interactions were not determined in this assay, which examined interactions based on RCR network predictions.

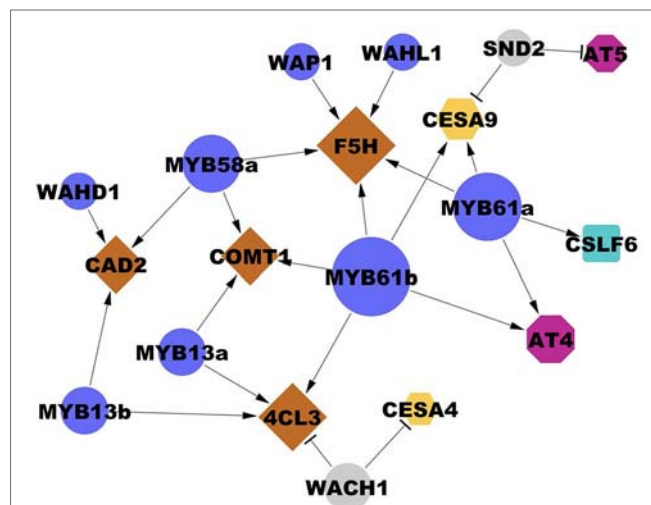


FIGURE 5 | Interactions between transcription factors and cell wall biosynthesis genes validated in the transient assay. Edges ending in a triangle (i.e., arrow head) indicate activation; edges ending in a bar indicate repression. Gene node size is proportional to degree (number of regulatory connections). 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 of 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 reference 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 ± 12 ^a	209 ± 31 ^b	95 ± 16 ^b	365 ± 34 ^b	164 ± 12 ^b	736 ± 49 ^b	164 ± 20 ^b	216 ± 19 ^b	289 ± 37 ^b	318 ± 19 ^a	371 ± 26 ^b
	OsCCR1	1.4 ± 0.1 ^a	1.7 ± 0.1 ^a	1.1 ± 0.1	1.3 ± 0.1	1.0 ± 0.1	0.4 ± 0.1 ^a	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 ± 0.1 ^a	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 ± 0.1 ^a	1.1 ± 0.1	1.0 ± 0.2	1.1 ± 0.1	0.9 ± 0.1	0.6 ± 0.1 ^a	0.9 ± 0.1	0.9 ± 0.2	1.0 ± 0.1	1.0 ± 0.1	1.1 ± 0.1
Secondary wall	OsCAD2	2.4 ± 0.5 ^a	1.3 ± 0.1	2.0 ± 0.2 ^a	1.4 ± 0.1 ^b	ND ^c	2.3 ± 0.3 ^a	0.8 ± 0.1	1.2 ± 0.1 ^a	1.0 ± 0.1	1.2 ± 0.2	1.1 ± 0.1
	OsCESA4	0.8 ± 0.1	1.4 ± 0.1 ^a	0.9 ± 0.1	ND ^c	1.8 ± 0.2 ^a	1.0 ± 0.2	0.8 ± 0.1	ND ^c	ND ^c	ND ^c	ND ^c
	OsCESA9	ND ^c	ND ^c	1.2 ± 0.2	1.3 ± 0.1	1.5 ± 0.3	0.2 ± 0.1 ^a	0.9 ± 0.2	1.3 ± 0.1	0.9 ± 0.1	1.0 ± 0.3	1.0 ± 0.1
	OsCSLF6	ND ^c	ND ^c	1.2 ± 0.3	ND ^c	1.0 ± 0.1	1.6 ± 0.4	1.2 ± 0.01	1.1 ± 0.1	0.6 ± 0.2	1.6 ± 0.2	1.1 ± 0.1
	OsCSLH1	ND ^c	ND ^c	1.1 ± 0.2	ND ^c	1.0 ± 0.1	0.9 ± 0.4	ND ^c	ND ^c	ND ^c	ND ^c	ND ^c
Hydroxycinnamic acid	OsAT4	ND ^c	ND ^c	1.4 ± 0.2	ND ^c	1.1 ± 0.1	1.1 ± 0.3	ND ^c	ND ^c	ND ^c	ND ^c	ND ^c
	OsAT5	1.0 ± 0.1	1.0 ± 0.1	1.0 ± 0.1	1.2 ± 0.1	1.2 ± 0.4	1.8 ± 0.1 ^a	0.8 ± 0.2	1.0 ± 0.2	ND ^c	1.2 ± 0.3	0.7 ± 0.3
		1.0 ± 0.2	0.7 ± 0.2	ND ^c	0.9 ± 0.2	ND ^c	1.2 ± 0.2	0.8 ± 0.1	0.9 ± 0.2	1.1 ± 0.1	0.9 ± 0.2	0.9 ± 0.3

^aTwo-tailed t-test p value < 0.05.

^bTwo-tailed t-test p value < 0.01.

^cND indicates the interactions were not determined in this assay, which examined relationships based on PCR 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 interfascicular 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

This study was supported by the Department of Energy Plant Feedstock Genomics Program under grant No. DE-SC0006904

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

ACKNOWLEDGMENTS

Thanks to Ms. Mary-Francis LaPorte for technical support. We appreciate intellectual input from Dr. Shin-han Shiu and

comments on the manuscript from Dr. Seung Rhee and the reviewers. Dr. M.S. Chern provided the GR-fusion construct.

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., Breikreutz, A., Boucher, L., Breikreutz, 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.1365-3113.2010.04363.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 A4or7 results 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., Mizrahi, 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., Mehrrens, 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
- Liepmann, 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., Ebenhoeh, 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., Shiota, 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. J., 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., Koshihara, 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.1365-3113X.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., Egboisiuba, 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

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Zhao, Lin, Romero-Gamboa, Saha, Goh, An, Jung, Hazen and Bartley. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Finding New Cell Wall Regulatory Genes in *Populus trichocarpa* Using Multiple Lines of Evidence

Anna Furches^{1,2†}, David Kainer^{1†}, Deborah Weighill^{1,2†}, Annabel Large^{1,3,4}, Piet Jones^{1,2}, Angelica M. Walker^{1,3,5,6}, Jonathon Romero^{1,2}, Joao Gabriel Felipe Machado Gazolla¹, Wayne Joubert⁷, Manesh Shah¹, Jared Streich¹, Priya Ranjan^{1,8}, Jeremy Schmutz^{9,10}, Avinash Sreedasyam¹⁰, David Macaya-Sanz¹¹, Nan Zhao⁸, Madhavi Z. Martin¹, Xiaolan Rao¹², Richard A. Dixon¹², Stephen DiFazio¹¹, Timothy J. Tschaplinski¹, Jin-Gui Chen¹, Gerald A. Tuskan¹ and Daniel Jacobson^{1,2*}

OPEN ACCESS

Edited by:

Mathew G. Lewsey,
La Trobe University, Australia

Reviewed by:

Jennifer R. Bromley,
British American Tobacco
(United Kingdom), United Kingdom
Amy Marshall-Colon,
University of Illinois at Urbana-
Champaign, United States

*Correspondence:

Daniel Jacobson
jacobsonda@ornl.gov

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 15 February 2019

Accepted: 09 September 2019

Published: 08 October 2019

Citation:

Furches A, Kainer D, Weighill D,
Large A, Jones P, Walker AM,
Romero J, Gazolla JGF, Joubert W,
Shah M, Streich J, Ranjan P,
Schmutz J, Sreedasyam A, Macaya-
Sanz D, Zhao N, Martin MZ, Rao X,
Dixon RA, DiFazio S, Tschaplinski TJ,
Chen J-G, Tuskan GA and
Jacobson D (2019) Finding New Cell
Wall Regulatory Genes in *Populus*
trichocarpa Using Multiple
Lines of Evidence.
Front. Plant Sci. 10:1249.
doi: 10.3389/fpls.2019.01249

¹ Biosciences Division, and The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ² The Brederes Center for Interdisciplinary Research and Graduate Education, University of Tennessee, Knoxville, TN, United States, ³ Oak Ridge Associated Universities (ORAU), Oak Ridge, TN, United States, ⁴ Department of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN, United States, ⁵ Department of Computer Science, Johns Hopkins University, Baltimore, MD, United States, ⁶ Department of Biology, Johns Hopkins University, Baltimore, MD, United States, ⁷ Oak Ridge Leadership Computing Facility, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ⁸ Department of Plant Sciences, The University of Tennessee Institute of Agriculture, University of Tennessee, Knoxville, TN, United States, ⁹ Joint Genome Institute, Walnut Creek, CA, United States, ¹⁰ HudsonAlpha Institute for Biotechnology, Huntsville, AL, United States, ¹¹ Department of Biology, West Virginia University, Morgantown, WV, United States, ¹² BioDiscovery Institute and Department of Biological Sciences, University of North Texas, Denton, TX, United States

Understanding the regulatory network controlling cell wall biosynthesis is of great interest in *Populus trichocarpa*, both because of its status as a model woody perennial and its importance for lignocellulosic products. We searched for genes with putatively unknown roles in regulating cell wall biosynthesis using an extended network-based Lines of Evidence (LOE) pipeline to combine multiple omics data sets in *P. trichocarpa*, including gene coexpression, gene comethylation, population level pairwise SNP correlations, and two distinct SNP-metabolite Genome Wide Association Study (GWAS) layers. By incorporating validation, ranking, and filtering approaches we produced a list of nine high priority gene candidates for involvement in the regulation of cell wall biosynthesis. We subsequently performed a detailed investigation of candidate gene GROWTH-REGULATING FACTOR 9 (*PtGRF9*). To investigate the role of *PtGRF9* in regulating cell wall biosynthesis, we assessed the genome-wide connections of *PtGRF9* and a paralog across data layers with functional enrichment analyses, predictive transcription factor binding site analysis, and an independent comparison to eQTN data. Our findings indicate that *PtGRF9* likely affects the cell wall by directly repressing genes involved in cell wall biosynthesis, such as *PtCCoAOMT* and *PtMYB.41*, and indirectly by regulating homeobox genes. Furthermore, evidence suggests that *PtGRF9* paralogs may act as transcriptional co-regulators that direct the global energy usage of the plant. Using our extended pipeline, we show multiple lines of evidence implicating the involvement of these genes in cell wall regulatory functions and demonstrate the value of this method for prioritizing candidate genes for experimental validation.

Keywords: lines of evidence, cell wall, regulation, Genome Wide Association Study, candidate gene identification, network analysis, multi-omic, *Populus trichocarpa*

INTRODUCTION

The biosynthesis and regulation of the plant cell wall has been the subject of a large body of research due to the industrial importance of lignocellulosic biomass, as well as the role of the cell wall in the function of other plant biological systems such as stress response, inter-cellular transport, and disease resistance. For industrially cultivated genera such as *Populus*, the primary cell wall constituents (i.e., cellulose, lignin, hemicellulose) provide feedstock for downstream products including biofuel, lumber, paper, and advanced lignin products (Sannigrahi et al., 2010; Porth et al., 2013). There is therefore broad interest in understanding the mechanisms that regulate the biosynthesis and modification of the cell wall, both from a yield and composition perspective.

A great variety of biopolymers are synthesized and incorporated into the primary and secondary cell wall, often in response to biotic and abiotic stress, nutrient availability, and developmental and temporal switches, all of which govern the macro-scale form of the plant. A highly complex network of genetic regulation has evolved to control the rate of biosynthesis of cell wall polymers, their intrinsic monomer composition, their transport to and subsequent deposition in the cell wall, and the expansion of the wall under changing intra-cellular conditions. In the model plant *Arabidopsis thaliana*, Bischoff et al. (2010) estimated that over 1,000 genes encode proteins related to the cell wall, while Cai et al. (2014) predicted a number closer to 3,000 based on clustering of gene co-expression (Bischoff et al., 2010; Cai et al., 2014). Furthermore, Taylor-Teeples et al. (2015) tested a library of 1,664 transcription factors in *A. thaliana* for interaction with the promoter regions of cell wall biosynthesis genes and found 413 such interactions in root vascular tissue alone (Taylor-Teeples et al., 2015). Studies such as these highlight the immense complexity involved in cell wall regulation, much of which is still to be elucidated.

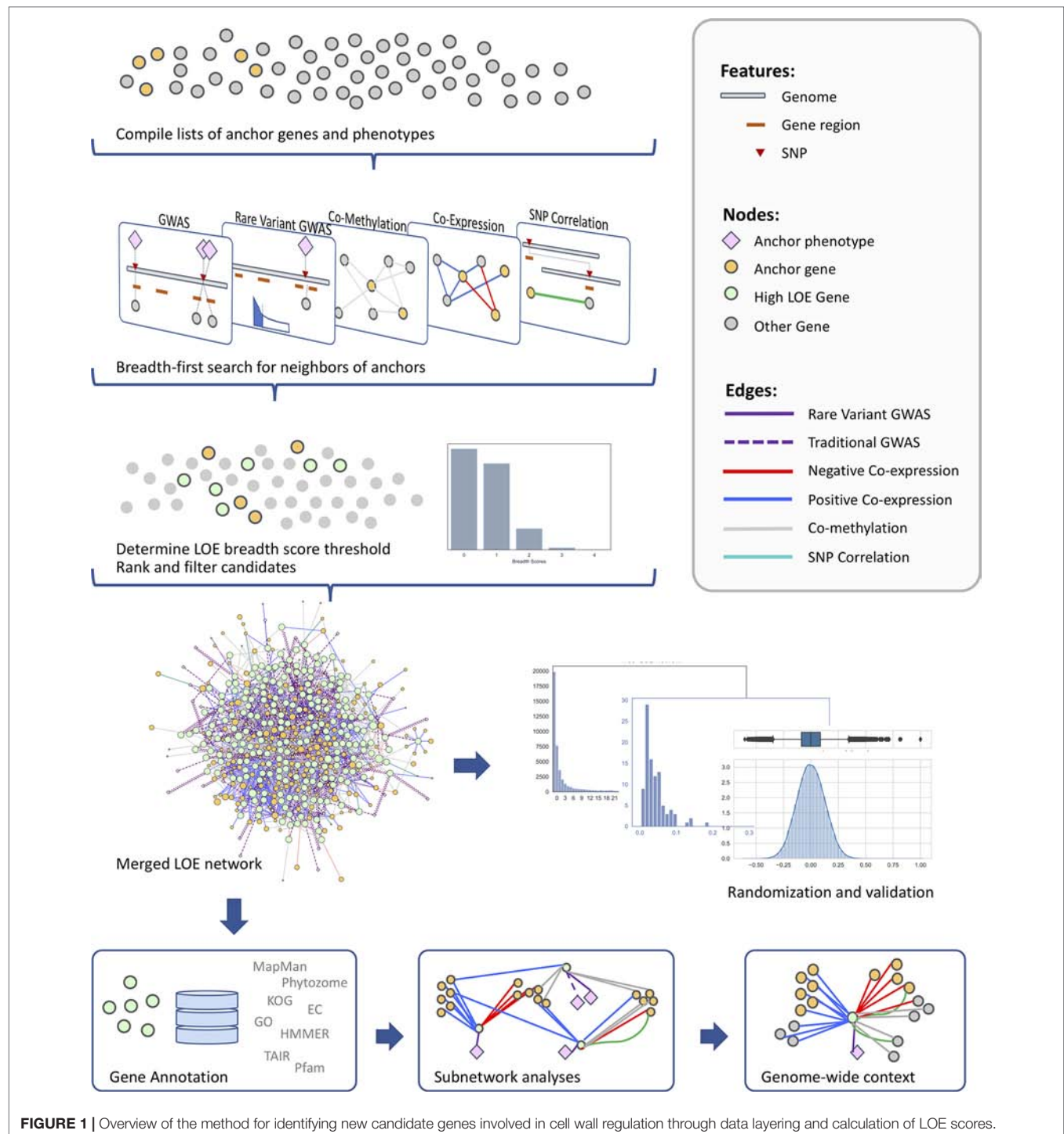
Due to poplar's status as a model woody plant and its importance for lignocellulosic products, many studies have investigated the regulatory network of the cell wall and its components in *Populus* species or in multiple genera in combination with *Populus* (Ohtani et al., 2011; Puzey et al., 2012; Lu et al., 2013; Porth et al., 2013; Yu et al., 2013; Ko et al., 2014; Wang et al., 2014; Zhong and Ye, 2014; Muchero et al., 2015; Lin et al., 2017; Shi et al., 2017; Xie et al., 2018a). Many of these studies have focused either on characterizing *Populus* homologs of genes that have been shown to have an effect on the cell wall chemistry or plant growth traits in mutant *Arabidopsis* lines, or perhaps were shown to be differentially expressed in comparisons of low and high growth genotypes. However, exploring the regulatory network controlling the cell wall in order to find new functional mechanisms is a challenging task due to the number of genes involved, extensive functional redundancy, and the multitude of transcriptional feedback loops. Such complex genetic architecture has contributed to the view that many quantitative traits are actually "omnigenic" (Boyle et al., 2017), such that virtually any expressed gene has a non-zero effect on the core biosynthetic genes at one or more transcriptional, post-transcriptional, post-translational, signaling or protein-protein interaction levels. Fisher (1919) predicted that rather

than a few core genes in biosynthetic pathways, the major portion of heritability is explained by a large number of loci across the entire genome that contribute small portions of the trait heritability. Under this omnigenic model, network-theory-based methods provide an elegant approach for mining omics datasets for regulatory relationships. Any biological entity (SNP, gene, protein, metabolite, etc.) can be modeled as a node and any relationship between those entities (association, co-expression, correlation, binding) can be modeled as an edge.

The network approach has been used in several studies of cell wall regulation to date, often focusing on finding clusters of genes that co-express with each other in certain tissues, thus finding putative functional units or networks. For example, Cai et al. (2014) performed co-expression network clustering in *Populus* and found major sub-clusters enriched for primary cell wall or secondary cell wall genes. Taylor-Teeples et al. (2015) produced networks based on *A. thaliana* transcription factors and their target binding sites, providing an expanded view of the multi-tiered regulatory system with respect to secondary cell wall (SCW) biosynthesis and xylem development. Yang et al. (2011) used 121 *A. thaliana* cell wall genes obtained from text mining followed by co-expression neighbor analysis to identify 694 *A. thaliana* genes and their 817 *Populus* orthologs as candidate genes for involvement in cell wall functions. Alejandro et al. (2012) identified the ABCG29 genes as transporting monolignol to the cell wall in *A. thaliana* by first analyzing co-expression networks followed by expression and functional analyses. These methods often produce an extensive list of candidate genes but with little more to support their involvement in cell wall regulation than the clustering or enrichment evidence.

Multi-omic approaches have also been performed, which include more data types to identify candidate genes. Porth et al. (2013) used a network-based multi-omic approach to find relationships between SNP, gene expression, and wood phenotype data from *P. trichocarpa*. They constructed six phenotypic-centric networks to identify genes that most influenced the expression of their related phenotype. From this study, they were able to identify candidate genes potentially related to cell wall biogenesis. Mizrachi et al. (2017) used a network-based approach to integrate known gene interactions and eQTN data in the form of a connectivity matrix with gene expression data through matrix multiplication in order to identify genes involved in lignin-related traits.

The use of multiple layers of omics data in the identification of candidate genes related to a particular phenotype provides an increased level of confidence and context surrounding the new candidate genes. In this study, we use an extended lines of evidence (LOE) pipeline for jointly mining multiple data layers to produce a curated short list of new candidate genes putatively involved in the regulation of cell-wall-related functions (Figure 1). We use an extensive set of "anchor" genes with documented roles in cell wall biosynthetic and regulatory processes and anchor metabolomic phenotypes measured in a Genome Wide Association Study (GWAS) population of *P. trichocarpa*. Multi-omic data layers (coexpression, comethylation, pairwise SNP correlation, and two SNP-metabolite GWAS data sets) are probed to find all genes in the genome with network connectivity



to the anchor set. A score is calculated for each gene with regards to the amount of evidence that the gene is involved in cell wall regulation and other cell-wall-related processes. The resulting merged LOE network of candidate genes is then subjected to validation, ranking, and filtering methods, as well as post-hoc analyses. The result is a set of 330 high-ranking candidate genes, which we then filter to a subset of regulatory genes not previously discussed in the context of the cell wall biosynthesis.

MATERIALS AND METHODS

This study makes use of various data accumulated for *P. trichocarpa* that have been used in previous investigations, including SNP data from a GWAS population, foliar metabolites measured in this GWAS population, and DNA methylation data across 10 different *P. trichocarpa* tissues (Vining et al., 2012), as well as the *P. trichocarpa* DOE Joint Genome Institute Plant

Gene Atlas (Sreedasyam et al, unpublished data; available from phytozome.jgi.doe.gov). Each data set was considered as a separate layer for this study, and integrated through the use of LOE scores. Below, the various layers are described as well as the network analysis methods used to merge layers and identify genes with high connectivity to cell wall systems.

Phenotypes

We made use of metabolite data previously obtained from leaf tissue and analyzed using GC-MS. Details can be found in (Tschaplinski et al., 2012; Li et al., 2012b; Weighill et al., 2018). To prevent spurious associations, we examined each phenotype for the presence of outliers using Median Absolute Deviation (MAD). If a sample's phenotype was more than six MADs from the population median it was removed from the GWAS for that phenotype.

Genotypes

SNP-based variant data (see DOI 10.13139/OLCF/1411410) were obtained from <https://bioenergycenter.org/besc/gwas/> and SNPs were filtered to the top 90% tranche (PASS SNPs) and call rate ≥ 0.5 using Plink (Purcell et al., 2007) and VCFtools (Danecek et al., 2011).

Genome Wide Association Layer

GWAS was performed using a linear mixed model (LMM), implemented in EMMAX (Kang et al., 2010) and leveraging ADIOS v1.13 (Lofstead et al., 2008) for scaling, to estimate the additive effect of each SNP while accounting for population structure and cryptic relatedness between samples. The tested SNPs excluded those with minor allele frequency (MAF) < 0.01 , and those with a population call rate above 0.75. In addition, we used linkage disequilibrium (LD) pruning on the main set of SNPs to produce a set of independent SNPs for estimating the genomic relationship matrix, used in the LMM. The resulting p-values were corrected for multi-hypotheses bias by applying the Benjamini-Hochberg approach (Benjamini and Hochberg, 1995) with a false-discovery rate (FDR) cutoff of 0.1.

Rare Variant GWAS Layer

While the GWAS Linear Mixed Model (LMM) tested common and less common SNPs (MAF ≥ 0.01) individually for significance, rarer SNPs were tested regionally in a joint fashion. Rare SNPs (MAF < 0.01) located within a given gene, or in the gene's 2-kb upstream and downstream flanking regions, were grouped as a region defined by that gene. RVtest (Zhan et al., 2016) was then used to apply the Sequence Kernel Association Test (SKAT) to each of the 41,335 regions defined from *P. trichocarpa* v3.0 annotations. SKAT tests each SNP in the region individually with an LMM and then forms a combined region score where each component SNP is weighted according to its MAF. Weights were drawn from a beta distribution with default shape parameters (1, 25), which produced a single P-value for the significance of association of each region, which were corrected for multiple testing with an FDR of 0.1.

Co-Expression Layer

A *P. trichocarpa* gene co-expression network was constructed as described in Weighill et al. (2018). RNA-seq reads from the *P. trichocarpa* DOE Joint Genome Institute Plant Gene Atlas (Sreedasyam et al., unpublished data; available from phytozome.jgi.doe.gov; see **Supplementary Table S1** for sample information) were trimmed using Skewer (Jiang et al., 2014), aligned to the version 3.0 *P. trichocarpa* reference (Tuskan et al., 2006) using Star (Dobin et al., 2013), and TPM (transcripts per million) values calculated for each gene and each sample. Star mapping was performed using the “-quantMode GeneCounts” option, which directs the program to count the number of reads per gene while performing the mapping. A read is counted if it overlaps one and only one gene. Both ends of the paired-end read are checked for overlaps. The counts coincide with those produced by htseq-count with default parameters. We then calculated the Spearman correlation coefficient between the expression profiles of all pairs of genes using the mcxarray package (Van Dongen, 2008) available from <https://micans.org/mcl/index.html>. An absolute threshold of 0.85 was applied in order to keep only those gene-gene pairs with strong co-expression.

Co-Methylation Layer

A *P. trichocarpa* gene co-methylation network was constructed as described in Weighill et al. (2018). MEDIP-Seq reads from the study by Vining et al. (2012) mapped to the *P. trichocarpa* V3 genome assembly, were obtained from Phytozome (Goodstein et al., 2011; Vining et al., 2012). The number of reads that mapped to each gene for each sample was determined using htseq-count (Anders et al., 2015). These counts were then converted to TPM values for each gene and each sample. Spearman correlation coefficients between the co-methylation profiles of all pairs of genes were then calculated in a manner similar to the co-expression layer, followed by an absolute threshold of 0.95.

Custom Correlation Coefficient Layer

After filtering the SNP set to remove those with MAF < 0.01 , the custom correlation coefficient (CCC) (Climer et al., 2014) between all pairs of remaining SNPs were calculated using a Parallel GPU implementation of the CCC (Joubert et al., 2017). In order to minimize correlation among SNPs due to linkage disequilibrium, only correlations from SNP pairs greater than 10 kb apart and with a CCC ≥ 0.7 were retained. SNPs were then mapped to the genes in which they were located, resulting in gene-gene correlations. Significantly correlated SNPs represent co-segregating and interacting cellular components.

Lines of Evidence Scoring and Network Analysis

The LOE method calculates a score for every gene in the genome by quantifying the connectivity of a given gene to anchor genes/phenotypes from the system of interest. Each data layer described above provides one possible line of evidence. For example, if Gene A co-expresses with one or more cell wall anchor genes, then this is counted as one line of evidence for

Gene A's involvement in the cell wall. A list of 295 anchor genes was compiled from the literature (Hao and Mohnen, 2014; Zhong and Ye, 2014; Nakano et al., 2015; Liu et al., 2017; Rao and Dixon, 2018) (**Supplementary Table S2**). Metabolites that affect cell wall development and composition, such as sugar substrates, lignin precursors, and lignin competitors, were also selected for use as cell wall anchor phenotypes (**Supplementary Table S3**).

To calculate LOE scores for each gene in the *P. trichocarpa* genome, each data layer was represented as a network. Each layer consisted of a list of source entities (cell wall anchor genes and phenotypes, or "anchor nodes"), target entities (potential candidate genes, or "target nodes"), and interactions between them (correlations/associations, or "edges"). From each layer's network, a breadth-first search was used to extract the neighbors of anchor nodes, resulting in a "one-hop" ("1-hop") network for each layer. LOE scores were calculated as per Weighill et al. (2018). Briefly, the LOE breadth score for a gene is the count of the different layers in which that gene has connections to anchor genes/phenotypes. An LOE depth score—the count of all connections to anchor genes/phenotypes across all data layers—was also calculated for each gene. After scoring, the 1-hop networks from all layers were thresholded based on the distribution of LOE breadth scores, then merged to form the LOE network containing cell wall anchor genes and phenotypes and all genes connected to them *via* one or more layers ("high LOE genes"). All genes in the merged LOE network were ranked based upon breadth and depth scores and genes with previously documented cell-wall-related roles were removed. Networks were visualized and manipulated with Cytoscape 3.6.1 (Shannon et al., 2003).

Gene Annotation, Functional Enrichment, and Expression Analyses

Functional annotations for *P. trichocarpa* genes were obtained from JGI Phytozome 12 (Goodstein et al., 2011) and MapMan using the Mercator tool (Lohse et al., 2014). A number of high LOE genes were not annotated in MapMan or Phytozome. To better understand the potential functions of those genes, protein sequences were extracted from the *P. trichocarpa* v3.1 primary transcript sequence (Tuskan et al., 2006) available from Phytozome and analyzed using HMMER v3.1b2 (Eddy, 1998) to annotate both Pfam v31.0 (Punta et al., 2011) and TIGRfam v15.0 (Haft et al., 2001) domains. Domains were thresholded using an independent E-value of 0.001. GO-term enrichment was performed on selected sets of genes using the BinGO Cytoscape app (Maere et al., 2005) using the Hypergeometric Test as well as Benjamini & Hochberg False Discovery Rate Correction at a significance level of 0.1.

A clustered heatmap of gene expression data was created using the Python (v3.6.2) package seaborn (v0.8.0; <https://seaborn.pydata.org/index.html>). Prior to analysis, six samples were removed from the data set that were outliers relative to their tissue type and treatment subgroups. Gene expression was normalized across tissues and genes were clustered using a Euclidean distance metric and Ward clustering method.

To assess orthology for a subset of genes during post-hoc analyses in Section 4.4.1, amino acid sequences containing

characteristic PFAM domains (<http://pfam.xfam.org/>) were obtained from UniProt (www.uniprot.org; KNOXI: PF03790 per Mukherjee et al., 2009; POX/BELL: PF07526 per Bellaoui et al., 2001) and reciprocal BLASTp searches were performed against *P. trichocarpa* and *A. thaliana* genomes using NCBI's BLAST (<https://blast.ncbi.nlm.nih.gov>) with default settings.

Network Validation

Randomizations of Expression and Methylation Data

We assessed whether our coexpression and comethylation networks contain greater biological signal than random networks by performing analyses on multiple randomized expression and methylation datasets. First we generated 100 randomized gene expression data sets by shuffling TPM values within genes across tissues, thereby preserving the observed range of expression for each gene but destroying the associations with tissue samples. We then generated a Spearman coexpression matrix for each random dataset and randomly subsampled 100,000 correlation values from each, resulting in a total pool of 10,000,000 random coexpression samples. We then collected 10,000,000 random subsamples from our observed coexpression data set and compared the distributions of our observed values to those of the shuffled data sets using the Wilcoxon rank-sum test using the Python package SciPy stats module (docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ranksums.html). We also performed this method with the comethylation data layer.

Functional Validation of LOE Network

To assess whether our observed LOE network captured a greater amount of biological function than random networks, we intersected the observed network as well as 100 randomized LOE networks with a GO-term functional network. We first constructed a functional network from GO Biological Process terms whereby genes that share GO terms are connected and are more likely to share biological function than unconnected genes. GO annotations for *P. trichocarpa* genes were obtained from PlantRegMap (Jin et al., 2017) and we removed any term present in over 1000 genes to avoid generating an overly dense network from highly generic functions. Furthermore, we weighted edges with a score inversely proportional to the number of genes with that GO term, such that between genes due to rarer GO terms were considered more functionally valuable than edges due to broader GO terms. If two genes shared multiple GO terms then we retained only the higher scoring edge. We then generated 100 randomized networks for each input data layer by holding anchor nodes and edges constant and replacing their 1-hop neighbors with gene labels randomly drawn from the genome, thereby ensuring that the size and structure of the randomized networks were comparable to the LOE input networks. For each set of random networks (consisting of one randomized network of each type: comethylation, coexpression, SNP correlation, traditional metabolite-GWAS, and rare variant metabolite-GWAS), LOE scoring and thresholding was performed. Each merged LOE network was then intersected with a GO-term functional network and an intersect score was recorded. The intersect score is calculated by summing the values of the GO-term network

edges that are also present in the LOE scored network. We then compared the intersect score of our observed LOE network to the distribution of randomized network intersect scores.

Expression Quantitative Trait Networks

We utilized eQTN data as an independent line of evidence for investigating the putative regulatory roles of the *PtGRF9* paralogs. RNAseq sequencing data from (Zhang et al., 2018) were obtained from the NCBI SRA database (SRA numbers: SRP097016–SRP097036; www.ncbi.nlm.nih.gov/sra). Reads were aligned to the *Populus trichocarpa* v.3.0 reference (Tuskan et al., 2006), using STAR (Dobin et al., 2013). Transcript per million (TPM) counts were then obtained for each genotype, resulting in a genotype-transcript matrix. For each gene transcript we determined outlier values, masking TPM values that exceeded a median absolute deviation from the non-zero median threshold of 5.0. Transcripts that had a non-outlier observed TPM value in more than 20% of the population were retained for further analysis. These expression profiles were then used as phenotypes in a GWAS, using EMMAX (Kang et al., 2010). Single nucleotide polymorphisms (SNPs) data, for the same population of *P. trichocarpa* genotypes, was obtained from (DOI 10.13139/OLCF/1411410). The SNPs were processed using VCFTOOLS (Danecek et al., 2011) and PLINK (Purcell et al., 2007), selecting for the 90% tranche and a minor allele frequency of 0.01. A hierarchical approach (Peterson et al., 2016) was used to correct for multiple hypotheses bias associated with the number of phenotypes. The procedure involved two rounds of false discovery rate (FDR) corrections, the initial using the Benjamini-Hochberg (Benjamini and Hochberg, 1995) procedure ($q_1 < 0.1$), followed by the Gavrilov-Benjamini-Sarkar stepdown approach (Gavrilov et al., 2009) ($q_2 < 5.1e-4$). SNP to phenotype association that passed the respective thresholds were determined to be statistically significant. 1-hop eQTN networks were then created around the *PtGRF9* paralogs.

RESULTS AND DISCUSSION

Evaluation of Expression and Methylation Data

The Wilcoxon rank-sum test was used to determine whether the distribution of correlation values differed between our observed data set and values from randomized datasets (Supplementary Figure S1). For both the expression and methylation data sets, the observed distributions were significantly different to random ($p < 0.01$ for both data types). Our coexpression data layer was thresholded to exclude correlation values below 0.85, resulting in 16,122 values (0.19%) being retained. In our shuffled data set, only 45 values (or $5.25e-04\%$) were above the 0.85 threshold. Our comethylation data layer was thresholded to exclude correlation values below 0.95, resulting in 87,458 values (0.88%) being retained. In our shuffled data set, only 1,090 values (0.01%) were above the 0.95 threshold.

Construction of LOE Network

The LOE method was used to identify new candidate genes involved in regulating the cell wall in *P. trichocarpa* by jointly probing five different omics data layers. LOE depth scores

were calculated for each gene, indicating the number of lines of evidence within each layer connecting that gene to an input set of cell wall anchor genes and metabolites. An LOE breadth score was also calculated for each gene, indicating the number of types of lines of evidence that connected the gene to input cell-wall-related targets. A merged LOE network was created after determining an appropriate LOE breadth score threshold and taking the union of all thresholded input networks. Threshold criteria dictated that candidate genes have a significant association with one or more metabolites in either the traditional or rare variant data layers as well as a total breadth score of three. We required a minimum of one GWAS association for retention in the merged network because metabolite-GWAS associations represent a measurable cell wall phenotype. A breadth score of three was selected in order to prioritize a small set of genes having strong evidence for involvement in cell-wall-related processes, and the distribution of breadth scores exhibits an inflection point at three (Supplementary Figure S2A). These criteria identified a list of 315 “high LOE genes” as potential candidates for involvement in cell-wall-related functions. Seven high LOE genes had a breadth score of four and 308 had a breadth score of three (Supplementary Figure S2B). Overall, high LOE genes were from a variety of functional categories (Supplementary Figure S3) and 80 of these genes were annotated with potential regulatory functions (Supplementary Table S4).

Candidate Gene Ranking

To prioritize candidates, we created three ranked tiers to which high LOE genes were assigned (Tier 1 is the highest priority, Tier 3 is the lowest priority). Genes were ranked by 1) breadth score and 2) total depth score minus co-methylation depth score. While our co-expression data vectors contain 64 data points per gene (64 tissues and experimental conditions), our co-methylation data vectors contain only 10 data points per gene (10 tissues and experimental conditions), resulting in an increased probability for spurious correlations in the co-methylation data layer. While the distribution of comethylation correlation values (Supplementary Figure S1) was significantly different than random, the shape of the distribution suggests a conservative approach is warranted. In order to avoid upwardly biasing gene rankings, co-methylation data was included in the first stage of the ranking process (overall rank by breadth score) but excluded from the second stage of the ranking process (ranking within breadth score bins by depth score). Genes with an LOE Breadth score of four were included in Tier 1 by default (seven genes). In addition, genes with an LOE Breadth score of three and total depth minus comethylation depth scores of five or greater were included in Tier 1, resulting assignment of 45 genes. Thirty-two genes were assigned to Tier 2 based on a total depth minus comethylation depth score of four. The remaining 238 high LOE genes had total depth minus comethylation depth scores of three or less and were assigned to Tier 3.

Functional Validation of LOE Network

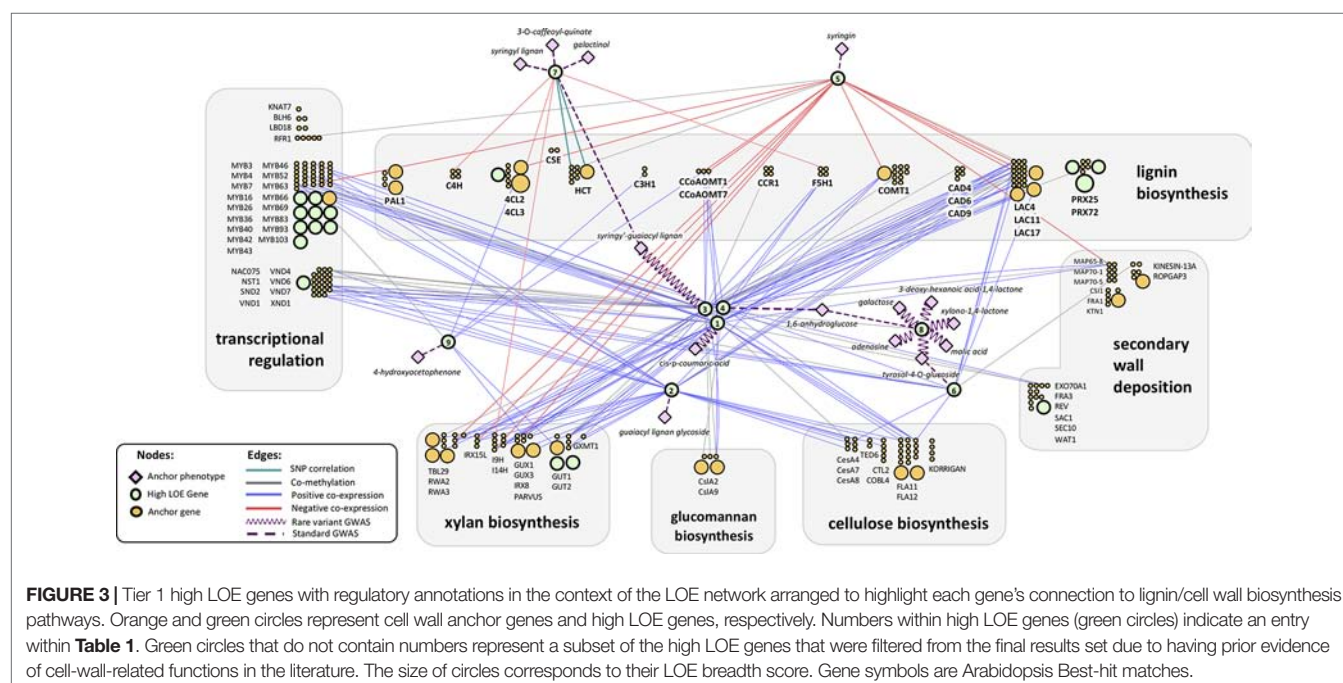
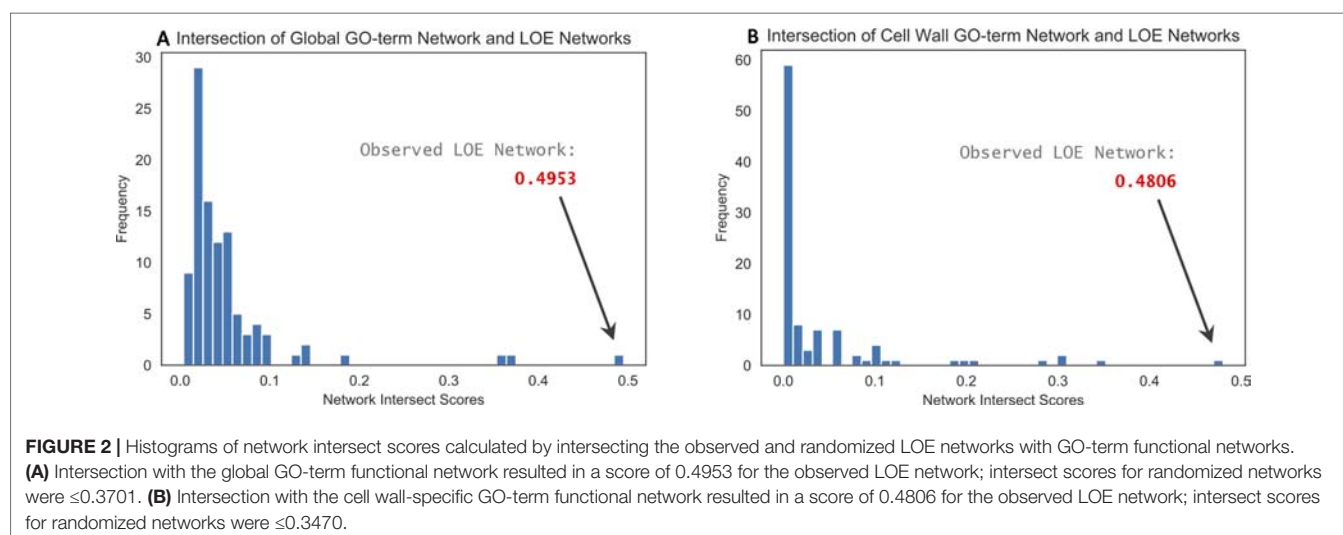
Intersection of the observed thresholded LOE network with the global GO-term functional network resulted in an intersect score of 0.4953, whereas intersect scores for the 100 randomized LOE

networks (also thresholded) ranged from 0 to 0.3701 (**Figure 2A**). Intersection of the observed LOE network with the cell wall-specific GO-term network resulted in a score of 0.4806; intersect scores for the 100 randomized networks ranged from 0 to 0.3470 (**Figure 2B**). These results imply that our observed LOE network captures a greater amount of biological signal than the randomized LOE networks.

Literature Evidence

Recovering genes for which cell-wall-related functions have been previously reported is an important internal validation for the LOE method. We performed an extensive literature review to find evidence of previously validated genes in our

results set. Forty-four genes were recovered with previous validation regarding cell-wall-related functions in *P. trichocarpa*, Arabidopsis, or other plant species and for which there is evidence of orthology in *P. trichocarpa* (see **Supplementary Table S5**). Fifteen of these high LOE genes were also in our anchor gene list. Genes with prior evidence of cell-wall-related functions were removed from our merged LOE network in order to present researchers with “new” candidate genes: 14 from Tier 1, four from Tier 2, and 11 from Tier 3; 17 of these genes are represented in **Figure 3**. However, the literature review process was not as thorough for Tiers 2 or 3, thus it is possible that some of the remaining genes in these tiers have prior connections to cell wall processes. The full ranked and filtered high LOE gene



list can be found in **Supplementary Table 6**. For the remainder of the manuscript, we focus on Tier 1 genes.

A notable example of a high LOE gene with prior evidence of a cell wall regulatory role is IQ-domain 10 calcium-signaling gene *PtIQD10* (Potri.011G096500). *PtIQD10* has a breadth score of three and a depth score of 48, including rare variant metabolite-GWAS associations with syringin, coniferin, and xylulose, and significant coexpression and comethylation with 41 cell wall anchor genes (**Supplementary Table S5**, **Supplementary Figure S4**, and **Supplementary Table S6**). The Arabidopsis ortholog *AtIQD10* (AT3G15050; orthology with *PtIQD10* and *P. deltoides* *PdIQD10* supported by phylogenetic analysis in Badmi et al., 2018) is differentially expressed in Arabidopsis lines overexpressing the transcription factor SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN 2 (*AtSND2*) (Hussey et al., 2011). Hussey et al. (2011) hypothesize *AtIQD10* activates *AtSND1* NAC, followed by activation of *SND2*, MYBs, and cell wall polymerization functions. Consistent with this model, orthologs of these genes are present in the *PtIQD10* one-hop neighborhood (**Supplementary Figure S4**). Additional evidence has recently been observed in *P. trichocarpa* congeners. An ortholog of *PtIQD10* in the *P. alba* x *P. glandulosa* hybrid “84k” is differentially expressed during the transition between primary and secondary growth phases in stems (Li et al., 2017). In addition, *P. deltoides* ortholog *PdIQD10* has higher expression levels in tension-stressed xylem tissues and secondary walled cells, and RNAi repression of *PdIQD10* results in altered phenotypes such as increased cellulose, wall glucose content, plant height, stem count, and stem density (Badmi et al., 2018; Macaya-Sanz et al., 2017). *PdIQD10* is coexpressed with secondary cell wall related genes such as *SUSY*, *CESAs*, and *KOR* (Badmi et al., 2018), orthologs of which are present in our *PtIQD10* subnetwork (Potri.018G103900 cellulose synthase/*PdCesA7-B/AtCESA7* and Potri.004G059600 *PtCESA.2/PdCESA8-B/AtCESA8*; see **Supplementary Table S7** for the *PtIQD10* one-hop subnetwork node information for **Supplementary Figure S4**).

In another example of a high LOE gene with prior evidence of a cell-wall-related role, Porth et al. (2013) found that a SNP in an exostosin family protein gene (Potri.019G044600) involved in xylogalacturonan biosynthesis was correlated with xylose (hemicellulose) content. In yet another example, Pomiès et al. (2017) found a berberine bridge enzyme gene (Potri.011G161500) with orthology to *AtEDA28/MEE23* (AT2G34790, shown to play a role in lignin monolignol metabolism) was highly up-regulated 72 h after mechanical perturbation of stems as plants modified cell wall properties in response. Another example with growing evidence of cell-wall-related regulatory functions is MADS-box transcription factor *PtAGL12* (Du et al., 2009; Du et al., 2011; Weighill et al., 2018; see **Supplementary Text S1**, **Supplementary Figure S5**, **Supplementary Table S8**, and **Supplementary Figure S6** for additional evidence regarding the putative role of *PtAGL12* in regulating cell wall biosynthesis).

Tier 1: Highest Priority Candidates for Cell Wall Regulation

Tier 1 genes have the strongest evidence of involvement in cell wall related processes (**Table 1**). Of these, nine genes had

regulatory annotations (via MapMan, www.arabidopsis.org, or PFAM; see **Supplementary Table S4** for categories considered regulatory). While the remaining 21 genes did not have regulatory annotations, our results suggest they play a role in cell wall biosynthesis.

Among Tier 1 regulatory genes, there were a total of 18 metabolite-GWAS associations, 8 of which were rare variant hits (**Figure 3**). Potri.013G093800 (Arabidopsis homolog AT1G71350, a eukaryotic translation initiation factor SUI1 family protein) has the highest number of rare variant metabolite-GWAS associations (six) of any high LOE gene as well as the highest number of total combined GWAS edges (seven). Most Tier 1 regulatory genes share edges with cell wall anchor genes from multiple process categories (in **Figure 3**, gray boxes indicating functional groupings of cell wall anchor genes). On average, Tier 1 genes were connected by multiple edges to four different functional groups, suggesting that Tier 1 genes influence multiple aspects of cell wall biosynthesis. Furthermore, eight Tier 1 regulatory genes shared edges with anchor cell wall transcriptional regulation genes (**Figure 3**).

Notably, coexpression edges for Tier 1 regulatory genes were either strictly negative for a given gene, or strictly positive, perhaps hinting at the regulatory mechanism of each gene. Two Tier 1 regulatory genes (Potri.015G006200: GROWTH-REGULATING FACTOR 9/*PtGRF9* and Potri.018G105600: NUCLEOID-ASSOCIATED PROTEIN YBAB) were negatively coexpressed with cell wall genes and six were positively co-expressed with cell wall genes. The negatively coexpressed genes (Potri.015G006200, Potri.018G105600) did not share any neighbor nodes, however they are both connected to lignin and xylan biosynthesis genes. In contrast, positively coexpressed Tier 1 regulatory genes had a large overlap in neighbor cell wall anchor genes. The overlap was even more pronounced among Potri.008G112300, Potri.001G216000, Potri.013G060500, and Potri.013G156300 despite a complete lack of overlap among metabolite-GWAS edges or MAPMAN functional annotations (**Supplementary Table S9**).

We conducted an in-depth investigation into the Tier 1 regulatory gene *PtGRF9* (Potri.015G006200) to assess support for *PtGRF9* playing a regulatory role in cell wall biosynthesis.

GROWTH-REGULATING FACTOR 9: Putative Master Regulator

The transcription factor gene GROWTH-REGULATING FACTOR 9 (*PtGRF9*/Potri.015G006200) had a breadth score of three and depth score of 17, including 13 negative coexpression edges (the highest negative coexpression depth score in our analysis). *PtGRF9* shared nine edges with lignin biosynthesis genes, four edges with xylan biosynthesis genes, two edges with transcriptional regulation genes, and one edge with a secondary cell wall deposition gene.

The *P. trichocarpa* genome annotation (v3.0; available on <https://phytozome.jgi.doe.gov>; Tuskan et al., 2006) indicates the best-hit Arabidopsis match for *PtGRF9* is AT5G53660 (growth-regulating factor 7, *AtGRF7*). To assess support for orthology, we performed reciprocal BLASTp searches of amino acid sequences

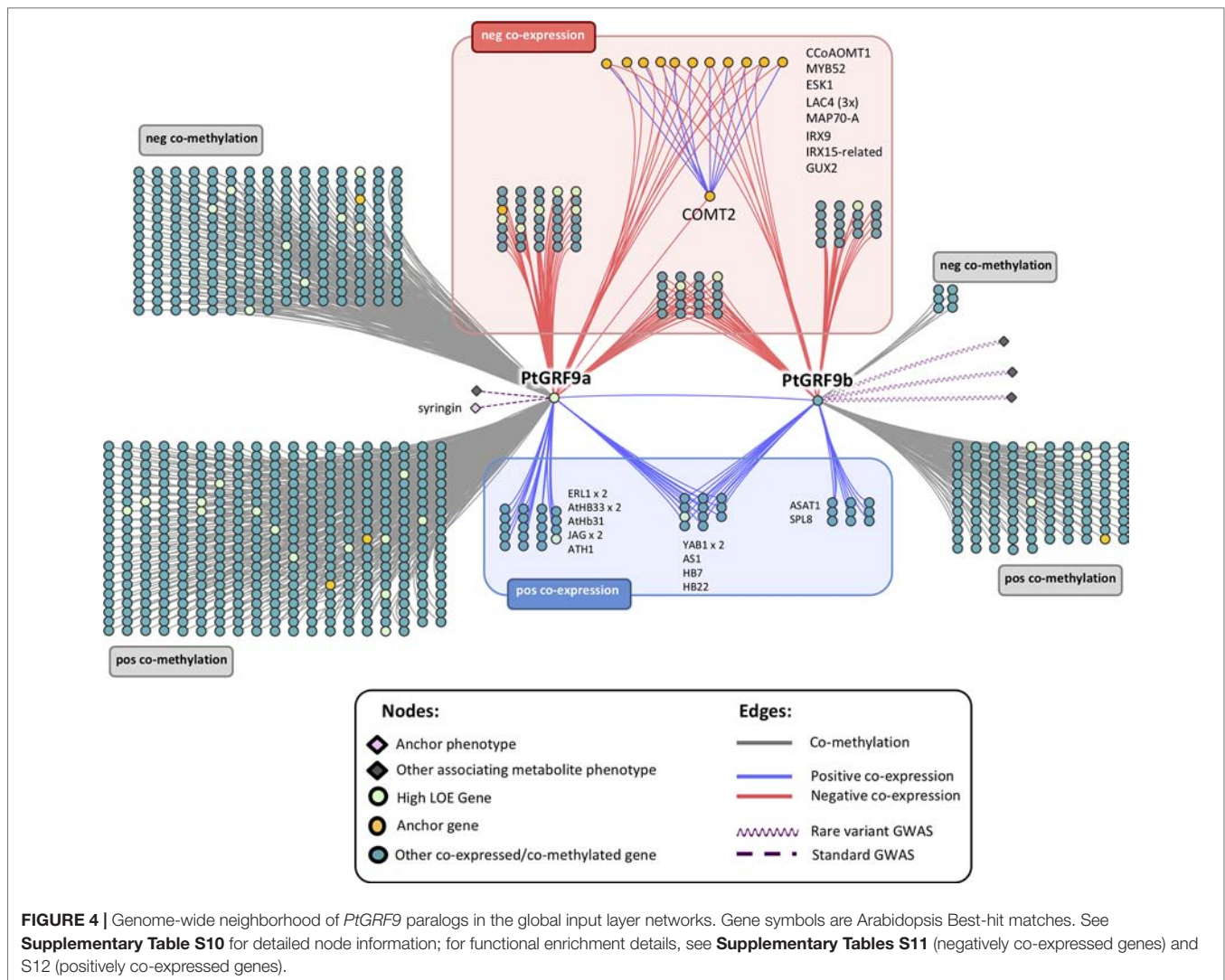
TABLE 1 | Tier 1 high LOE genes. See **Supplementary Table S9** for additional score and annotation information.

Node #	Gene ID	Arabidopsis gene/domain symbol	Description	Arabidopsis ortholog
Regulatory genes				
1	Potri.008G112300		DNA glycosylase superfamily protein	AT1G13635.2
2	Potri.001G216000	EAR1	ENHANCER OF ABA CO-RECEPTOR 1	AT5G22090.1
3	Potri.013G060500	ATCRT1	RING/U-box superfamily protein	AT5G56340.1
4	Potri.013G156300	Shisa	Wnt and FGF inhibitory regulator	
5	Potri.015G006200	AtGRF7, GRF7	growth-regulating factor 7	AT5G53660.1
6	Potri.017G053000	AMC1, ATMC1, ATMCPB1	metacaspase 1	AT1G02170.1
7	Potri.018G105600	YbaB_DNA_bd	YbaB/EbfC DNA-binding family	AT2G24020.1
8	Potri.013G093800		eukaryotic translation initiation factor SUI1 family protein	AT1G71350.1
9	Potri.010G072700		RING/U-box superfamily protein	AT5G43420.1
Other genes				
	Potri.004G085400	ATGLN1;1, ATGSR1, GLN1;1, GSR 1	glutamine synthase clone R1	AT5G37600.1
	Potri.006G256000		Phox (PX) domain-containing protein	AT4G32160.1
	Potri.012G093800	ATNDPK2, NDPK1A, NDPK2	nucleoside diphosphate kinase 2	AT5G63310.1
	Potri.010G155600		Leucine-rich repeat transmembrane protein kinase	AT1G53440.1
	Potri.001G340400	SEO_N	SEO_N-Sieve element occlusion N-terminus	
	Potri.006G153300		N-acetylated-alpha-linked acidic dipeptidase (NAALAD)	AT5G19740.1
	Potri.008G156600	AST12, SULTR3;1	sulfate transporter 3;1	AT3G51895.1
	Potri.003G079900	AW: HRGP	hydroxyproline-rich glycoprotein family protein	AT4G16790.1
	Potri.T135500	CYCP4;1	cyclin p4;1	AT2G44740.1
	Potri.018G090300	AHA1, HA1, OST2, PMA	H(+)-ATPase 1	AT2G18960.1
	Potri.017G059300	SHM4	serine hydroxymethyltransferase 4	AT4G13930.1
	Potri.004G059900		Protein of unknown function DUF1685	AT2G42760.1
	Potri.016G115200	LHCB4.2	light harvesting complex photosystem II	AT3G08940.2
	Potri.015G063400	ATATH2, ATH2	ABC2 homolog 2	AT3G47740.1
	Potri.019G087700	ATSERK1, SERK1	somatic embryogenesis receptor-like kinase 1	AT1G71830.1
	Potri.007G027400		anti-muellerian hormone type-2 receptor	AT3G50685.1
	Potri.005G067000		Protein kinase protein with adenine nucleotide alpha hydrolases-like domain	AT1G77280.1
	Potri.001G352200	ATPUP10, PUP10	purine permease 10	AT4G18210.1
	Potri.011G142200	PSBR	photosystem II subunit R	AT1G79040.1
	Potri.006G060100	CRR6	chlororespiratory reduction 6	AT2G47910.1
	Potri.010G113700	FAB1C	FORMS APLOID AND BINUCLEATE CELLS 1C	AT1G71010.1

containing the WRC (PF08879) and QLQ (PF08880) domains from *A. thaliana* and *P. trichocarpa* (obtained from UniProt; www.uniprot.org) and a phylogenetic analysis (see **Supplementary Figure S7**). Our results support an orthologous relationship between *PtGRF9* and *AtGRF7*, which is consistent with the phylogenetic analysis of Cao et al. (2016). While investigating support for orthology between *PtGRF9* and *AtGRF7*, we discovered a second *AtGRF7* ortholog in the *P. trichocarpa* genome (Potri.012G022600; hereafter, Potri.015G006200 is referred to as “*PtGRF9a*” and Potri.012G022600 as “*PtGRF9b*”; **Supplementary Figure S7**). *PtGRF9b* was not present in our set of high LOE genes because it has a breadth score of 2 and was not associated with any cell wall phenotypes through GWAS analyses. Because *PtGRF9b* had strong positive coexpression with *PtGRF9a* and shared edges with many cell wall genes, we included *PtGRF9b* in further analyses.

We constructed genome-wide 1-hop networks around each *PtGRF9* paralog across all data layers to assess the functional annotations of nearest neighbors (**Figure 4**; see **Supplementary**

Table S10 for detailed information about nodes). *PtIQD10* is present in the 1-hop network, along with many other genes with documented roles in cell wall processes. *PtGRF9a* and *PtGRF9b* are jointly positively co-expressed with 14 genes (one of which is a high LOE gene related to cell wall processes) and are jointly negatively co-expressed with 27 genes (including 7 cell wall anchor genes and 2 high LOE genes), implying an overlap in function. However, the bulk of neighbor genes are unique to each paralog, indicating divergence and perhaps specialization for specific tissues and conditions. GO-term functional enrichment analysis of the negative co-expression nodes in the 1-hop network showed significant enrichment for cell wall biological processes, including lignin biosynthesis, xylan biosynthesis and cell wall organization or biogenesis (**Supplementary Table S11**). In addition, the metabolite-GWAS association between *PtGRF9a* and syringin (a monolignol glucoside) indicated this SNP is associated with an allelic effect on syringin concentration (**Supplementary Figure S8**), further implicating *PtGRF9a* and *PtGRF9b* as repressors of secondary cell wall formation.



In Arabidopsis, *AtGRF7* is one of nine members of the *GRF* family of transcription factors (there are 20 *GRF* homologs in *P. trichocarpa*) that affect growth *via* multiple mechanisms (Omidbakhshfard et al., 2015). *AtGRF7* has specifically been shown to modulate drought response by repressing *DREB2A* (Joshi et al., 2016) which ensures that drought response genes normally activated by *DREB2A* are not expressed under non-drought conditions, thus avoiding reduced growth. In addition to stress response, *GRF* genes are involved in regulating cell proliferation and differentiation in the shoot apical meristem (SAM). *GRF* genes therefore impact the elongation of stems, new leaf initiation, and the size and shape of leaves (Gonzalez et al., 2012). The phenotypic penetrance may occur as part of a complex formed with *GRF* Interacting Factor (*GIF1/AN3*) proteins (Hoe Kim and Tsukaya, 2015), where the *GRF*-*GIF* complex serves as a transcriptional activator, recruits chromatin remodeling complexes, and regulates the meristematic state of a tissue.

GO-term enrichment analysis of the positive coexpression nodes in the *PtGRF9* 1-hop network was consistent with roles

reported in the literature for *GRF* genes (**Supplementary Table S12**). The most significantly enriched Biological Process GO terms include specification of axis polarity, shoot system development, shoot system morphogenesis and negative regulation of cell proliferation. Numerous osmotic-stress related genes are also found in the *PtGRF9* network (e.g., *AHA1/OST2*, *ERL1*, *PIP2;2*, *TIP4;1*, and *AREB3*), reflecting the well-documented relationship between *AtGRF7* and drought response. Significant connections between the *PtGRF9* paralogs and *PtGIF1* or *PtDREB2A* are not present in our LOE network. On closer inspection of co-expression values across tissues we see that *PtGRF9a* and *PtGIF1* do coexpress strongly in bud and immature leaf, but expression diverges in mature leaf and roots which causes the strength of coexpression to fall just below our 0.85 threshold (**Supplementary Figure S9**). The case is less clear for *PtDREB2A* as it shows little expression in most tissues.

Evidence that the *PtGRF9* paralogs play roles in regulating growth, defense, stress response, secondary growth, and cell wall biosynthesis suggest that *PtGRF9a* and *PtGRF9b* could be

transcriptional co-regulators as described by Xie et al. (2018b), acting as master regulators that direct the global allocation of energy within a plant.

Evidence for Regulation of the Cell Wall by *PtGRF9*

To date, a role for the *GRF* family in cell wall regulation has not been reported, though it has been noted that cell proliferation and timing of differentiation must require control or delay of secondary cell wall deposition (Mele et al., 2003). Barros et al. (2015) noted that lignin cannot be removed once deposited, thus, specific regulatory mechanisms are required to control lignin biosynthesis and deposition at specific stages during cell differentiation. The contrasting patterns of coexpression between cell wall biosynthesis and meristematic control in our *PtGRF9* 1-hop network (Figure 4) suggest that it could be involved in such a mechanism. Furthermore, the GWAS association with syringin suggests that allelic variation in *PtGRF9a* in this population may have an additive effect on the amount of sinapyl alcohol stored or released for cell wall lignification.

Knowledge regarding downstream targets of *GRF* genes is incomplete (see Omidbakhshfard et al., 2015 for a comprehensive review). AtGRF7 has been shown to repress *AtDREB2A* by binding to the motif *TGTCAGG* (Kim et al., 2012). Additionally, the central *CAG* sub-motif is enriched in the promoter of *KNOX* genes that are targeted by *GRFs* (Kuijt et al., 2014). We searched for the complete *TGTCAGG* motif in the promoter regions of Arabidopsis homologs of the genes that coexpress with *PtGRF9a* using the online *athamap.de* tool, revealing two potential AtGRF7 targets in our 1-hop network: caffeoyl coenzyme A O-methyltransferase 1 (*AT4G34050/AtCCoAOMT1*) and MADS-box transcription factor *AtAGL12* (*AT1G71692*). Both genes are relevant to the cell wall, and *P. trichocarpa* homologs of these genes are negatively co-expressed with *PtGRF9a*. To further investigate these genes as potential *PtGRF9a* targets, we used Analysis of Motif Enrichment (AME) (McLeay and Bailey, 2010), but found no evidence for enrichment of the *TGTCAGG* motif in the 2-kb upstream or CDS regions of *PtCCoAOMT* (Potri.001G304800 and Potri.009G099800) or *PtAGL12* (Potri.013G102600). Manual examination revealed that the *TGTCAGG* motif appears inexact in the upstream regions of *PtCCoAOMT1* and *PtAGL12* (*TGTTTCAGG* in *CCoAOMT1* Potri.009G099800; *TGTCAGC* in *PtCCoAOMT* Potri.001G304800 and *PtAGL12*). Consistent with the findings of Franco-Zorrilla et al. (2014), who show that repressor TFs such as *PtGRF9a* are more likely than activator TFs to bind downstream of a target gene, we found 27 *Populus* genes significantly enriched for *TGTCAGG* in the 1-kb downstream region, including *PtMYB41* (Potri.012G039400, a homolog of *AtMYB52*), which is negatively coexpressed with *PtGRF9a*. *AtMYB52* is a TF known to induce secondary cell wall biosynthesis genes and its repression reduces secondary wall thickening in fibers (Zhong et al., 2008). Furthermore, *AtMYB52* overexpression has been linked with drought tolerance (Park et al., 2011). Given the established role of AtGRF7 in drought response, repression of *PtMYB41* is a potential avenue for *PtGRF9a* to regulate both lignification and drought tolerance, although further experimental evidence is required.

Analysis of our 1-hop network suggests that *PtGRF9* also affects cell wall biosynthesis by regulating a host of homeobox genes. Twenty homeobox genes were present in the *PtGRF9* network, including *PtATHB12* (Potri.001G188800; homolog of *AtHB15/AT1G52150*), which has been shown to influence secondary wall formation and cambial production of xylem (Schrader, 2004; Cassan-Wang et al., 2013), and *PtAGL12* (Du et al., 2009; Du et al., 2011; Weighill et al., 2018) (Supplementary Figure S5). There was also indirect evidence in the *PtGRF9* network suggesting *PtGRF9* interacts with *PtKNOX* genes. *KNOX* genes are involved in meristem maintenance and are downregulated to facilitate lateral primordia development and the differentiation of cambium into xylem and phloem (Hertzberg et al., 2001; Schrader, 2004; Hay and Tsiantis, 2010). *GRF* genes are involved in specification of primordia cells and have been shown to repress *KNOX* genes by forming hairpins in targeted regions (Kuijt et al., 2014; Hoe Kim and Tsukaya, 2015). Interactions between AtGRF7 and *KNOX* genes have yet to be investigated, but the primary motif of the target sequence by which AtGRF7 binds *AtDREB2A* was shown to be enriched in several *KNOX* genes, and experiments in rice, barley, and Arabidopsis have confirmed that multiple *GRF* genes bind these motifs in *KNOX* genes (Kim et al., 2012; Kuijt et al., 2014). The presence of several genes that exclusively or directly interact with *KNOX* genes in the 1-hop network strongly implies that *PtGRF9* proteins influence the cell wall via interactions with the *PtKNOX1* genes *PtSTM* and *PtBP*, and likely other *PtKNOX* genes as well (Supplementary Table S13 and Figure 5). Although *KNOX* family genes were not present in the *PtGRF9* network, this was likely due to highly tissue-specific expression patterns which our coexpression analysis methods were not designed to detect (see Supplementary Figure S10).

The *PtKNOX*-associated genes in the *PtGRF9* network have documented roles in cell wall and secondary growth phenotypes (Figure 5). SHOOT-MERISTEMLESS (*PtSTM*) downregulates gibberellic acid levels by repressing gibberellin 20-oxidase (*PtGA20ox*) biosynthesis genes and upregulating catabolism genes such as *PtGA2ox4* (positively co-methylated with *PtGRF9a*), which inhibits xylem production (Eriksson et al., 2000; Jasinski et al., 2005). Overexpression of *PtSTM/ARBORKNOX1* (*PtSTM/PtARK1*) in *P. tremula* × *P. alba* has been shown to inhibit differentiation of leaf primordia, elongation of internodes, and differentiation of secondary vascular cells (Groover et al., 2006). Counterintuitively, overexpression of *PtSTM/PtARK1* in secondary meristems also results in upregulation of some lignin biosynthesis genes and increased lignin content. Long-term transcriptional repression of *BREVIPEDICELLUS* (*AtBP*), *KNOTTED-like 2* from *A. thaliana* (*AtKNAT2*) and *AtKNAT6* outside the meristem is facilitated by chromatin remodeling carried out by the protein encoded by *ASYMMETRIC LEAVES 1* (*AtAS1*; *PtAS1* is positively co-expressed with *PtGRF9a* and *PtGRF9b*), which dimerizes with *AtAS2* and recruits the histone chaperone protein encoded by *AtHIRA* (*PtHIRA* is negatively co-methylated with *PtGRF9a*) (Guo et al., 2008; Hay and Tsiantis, 2010). *AS2* is involved in controlling seasonal lignification in spruce, likely through its role in repressing *BP*

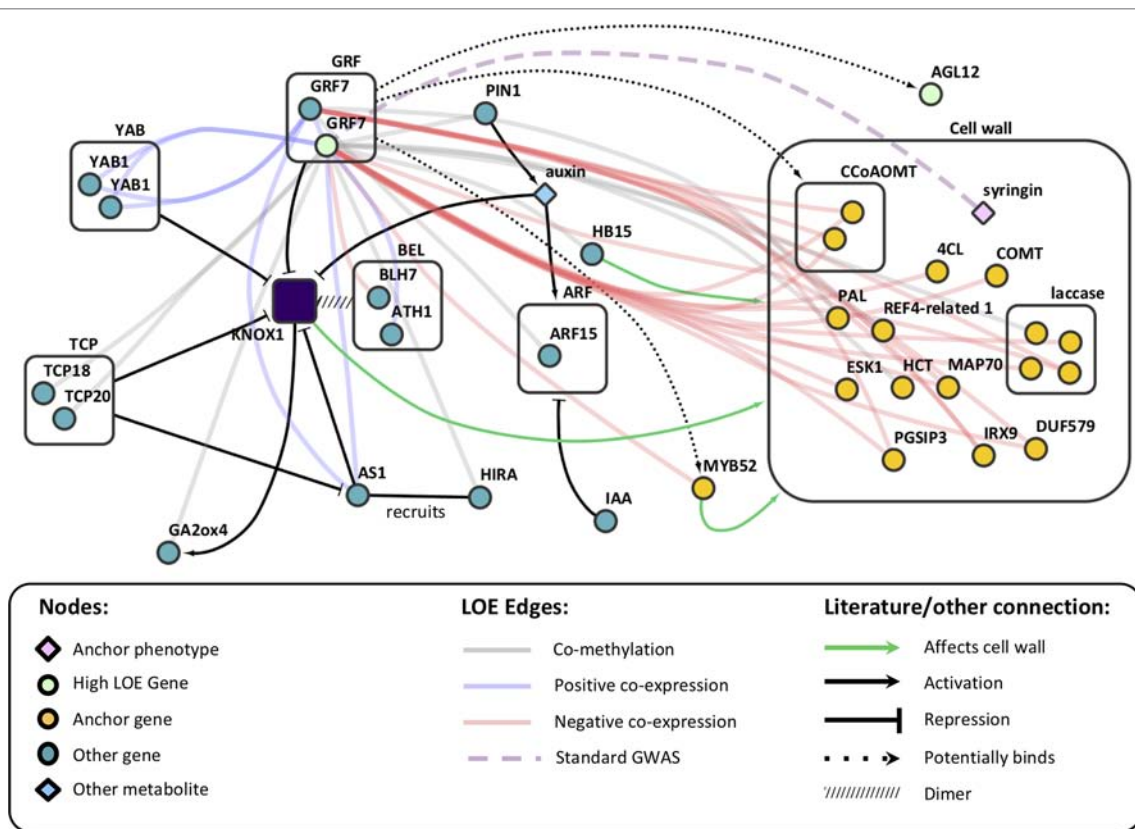


FIGURE 5 | Literature evidence linking *AtGRF7* genes to *AtKNOX* genes and the cell wall, overlaid on LOE edges from the *PtGRF9* subnetwork. See **Supplementary Table S13** for *P. trichocarpa* orthologs.

(Jokipii-Lukkari et al., 2018). BP decreases lignin deposition and regulates the localization of lignification by binding the promoters of *AtCOMT1*, *AtCCoAOMT1*, laccases, and peroxidases (putative orthologs of which are all negatively co-expressed with *PtGRF9a* and *PtGRF9b*) (Mele et al., 2003). The *PtGRF9* network includes many of the cell wall biosynthesis-related genes that Mele et al. (2003) found to be differentially expressed in *bp* mutants, including five putative orthologs (*PAL1*, *OMT1*, two *CCoAOMT1* paralogs, *PME3*, and *GH9B5*; see **Supplementary Table S13**) and an additional 23 genes belonging to the same families as differentially expressed genes in *bp* mutants (*4CL2*, five *PMEs*, *KCS19*, four peroxidases, four laccases, *ERD4*, *GAUT4*, *PUB24*, *MEE23*, *ERF1-3*, and three R2R3 MYBs: *MYB52*, *MYB93*, *MYB111*). Consistent with these observations in Arabidopsis, overexpression of *AtBP/ARBORKNOX2* (*AtBP/AtARK2*) in *P. alba* x *P. tremula* results in downregulation of ABNORMAL FLORAL ORGANS (*PtAFO/PtYAB1*), PIN-FORMED 1 (*PtPIN1*), *PtAGL12* (all negatively co-expressed with *PtGRF9a*) and *PtGA2ox* genes, leading to inhibition of cellular differentiation and division and decreases in biomass (Du et al., 2009). Furthermore, overexpression of *PtBP/PtARK2* results in downregulation of cell wall biosynthesis genes, decreased lignin content, reduced phloem fibers, and reduced secondary xylem in stems.

We did not find a connection between the *PtGRF9* genes and cell wall anchor genes *KNAT7* (Potri.001G112200, a *PtKNOX2* gene) and *BEL1*-like homeodomain 6 genes (*PtBLH6*, Potri.004G159300 and Potri.009G120800). These genes have well-documented roles in cell wall regulation (Li et al., 2012a; Cassan-Wang et al., 2013). However, the *PtGRF9* genes do not appear to be involved in their regulation, perhaps because *PtKNOX2* genes are generally more functionally diverse and broadly expressed than *PtKNOX1* genes (consistent with expression data in **Supplementary Figure S10**) (Furumizu et al., 2015). In addition, they are not involved in meristematic maintenance, and in some cases seem to overlap in function with genes that promote differentiation (consistent with what we recovered in our LOE analysis).

PtGRF9 eQTN Network: An Independent Line of Evidence

As a means of independently evaluating support for the hypothesis that *PtGRF9* paralogs regulate cell wall biosynthesis, we constructed 1-hop eQTN networks around *PtGRF9a* and *PtGRF9b* (**Figure 6**; detailed node information available in **Supplementary Table S14**). SNPs in both the *PtGRF9a* and *PtGRF9b* 1-hop networks were associated with cell wall expression phenotypes in leaf and xylem tissues, as well as expression

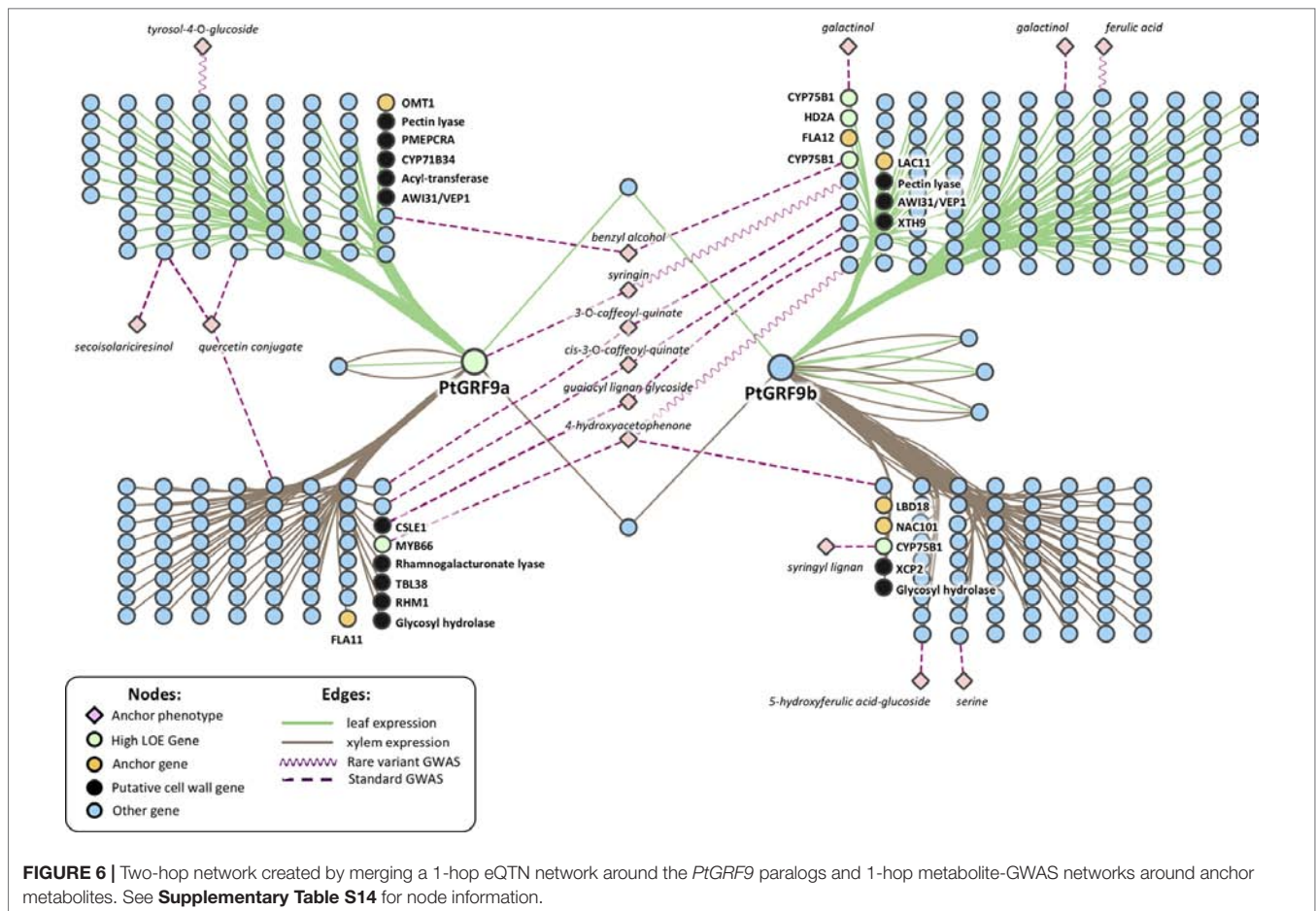


FIGURE 6 | Two-hop network created by merging a 1-hop eQTN network around the *PtGRF9* paralogs and 1-hop metabolite-GWAS networks around anchor metabolites. See **Supplementary Table S14** for node information.

phenotypes consistent with the previously documented roles of *AtGRF7* and other *GRF* orthologs in regulating functions such as growth, defense, and stress response. In agreement with the multi-omic 1-hop network described in the section *GROWTH-REGULATING FACTOR 9: Putative Master Regulator* (Figure 4), the eQTN network indicated each paralog has connectivity with cell-wall-related genes affecting multiple facets of cell wall biosynthesis, including transcriptional regulation, cellulose biosynthesis, lignin biosynthesis, xylan biosynthesis, and secondary cell wall deposition. Also consistent with the multi-omic 1-hop network, the eQTN analysis indicated that despite a low degree of topological overlap between the *PtGRF9a* and *PtGRF9b* neighborhoods, the paralogs still largely overlap in function.

To gain an understanding of how the *PtGRF9* paralogs potentially affect cell wall metabolites, the 1-hop eQTN network was merged with 1-hop anchor metabolite networks generated from traditional and rare variant metabolite-GWAS data layers. Beyond the direct GWAS association of *PtGRF9a* with syringin, 14 additional anchor metabolites are present in the 2-hop eQTN to metabolite-GWAS network (Figure 6), 6 of which are indirectly associated with both paralogs through various intermediate genes. There appears to be a pattern of segregation regarding metabolite associations between tissue

types and *PtGRF9* paralogs, perhaps indicating that these genes are diverging to fulfill different tissue-specific regulatory roles.

Future Directions

Our extended network analysis pipeline has provided a short list of putative cell wall regulatory genes to the scientific community for experimental validation. We performed an in-depth investigation of the *PtGRF9* paralogs, which are particularly promising candidates for regulation of cell wall biosynthesis and secondary growth. Furthermore, we show the *PtGRF9* paralogs are potential transcriptional co-regulators that coordinate the flow of energy among growth, defense, stress response, and lignification, in a manner consistent with the hypothesis of Xie et al. (2018b). The ability to manipulate transcriptional co-regulators such as these *via* genetic engineering and breeding programs would provide a powerful tool for shaping bioenergy crops.

Incorporating a rare variant metabolite-GWAS data layer in the LOE analysis has proven to be a valuable asset in the identification of new candidate genes. Incorporating a genome-wide eQTN (SNP-to-expression-phenotype GWAS) data layer in future analyses would provide greater clarity regarding the mechanisms through which these genes regulate cell-wall-related functions. Furthermore, DNA affinity purification sequencing

(DAP-seq) could provide further support for hypothesized transcription factor binding sites, and thus help elucidate relevant transcription factor regulatory networks. Tissue-specific expression analysis across a GWAS population would allow for increased “tissue level resolution” of the regulatory networks. The extended network analysis pipeline will be a valuable tool to integrate these new layers with the previous networks to produce a holistic model of cell wall regulation.

DATA AVAILABILITY STATEMENT

Populus trichocarpa genome sequence, annotation, and Gene Atlas expression data sets are available on Phytozome (<http://phytozome.jgi.doe.gov>). *P. trichocarpa* variant data (DOI 10.13139/OLCF/1411410) is available from <https://doi.org/10.13139/OLCF/1411410>. Scripts used to calculate LOE scores, create GO-term networks, and calculate weighted intersect scores are available on GitHub: https://github.com/afurches/Cell_Wall_LOE.

AUTHOR CONTRIBUTIONS

XR and RD provided anchor gene IDs. TT provided anchor metabolite IDs. WJ developed the Parallel GPU CCC application code. DK performed the rare variant GWAS analysis. PJ performed the standard GWAS and outlier analysis and constructed the standard GWAS and eQTN networks. DW calculated methylation TPM values and constructed the observed co-expression, co-methylation and SNP correlation networks and calculated LOE scores for the observed LOE network. AF performed expression and methylation randomizations. JG performed random subsampling. AF calculated correlation values and performed significance tests and performed LOE input layer randomizations, constructed randomized LOE networks, and calculated randomized LOE scores. DK and JR created the GO-term network, and AF calculated GO intersect scores. AF performed candidate ranking, generated the clustered expression heatmap, and performed the phylogenetic analyses. DK performed the functional enrichment and TF-binding analyses. AF created eQTN subnetworks. MS mapped gene expression atlas reads and RNAseq reads and calculated gene expression TPM values. SD and GT led the effort on constructing the GWAS population. TT led the leaf sample collection for GCMS-based metabolomic analysis, identified the peaks, and summarized the metabolomics data. PR did automated extraction of metabolite intensity from GCMS. MM collected the leaf samples and manually extracted the metabolite data. NZ performed leaf sample preparation, extracted, derivatized, and analyzed the metabolites by GCMS. JSch and AS generated the gene expression atlas data. SD and DM-S generated the SNP calls. J-GC provided RNAseq data. AF, AW, DK, and AL performed the in-depth literature searches. AL performed the domain annotation. AF, DK, DW, AL, and JStr wrote the manuscript. All authors edited the manuscript. DJ conceived of

and supervised the study, participated in the network analysis, and generated MapMan annotations.

FUNDING

Funding was provided by The Center for Bioenergy Innovation (CBI), U.S. Department of Energy Bioenergy Research Centers supported by the Office of Biological and Environmental Research in the DOE Office of Science.

An award of computer time was provided by the INCITE program. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725. This research also used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

Support for the Poplar GWAS dataset was provided by The BioEnergy Science Center (BESC) and The Center for Bioenergy Innovation (CBI), U.S. Department of Energy Bioenergy Research Centers supported by the Office of Biological and Environmental Research in the DOE Office of Science. The Poplar GWAS Project used resources of the Oak Ridge Leadership Computing Facility and the Compute and Data Environment for Science at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

The JGI Plant Gene Atlas project conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

AL and AW acknowledge that this project was supported in part by appointments to the Higher Education Research Experiences (HERE) Program at Oak Ridge National Laboratory, administered by the Oak Ridge Institute for Science and Education (ORISE). ORISE is managed by Oak Ridge Associated Universities (ORAU) for the U.S. Department of Energy (DOE).

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

ACKNOWLEDGMENTS

The authors would like to acknowledge Nancy Engle, David Weston, Ryan Aug, KC Cushman, Lee Gunter, and Sara Jawdy for metabolomics sample collection. We thank the Department of Energy Joint Genome Institute (JGI) and collaborators for prepublication access to the Plant Gene Atlas Data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2019.01249/full#supplementary-material>

REFERENCES

- Aguilar-Martínez, J. A., and Sinha, N. (2013). Analysis of the role of Arabidopsis class I TCP genes AtTCP7, AtTCP8, AtTCP22, and AtTCP23 in leaf development. *Front. Plant Sci.* 4, 1–13. doi: 10.3389/fpls.2013.00406
- Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Park, J., et al. (2012). AtABCG29 is a monolignol transporter involved in lignin biosynthesis. *Curr. Biol.* 22, 1207–1212. doi: 10.1016/j.cub.2012.04.064
- Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638
- Badmi, R., Payyavula, R. S., Bali, G., Guo, H.-B., Jawdy, S. S., Gunter, L. E., et al. (2018). A new calmodulin-binding protein expresses in the context of secondary cell wall biosynthesis and impacts biomass properties in populus. *Front. Plant Sci.* 9, 1669. doi: 10.3389/fpls.2018.01669
- Barros, J., Serk, H., Granlund, I., and Pesquet, E. (2015). The cell biology of lignification in higher plants. *Ann. Bot.* 115, 1053–1074. doi: 10.1093/aob/mcv046
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Bellaoui, M., Pidkowich, M. S., Samach, A., Kushalappa, K., Kohalmi, S. E., Modrusan, Z., et al. (2001). The Arabidopsis BELL1 and KNOX TALE homeodomain proteins interact through a domain conserved between plants and animals. *Plant Cell* 13, 2455–70. doi: 10.1105/tpc.010161
- Bischoff, V., Nita, S., Neumetzler, L., Schindelasch, D., Urbain, A., Eshed, R., et al. (2010). TRICHOME BIREFRINGENCE and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in Arabidopsis. *Plant Physiol.* 153, 590–602. doi: 10.1104/pp.110.153320
- Boyle, E. A., Li, Y. I., and Pritchard, J. K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169, 1177–1186. doi: 10.1016/j.cell.2017.05.038
- Cai, B., Li, C.-H., and Huang, J. (2014). Systematic identification of cell-wall related genes in Populus based on analysis of functional modules in co-expression network. *PLoS One* 9, e95176. doi: 10.1371/journal.pone.0095176
- Cao, Y., Han, Y., Jin, Q., Lin, Y., and Cai, Y. (2016). Comparative genomic analysis of the GRF genes in Chinese pear (*Pyrus bretschneideri* Rehd), poplar (*Populus*), grape (*Vitis vinifera*), Arabidopsis and rice (*Oryza sativa*). *Front. Plant Sci.* 7, 1–14. doi: 10.3389/fpls.2016.01750
- Cassan-Wang, H., Goué, N., Saidi, M. N., Legay, S., Sivadon, P., Goffner, D., et al. (2013). Identification of novel transcription factors regulating secondary cell wall formation in Arabidopsis. *Front. Plant Sci.* 4, 1–14. doi: 10.3389/fpls.2013.00189
- Climer, S., Yang, W., Fuentes, L., Dávila-Román, V. G., and Gu, C. C. (2014). A custom correlation coefficient (CCC) approach for fast identification of multi-SNP association patterns in genome-wide SNPs data. *Genet. Epidemiol.* 38, 610–621. doi: 10.1002/gepi.21833
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330
- Ding, L., Yan, S., Jiang, L., Zhao, W., Ning, K., Zhao, J., et al. (2015). Hanaba taranu (han) Bridges meristem and organ primordia boundaries through Pinhead, jagged, blade-on-petiole2 and cytokinin oxidase 3 during flower development in Arabidopsis. *PLoS Genet.* 11, 1–23. doi: 10.1371/journal.pgen.1005479
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- Du, J., Mansfield, S. D., and Groover, A. T. (2009). The Populus homeobox gene ARBORKNOX2 regulates cell differentiation during secondary growth. *Plant J.* 60, 1000–1014. doi: 10.1111/j.1365-3113X.2009.04017.x
- Du, J., Miura, E., Robischon, M., Martinez, C., and Groover, A. (2011). The Populus Class III HD ZIP transcription factor POPCORONA affects cell differentiation during secondary growth of woody stems. *PLoS One* 6, e17458. doi: 10.1371/journal.pone.0017458
- Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics* 14, 755–763. doi: 10.1093/bioinformatics/14.9.755
- Eriksson, M. E., Israelsson, M., Olsson, O., and Moritz, T. (2000). Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nat. Biotechnol.* 18, 784–788. doi: 10.1038/77355
- Fisher, R. A. (1919). XV.—The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinburgh* 52, 399–433. doi: 10.1017/S0080456800012163
- Franco-Zorrilla, J. M., López-Vidriero, I., Carrasco, J. L., Godoy, M., Vera, P., and Solano, R. (2014). DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc. Natl. Acad. Sci.* 111, 2367–2372. doi: 10.1073/pnas.1316278111
- Furumizu, C., Alvarez, J. P., Sakakibara, K., and Bowman, J. L. (2015). Antagonistic roles for KNOX1 and KNOX2 genes in patterning the land plant body plan following an ancient gene duplication. *PLoS Genet.* 11, 1–24. doi: 10.1371/journal.pgen.1004980
- Gavrilov, Y., Benjamini, Y., and Sarkar, S. K. (2009). An adaptive step-down procedure with proven FDR control under independence. *Ann. Stat.* 37, 619–629. doi: 10.1214/07-AOS586
- Gonzalez, N., Vanhaeren, H., and Inzé, D. (2012). Leaf size control: complex coordination of cell division and expansion. *Trends Plant Sci.* 17, 332–340. doi: 10.1016/j.tplants.2012.02.003
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2011). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944
- Groover, A. T., Mansfield, S. D., DiFazio, S. P., Dupper, G., Fontana, J. R., Millar, R., et al. (2006). The Populus homeobox gene ARBORKNOX1 reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium. *Plant Mol. Biol.* 61, 917–932. doi: 10.1007/s11103-006-0059-y
- Guo, M., Thomas, J., Collins, G., and Timmermans, M. C. P. (2008). Direct repression of KNOX loci by the asymmetric leaves1 complex of Arabidopsis. *Plant Cell Online* 20, 48–58. doi: 10.1105/tpc.107.056127
- Haft, D. H., Loftus, B. J., Richardson, D. L., Yang, F., Eisen, J. A., Paulsen, I. T., et al. (2001). TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res.* 29, 41–43. doi: 10.1093/nar/29.1.41
- Hamant, O., and Pautot, V. (2010). Plant development: a tale story. *Comptes Rendus Biol.* 333, 371–381. doi: 10.1016/j.crv.2010.01.015
- Hao, Z., and Mohnen, D. (2014). A review of xylan and lignin biosynthesis: foundation for studying Arabidopsis irregular xylem mutants with pleiotropic phenotypes. *Crit. Rev. Biochem. Mol. Biol.* 49, 212–241. doi: 10.3109/10409238.2014.889651
- Hay, A., and Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. *Development* 137, 3153–3165. doi: 10.1242/dev.030049
- Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., et al. (2001). A transcriptional roadmap to wood formation. *Proc. Natl. Acad. Sci.* 98, 14732–14737. doi: 10.1073/pnas.261293398
- Hoe Kim, J., and Tsukaya, H. (2015). Regulation of plant growth and development by the growth-regulating factor and grf-interacting factor duo. *J. Exp. Bot.* 66, 6093–6107. doi: 10.1093/jxb/erv349
- 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
- Jasinski, S., Piazza, P., Craft, J., Hay, A., Woolley, L., Rieu, I., et al. (2005). Knox action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. *Curr. Biol.* 15, 1560–1565. doi: 10.1016/j.cub.2005.07.023
- Jiang, H., Lei, R., Ding, S.-W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* 15, 182. doi: 10.1186/1471-2105-15-182
- Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982
- Jokipii-Lukkari, S., Delhomme, N., Schiffthaler, B., Mannapperuma, C., Prestele, J., Nilsson, O., et al. (2018). Transcriptional roadmap to seasonal variation in wood formation of Norway spruce. *Plant Physiol.* 176, 2851–2870. doi: 10.1104/pp.17.01590

- Joshi, R., Wani, S. H., Singh, B., Bohra, A., Dar, Z. A., Lone, A. A., et al. (2016). Transcription factors and plants response to drought stress: current understanding and future directions. *Front. Plant Sci.* 7, 1029. doi: 10.3389/fpls.2016.01029
- Joubert, W., Nance, J., Climer, S., Weighill, D., and Jacobson, D. (2017). Parallel accelerated custom correlation coefficient calculations for genomics applications. *Parallel Comput.* 75, 130–145. doi: 10.1016/j.parco.2018.03.009
- Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548
- Kim, J.-S., Mizoi, J., Kidokoro, S., Maruyama, K., Nakajima, J., Nakashima, K., et al. (2012). Arabidopsis growth-regulating factor7 functions as a transcriptional repressor of abscisic acid- and osmotic stress-responsive genes, including DREB2A. *Plant Cell* 24, 3393–3405. doi: 10.1105/tpc.112.100933
- Ko, J.-H., Jeon, H.-W., Kim, W.-C., Kim, J.-Y., and Han, K.-H. (2014). The MYB46/MYB83-mediated transcriptional regulatory programme is a gatekeeper of secondary wall biosynthesis. *Ann. Bot.* 114, 1099–1107. doi: 10.1093/aob/mcu126
- Kuijt, S. J. H., Greco, R., Agalou, A., Shao, J., t Hoen, C. C. J., Overnas, E., et al. (2014). Interaction between the growth-regulating factor and knotted1-like homeobox families of transcription factors[W]. *Plant Physiol.* 164, 1952–1966. doi: 10.1104/pp.113.222836
- Li, E., Bhargava, A., Qiang, W., Friedmann, M. C., Forneris, N., Savidge, R. A., et al. (2012a). The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in Arabidopsis and is functionally conserved in Populus. *New Phytol.* 194, 102–115. doi: 10.1111/j.1469-8137.2011.04016.x
- Li, H., Yang, Y., Wang, Z., Guo, X., Liu, F., Jiang, J., et al. (2016). BpMADS12 gene role in lignin biosynthesis of Betula platyphylla Suk by transcriptome analysis. *J. For. Res.* 27, 1111–1120. doi: 10.1007/s11676-016-0229-y
- Li, Y., Jin, F., Chao, Q., and Wang, B.-C. (2017). Proteomics analysis reveals the molecular mechanism underlying the transition from primary to secondary growth of poplar. *J. Plant Physiol.* 213, 1–15. doi: 10.1016/j.jplph.2017.02.009
- Li, Y., Tschaplinski, T. J., Engle, N. L., Hamilton, C. Y., Rodriguez, M., Liao, J. C., et al. (2012b). Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. *Biotechnol. Biofuels* 5, 2. doi: 10.1186/1754-6834-5-2
- Lin, Y.-C. J., Chen, H., Li, Q., Li, W., Wang, J. P., Shi, R., et al. (2017). Reciprocal cross-regulation of VND and SND multigene TF families for wood formation in *Populus trichocarpa*. *Proc. Natl. Acad. Sci.* 114, E9722–E9729. doi: 10.1073/pnas.1714422114
- Liu, Y., Wei, M., Hou, C., Lu, T., Liu, L., Wei, H., et al. (2017). Functional characterization of Populus PsnSHN2 in coordinated regulation of secondary wall components in tobacco. *Sci. Rep.* 7, 1–11. doi: 10.1038/s41598-017-00093-z
- Lofstead, J. F., Klasky, S., Schwan, K., Podhorski, N., and Jin, C. (2008). “Flexible IO and integration for scientific codes through the adaptable IO system (ADIOS),” in *Proceedings of the 6th international workshop on Challenges of large applications in distributed environments - CLADE '08*, vol. 15. (New York, New York, USA: ACM Press). doi: 10.1145/1383529.1383533
- Lohse, M., Nagel, A., Herter, T., May, P., Schroda, M., Zrenner, R., et al. (2014). Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. *Plant Cell Environ.* 37, 1250–1258. doi: 10.1111/pce.12231
- Lu, S., Li, Q., Wei, H., Chang, M.-J., Tunlaya-Anukit, S., Kim, H., et al. (2013). Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in *Populus trichocarpa*. *Proc. Natl. Acad. Sci.* 110, 10848–10853. doi: 10.1073/pnas.1308936110
- Macaya-Sanz, D., Chen, J. G., Kalluri, U. C., Muchero, W., Tschaplinski, T. J., Gunter, L. E., et al. (2017). Agronomic performance of Populus deltoides trees engineered for biofuel production. *Biotechnol. Biofuels* 10, 253. doi: 10.1186/s13068-017-0934-6
- Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics* 21, 3448–3449. doi: 10.1093/bioinformatics/bti551
- McLeay, R. C., and Bailey, T. L. (2010). Motif enrichment analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* 11, 165. doi: 10.1186/1471-2105-11-165
- Mele, G., Ori, N., Sato, Y., and Hake, S. (2003). The knotted1-like homeobox gene BREVIPEDICELLUS regulates cell differentiation by modulating metabolic pathways. *Genes Dev.* 17, 2088–2093. doi: 10.1101/gad.1120003
- Mizrachi, E., Verbeke, L., Christie, N., Fierro, A. C., Mansfield, S. D., Davis, M. F., et al. (2017). Network-based integration of systems genetics data reveals pathways associated with lignocellulosic biomass accumulation and processing. *Proc. Natl. Acad. Sci.* 114, 1195–1200. doi: 10.1073/pnas.1620119114
- Muchero, W., Guo, J., DiFazio, S. P., Chen, J. G., Ranjan, P., Slavov, G. T., et al. (2015). High-resolution genetic mapping of allelic variants associated with cell wall chemistry in Populus. *BMC Genomics* 16, 24. doi: 10.1186/s12864-015-1215-z
- Mukherjee, K., Brocchieri, L., and Bürglin, T. R. (2009). A comprehensive classification and evolutionary analysis of plant homeobox genes. *Mol. Biol. Evol.* 26, 2775–2794. doi: 10.1093/molbev/msp201
- Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N. A., and Ohtani, M. (2015). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* 6, 1–18. doi: 10.3389/fpls.2015.00288
- Ohtani, M., Nishikubo, N., Xu, B., Yamaguchi, M., Mitsuda, N., Goué, N., et al. (2011). A NAC domain protein family contributing to the regulation of wood formation in poplar. *Plant J.* 67, 499–512. doi: 10.1111/j.1365-313X.2011.04614.x
- Omidbakhshfard, M. A., Proost, S., Fujikura, U., and Mueller-Roeber, B. (2015). Growth-regulating factors (GRFs): a small transcription factor family with important functions in plant biology. *Mol. Plant* 8, 998–1010. doi: 10.1016/j.molp.2015.01.013
- Park, M. Y., Kang, J. Y., and Kim, S. Y. (2011). Overexpression of AtMYB52 confers ABA hypersensitivity and drought tolerance. *Mol. Cells* 31, 447–454. doi: 10.1007/s10059-011-0300-7
- Peterson, C. B., Bogomolov, M., Benjamini, Y., and Sabatti, C. (2016). Many phenotypes without many false discoveries: error controlling strategies for multitrait association studies. *Genet. Epidemiol.* 40, 45–56. doi: 10.1002/gepi.21942
- Pomiès, L., Decourteix, M., Franchel, J., Moulia, B., and Leblanc-Fournier, N. (2017). Poplar stem transcriptome is massively remodelled in response to single or repeated mechanical stimuli. *BMC Genomics* 18, 300. doi: 10.1186/s12864-017-3670-1
- Porth, I., Klápště, J., Skyba, O., Friedmann, M. C., Hannemann, J., Ehling, J., et al. (2013). Network analysis reveals the relationship among wood properties, gene expression levels and genotypes of natural *Populus trichocarpa* accessions. *New Phytol.* 200, 727–742. doi: 10.1111/nph.12419
- Punta, M., Coghill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., et al. (2011). The Pfam protein families database. *Nucleic Acids Res.* 40, D290–D301. doi: 10.1093/nar/gkr1065
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Puzey, J. R., Karger, A., Axtell, M., and Kramer, E. M. (2012). Deep annotation of *Populus trichocarpa* microRNAs from diverse tissue sets. *PLoS One* 7, e33034. doi: 10.1371/journal.pone.0033034
- 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
- Sannigrahi, P., Ragauskas, A. J., and Tuskan, G. A. (2010). Poplar as a feedstock for biofuels: a review of compositional characteristics Poulomi. *Biofuels Bioprod. Biorefining* 4, 209–226. doi: 10.1002/bbb.206
- Schrader, J. (2004). A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. *Plant Cell Online* 16, 2278–2292. doi: 10.1105/tpc.104.024190
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Shi, R., Wang, J. P., Lin, Y.-C., Li, Q., Sun, Y.-H., Chen, H., et al. (2017). Tissue and cell-type co-expression networks of transcription factors and wood component genes in *Populus trichocarpa*. *Planta* 245, 927–938. doi: 10.1007/s00425-016-2640-1
- Taylor-Teeple, 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. doi: 10.1038/nature14099
- Tschaplinski, T. J., Standaert, R. F., Engle, N. L., Martin, M. Z., Sangha, A. K., Parks, J. M., et al. (2012). Down-regulation of the caffeic acid O-methyltransferase gene in switchgrass reveals a novel monolignol analog. *Biotechnol. Biofuels* 5, 71. doi: 10.1186/1754-6834-5-71

- Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* (80-.). 313, 1596–1604. doi: 10.1126/science.1128691
- Van Dongen, S. (2008). Graph clustering *via* a discrete uncoupling process. *SIAM J. Matrix Anal. Appl.* 30, 121–141. doi: 10.1137/040608635
- Vining, K. J., Pomraning, K. R., Wilhelm, L. J., Priest, H. D., Pellegrini, M., Mockler, T. C., et al. (2012). Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics* 13, 27. doi: 10.1186/1471-2164-13-27
- Wang, J., Yan, D. W., Yuan, T. T., Gao, X., and Lu, Y. T. (2013). A gain-of-function mutation in IAA8 alters Arabidopsis floral organ development by change of jasmonic acid level. *Plant Mol. Biol.* 82, 71–83. doi: 10.1007/s11103-013-0039-y
- Wang, S., Li, E., Porth, I., Chen, J.-G., Mansfield, S. D., and Douglas, C. J. (2014). Regulation of secondary cell wall biosynthesis by poplar R2R3 MYB transcription factor PtrMYB152 in Arabidopsis. *Sci. Rep.* 4, 5054. doi: 10.1038/srep05054
- Weighill, D., Jones, P., Shah, M., Ranjan, P., Muchero, W., Schmutz, J., et al. (2018). Pleiotropic and epistatic network-based discovery: integrated networks for target gene discovery. *Front. Energy Res.* 6, 30. doi: 10.3389/fenrg.2018.00030
- Xie, M., Muchero, W., Bryan, A. C., Yee, K. L., Guo, H., Zhang, J., et al. (2018a). A 5-enolpyruvylshikimate 3-phosphate synthase functions as a transcriptional repressor in *Populus*. *Plant Cell* 30, 1645–1660. doi: 10.1105/tpc.18.00168
- Xie, M., Zhang, J., Tschaplinski, T. J., Tuskan, G. A., Chen, J.-G., and Muchero, W. (2018b). Regulation of lignin biosynthesis and its role in growth-defense tradeoffs. *Front. Plant Sci.* 9, 1427. doi: 10.3389/fpls.2018.01427
- Yang, X., Ye, C.-Y., Bisaria, A., Tuskan, G. A., and Kalluri, U. C. (2011). Identification of candidate genes in Arabidopsis and *Populus* cell wall biosynthesis using text-mining, co-expression network analysis and comparative genomics. *Plant Sci.* 181, 675–687. doi: 10.1016/j.plantsci.2011.01.020
- Yu, L., Chen, H., Sun, J., and Li, L. (2013). PtrKOR1 is required for secondary cell wall cellulose biosynthesis in *Populus*. *Tree Physiol.* 34, 1289–1300. doi: 10.1093/treephys/tpu020
- Zhan, X., Hu, Y., Li, B., Abecasis, G. R., and Liu, D. J. (2016). RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics* 32, 1423–1426. doi: 10.1093/bioinformatics/btw079
- Zhang, J., Yang, Y., Zheng, K., Xie, M., Feng, K., Jawdy, S. S., et al. (2018). Genome-wide association studies and expression-based quantitative trait loci analyses reveal roles of HCT2 in caffeoylquinic acid biosynthesis and its regulation by defense-responsive transcription factors in *Populus*. *New Phytol.* 220, 502–516. doi: 10.1111/nph.15297
- 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 Online* 20, 2763–2782. doi: 10.1105/tpc.108.061325
- Zhong, R., and Ye, Z.-H. (2014). Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant Cell Physiol.* 56, 195–214. doi: 10.1093/pcp/pcu140

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Furches, Kainer, Weighill, Large, Jones, Walker, Romero, Gazolla, Joubert, Shah, Streich, Ranjan, Schmutz, Sreedasyam, Macaya-Sanz, Zhao, Martin, Rao, Dixon, DiFazio, Tschaplinski, Chen, Tuskan and Jacobson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Secondary Wall Regulating NACs Differentially Bind at the Promoter at a *CELLULOSE SYNTHASE A4* Cis-eQTL

Jennifer R. Olins^{1†}, Li Lin^{1†}, Scott J. Lee^{1,2}, Gina M. Trabucco^{1,3}, Kirk J.-M. MacKinnon^{1,3} and Samuel P. Hazen^{1*}

¹ Biology Department, University of Massachusetts, Amherst, MA, United States, ² Plant Biology Graduate Program, University of Massachusetts, Amherst, MA, United States, ³ Molecular and Cellular Biology Graduate Program, University of Massachusetts, Amherst, MA, United States

OPEN ACCESS

Edited by:

Charles T. Anderson,
Pennsylvania State University,
United States

Reviewed by:

Nobutaka Mitsuda,
National Institute of Advanced
Industrial Science and Technology
(AIST), Japan
Kanwarpal Singh Dhugga,
Consultative Group on International
Agricultural Research (CGIAR),
United States

*Correspondence:

Samuel P. Hazen
hazen@bio.umass.edu

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 11 October 2018

Accepted: 06 December 2018

Published: 21 December 2018

Citation:

Olins JR, Lin L, Lee SJ,
Trabucco GM, MacKinnon KJ-M and
Hazen SP (2018) Secondary Wall
Regulating NACs Differentially Bind
at the Promoter at a *CELLULOSE*
SYNTHASE A4 Cis-eQTL.
Front. Plant Sci. 9:1895.
doi: 10.3389/fpls.2018.01895

Arabidopsis thaliana *CELLULOSE SYNTHASE A4/7/8* (*CESA4/7/8*) are three non-redundant subunits of the secondary cell wall cellulose synthase complex. Transcript abundance of these genes can vary among genotypes and expression quantitative trait loci (eQTL) were identified in a recombinant population of the accessions Bay-0 and Shahdara. Genetic mapping and analysis of the transcript levels of *CESAs* between two distinct near isogenic lines (NILs) confirmed a change in *CESA4* expression that segregates within that interval. We sequenced the promoters and identified 16 polymorphisms differentiating *CESA4^{Sha}* and *CESA4^{Bay}*. In order to determine which of these SNPs could be responsible for this eQTL, we screened for transcription factor protein affinity with promoter fragments of *CESA4^{Bay}*, *CESA4^{Sha}*, and the reference genome *CESA4^{Col}*. The wall thickening activator proteins NAC SECONDARY WALL THICKENING PROMOTING FACTOR2 (NST2) and NST3 exhibited a decrease in binding with the *CESA4^{Sha}* promoter with a tracheary element-regulating *cis*-element (TERE) polymorphism. While NILs harboring the TERE polymorphisms exhibited significantly different *CESA4* expression, cellulose crystallinity and cell wall thickness were indistinguishable. These results suggest that the TERE polymorphism resulted in differential transcription factor binding and *CESA4* expression; yet *A. thaliana* is able to tolerate this transcriptional variability without compromising the structural elements of the plant, providing insight into the elasticity of gene regulation as it pertains to cell wall biosynthesis and regulation. We also explored available DNA affinity purification sequencing data to resolve a core binding site, C(G/T)TNNNNNNA(A/C)G, for secondary wall NACs referred to as the VNS element.

Keywords: *CELLULOSE SYNTHASE A4*, NAC transcription factor, expression QTL, VNS element, tracheary element-regulating *cis*-element

INTRODUCTION

While a primary cell wall surrounds all plant cells, a secondary cell wall is also found in xylem cells responsible for water transportation, structural fibers, and cells that serve as an outside barrier to the external environment. These thick and relatively inflexible walls are composed of a complex of cellulose, hemicelluloses, and the polyphenolic polymer lignin. Cellulose is the most abundant

fraction in the majority of tissues and exists as long unbranched β -1,4-linked glucan chains. Cellulose chains coalesce in parallel to form a single microfibril via hydrogen bonding and van der Waals forces. Depending on their density and nature of the commingling polymers and linkages, microfibrils can contribute to a matrix that ranges from fairly elastic to extremely rigid. Aside from the obvious functional virtues to plants, secondary cell wall rich in cellulose are a valuable feedstock for the pulp and paper and the biofuel industries (Carroll and Somerville, 2009). Understanding the regulation of secondary cell wall composition, especially the effects of natural genetic variation, will facilitate enhanced gene modification and plant breeding for more efficient biomass production.

Cellulose is synthesized at the plasma membrane by multiple Cellulose Synthase A (CESA) proteins organized into rosette shaped complexes (Mueller and Brown, 1980). The rosette is composed of at least 18 CESA subunits organized into six globules, termed cellulose synthase complexes (CSCs) (Jarvis, 2013). There are ten proteins in the *A. thaliana* CESA family, while direct polymerization activity remains to be documented, all but CESA10 have been shown to be associated with cellulose biosynthesis (Guerriero et al., 2010). Mutations in a number of non-CESA genes, including members of the COBRA family, also exhibit polymerization defects (Ching et al., 2006; Sindhu et al., 2007; Zhang et al., 2010; Kotake et al., 2011). All members of the CESA superfamily, which includes the CESAs and seven CESA-like (CSL) families in eudicots, are integral membrane proteins; CESAs have two transmembrane domains in the N-terminus and six at the C-terminus, while CSLs are more variable (Somerville, 2006; Nixon et al., 2016; Little et al., 2018). Additionally, all ten CESA family members contain a LIM-like Zn-binding domain/RING finger, which is known to be involved in protein-protein interactions (Somerville, 2006). These attributes are in line with the finding that CESAs form a complex embedded in the cell membrane (Arioli et al., 1998; Fagard et al., 2000; Scheible et al., 2001; Taylor et al., 2003; Timmers et al., 2009). Across plant species and under most circumstances, three distinct non-redundant CESAs are required for optimal production of cellulose, and in plants, CESA1, 3, and one or more of 2/5/6/9 are involved in primary cell wall development, while CESA4, 7, and 8 are responsible for secondary cell wall biosynthesis (Kumar and Turner, 2015).

Early studies of knock-out mutations in CESA genes in *A. thaliana* first revealed the non-redundant nature of the secondary CESAs, as mutants harboring null mutations in any of the three secondary CESAs exhibited irregular cell walls and weaker stems (Taylor et al., 2000). To further understand the intricacies of the CESA genes, subsequent studies investigated the effects of more subtle changes in CESA expression. Virus-induced gene silencing (VIGS) of CESA genes in *Nicotiana benthamiana* led to dwarfed phenotypes and reduced cellulose content (Burton et al., 2000). A similar VIGS study in flax again resulted in plants of shorter stature that only displayed a slight reduction in sugar content (Chantreau et al., 2015). Comparable phenotypes were also observed in *Brachypodium distachyon* with compromised expression of CESA genes by use of artificial microRNAs, and more recently, it was shown that

both CESA knock-down and overexpression induced similar biomass-compromised phenotypes in *Panicum virgatum* L. (Handakumbura et al., 2013; Mazarei et al., 2018).

Non-redundancy within secondary cell wall CESAs suggests the sensitive nature of cellulose synthesis and cell wall growth. In conjunction, the 18 CESAs in each rosette likely synthesize an individual β -1-4-glucan cellulose chain that coalesce in parallel to form a single microfibril via hydrogen bonding and van der Waals forces, though cases of 24-chain fibrils have been reported (Thomas et al., 2013). Variation among species in the orientation and size of the CSCs correlates with the size and thickness of microfibrils. Though the microfibrils have an organized, crystalline structure, the inner chains in the bundle tend to exhibit higher degrees of crystallinity, while sheath fibers are more disordered (Nixon et al., 2016). The degree of crystallinity within the microfibril has been found to be inversely correlated with the rate of polymerization. Moreover, disruptions to CESA domains that either encourage microfibril aggregation or membrane complex subunit associations show an overall reduction in crystallinity (Harris et al., 2012).

The expression of CESA genes and many others that encode additional cell wall components are highly co-regulated (Brown et al., 2005; Persson et al., 2005). A current model for transcriptional regulation of cell wall genes is a series of feed forward loops (Taylor-Teeple et al., 2015; Zhang et al., 2018). Transcription factors most commonly bind to the promoters of secondary cell wall genes associated with different classes of wall components as well as the promoters of other wall regulating transcription factors. Three phylogenetically distinct groups of NACs play the role of direct and indirect activators of cell wall gene expression: VNDs (VASCULAR-RELATED NAC-DOMAIN), NSTs (NAC SECONDARY WALL THICKENING), and SNDs (SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN) (Yao et al., 2012; Nakano et al., 2015). There are seven *A. thaliana* VNDs and all can activate xylem vessel cell differentiation (Kubo et al., 2005). Similarly, the NSTs positively activate cell wall thickening, but in fiber cells rather than vasculature (Mitsuda et al., 2007; Zhong et al., 2010). In *A. thaliana*, these include NST1, NST2, and NST3/SND1. The third class, consists of two *A. thaliana* genes, SND2 and SND3. Similar to the NSTs, SND2 and its orthologs in poplar, rice, and switchgrass can also directly and indirectly activate secondary cell wall thickening (Zhong et al., 2008; Hussey et al., 2011; Rao and Dixon, 2018; Ye et al., 2018). While these three classes of NAC transcription factors are distinct in their amino acid sequence as well as their role in wall development, they are reported to bind similar DNA sequences. The consensus sequence CTTNAAAGCNA, named the tracheary element-regulating cis-element (TERE), was initially identified in the promoters of genes associated with secondary cell wall formation and programmed cell death of vasculature (Pyo et al., 2007). Subsequently, VNDs, NSTs, and SNDs have been shown to bind directly and specifically to the TERE motif and a similar target, the secondary wall NAC-binding element [SNBE, (T/A)NN(C/T)(T/C/G)TNNNNNNNA(A/C)GN(A/C/T)(A/T)] (Pyo et al., 2007; Ohashi-Ito et al., 2010; Zhong et al., 2010;

Taylor-Teeple et al., 2015; Ye et al., 2018). Thus, these NACs have both overlapping and distinct roles in the regulation of cell differentiation and secondary cell wall biosynthesis.

Naturally occurring genetic diversity offers a rich source from which to identify promising genes and variants for bioenergy crop breeding. Identification of markers associated with advantageous biomass accumulation traits has been carried out in chickpea, maize, sorghum, and other species, and such studies frequently pinpoint loci of *CESA* genes as potential targets (Shiringani and Friedt, 2011; Barrière et al., 2012; Zhao et al., 2013; Kujur et al., 2015). In addition, analysis of variable expression profiles among accessions has also been highlighted as a plant breeding tool, as exhibited by studies in sugarcane, loblolly pine, and shrub willow (Serapiglia et al., 2012; Palle et al., 2013; Kasirajan et al., 2018). Such studies contribute to the identification of candidate biomass accumulation genes, and also present a method of early selection in the breeding process. In general, transcriptome analysis suggests upregulation of *CESA* genes correlates with increased biomass (Serapiglia et al., 2012; Kujur et al., 2015; Kasirajan et al., 2018). In the present study, we investigated the causes and consequences of an eQTL at the *CESA4 cis*-eQTL in the *A. thaliana* Bayreuth (Bay-0) and Shakhara (Sha) recombinant inbred line (RIL) population (Loudet et al., 2002; West et al., 2007). We aimed to take advantage of the natural variation of *A. thaliana* to better understand the genotypic and phenotypic diversity for biofuel-relevant traits, specifically by studying the regulation of and variation in crystalline cellulose content within the plant cell wall.

MATERIALS AND METHODS

Yeast One-Hybrid Protein-DNA Interaction Assays

Yeast one-hybrid protein-DNA interaction assays were conducted as previously described (Taylor-Teeple et al., 2015). The transcription factors were transformed into each yeast strain and the β -galactosidase activity was determined as previously described (Pruneda-Paz et al., 2009). Positive interactions were visually identified as incidence of yellow caused by the presence of ortho-nitrophenyl cleavage from colorless ortho-nitrophenyl- β -D-galactoside by β -galactosidase. The DNA bait strains were, similarly, tested for self-activation prior to screening, under selection but in the absence of any prey vector. A total of 34 *E. coli* strains harboring different *A. thaliana* transcription factors (Supplementary Table 1) were arrayed in 96-well plates and plasmids were prepared. proCESA4^{Col} (539 bp), proCESA4^{Bay} (539 bp), and proCESA4^{Sha} (540 bp) were cloned and recombined with reporter genes. Promoter sequences and primers used are described in Supplementary Table 2. Nine of overlapping fragments of CESA4^{Col} were independently cloned according to Pruneda-Paz et al. (2009). The oligonucleotides used to amplify promoter fragments are described in Supplementary Table 2. The screen was replicated in full to confirm the results and each clone was sequenced to re-confirm identity.

Electrophoretic Mobility Shift Assays

To express recombinant NST2 or SND1 protein, coding sequences were cloned and fused to glutathione S-transferase tag in the pDONR211 vector and then transferred into pDEST15 (Invitrogen). *E. coli* strain BL21-AI (Invitrogen) transformed with pDEST15-GST:NST2 were grown in liquid media to an OD₆₀₀ of 0.4, treated with 0.2% L-arabinose to induce expression overnight and harvested by centrifugation the following day. Cells were treated with 1 mg/mL lysozyme on ice for 30 min in minimal volume of 1X PBS buffer and lysed by sonication. Cell lysates were clarified by centrifugation and incubated with 100 μ L of glutathione sepharose beads (GE Healthcare, Pittsburgh, PA, United States) for 30 min at 4°C with rotation. The beads were transferred to a column, washed with 10 volumes of 1X PBS. Protein was eluted in 100 mM Tris-HCl pH8.0, 100 mM NaCl and 3 mg/mL glutathione buffer and purified protein was re-suspended in 50% glycerol and stored at -80°C.

Three overlapping probes were generated for CESA4^{Bay}-C, CESA4^{Bay}-D, CESA4^{Bay}-E, CESA4^{Sha}-C, CESA4^{Sha}-D, CESA4^{Sha}-E promoter fragments using the same oligonucleotides described in Supplementary Table 2. Reactions were carried out in binding buffer (10 mM Tris, pH7.5, 50 mM KCl, 1 mM DTT, 2.5% glycerol, 5 mM MgCl₂, 0.1% IGEPAL CA-630, and 0.05 μ g/ μ L calf thymus DNA). Following the addition of 150 ng of protein from the GST purification eluate, reactions were incubated at room temperature for 30 min. Protein-DNA complexes were separated from the free DNA on 1% agarose/1X TAE gels at 4°C. The agarose gels were stained with ethidium bromide and bands visualized under UV light.

Characterization of Near Isogenic Lines

To develop NILs, RILs maintaining heterozygosity at the *CESA4* locus in the F6 generation were identified (Loudet et al., 2002). One plant per RIL carrying a heterozygous *CESA4* locus was identified via genotyping and selfed to obtain the F7 seeds. Within the F7 plants, pairs of lines were identified that were either homozygous for the CESA4^{Bay} and CESA4^{Sha} alleles to generate the NILs used in this study. The nearly isogenic lines analyzed in this study were developed from RIL93 and RIL350 segregating for the *CESA4* region. Individual HIF plants were genotyped for the Bay-0 or Sha *CESA4* allele by PCR of a 543 bp of *CESA4* promoter with primers described in Supplementary Table 2. PCR products were subjected to restriction enzyme digestion with BsrDI (New England BioLabs, Ipswich, MA, United States) at 65°C for 2 h. The Sha allele includes a BsrDI cut site (CGTTAC|NN) resulting in 401 and 142 bp fragments. The 543 bp Bay-0 allele contains no BsrDI restriction sites and remains undigested. Transcript abundance of *CESA4* in the near isogenic lines (NILs) was quantified in 5 cm stems. Stem tissue was frozen in liquid nitrogen and pulverized with metal beads in a Retsch (Haan, Germany) Mixer Mill MM400. RNA extraction and cDNA synthesis was conducted as described above. Primers for *CESA4* real-time PCR are described in Supplementary Table 2.

Quantification of Crystalline Cellulose

To quantify and compare levels of crystalline cellulose, the Updegraff assay as adapted and described by Kumar and Turner was used (Updegraff, 1969; Kumar and Turner, 2015). Briefly, alcohol-insoluble residue (AIR) samples were first prepared from senesced stem tissue; samples were weighed at this stage for later calculations. Then, 3 mL of an acetic/nitric acid solution (8:1:2, acetic acid: nitric acid: water) was added to AIR samples and incubated in a boiling water bath for 30 min. This step removes hemicellulose and lignin while leaving cellulose microfibrils intact. The remaining cellulosic material was then swelled in 67% sulfuric acid in a boiling water bath for 5 min to disorganize the polymers, and the subsequent monomers were finally analyzed with a sulfuric acid/anthrone colorimetric assay. Absorbance was measured at a wavelength of 620 nm with a *SpectraMax M5* and corrected to a known glucose standard to calculate the percent cell wall composition of cellulose.

Quantification of Stem Thickness

To investigate the histological effects of expression variation on the cell wall, 100 μm cross sections were sliced using a *Leica Biosystems* VT1000S Vibrating-blade microtome. Sections were incubated on the bench top for 2 min in a 2% w/v phloroglucinol/ethanol solution (Fisher Scientific, Waltham, MA), mounted in a 1:1 concentrated hydrochloric acid: water solution, and immediately imaged using a *Nikon Eclipse E200* microscope and a *PixLink* scope camera. For each stem, three sections were imaged, and five cells from three different portions of each stem were measured to get an average of 45 measurements per biological sample.

Polarized Light Microscopy

To qualitatively analyze cellulose crystallinity under polarized light, internode segments were first cut on a vibratome into 100 μm sections, fixed in 2% glutaraldehyde and then embedded in Epon/Araldite (Sigma-Aldrich, St. Louis, MO, United States) before slicing into 0.5 μm sections on an Reichert-Jung, Ultracut E microtome (Vienna, Austria). Polarized light microscopy takes advantage of variation in the passage of light through crystalline structures to uncover discrete differences in crystallinity configurations. Specifically, the LC-PolScope (CRI, Cambridge, MA, United States) employed for this assay uses polarized light to measure birefringent retardance and the intensity of the images generated directly correlates to the retardance value, thus qualitative differences in crystallinity can be observed. Several stem samples from each HIF were imaged and analyzed.

DNA Affinity Purification Sequence Analysis

In order to determine a consensus VND, NST, and SND (VNS) protein binding motif, the MEME.txt motif files from O'Malley et al. (2016) were visually aligned based on nucleotide similarity and trimmed to a length of 13 bases. An average for each nucleotide, at each motif position was calculated using a bash script that referenced the aligned and trimmed motif files. The

consensus matrix was then loaded into R and the Bioconductor package SeqLogo was used to generate the motif logo. A bash shell script counted each of the VNS motif variant across the *A. thaliana* TAIR 10 sequence assembly and for each DAP-seq set of binding sites. Sum of the counts determined the total number of VNS motif sites and percentages were calculated as the proportion of each motif variant relative to the sum and multiplied by 100. The process was repeated for the DAP-seq binding site data using the reference genome file as a guide to extract DAP-seq peak fasta sequences. ThenarrowPeak DAP-seq peak data files were then searched iteratively through each of the peak binding sites. The percentage shown for DAP-seq data represents the mean percentage across all binding site files.

Statistical Analysis

Correlation coefficients and tests of significance were calculated using Pearson's correlation tests for all *CESA* genes pairwise using replicate gene expression data for Bay-0 and Sha parental lines and RILs using the rcorr function from the Hmisc library in R 3.2.5. Two tailed Student's *t*-tests were carried out in Excel to assess gene expression, cell thickness, and cellulose content data sets. BoxPlots were generated with the BoxPlotR Web Tool (Spitzer et al., 2014).

Accession Numbers

CESA1 (At4g32410), CESA2 (At4g39350), CESA3 (At5g05170), CESA4 (At5g44030), CESA5 (At5g09870), CESA6 (At5g64740), CESA7 (At5g17420), CESA8 (At4g18780), CESA9 (At2g21770), CESA10 (At2g25540), NST1 (At2g46770), NST2 (At3g61910), NST3/SND1 (At1g32770), PP2AA3 (At1g13320), TIP41 (At4g34270).

RESULTS

Natural Variation and Co-expression of the Secondary Wall CESAs

We explored the observation that *CESA* gene expression varied among different accessions of *A. thaliana*. West et al. (2007) measured the expression of 22,746 genes in replicated samples of the Bay-0/Sha RIL population. Above ground tissue of 211 short-day grown plants was assayed after 6 weeks of growth using the Affymetrix ATH1 GeneChip microarray. There was a continuous range of values and normal distributions were observed for the ten *CESA* genes known to play a role in the biosynthesis of cellulose in secondary cell walls (Figure 1A and Supplementary Figure 1). The midparent values (i.e., the values halfway between the two parents) and the median of the RILs for *CESA4*, *CESA7*, and *CESA8* (Figure 1A) and the seven other *CESA* genes (Supplementary Figure 1) were very similar.

The *CESA* genes have been shown to be highly co-regulated within their functional classes. Three of the primary CESAs, *CESA1*, 3, and 6, had the most similar gene expression to each other in a meta analysis of gene expression and the same is true among the secondary CESAs: *CESA4*, 7, and 8 (Persson et al., 2005). Candidate genes for the transcriptional regulation

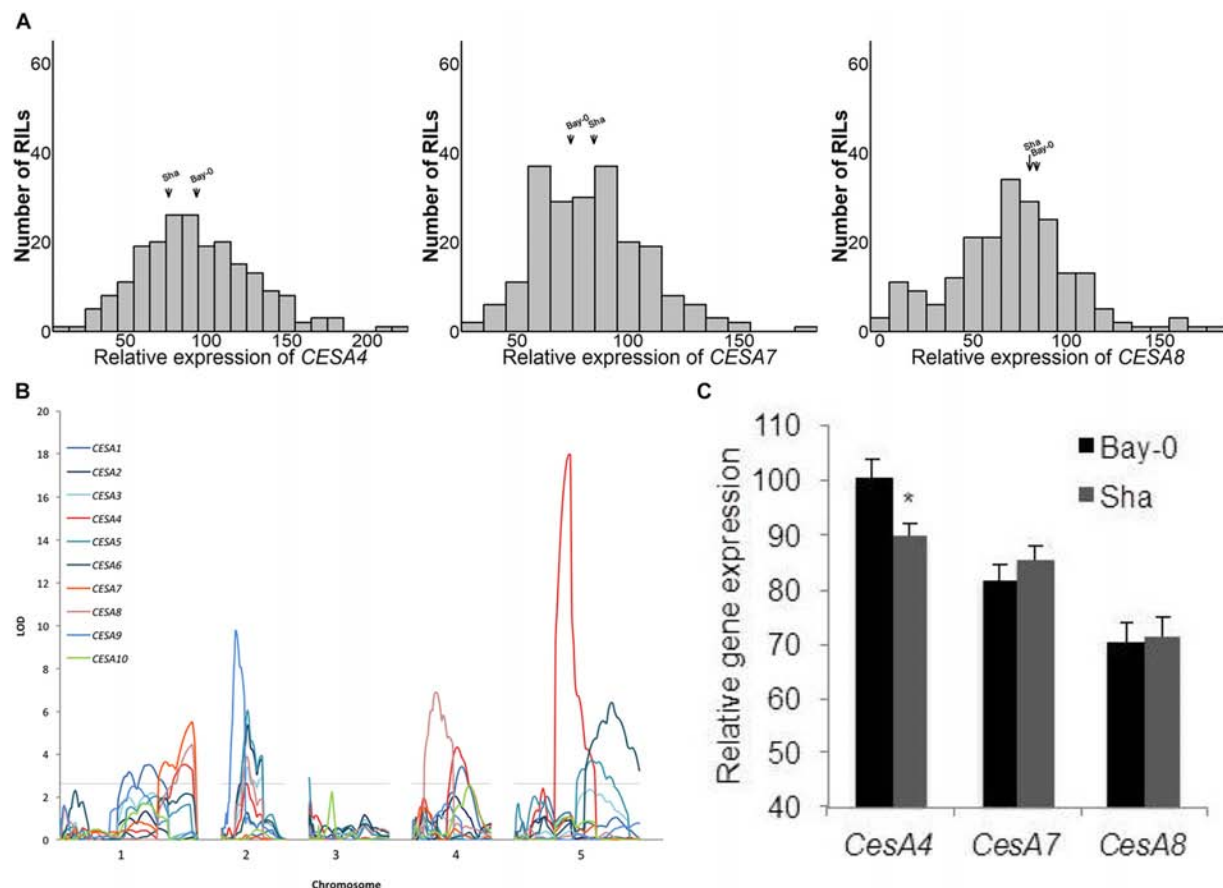


FIGURE 1 | (A) Histograms of relative expression of *CESA4/7/8* within the Bay-0 × Sha recombinant inbred line population (RIL). **(B)** Genetic mapping identified a *CESA4* expression quantitative trait locus coincident with the physical position of the *CESA4* locus on chromosome five. **(C)** Relative expression of *CESA4/7/8* from RIL parents carrying Bay-0 or Sha allele. Lines carrying Bay-0 allele had significantly higher expression of *CesA4*, while no differences were observed between RILs for the expression of *CESA7* or *CESA8*. * $P < 0.01$.

TABLE 1 | Pairwise Pearson's correlation coefficients of means of relative gene expression of the Bay-0/Sha recombinant inbred line population.

Transcript	<i>CESA4</i>	<i>CESA7</i>	<i>CESA8</i>	<i>CESA1</i>	<i>CESA3</i>	<i>CESA6</i>	<i>CESA9</i>	<i>CESA2</i>	<i>CESA5</i>
<i>CESA4</i>	–	–	–	–	–	–	–	–	–
<i>CESA7</i>	0.35	–	–	–	–	–	–	–	–
<i>CESA8</i>	0.35	0.32	–	–	–	–	–	–	–
<i>CESA1</i>	0.46	0.22	0.23	–	–	–	–	–	–
<i>CESA3</i>	0.44	0.22	0.23	0.80	–	–	–	–	–
<i>CESA6</i>	0.41	0.22	0.26	0.68	0.76	–	–	–	–
<i>CESA9</i>	0.14	0.19	0.02	0.12	0.2	0.14	–	–	–
<i>CESA2</i>	0.35	ns	0.28	0.57	0.73	0.50	0.17	–	–
<i>CESA5</i>	0.35	0.23	0.24	0.70	0.75	0.58	0.14	0.61	–
<i>CESA10</i>	ns	0.19	ns	ns	ns	ns	0.11	ns	ns

All values are significant at $P < 0.01$ unless identified as not significant (ns).

of *CESA* gene expression have been successfully identified using the same type of analysis in various species (Brown et al., 2005; Persson et al., 2005; Yamaguchi et al., 2010; Ruprecht et al., 2011; Handakumbura et al., 2018). Among the Bay-0/Sha RILs, the expression of *CESA4*, 7, and 8 were significantly correlated as were the primary wall *CESAs* (Table 1).

Secondary Wall *CESA* eQTL

Among the genes whose expression was measured using the ATH1 microarray, 69% were associated with an eQTL and each transcript was mapped to an average of 2.34 loci (West et al., 2007). To pinpoint the cause of *CESA* gene expression variation among RILs, we searched for eQTL for these genes (Figure 1B).

TABLE 2 | Summary of CESA eQTL.

Chromosome	Map position (cM)	Peak (bp)	Gene	eQTL type
1	69.5–74.5	24,374,008	<i>CESA1</i>	<i>Trans</i>
1	76.7–87.52	27,088,847	<i>CESA1</i>	<i>Trans</i>
1	85.5–100.9	29,016,886	<i>CESA7</i>	<i>Trans</i>
1	92.4–99.9	29,016,886	<i>CESA4</i>	<i>Trans</i>
1	92.4–100.9	29,016,886	<i>CESA8</i>	<i>Trans</i>
2	24.5–41.4	11,095,452	<i>CESA9</i>	<i>Cis</i>
2	38.5–55.9	13,192,607	<i>CESA2</i>	<i>Trans</i>
2	38.5–55.9	13,192,607	<i>CESA5</i>	<i>Trans</i>
2	39.5–47.9	13,192,607	<i>CESA8</i>	<i>Trans</i>
2	42.4–47.9	13,192,607	<i>CESA3</i>	<i>Trans</i>
2	43.4	13,192,607	<i>CESA4</i>	<i>Trans</i>
3	0.0–4.2	786,303	<i>CESA5</i>	<i>Trans</i>
4	28.7–54.6	11,524,362	<i>CESA8</i>	<i>Cis</i>
4	52.6–64.9	15,790,523	<i>CESA4</i>	<i>Trans</i>
4	57.4–63.9	17,261,718	<i>CESA1</i>	<i>Cis</i>
5	56.5–78.5	18,662,765	<i>CESA4</i>	<i>Cis</i>
5	68.7–86.5	23,879,425	<i>CESA5</i>	<i>Trans</i>
5	73.5–96.9	25,924,795	<i>CESA6</i>	<i>Cis</i>

Position was calculated using the global 5% significance threshold 2.616 (West et al., 2007).

An eQTL for each of the *CESAs* was mapped to a position of the genome outside of those genes, i.e., *trans*-eQTL (Figure 1B and Table 2). A *trans*-eQTL common to all three secondary wall *CESA* genes was found near the bottom of chromosome 1. Overlapping *trans*-eQTLs were identified for *CESA4* and 8 on chromosomes 2 and 4. An eQTL unique to *CESA4* with the greatest LOD score of 17.95 was found on chromosome 5, coincident with the *CESA4* genomic locus. *CESA4* transcript abundance in RILs varied significantly depending on the presence of the Bay-0 *CESA4* promoter allele (*CESA4^{Bay}*) or the Sha *CESA4* promoter allele (*CESA4^{Sha}*) (Figure 1C). No differences were observed for the expression of other secondary wall *CESA* genes, *CESA7* or *CESA8*, between RILs with either *CESA4^{Bay}* or *CESA4^{Sha}* (Figure 1C).

A SNP in the CESA4 Promoter at the CESA4 Cis-eQTL Induces Differential Binding of Cell Wall Thickening NAC Regulators

Possible mechanisms for such a *cis*-eQTL include functional polymorphisms in the promoter of the gene in question, and we hypothesized that this change may disrupt interaction of SND1 or NST2, which we previously identified to interact with the *CESA4* promoter using yeast one-hybrid (Taylor-Teeple et al., 2015). As such, we sequenced the promoters of Bay-0, Sha, and Col-0 and identified 16 single nucleotide polymorphisms (SNPs) between *CESA4^{Sha}* and *CESA4^{Col}* and only two SNPs between *CESA4^{Bay}* and *CESA4^{Col}* (Figure 2). To specify which of the 16 SNPs were likely responsible for the *CESA4 cis*-eQTL, we screened nine overlapping fragments (A–I) of *CESA4^{Bay}* and *CESA4^{Sha}* promoters by yeast one-hybrid with *CESA4^{Col}* promoter interacting proteins (Figures 2, 3A). In this assay,

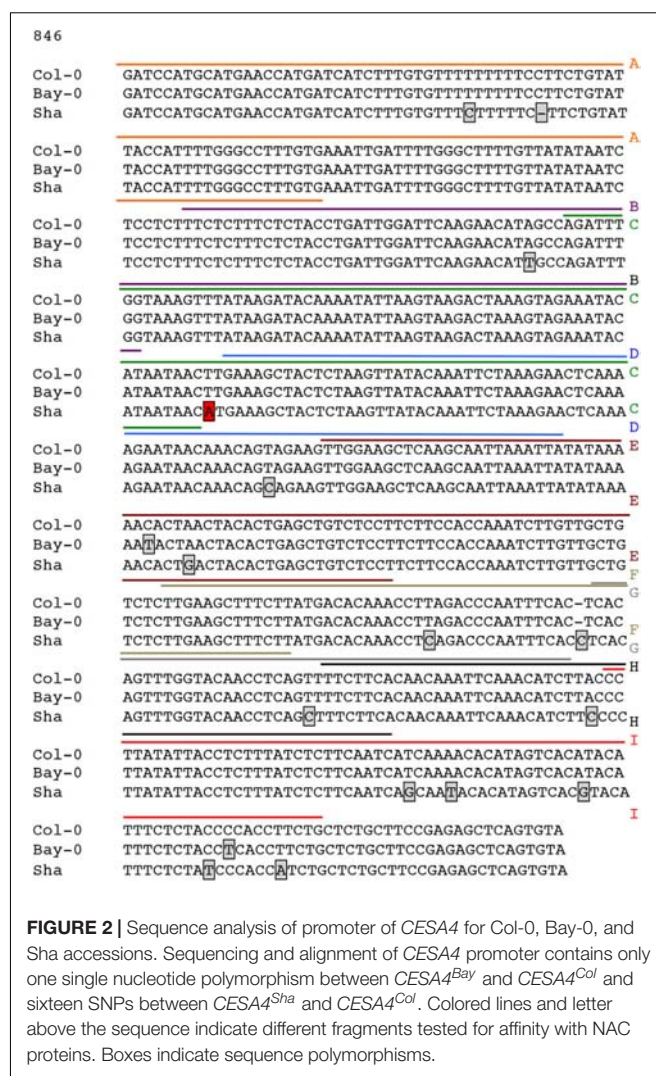
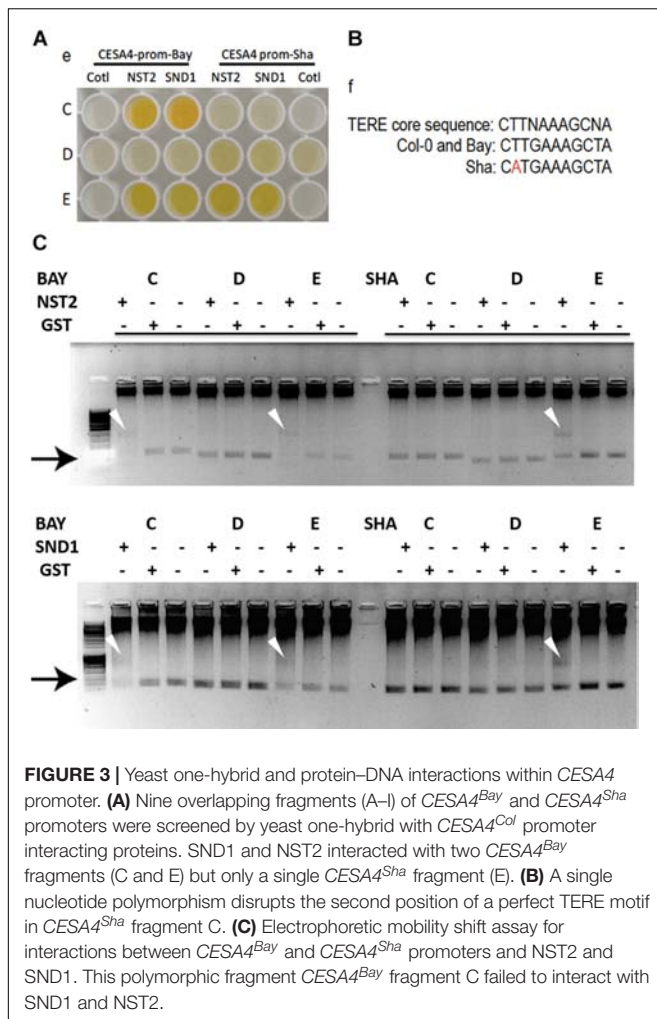


FIGURE 2 | Sequence analysis of promoter of *CESA4* for Col-0, Bay-0, and Sha accessions. Sequencing and alignment of *CESA4* promoter contains only one single nucleotide polymorphism between *CESA4^{Bay}* and *CESA4^{Col}* and sixteen SNPs between *CESA4^{Sha}* and *CESA4^{Col}*. Colored lines and letter above the sequence indicate different fragments tested for affinity with NAC proteins. Boxes indicate sequence polymorphisms.

transcription factor proteins were fused to the Gal4 activation domain and each protein is tested for an interaction with the *CESA4* promoter fragments immediately upstream of the *lacZ* reporter. A positive interaction can result in the cleavage from colorless ortho-nitrophenyl-β-D-galactoside by β-galactosidase resulting in a yellow color. The well-characterized wall thickening regulators SND1 and NST2 interacted with two *CESA4^{Bay}* fragments (C and E) but only a single *CESA4^{Sha}* fragment (E). One SNP in this region disrupts the second position of a perfect *TERE* motif in *CESA4^{Sha}* fragment C (Figure 3B). This polymorphic fragment failed to interact with SND1 and NST2 (Figure 3A). This suggests the expression differences between *CESA4^{Bay}* and *CESA4^{Sha}* could have occurred as a consequence of differential binding of NST2 or SND1 proteins to the *TERE* motif in fragment C.

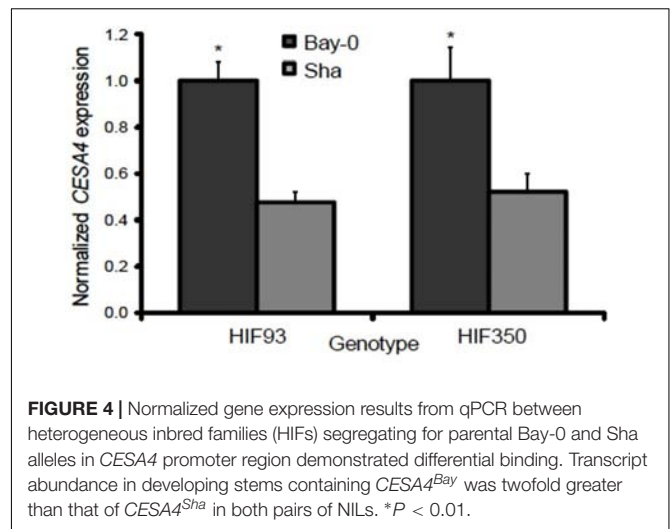
To further explore the possibility of this regulatory mechanism, EMSA was performed to confirm the differential protein-DNA with probes corresponding to *CESA4* promoter fragments C, D, and E in the presence or absence of extracts of *Escherichia coli* expressing GST-NST2 and GST-SND1



(Figure 3C). As anticipated, differences in mobility were observed with the *TERE* motif-containing fragment C of *CESA4*^{Bay} but not with the corresponding fragment of *CESA4*^{Sha}, confirming our yeast one-hybrid observations. Also consistent with our yeast one-hybrid results, fragment D from both accessions did not produce a DNA species with retarded mobility, but the *TERE* motif-containing fragment E did in both cases. Bacterial extracts harboring the empty GST vector did not produce any comparable shifted species. Taken together, these data suggest that the *CESA4*^{Sha} *TERE* motif polymorphism may result in diminished binding by SND1 and NST2 proteins.

CESA4 Is Differentially Expressed in *CESA4*^{Bay} and *CESA4*^{Sha} Near Isogenic Lines

Resolving the effects of the *CESA4*^{Bay} and *CESA4*^{Sha} *cis*-regulatory region is confounded by the entirety of the sequence variation between Bay-0 and Sha, which are maintained in different combinations among the RILs. To isolate the influence of the *CESA4* locus, we identified RILs with residual heterozygosity at *CESA4* and tested derived NILs. We isolated the

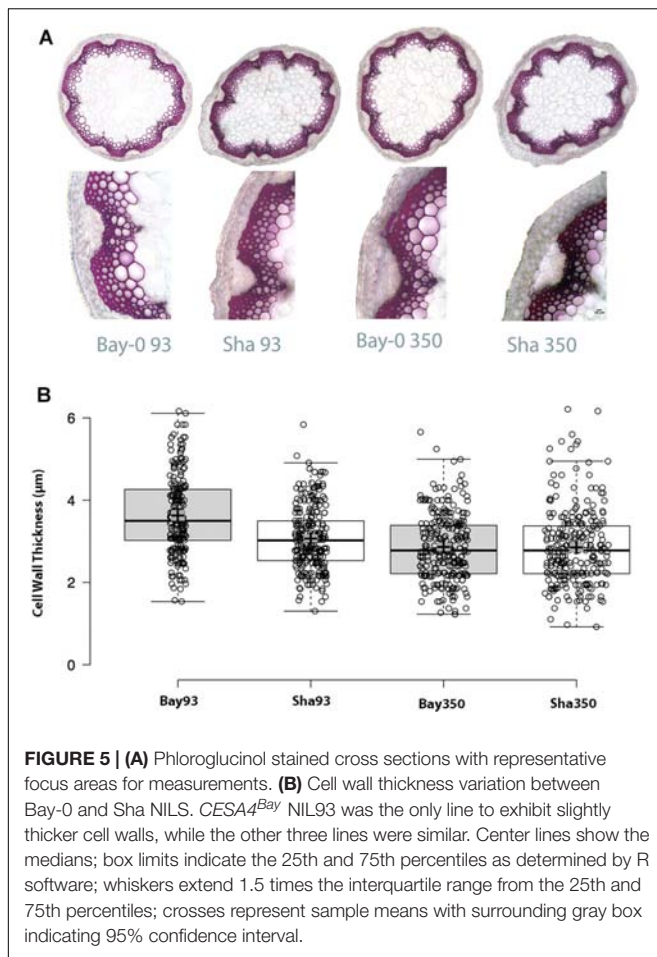


NILs from heterozygous inbred family 93 (HIF93) and HIF350, segregating for the *CESA4* promoter interval that demonstrated differential binding. Transcript abundance in developing stems containing *CESA4*^{Bay} was twofold greater than that of *CESA4*^{Sha} in both pairs of NILs (Figure 4). These results further support the differential binding between Bay-0 and Sha by NST2 and SND1, thus changing expression of *CESA4*.

No Differences in Cell Wall Properties Were Observed Between *CESA4*^{Bay} and *CESA4*^{Sha} Near Isogenic Lines

Considering the critical role of *CESA4* in plant structure, we wished to evaluate the phenotypic effect of the *cis*-eQTL. Bay-0 and Sha *CESA4* NILs were indistinguishable at the macro level, but a subtle consequence at the cellular level remained a possibility. To explore the effect of the *CESA4* *cis*-eQTL on cellulose, Bay-0 and Sha were first examined with stem histology to measure cell wall thickness. Stem cross sections were indistinguishable between NILs (Figure 5A). All xylem and interfascicular cells had well-defined edges and equivalent phloroglucinol staining. The *CESA4*^{Bay} NIL93 exhibited an average cell wall thickness of 3.57 μ m, which was slightly greater than the 3.07 μ m thickness of their *CESA4*^{Sha} NIL93 counterparts. The *CESA4*^{Bay} NIL93 samples were quite variable with a standard deviation of 0.45 μ m. Cell wall thickness of *CESA4*^{Bay} and *CESA4*^{Sha} NIL350 were almost identical, 2.86 and 2.85 μ m, respectively (Figure 5B).

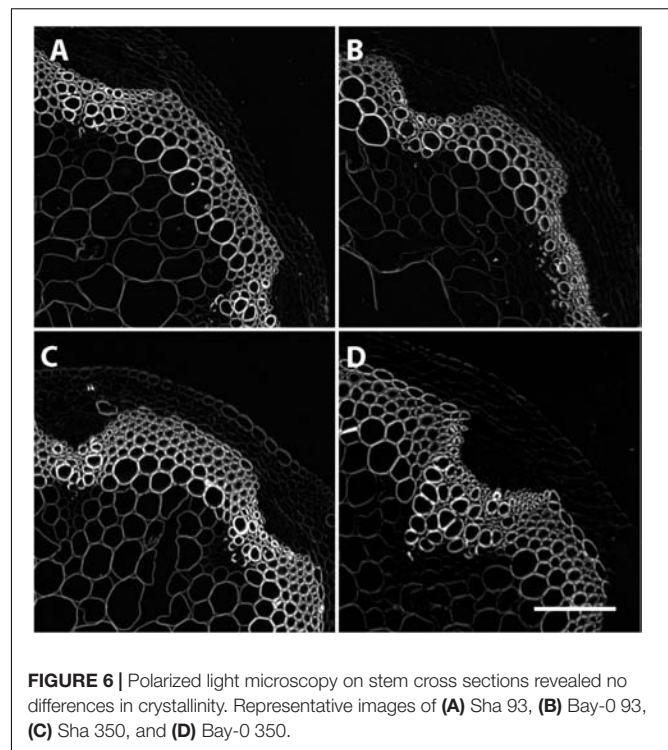
While no significant differences in overall lignin staining or cell thickness was observed, we hypothesized that the reduction in *CESA4* expression may still compromise the cellulose crystalline structure of the NILs carrying the *CESA4*^{Sha} allele. To evaluate this, we employed Updegraff assay to quantify crystalline cellulose and polarized light microscopy as a secondary indicator of crystallinity. Percent composition of crystalline cellulose was slightly higher in *CESA4*^{Bay} NIL93 than *CESA4*^{Sha} NIL93, 28.2% vs. 26.8%, but lower in *CESA4*^{Bay} NIL350 than *CESA4*^{Sha} NIL350, 27.0% vs. 29.2%. Percent cellulose crystallinity was



marginally greater in the Bay-0 parental accession than Sha, 27.0% vs. 25.3% (**Supplementary Figure 2**). However, the NILs were not significantly different. Polarized light images of all samples mirrored brightfield images and were indistinguishable between NILs (**Figure 6**). Birefringent retardance was strong and consistent in xylem and interfascicular cells, revealing no difference in quantities or order of crystalline cellulose. The comparable phenotypes between NILs revealed in both the histological and chemical assays suggest the decreased *CESA4* expression observed in *CESA4^{Sha}* NILs does not disrupt cellulose abundance or crystallinity.

The Secondary Wall Associated NAC Proteins Bind the Same Sequence

The TERE and SNBE are compatible sequences, independently identified as binding sites of VND and NST NAC proteins (Pyo et al., 2007; Ohashi-Ito et al., 2010; Zhong et al., 2010). The TERE sequence, CTTNAAAGCNA, is consistent with the following internal sequence of the SNBE: CTTNNNNNNA. To further resolve this binding site we explored the DNA affinity purification sequencing (DAP-seq) data previously generated for *A. thaliana* transcription factors and Col-0 genomic DNA



(O'Malley et al., 2016). We searched the binding peaks of DAP-seq data for available VND, SND, and NST proteins: VND1, 2, 3, 4, 6, SND2, SND3, and NST1. The data included DNA sequence from libraries of DNA where methylcytosines were removed by PCR and unamplified libraries. There is a striking similarity between the top enriched motif described by the DAP-seq and the SNBE/TERE motif (**Supplementary Table 3**). The core DAP-seq derived motif similarity across all three groups of NACs is three positions flanking seven nucleotides of any sequence: C(G/T)TNNNNNNA(A/C)G. We refer to this as the VND/NST/SND element (VNS, **Figure 7**). Interestingly, the motif appears to be palindromic depending on the variable positions. We searched the sequences of the binding sites and found the prefix CTT and AAG suffix to represent more than half of the total occurrences of the VNS, however, each possible sequence was similar to the frequency in the genome (**Table 3**). Therefore, there does not appear to be a preference for the variable positions. There was a low frequency of an A or a T in the first and last position of the binding site, respectively. The NAC binding site in the *CESA4* promoter is consistent with both the TERE and a VNS, but with a T at the last position.

DISCUSSION

We investigated the variable expression of *CESA4* in a RIL population of *A. thaliana*. After observation of an eQTL at the *CESA4* locus, SNP analysis and transcription factor binding assays revealed a disruption in an otherwise perfect TERE motif in the *CESA4^{Sha}* allele. Chemical and histological analysis of

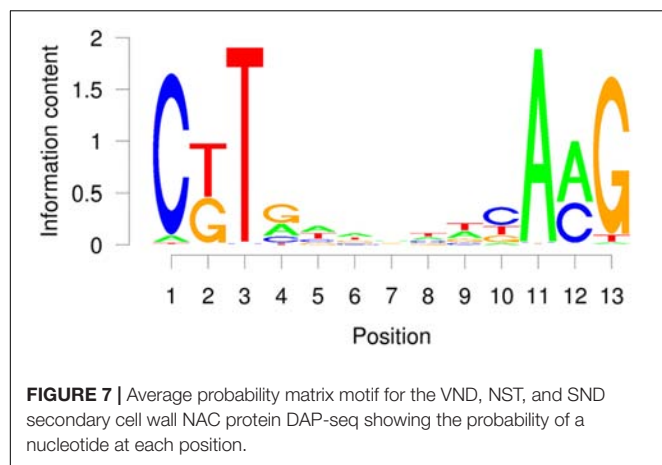


TABLE 3 | Comparison of VND/NST/SND (VNS) NAC protein binding motif variants by percentage in the *Arabidopsis thaliana* genome.

VNS element variants	VNS sites in the Col-0 genome sequence	VND/NST/SND DAP-seq VNS peak motifs
	%	
C _T TNNNNNNNAAG	57.0	55.9
C _T TNNNNNNNACG	18.0	23.4
CG _T NNNNNNNAAG	18.0	23.4
CG _T NNNNNNNACG	8.5	11.5

the cell wall composition was unable to detect any differences between NILs carrying the *CESA4*^{Bay} or *CESA4*^{Sha} allele.

Natural variation in transcript abundance and its association with growth and biomass traits has been documented in a number of instances. Analysis of transcript abundance in a *Eucalyptus* backcross population revealed downregulation of lignin biosynthesis genes is associated with an increase in growth rate (Kirst et al., 2004). Meanwhile, upregulation of genes may confer either positive or negative effects on biomass accumulation; eQTL analysis of high and low biomass pools of *Poplar* revealed that of the identified loci of differential expression, half were upregulated in high biomass trees and half were upregulated in low biomass trees (Du et al., 2015). Alternatively, genotypic variation of patterns of differential expression between mature and immature tissues can also be an indicator of advantageous biomass traits, as presented in a recent study of the sugarcane transcriptome, in which *ShCESA4* and *ShCESA7* were differentially expressed between top and bottom internodes in high fiber genotypes only (Kasirajan et al., 2018). Notably, increased expression of the *CESAs* is frequently associated with higher sugar content or biomass. In alfalfa, *CESA4* was upregulated in genotypes exhibiting greater cellulose content, and similarly, upregulation of *CESA4* in shrub willow was associated with increased total polysaccharide content (Yang et al., 2010; Serapiglia et al., 2012). Though the present study concurs with others that underline the occurrence of naturally variable expression patterns, it differs with the lack of an association between *CESA4* expression and cellulose content. This study also diverges from previous reports of *CESA*

expression, as the correlation coefficients among the secondary cell wall *CESAs* were relatively low (Appenzeller et al., 2004). This discrepancy is likely due to the low enrichment of secondary cell walls in the tissue tested.

Variation in transcript abundance among genotypes may be caused by non-synonymous or synonymous SNPs in coding regions, as well as SNPs in introns and 3' and 5'UTR, as was reported in *Pinus taeda*, *Eucalyptus*, and *Picea glauca* (Kirst et al., 2004; Beaulieu et al., 2011; Palle et al., 2013). Such SNPs can induce either *trans*-eQTLs, functioning by affecting the expression of transcription factor targets or other regulatory interactions, or *cis*-eQTLs, modulating expression by discrepancies in promoter sequence, as presented here. Though non-coding regions face less selective pressure than coding sequences, development-specific expression profiles and regions responsible for transcription factor binding have been shown to have greater conservation of motifs in promoters across species (Creux et al., 2008; Ding et al., 2012). We describe a loss-of-function mutation to the *TERE* motif, complementing previous reports of the effects of aberrations to this 11 bp sequence (Pyo et al., 2007).

A number of studies have investigated the effects of perturbed *CESA* expression on biomass production, and several studies have underlined the flexibility of species to tolerate moderate discrepancies in expression. amiRNA inhibition of *CESA4* expression in *B. distachyon*, in which *BdCESA4* expression was reduced almost 10-fold, resulted in compromised cellulose production, yet the same study found that reduction in *CESA7* expression by only 1.5-fold resulted in only moderate effects on cell wall composition (Handakumbura et al., 2013; Mazarei et al., 2018). A similar study of *Panicum virgatum* highlighted that modest reductions in expression caused no detectable changes to plant structure or biomass production, and only samples with a >40% reduction in expression were compromised in cell wall traits (Mazarei et al., 2018). Notably, studies reporting the most profound phenotypes typically discuss mutations in the coding region of genes or complete loss-of-function alleles (Tanaka et al., 2003; Taylor et al., 2003; Joshi et al., 2011).

The presence of loss-of-function *cis*-eQTLs across viable accessions found in variable geographic locations also underlines the ability of the plant genome to tolerate expression irregularities and perhaps poses more questions about posttranscriptional and translational regulation, as well as the role of local environment in trait variance (Vuylsteke and van Eeuwijk, 2008; Zan et al., 2016). The results suggest that regulatory mechanisms may be at play, highlighting the elasticity of the plant genome and proteome. A plethora of both naturally occurring and mutagenesis-induced transcript discrepancies have been associated with a structural or developmental phenotype, and the same was expected in this study. However, resistance to cellulose perturbation has also been reported. For example, co-suppression of several *CESAs* in barley (*Hordeum vulgare*) caused by constitutive expression using the CaMV35S promoter was only sometimes (25% of the time) associated with a compromised cell wall phenotype (Tan et al., 2015). Transgenic overexpression of each secondary cell wall *CESA*, *CESA4*, 7, and 8, resulted in lower transcript levels for all three endogenous secondary cell wall

CESAs. Moreover, transcript levels of all three CESAs tended to correlate, regardless of which CESA cDNA was driven by the CaMV35S promoter, suggesting that regulatory mechanisms limited by synthase complex stoichiometry are at play (Tan et al., 2015).

Co-regulation of the CSC subunits at the transcript level was not reported in HIFs for this study, but posttranscriptional regulation poses another possible mechanism to explain the lack of phenotypic effect. While transcript analysis is a powerful tool and can provide invaluable data on genome regulation, stress response, and more, mRNA accumulation alone is not enough to draw definitive conclusions about protein expression. Indeed, the global correlation between the transcriptome and proteome in both prokaryotes and eukaryotes has been found to be weak, at best (Gygi et al., 1999; Washburn et al., 2003; Soto-Suárez et al., 2016). Posttranscriptional regulation mechanisms, half-life, and localization and interactions all may play a role in protein expression levels, causing them to differ from transcript abundance.

A number of studies have identified cases where protein expression of complex subunits is posttranscriptionally controlled and seemingly limited by complex stoichiometry; generally, mRNA levels of individual subunits have not been found to correlate with complex expression (Washburn et al., 2003; Hajdich et al., 2010; Lalanne et al., 2018). For example, it was found that the alpha and beta plastidial pyruvate kinase subunits had similar protein expression but discrepant transcript expression, suggesting that the requirements of complex assembly can be a regulatory factor in protein accumulation. A similar scenario could likely be occurring in the CSC, and

further studies would benefit from an investigation of CESA protein expression levels.

AUTHOR CONTRIBUTIONS

JO, LL, and SH conceived and designed the study. JO, LL, and GT acquired the data. JO, LL, SL, KM, and SH analyzed and interpreted the data. JO, LL, and SH drafted the manuscript.

FUNDING

This research was supported by the Office of Science Department of Energy Grant DE-SC0006621 and DE-FG02-08ER64700DE to SH.

ACKNOWLEDGMENTS

We would like to thank Dr. Tobias Baskin for his assistance in collecting the polarized light images. We thank Olivier Loudet (INRA–Genetics and Plant Breeding, Versailles, France) for HIF lines.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.01895/full#supplementary-material>

REFERENCES

- Appenzeller, L., Doblin, M., Barreiro, R., Wang, H., Niu, X., Carrigan, L., 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
- Arioli, T., Peng, L., Betzner, A. S., Burn, J., Wittke, W., Herth, W., et al. (1998). Molecular analysis of cellulose biosynthesis in *Arabidopsis*. *Science* 279, 717–720. doi: 10.1126/science.279.5351.717
- Barrière, Y., Méchin, V., Lefevre, B., and Maltese, S. (2012). QTLs for agronomic and cell wall traits in a maize RIL progeny derived from a cross between an old Minnesota13 line and a modern Iodent line. *Theor. Appl. Genet.* 125, 531–549. doi: 10.1007/s00122-012-1851-5
- Beaulieu, J., Doerken, T., Boyle, B., Clément, S., Deslauriers, M., Beauseigle, S., et al. (2011). Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics* 188:197. doi: 10.1534/genetics.110.125781
- Brown, D. M., Zeef, L. A. H., Ellis, J., Goodacre, R., and Turner, S. R. (2005). Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* 17, 2281–2295. doi: 10.1105/tpc.105.031542
- Burton, R. A., Gibeaut, D. M., Bacic, A., Findlay, K., and Roberts, K. (2000). Virus-induced silencing of a plant cellulose synthase gene. *Plant Cell* 12:691. doi: 10.1105/tpc.12.5.691
- Carroll, A., and Somerville, C. (2009). Cellulosic biofuels. *Annu. Rev. Plant Biol.* 60, 165–182. doi: 10.1146/annurev.arplant.043008.092125
- Chantreau, M., Chabbert, B., Billiard, S., Hawkins, S., and Neutelings, G. (2015). Functional analyses of cellulose synthase genes in flax (*Linum usitatissimum*) by virus-induced gene silencing. *Plant Biotechnol. J.* 13, 1312–1324. doi: 10.1111/pbi.12350
- Ching, A., Dhugga, K. S., Appenzeller, L., Meeley, R., Bourett, T. M., Howard, R. J., et al. (2006). Brittle stalk 2 encodes a putative glycosylphosphatidylinositol-anchored protein that affects mechanical strength of maize tissues by altering the composition and structure of secondary cell walls. *Planta* 224, 1174–1184. doi: 10.1007/s00425-006-0299-8
- Creux, N. M., Ranik, M., Berger, D. K., and Myburg, A. A. (2008). Comparative analysis of orthologous cellulose synthase promoters from *Arabidopsis*, *Populus* and *Eucalyptus*: evidence of conserved regulatory elements in angiosperms. *New Phytol.* 179, 722–737. doi: 10.1111/j.1469-8137.2008.02517.x
- Ding, J., Hu, H., and Li, X. (2012). Thousands of cis-regulatory sequence combinations are shared by *Arabidopsis* and poplar. *Plant Physiol.* 158:145. doi: 10.1104/pp.111.186080
- Du, Q., Gong, C., Wang, Q., Zhou, D., Yang, H., Pan, W., et al. (2015). Genetic architecture of growth traits in *Populus* revealed by integrated quantitative trait locus (QTL) analysis and association studies. *New Phytol.* 209, 1067–1082. doi: 10.1111/nph.13695
- Fagard, M., Desnos, T., Desprez, T., Goubet, F., Refregier, G., Mouille, G., et al. (2000). PROCUSTE1 encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of *Arabidopsis*. *Plant Cell* 12, 2409–2423. doi: 10.1105/tpc.12.12.2409
- Guerrero, G., Fugelstad, J., and Bulone, V. (2010). What do we really know about cellulose biosynthesis in higher plants? *J. Integr. Plant Biol.* 52, 161–175. doi: 10.1111/j.1744-7909.2010.00935.x
- Gygi, S. P., Rochon, Y., Franza, B. R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* 19:1720. doi: 10.1128/MCB.19.3.1720
- Hajdich, M., Hearne, L. B., Miernyk, J. A., Casteel, J. E., Joshi, T., Agrawal, G. K., et al. (2010). Systems analysis of seed filling in *Arabidopsis*: using general linear modeling to assess concordance of transcript and protein expression. *Plant Physiol.* 152, 2078. doi: 10.1104/pp.109.152413

- Handakumbura, P., Matos, D., Osmont, K., Harrington, M., Heo, K., Kafle, K., et al. (2013). Perturbation of *Brachypodium distachyon* CELLULOSE SYNTHASE A4 or 7 results in abnormal cell walls. *BMC Plant Biol.* 13:131. doi: 10.1186/1471-2229-13-131
- Handakumbura, P. P., Brow, K., Whitney, I. P., Zhao, K., Sanguinet, K. A., Lee, S. J., et al. (2018). SECONDARY WALL ASSOCIATED MYB1 is a positive regulator of secondary cell wall thickening in *Brachypodium distachyon* and is not found in the *Brassicaceae*. *Plant J.* 96, 532–545. doi: 10.1111/tpj.14047
- Harris, D. M., Corbin, K., Wang, T., Gutierrez, R., Bertolo, A. L., Petti, C., et al. (2012). Cellulose microfibril crystallinity is reduced by mutating C-terminal transmembrane region residues CESA1(A903V) and CESA3(T942I) of cellulose synthase. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4098–4103. doi: 10.1073/pnas.1200352109
- 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
- Jarvis, M. C. (2013). Cellulose biosynthesis: counting the chains. *Plant Physiol.* 163, 1485–1486. doi: 10.1104/pp.113.231092
- Joshi, C. P., Thammannagowda, S., Fujino, T., Gou, J.-Q., Avci, U., Haigler, C. H., et al. (2011). Perturbation of wood cellulose synthesis causes pleiotropic effects in transgenic aspen. *Mol. Plant* 4, 331–345. doi: 10.1093/mp/ssq081
- Kasirajan, L., Hoang, N. V., Furtado, A., Botha, F. C., and Henry, R. J. (2018). Transcriptome analysis highlights key differentially expressed genes involved in cellulose and lignin biosynthesis of sugarcane genotypes varying in fiber content. *Sci. Rep.* 8:11612. doi: 10.1038/s41598-018-30033-4
- Kirst, M., Myburg, A., De Leon, J., Kirst, M., Scott, J., and Sederoff, R. (2004). Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *eucalyptus*. *Plant Physiol.* 135, 2368–2378. doi: 10.1104/pp.103.037960
- Kotake, T., Aohara, T., Hirano, K., Sato, A., Kaneko, Y., Tsumura, Y., et al. (2011). Rice brittle culm 6 encodes a dominant-negative form of CesA protein that perturbs cellulose synthesis in secondary cell walls. *J. Exp. Bot.* 62, 2053–2062. doi: 10.1093/jxb/erq395
- Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19, 1855–1860. doi: 10.1101/gad.1331305
- Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., et al. (2015). A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. *Sci. Rep.* 5:11166. doi: 10.1038/srep11166
- Kumar, M., and Turner, S. (2015). Plant cellulose synthesis: CESA proteins crossing kingdoms. *Phytochemistry* 112, 91–99. doi: 10.1016/j.phytochem.2014.07.009
- Lalanne, J.-B., Taggart, J. C., Guo, M. S., Herzel, L., Schieler, A., Li, G.-W. (2018). Evolutionary convergence of pathway-specific enzyme expression stoichiometry. *Cell* 173, 749.e738–761.e738. doi: 10.1016/j.cell.2018.03.007
- Little, A., Schwerdt, J. G., Shirley, N. J., Khor, S. F., Neumann, K., Donovan, L. A., et al. (2018). Revised phylogeny of the cellulose synthase gene superfamily: insights into cell wall evolution. *Plant Physiol.* 177, 1124–1141. doi: 10.1104/pp.17.01718
- Loudet, O., Chaillou, S., Camilleri, C., Bouchez, D., and Daniel-Vedele, F. (2002). Bay-0 × Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in *Arabidopsis*. *Theor. Appl. Genet.* 104, 1173–1184. doi: 10.1007/s00122-001-0825-9
- Mazarei, M., Baxter, H. L., Li, M., Biswal, A. K., Kim, K., Meng, X., et al. (2018). Functional analysis of cellulose synthase Cesa4 and Cesa6 genes in switchgrass (*Panicum virgatum*) by overexpression and RNAi-Mediated gene silencing. *Front. Plant Sci.* 9:114. doi: 10.3389/fpls.2018.01114
- 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
- Mueller, S. C., and Brown, R. M. (1980). Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants. *J. Cell Biol.* 84, 315. doi: 10.1083/jcb.84.2.315
- Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N. A., and Ohtani, M. (2015). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* 6:288. doi: 10.3389/fpls.2015.00288
- Nixon, B. T., Mansouri, K., Singh, A., Du, J., Davis, J. K., Lee, J.-G., et al. (2016). Comparative structural and computational analysis supports eighteen cellulose synthases in the plant cellulose synthesis complex. *Sci. Rep.* 6:28696. doi: 10.1038/srep28696
- Ohashi-Ito, K., Oda, Y., and Fukuda, H. (2010). *Arabidopsis* VASCULAR-RELATED NAC-DOMAIN6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. *Plant Cell* 22, 3461–3473. doi: 10.1105/tpc.110.075036
- O'Malley, R. C., Huang, S. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., et al. (2016). Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* 165, 1280–1292. doi: 10.1016/j.cell.2016.04.038
- Palle, S. R., Seeve, C. M., Eckert, A. J., Wegrzyn, J. L., Neale, D. B., and Loopstra, C. A. (2013). Association of loblolly pine xylem development gene expression with single-nucleotide polymorphisms. *Tree Physiol.* 33, 763–774. doi: 10.1093/treephys/tp054
- Persson, S., Wei, H., Milne, J., Page, G. P., and Somerville, C. R. (2005). Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8633–8638. doi: 10.1073/pnas.0503392102
- Pruneda-Paz, J. L., Breton, G., Para, A., and Kay, S. A. (2009). A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock. *Science* 323, 1481–1485. doi: 10.1126/science.1167206
- Pyo, H., Demura, T., and Fukuda, H. (2007). TERE; a novel cis-element responsible for a coordinated expression of genes related to programmed cell death and secondary wall formation during differentiation of tracheary elements. *Plant J.* 51, 955–965. doi: 10.1111/j.1365-3113X.2007.03180.x
- 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
- Ruprecht, C., Mutwil, M., Saxe, F., Eder, M., Nikoloski, Z., and Persson, S. (2011). Large-scale co-expression approach to dissect secondary cell wall formation across plant species. *Front. Plant Sci.* 2:23. doi: 10.3389/fpls.2011.00023
- Scheible, W. R., Eshed, R., Richmond, T., Delmer, D., and Somerville, C. R. (2001). Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in *Arabidopsis* Ixr1 mutants. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10079–10084. doi: 10.1073/pnas.191361598
- Serapiglia, M. J., Cameron, K. D., Stipanovic, A. J., and Smart, L. B. (2012). Correlations of expression of cell wall biosynthesis genes with variation in biomass composition in shrub willow (*Salix* spp.) biomass crops. *Tree Genet. Genomes* 8, 775–788. doi: 10.1007/s11295-011-0462-7
- Shiringani, A. L., and Friedt, W. (2011). QTL for fibre-related traits in grain × sweet sorghum as a tool for the enhancement of sorghum as a biomass crop. *Theor. Appl. Genet.* 123:999. doi: 10.1007/s00122-011-1642-4
- Sindhu, A., Langewisch, T., Olek, A., Multani, D. S., McCann, M. C., Vermerris, W., et al. (2007). Maize brittle Stalk2 encodes a COBRA-like protein expressed in early organ development but required for tissue flexibility at maturity. *Plant Physiol.* 145, 1444–1459. doi: 10.1104/pp.107.102582
- Somerville, C. (2006). Cellulose synthesis in higher plants. *Annu. Rev. Cell Dev. Biol.* 22, 53–78. doi: 10.1146/annurev.cellbio.22.022206.160206
- Soto-Suárez, M., Serrato, A. J., Rojas-González, J. A., Bautista, R., and Sahrawy, M. (2016). Transcriptomic and proteomic approach to identify differentially expressed genes and proteins in *Arabidopsis thaliana* mutants lacking chloroplastic 1 and cytosolic FBPs reveals several levels of metabolic regulation. *BMC Plant Biol.* 16:258. doi: 10.1186/s12870-016-0945-7
- Spitzer, M., Wildenhain, J., Rappsilber, J., and Tyers, M. (2014). BoxPlotR: a web tool for generation of box plots. *Nat. Methods* 11, 121–122. doi: 10.1038/nmeth.2811
- Tan, H.-T., Shirley, N. J., Singh, R. R., Henderson, M., Dhugga, K. S., Mayo, G. M., et al. (2015). Powerful regulatory systems and post-transcriptional gene silencing resist increases in cellulose content in cell walls of barley. *BMC Plant Biol.* 15:62. doi: 10.1186/s12870-015-0448-y
- Tanaka, K., Murata, K., Yamazaki, M., Onosato, K., Miyao, A., and Hirochika, H. (2003). Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. *Plant Physiol.* 133, 73–83. doi: 10.1104/pp.103.022442
- Taylor, N. G., Howells, R. M., Huttly, A. K., Vickers, K., and Turner, S. R. (2003). Interactions among three distinct CesA proteins essential for cellulose synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1450–1455. doi: 10.1073/pnas.0337628100

- Taylor, N. G., Laurie, S., and Turner, S. R. (2000). Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in *Arabidopsis*. *Plant Cell* 12, 2529–2540. doi: 10.1105/tpc.12.12.2529
- 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
- Thomas, L., Forsyth, T., Šturcová, A., Kennedy, C., May, R., Altaner, C., et al. (2013). Structure of cellulose microfibrils in primary cell walls from collenchyma1. *Plant Physiol.* 161, 465–476. doi: 10.1104/pp.112.206359
- Timmers, J., Vernhettes, S., Desprez, T., Vincken, J.-P., Visser, R. G. F., and Trindade, L. M. (2009). Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. *FEBS Lett.* 583, 978–982. doi: 10.1016/j.febslet.2009.02.035
- Updegraff, D. (1969). Semimicro determination of cellulose in biological materials. *Anal. Biochem.* 32, 420–424. doi: 10.1016/S0003-2697(69)80009-6
- Vuylsteke, M., and van Eeuwijk, F. (2008). The use of general and specific combining abilities in a context of gene expression relevant to plant breeding. *Euphytica* 161, 115–122. doi: 10.1007/s10681-007-9452-0
- Washburn, M. P., Koller, A., Oshiro, G., Ulaszek, R. R., Plouffe, D., Deciu, C., et al. (2003). Protein pathway and complex clustering of correlated mRNA and protein expression analyses in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3107–3112. doi: 10.1073/pnas.0634629100
- West, M., Kim, K., Kliebenstein, D., van Leeuwen, H., Micheltore, R., Doerge, R., et al. (2007). Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in *Arabidopsis*. *Genetics* 175, 1441–1450. doi: 10.1534/genetics.106.064972
- Yamaguchi, M., Goue, N., Igarashi, H., Ohtani, M., Nakano, Y., Mortimer, J. C., et al. (2010). VASCULAR-RELATED NAC-DOMAIN6 (VND6) and VND7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant Physiol.* 153, 906–914. doi: 10.1104/pp.110.154013
- Yang, S. S., Xu, W. W., Tesfaye, M., Lamb, J. F. S., Jung, H.-J. G., VandenBosch, K. A., et al. (2010). Transcript profiling of two alfalfa genotypes with contrasting cell wall composition in stems using a cross-species platform: optimizing analysis by masking biased probes. *BMC Genomics* 11:323. doi: 10.1186/1471-2164-11-323
- Yao, D., Wei, Q., Xu, W., Syrenne, R., Yuan, J., and Su, Z. (2012). Comparative genomic analysis of NAC transcriptional factors to dissect the regulatory mechanisms for cell wall biosynthesis. *BMC Bioinformatics* 13:S10. doi: 10.1186/1471-2105-13-S15-S10
- Ye, Y., Wu, K., Chen, J., Liu, Q., Wu, Y., Liu, B., et al. (2018). OsSND2, a NAC family transcription factor, is involved in secondary cell wall biosynthesis through regulating MYBs expression in rice. *Rice* 11:36. doi: 10.1186/s12284-018-0228-z
- Zan, Y., Shen, X., Forsberg, S. K. G., Carlborg, Ö. (2016). Genetic regulation of transcriptional variation in natural *Arabidopsis thaliana* accessions. *G3* 6:2319. doi: 10.1534/g3.116.030874
- Zhang, M., Zhang, B., Qian, Q., Yu, Y., Li, R., Zhang, J., et al. (2010). Brittle culm 12, a dual-targeting kinesin-4 protein, controls cell-cycle progression and wall properties in rice. *Plant J.* 63, 312–328. doi: 10.1111/j.1365-313X.2010.04238.x
- Zhang, J., Xie, M. A., Tuskan, G., Muchero, W., and Chen, J.-G. (2018). Recent advances in the transcriptional regulation of secondary cell wall biosynthesis in the woody plants. *Front. Plant Sci.* 9:135. doi: 10.3389/fpls.2018.01535
- Zhao, H., Wang, B., He, J., Yang, J., Pan, L., Sun, D., et al. (2013). Genetic diversity and population structure of *Miscanthus sinensis* germplasm in China. *PLoS One* 8:e75672. doi: 10.1371/journal.pone.0075672
- Zhong, R., Lee, C., and Ye, Z. H. (2010). Global analysis of direct targets of secondary wall NAC master switches in *Arabidopsis*. *Mol. Plant* 3, 1087–1103. doi: 10.1093/mp/ssq062
- 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

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Olins, Lin, Lee, Trabucco, MacKinnon and Hazen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Post-Golgi Trafficking and Transport of Cell Wall Components

Rosalie Sinclair*, Michel Ruiz Rosquete and Georgia Drakakaki*

Department of Plant Sciences, University of California, Davis, Davis, CA, United States

OPEN ACCESS

Edited by:

Diane C. Bassham,
Iowa State University, United States

Reviewed by:

Harriet T. Parsons,
University of Copenhagen, Denmark
Gian Pietro Di Sansebastiano,
University of Salento, Italy

*Correspondence:

Rosalie Sinclair
rmsinclair@ucdavis.edu
Georgia Drakakaki
gdrakakaki@ucdavis.edu

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 02 July 2018

Accepted: 16 November 2018

Published: 07 December 2018

Citation:

Sinclair R, Rosquete MR and
Drakakaki G (2018) Post-Golgi
Trafficking and Transport of Cell Wall
Components.
Front. Plant Sci. 9:1784.
doi: 10.3389/fpls.2018.01784

The cell wall, a complex macromolecular composite structure surrounding and protecting plant cells, is essential for development, signal transduction, and disease resistance. This structure is also integral to cell expansion, as its tensile resistance is the primary balancing mechanism against internal turgor pressure. Throughout these processes, the biosynthesis, transport, deposition, and assembly of cell wall polymers are tightly regulated. The plant endomembrane system facilitates transport of polysaccharides, polysaccharide biosynthetic and modifying enzymes and glycoproteins through vesicle trafficking pathways. Although a number of enzymes involved in cell wall biosynthesis have been identified, comparatively little is known about the transport of cell wall polysaccharides and glycoproteins by the endomembrane system. This review summarizes our current understanding of trafficking of cell wall components during cell growth and cell division. Emerging technologies, such as vesicle glycomics, are also discussed as promising avenues to gain insights into the trafficking of structural polysaccharides to the apoplast.

Keywords: post-Golgi trafficking, trans-Golgi Network, endosome, polysaccharide trafficking, cell wall, endomembrane trafficking, glycome analysis, SNARE

THE PLANT CELL WALL

Consisting of a complex weaving of macromolecules, the cell wall is essential for many cellular processes such as development, cell integrity, signal transduction, defense, and maintenance of turgor pressure (Collins et al., 2003; Cosgrove, 2005, 2016). Roughly 40 cell types make up a plant with their cell walls determining their unique shape and function (Somerville et al., 2004; De Souza et al., 2015; Cosgrove, 2016; Chebli and Geitmann, 2017). The structurally dynamic and heterogeneous primary walls of young plant cells are predominantly comprised of cellulose microfibrils embedded in a matrix of pectin, hemicellulose, and glycoproteins (McCann et al., 1992; Somerville et al., 2004; Burton et al., 2010). Although during the last 20 years many cell wall biosynthetic enzymes have been identified, the understanding of the mechanisms facilitating polysaccharide transport is far from comprehensive. With a better understanding of cell wall component trafficking pathways, detailed models of cell wall deposition and maturation can be constructed, providing insights into the dynamic organization of cell wall during plant growth and in response to environmental cues. This review focuses on transport and deposition of cell wall components during primary cell wall formation. For a comprehensive review of secondary cell wall biosynthesis see (Meents et al., 2018).

Xyloglucan (XyG) represents the predominant hemicellulose in the primary cell walls of eudicots and non-graminaceous monocots. XyG is a β -1,4 glucan featuring a regular pattern of substitutions

occurring along the xylose residue on the glucan backbone. Xylose substitutions consist of galactose, O-acetylated galactose and fucose residues (Keegstra, 2010; Scheller and Ulvskov, 2010; Pauly and Keegstra, 2016; Kim et al., 2018). Pectin, a complex heterogeneous assembly of polysaccharides whose structural backbones contain galacturonic acid residues, constitutes a major part of the matrix into which cellulose microfibrils are embedded (Atmodjo et al., 2013). Both pectin and hemicellulose are synthesized by Golgi localized enzymes and thus require transport to the plasma membrane (PM), a critical yet poorly understood role of the endomembrane system. In contrast, cellulose, another key component of cell walls, is synthesized at the PM by Cellulose Synthase Complexes (CSCs) (McFarlane et al., 2014). Although the transport of pectin, hemicellulose and CSCs to the PM is currently thought to utilize the conventional ER–Golgi–*trans*-Golgi Network (TGN)–PM traffic route, unconventional pathways have been postulated. This review primarily focuses on TGN dependent pathways. For recent reviews on unconventional protein secretion (UPS) see (Davis et al., 2016; van de Meene et al., 2017; Wang X. et al., 2017).

MULTIPLE FACTORS MAY COORDINATE TGN-MEDIATED TRANSPORT OF CELL WALL COMPONENTS

Accurate spatial and temporal delivery of cell wall material is essential in choreographing cellular responses to the environment such as those elicited by pathogens. The endomembrane system's intricate array of molecular players orchestrates the timely delivery of cargos crucial to cell functions, and cell wall components are no exception. The TGN is the membrane compartment on the *trans*-side of Golgi responsible for sorting and packaging cargo molecules targeted to the PM or vacuoles (Roth et al., 1985; Griffiths and Simons, 1986; Kang et al., 2011; Rosquete et al., 2018). Unlike in other eukaryotes, plant TGN also serves as an early endosome (Dettmer et al., 2006; Viotti et al., 2010). A specialized role of the Golgi apparatus and the TGN in plants is the biosynthesis and sorting of cell wall components including biosynthetic enzymes, structural proteins and the matrix polysaccharides hemicellulose and pectin (Cosgrove, 2005; Worden et al., 2012; Kim and Brandizzi, 2016; van de Meene et al., 2017).

The function of the TGN is regulated by a plethora of factors including RAB GTPases, soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs), tethers, accessory proteins, as well as vesicle pH and lipid composition (Ebine and Ueda, 2015; Zhen and Stenmark, 2015; Rosquete et al., 2018). SNARE proteins mediate membrane fusion (Ueda et al., 2012; Bombardier and Munson, 2015), with syntaxins representing a sub-family of SNAREs. *Arabidopsis* contains distinct syntaxins of plants (SYPs) localized to different compartments of the endomembrane system (Sanderfoot et al., 2000). SYP61, a plant SNARE interacting with SYP42/3 and VTI12 members in the TGN, appears central to the organelle's trafficking functions (Drakakaki et al., 2012; Hofmann, 2017; Li et al., 2017), discussed in depth in this review.

RAB GTPases have been shown to confer specificity to vesicle traffic and mediate membrane fusion between a donor and an acceptor compartment (Woollard and Moore, 2008; Ebine et al., 2011; Lunn et al., 2013b; Bhuin and Roy, 2014). There are 57 *Arabidopsis* RAB genes, split into 8 clades and further classified into subclades of various sizes (Zhang et al., 2007). Vernoud et al. (2013) in trafficking of cell wall components. In a study that used Fourier transformed infrared spectroscopy to evaluate cell wall composition, the percentage of pectin, cellulose and hemicellulose within the cell wall was affected by single mutants of RABA1, RABA2, and RABA4, respectively (Lunn et al., 2013a). Based on this, the authors hypothesized specialized roles for those GTPases in the transport of specific types of polysaccharides from Golgi to the cell surface (Lunn et al., 2013a). Interestingly, RABA2, together with RABA3, has also been implied in trafficking of material to forming cell plates during cytokinesis (Chow et al., 2008) opening the question whether these two RABs are involved in the trafficking and deposition of polysaccharides at the cell plate. Though exciting, further investigation is needed into the possible specificity of these RAB GTPases (see Figure 1).

Although the roles of TGN resident tethers, adaptor proteins and vesicle lipid composition in post-Golgi protein trafficking and TGN functional compartmentalization have started to emerge, very little or nothing is known on how they regulate/impact TGN-mediated polysaccharide trafficking (Kim and Bassham, 2011; Bashline et al., 2013; Di Rubbo et al., 2013; Vukasinovic and Zarsky, 2016; Wattleit-Boyer et al., 2016; Ravikumar et al., 2017; Boutte, 2018).

TRAFFICKING OF CELL WALL STRUCTURAL PROTEINS

Structural cell wall proteins such as arabinogalactan proteins (AGPs), extensins and proline-rich proteins (PRPs) play a role in defining the cell wall's physical functional properties (Showalter and Basu, 2016a; Johnson et al., 2017a,b). However, we know very little about their location within the cell wall, their specific roles and even less about their intracellular trafficking routes. Regarding the latter, both Golgi-dependent and Golgi-independent pathways have been suggested to contribute to the secretion of cell wall proteins (Ellis et al., 2010; Tan et al., 2012; Poulsen et al., 2015; Showalter and Basu, 2016a,b).

Glycosylation is a common post-translational modification amongst PM and cell wall proteins (Nguema-Ona et al., 2014). Interestingly, a recent work combining high-resolution tandem mass spectrometry and available subcellular localization data, established links between the most prominent N-glycan structures present in a given glycoprotein and the protein's distribution in the endomembrane system, in turn reflecting the glycoprotein's trafficking pattern. The presence of complex N-glycans correlated with Golgi/PM localization while paucimannose structures were associated with extracellular glycoproteins (Zeng et al., 2018).

Arabinogalactan proteins are thought to be secreted through a Golgi-dependent pathway, supported by proteomic analysis

of Golgi-enriched fractions (Ford et al., 2016). Contrary to this model, a study using tobacco cells detected transiently expressed AGPs glycosyltransferases in a double-membraned, Exocyst-Positive Organelle (EXPO) (Poulsen et al., 2014). EXPO has been implicated in Golgi-independent, unconventional secretion of, mostly, proteins that do not carry a signal peptide (leaderless proteins) (Wang et al., 2010; Poulsen et al., 2014, 2015; Davis et al., 2016). These findings suggest that Golgi-independent pathways may also mediate secretion of AGPs. However, such a view requires further examination as both the glycosyltransferases and the AGP protein core are not leaderless and are thus expected to follow the canonical secretory pathway.

TRAFFICKING OF CELL WALL BIOSYNTHETIC AND MODIFYING ENZYMES

The Cellulose Synthase Complexes assemble in the Golgi apparatus and are then transported to the PM, both events likely being assisted by STELLO proteins (McFarlane et al., 2014; Lampugnani et al., 2018). Tracking of CSCs at the PM has shown them moving linearly along cortical microtubules (MT) supporting the MT-cellulose alignment hypothesis (Paredes et al., 2006).

Perhaps the most studied cell wall-related trafficking process, the transport of CESAs has been proposed to be mediated by the plant specific endomembrane compartments SmaCC (small CESA compartment) and MASCs (microtubule-associated cellulose synthase compartments) (Crowell et al., 2009; Gutierrez et al., 2009). The role of MASCs/SmaCCs are involved in secretory and endocytic/ recycling routes of CESAs (Crowell et al., 2009; Gutierrez et al., 2009). Recently, the plant-specific protein PATROL1 (PTL1) and members of the exocyst complex have been associated with vesicle docking and secretion of CSCs during primary cell wall formation (Zhu et al., 2018).

Clathrin-mediated endocytosis and recycling rates influence the steady-state levels of CSCs at the PM (Bashline et al., 2013; Lei et al., 2015; Sanchez-Rodriguez et al., 2018). The plant-specific T-PLATE complex, a major adaptor module for clathrin-mediated endocytosis, was recently shown to recognize CSCs for their internalization (Gadeyne et al., 2014; Sanchez-Rodriguez et al., 2018).

CESAs are cargo of the SYP61 vesicles, implicating the SYP61 TGN compartment in the post-Golgi trafficking of these enzymes (Gutierrez et al., 2009; Drakakaki et al., 2012). This is supported by the finding that CESTRIN, a small molecule that reduces the motility of CSCs at the PM, increases the association of CESAs with SYP61 vesicles (Worden et al., 2015). However, it remains unclear whether the CESAs found in the SYP61 compartment represent endocytic or secretory pools, or both. Further, CESAs secretion is affected in mutants of VHA-a1, a TGN specific proton pump isoform that partially localizes to the SYP61 compartment where it plays a prominent role in the establishment of the vesicle luminal pH (Luo et al., 2015).

Similarly to CSCs, callose biosynthesis enzymes such as Glucan Synthase-Like (GSLs) are trafficked to the PM prior

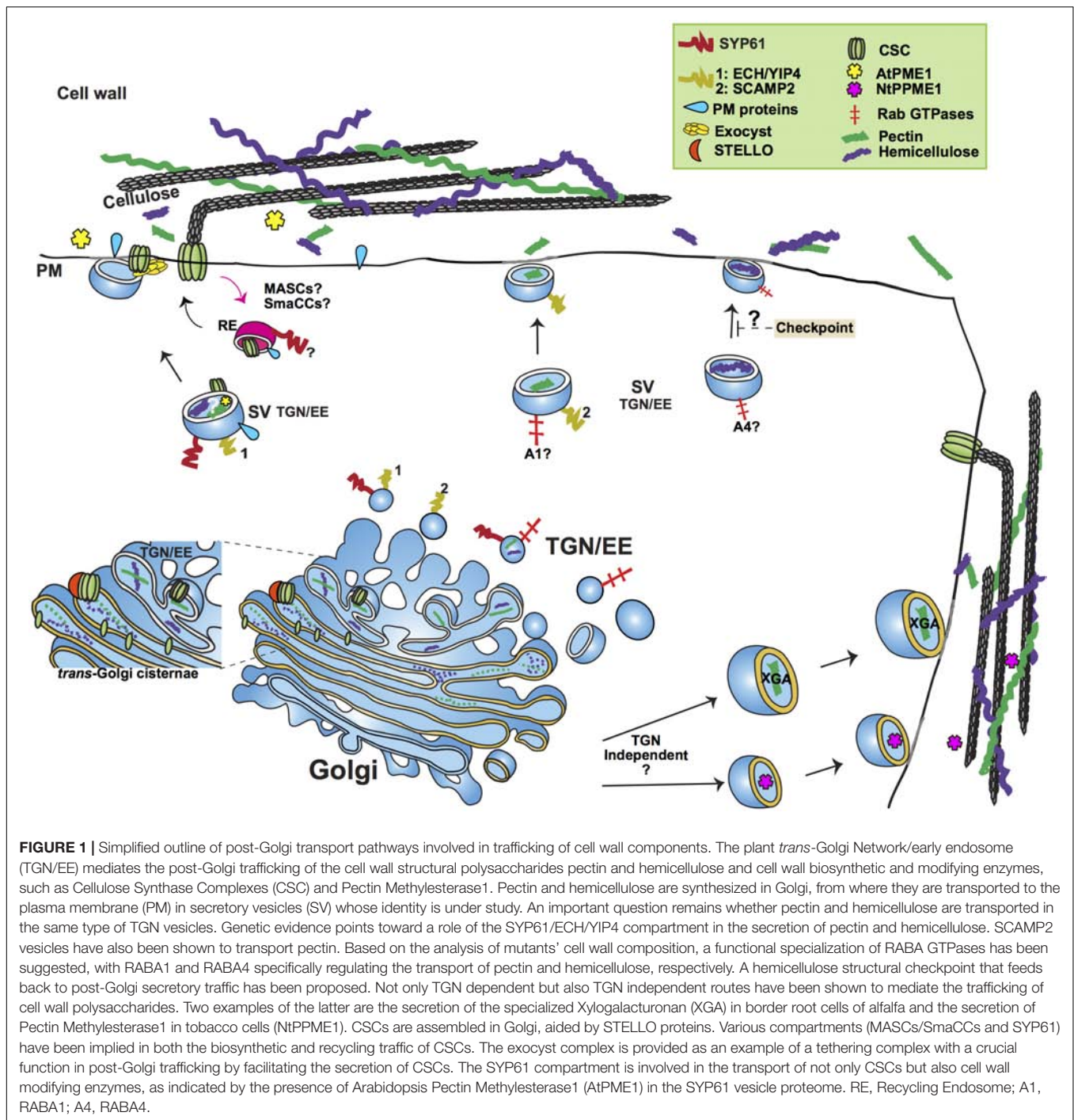
to the initiation of polysaccharide synthesis (Brownfield et al., 2007, 2008; Toller et al., 2008). Not much is known about the trafficking routes of the twelve Arabidopsis GSL isoforms although several of them were identified in the proteome of SYP61 TGN/EE vesicles (Drakakaki et al., 2012). The latter suggests a canonical secretory route to the PM, with the likely involvement of the EXOCYST tethering complex in the final event of fusion to the PM, as indicated by callose deposition studies in Arabidopsis trichomes (Kulich et al., 2018). Interestingly, trafficking of the GSL isoform PMR4 to sites of callose accumulation during the plant response to the pathogen *Blumeria graminis* f. sp. *hordei* has been suggested to occur via unconventional pathways, with the involvement of either multivesicular bodies (Bohlenius et al., 2010) or exosomes (Ellinger et al., 2013). Despite these observations, the trafficking of GSLs during different growth stages and stress conditions needs further investigation.

Cell wall associated enzymes such as apoplastic glycosidases contribute to the modification of polysaccharides during cell wall assembly creating cell wall structural diversity (Showalter, 1993; De Caroli et al., 2011; Gunl et al., 2011a,b; Sampedro et al., 2012; Frankova and Fry, 2013; Pauly and Keegstra, 2016). These proteins are thought to traffic through the secretory pathway; however, recent evidence indicates that multiple pathways may be involved. The main known cell wall modifying enzyme acting on XyG, β -GALACTOSIDASE 10 (AtBGAL10), has three distinct N-glycosites, two with multiple high-mannose structures and the third containing paucimannose structures (Sampedro et al., 2012). Based on the presence of high-mannose structures, Zeng et al. (2018) speculated that an UPS pathway could be involved in the trafficking of β -GALACTOSIDASE 10.

Pectin methylesterase1 (AtPME1), a pectin modifying enzyme, was identified in the proteome of Golgi and also in those of SYP61 and VHA-a1 TGN vesicles, indicating that a Golgi-TGN route is involved in its trafficking to the PM (Drakakaki et al., 2012; Nikolovski et al., 2012; Groen et al., 2014; Heard et al., 2015). However, a different, Golgi-PM pathway that bypasses TGN was identified for the trafficking of tobacco's pollen-specific NtPPME1 (Wang H. et al., 2016).

TRANSPORT OF STRUCTURAL POLYSACCHARIDES

Compared to the number of studies that have provided insights into the secretion of cell wall biosynthetic proteins, less is known on the secretion of structural polysaccharides. Our current knowledge of cell wall polysaccharide transport results primarily from immunoelectron microscopy studies (EM) (Moore et al., 1986, 1991; Moore and Staehelin, 1988; Lynch and Staehelin, 1992; Young et al., 2008; Kang et al., 2011). However, limitations arise from the incompatibility of staining with traditional antibodies, and electron microscopy itself, with live imaging, restricting the experiments to sections of embedded tissue. Despite such limitations, a few studies have shed light into the intracellular distribution of plant polysaccharides.



A seminal study in sycamore maple (*Acer pseudoplatanus*) cells detected the XyG backbone in *trans*-Golgi cisternae, whereas fucosylated XyG side chains were identified in both the *trans* cisternae and the TGN (Zhang and Staehelin, 1992). This suggests a developmental “assembly line” consisting of the initial biosynthesis of the backbone followed by the addition of side chains in Golgi sub-compartments, with the TGN transporting, mostly, fully substituted XyG (Zhang and Staehelin, 1992). In the same study, low-methylesterified pectin backbone, detected

by the antibody JIM5, was found distributed in the *cis*- and *medial*-Golgi and at the cell wall whereas high-methylesterified pectin, detected by JIM7, was localized to the *medial*- and *trans*-Golgi, in secretory vesicles and at the cell wall. These observations indicate that pectins undergo maturation while they are delivered to the *trans*-Golgi, and that high-methylesterified pectin is the predominantly secreted form (Zhang and Staehelin, 1992). They also suggest that conventional post-Golgi trafficking pathways are used by both XyG and pectin. Importantly, the colocalization of

XyG and pectin epitopes in transport vesicles of clover root tips indicate that TGN vesicles can potentially carry both polymers (Lynch and Staehelin, 1992).

Recent studies point to a role of cortical MTs in pectin deposition at the cell wall. Mucilage secretion was shown to be targeted to PM domains lined by abundant cortical MTs. Corroborating this observation, the temperature-sensitive MT mutant *mor1-1* exhibited a decreased mucilage secretion in seeds, in the same study (McFarlane et al., 2008). In addition, the fragile Fiber1 (FRA1) kinesin has been implicated in pectin deposition (Kong Z. et al., 2015; Zhu et al., 2015).

Not surprisingly, the cell type has been shown to determine the trafficking fate of vesicles transporting cell wall polysaccharides. Xylogalacturonan (XGA), a pectin variant secreted by root border cells, is transported by distinct large vesicles, released from the *trans*-Golgi cisternae. XGA-loaded vesicles were shown to fuse with the PM in alfalfa border root cells but not in peripheral cells indicating the existence of regulatory mechanisms conferring cell type specificity to the trafficking and secretion of specialized polysaccharides (Wang P. et al., 2017; see **Figure 1**).

As aforementioned, proteomic analysis of SYP61 vesicles identified proteins involved in cell wall development (Drakakaki et al., 2012). These include the TGN-resident complex formed by ECHIDNA (ECH) and the YPT/RAB GTPase interacting Proteins 4a and 4b (YIP4a and YIP4b), implicated in the secretion of pectin and XyG (Gendre et al., 2011, 2013). Inhibition of cell elongation and altered secretion of XyG and RGI pectins were shown in mutants of YIP4A, YIP4B, and ECH (Gendre et al., 2011, 2013). Further, antibodies for fucosylated XyG have been shown to label a RABA4b TGN compartment in *Arabidopsis* (Kang et al., 2011). Because RABA4b and SYP61 colocalize at TGN, this observation further supports a role for the SYP61 compartment in the trafficking of structural polysaccharides. The PM resident syntaxin SYNTAXIN OF PLANTS121 (SYP121) has been shown to form a SNARE complex with SYP61 mediating the secretion of PM protein cargo (Geelen et al., 2002; Hachez et al., 2014). A role for SYP121 in the secretion of SYP61 polysaccharide cargo is thus likely. Interestingly, the AtSYP121 homolog, NtSyr1, seems dispensable for polysaccharide transport in tobacco cells based on transient studies, an observation that awaits thorough testing in *Arabidopsis* and other plant systems (Leucci et al., 2007; see **Figure 1**).

Anti-pectin antibodies stain SCAMP2 (Secretory Carrier Membrane Protein 2) vesicles in tobacco BY-2 cells, suggesting the involvement of that vesicle population in the transport of pectins (Toyooka et al., 2009). The SCAMP protein structure is well conserved in eukaryotes (Law et al., 2012). In humans, SCAMP2 regulates exocytosis by forming a membrane fusion complex with the small GTPase Arf6 (ADP-ribosylation factor 6), phospholipase D1 (PLD1), and SYNTAXIN 1 (Liu et al., 2002, 2005). It is plausible that SCAMP2, via a similar mechanism, regulates exocytosis of pectin and other polysaccharide cargo in plant cells.

A role for the Exocyst complex in pectin deposition has been suggested, based on genetic evidence. Mutants of the Exocyst subunits SEC8 and EXO70A1 show reduced pectin accumulation in the seed coat. Further, reduced pectin deposition was also observed in a gain-of-function mutation of ROH1, an interactor of the exocyst subunit Exo70A1, supporting the involvement of this tethering factor in polysaccharide transport (Kulich et al., 2010).

Intriguingly, the structure of cell wall polysaccharides, rather than their levels, seems to influence their trafficking, as suggested by the formation of intracellular aggregates containing xyloglucan and deesterified homogalacturonan in mutants of the XyG biosynthetic enzyme galactosyltransferase MUR3 (Kong Y. et al., 2015). It is thus likely that XyG structure checkpoints that feedback to post-Golgi secretory traffic exist. Interestingly, a role for MUR3 in maintaining the organization of Golgi via its interaction with actin filaments has been suggested, hinting at a link between actin cytoskeleton and structural polysaccharides transport (Tamura et al., 2005).

POLYSACCHARIDE TRANSPORT TO THE CELL PLATE

During plant cytokinesis, a cell plate that partitions the cytoplasm of the dividing cell is formed *de novo* (Samuels et al., 1995; Staehelin and Hepler, 1996; Jurgens, 2005; Drakakaki, 2015). Such event requires the coordinated action of cytoskeletal transitions and endomembrane trafficking (Samuels et al., 1995; Otegui et al., 2001; Otegui and Staehelin, 2004; Segui-Simarro et al., 2004; Lee and Liu, 2013; Smertenko et al., 2017). Cell plate development occurs in four stages that exist simultaneously. It requires the directed and choreographed accumulation of post-Golgi vesicles to the phragmoplast at the division plane and removal/recycling of excess material (Samuels et al., 1995; Segui-Simarro et al., 2004; Drakakaki, 2015; Smertenko et al., 2017). The deposition of cell wall polymers transforms the lumen of this membrane compartment into a new cross wall, physically separating the daughter cells (Drakakaki, 2015; Smertenko et al., 2017). Whereas a number of studies have investigated membrane dynamics (van Oostende-Triplett et al., 2017), few reports exist on polysaccharide deposition and its explicit role during cell plate maturation as summarized in recent reviews (Drakakaki, 2015; Chen et al., 2018).

Vesicle trafficking during cell plate formation is controlled by many molecular players, including Rab GTPases, SNAREs, tethering factors and other regulatory proteins (reviewed in (McMichael and Bednarek, 2013; Boruc and Van Damme, 2015; Drakakaki, 2015; Smertenko et al., 2017). Two well-studied factors are RABA2A, which regulates the delivery of TGN derived vesicles to the leading edge of the cell plate (Chow et al., 2008) and the cytokinesis specific SNARE KNOLLE, which catalyzes homotypic fusion of vesicles at the cell plate (Lauber et al., 1997; Assaad et al., 2001; Heese et al., 2001; Zheng et al.,

2002; Zhang et al., 2011; El Kasmi et al., 2013; Karnahl et al., 2018).

The delivery and deposition of cell wall materials to the cell plate has been primarily studied with electron and fluorescence microscopy utilizing polysaccharide-specific antibodies. The current notion is that structural polysaccharides such as hemicellulose and pectins are transported in trans-Golgi derived secretory vesicles to the expanding and maturing cell plate (Moore and Staehelin, 1988; Samuels et al., 1995; Toyooka et al., 2009; Drakakaki, 2015). The presence of XyG is detected at early stages with enrichment in later stages (Moore and Staehelin, 1988). In *Arabidopsis*, the pectin backbone has been detected at the cell plate in methylesterified form (Clausen et al., 2003; Rybak et al., 2014). In red clover root tips, RGI and polygalacturonic acid labeling were observed at the middle lamella, the mature central layer of the cell plate that serves as a glue between adjacent cells. However, RGI was not detected at the early cell plate, suggesting that acidic polysaccharides may be deposited at later stages of cross wall development (Moore and Staehelin, 1988). In addition to Golgi/TGN derived polysaccharides (Moore and Staehelin, 1988), internalized pectin glycans have been implicated in cell plate formation (Baluska and Volkmann, 2002; Dhonukshe et al., 2006), a notion that awaits further investigation.

Callose and cellulose are vital luminal polysaccharides of the cell plate as supported by genetic evidence (Zuo et al., 2000; Beeckman et al., 2002; Chen et al., 2009; Thiele et al., 2009; Guseman et al., 2010; Gu et al., 2016). The relative spatiotemporal distribution of callose, cellulose and their biosynthetic enzymes during the different stages of cell plate formation is not fully elucidated. According to current thinking, callose is transiently incorporated for mechanical support during the middle/late stages of cell plate formation and is ultimately replaced by cellulose, for a more rigid luminal network (Samuels et al., 1995; Thiele et al., 2009). However, recent live imaging of cellulose synthase has shown that it accumulates at the early tubulo-vesicular network stage, concomitant with cellulose (Miart et al., 2014). Cell wall biosynthetic enzymes in the Cellulose Synthase Like-D family (CslD) exhibit high homology to CESAs and are also involved in cell plate formation (Gu et al., 2016). The cell cycle-regulated CslD5 is localized at early cell plate stages where it presumably produces a cellulose-like molecule, as previously shown for CslD3 in polarized root hair growth (Park et al., 2011).

Stains with both the synthetic chemical dye β -glycosyl Yariv and the monoclonal antibody LM14 have shown that AGPs, together with polysaccharides, contribute to cell plate expansion (Yu and Zhao, 2012; Rybak et al., 2014). In addition, EXTENSIN3, a hydroxyl-proline-rich glycoprotein has been implicated in cytokinesis (Hall and Cannon, 2002; Cannon et al., 2008). It is hypothesized that the self-assembled EXTENSIN network can provide mechanical support during the expansion of the cell plate, presumably via interaction with pectins (Cannon et al., 2008).

Although several endomembrane proteins have been associated with cell plate assembly, including the aforementioned SCAMP2 and the exocyst complex, little is known on their

direct involvement in polysaccharide transport to the cell plate (Toyooka et al., 2009; Rybak et al., 2014).

EMERGING TECHNOLOGIES TO DISSECT POLYSACCHARIDE TRANSPORT

Live imaging of polysaccharides remains technically very challenging, making it difficult to assess the colocalization of a particular polysaccharide cargo with subcellular protein markers for specific vesicle populations. However, thanks to recent light microscopy advances that allow the use of photoactivatable and photoconvertible forms of cell wall associated proteins, combined with improved resolution imaging of carbohydrates, our knowledge of cell wall components trafficking and deposition is expected to quickly expand (Fernandez-Suarez and Ting, 2008; Toyooka and Kang, 2014; Mishin et al., 2015; Wang B. et al., 2016; Komis et al., 2018; Voiniciuc et al., 2018). In addition, field emission scanning electron microscopy (FESEM) with nanogold affinity tags affords resolution of spatial location and conformation of cell wall polymers and has proved useful to study XyG-cellulose interactions at the cell wall although this approach does not allow for live imaging (Zheng et al., 2018).

Live imaging of cell wall glycans using small oligosaccharide probes modified via click chemistry, together with polysaccharide dyes, can also contribute useful insights (Mravec et al., 2014). However, the toxicity of click chemistry reagents limits their use in live imaging experiments (Anderson et al., 2010, 2012; Wallace et al., 2012; Wang B. et al., 2016). The ever-expanding palette of metabolically labeled glycans could become a great asset for the dissection of cell wall metabolism once adapted for live imaging (Hoogenboom et al., 2016; Zhu and Chen, 2017). Cell wall glycan-directed antibodies are an elegant alternative for the identification of plant cell carbohydrates, and can be arrayed on automated large scale enzyme-linked immunosorbent assay (ELISA) platforms (glycome profiling) (Moller et al., 2008; Pattathil et al., 2010, 2012; Ruprecht et al., 2017). Cell-permeable, live imaging-compatible nanobodies represent another promising tool (Herce et al., 2017). Exemplifying their potential, a very recent study using a nanobody-epitope interaction-based protein labeling and tracking approach helped dissect a TGN/EE-to-*cis*-Golgi recycling pathway for vacuolar sorting receptors in *Nicotiana tabacum* cells (Fruholz et al., 2018).

To date, there are no suitable glycomic approaches that capture both the polysaccharide contents of specific vesicle populations and the detailed polysaccharide structures. By combining a vesicle isolation methodology, such as that established for the SYP61 compartment (Drakakaki et al., 2012), with vesicle glycome profiling, the roles of different vesicle populations in polysaccharide transport, in relation with developmental stages and responses to environmental stimuli, can be defined. In addition, glycomes of isolated vesicles, as described above, can be coupled with their respective proteomes, obtained with advanced mass spectrometry analysis (Parsons and Lilley, 2018), and with vesicle lipid composition profiling (Haraszti et al., 2016; Wattelet-Boyer et al., 2016; Boutte, 2018).

for a better understanding of how the endomembrane system regulates cell wall transport and deposition.

Oligosaccharide mass profiling (OLIMP) utilizes specific glycosyl-hydrolases to digest cell wall polysaccharides to soluble oligosaccharides detectable by MALDI-TOF mass spectrometry (Obel et al., 2009; Gunl et al., 2010, 2011a). Analyzing Arabidopsis Golgi-enriched microsomal fractions by OLIMP showed that in the Golgi apparatus XyG oligosaccharides (XyGOs) with a lower level of xylose residue substitution by galactose and fucose are more abundant than XyGOs with a higher degree of substitution (Obel et al., 2009; Gunl et al., 2010, 2011a). However, overall this approach has limitations since it is not possible to separate contributions of the Golgi from the TGN or from endoplasmic reticulum contamination. The adequacy of OLIMP to characterize the polysaccharide cargo of specific vesicle populations has yet to be demonstrated.

Proximity tagging methods for protein localization at subcompartmental resolution, such as APEX, BioID, and SPPLAT, have the potential of not only solving the components of large protein complexes involved in cell wall biosynthesis and deposition but also their spatial distribution in membrane microdomains of subcellular compartments (Parsons and Lilley, 2018).

Further, a number of methodologies that have proved useful to study post-Golgi trafficking in other eukaryotes could be successfully implemented in the Plant field. One such method assessed the effect of ectopic intracellular localization of tethering factors on the trafficking fate of cognate vesicles

(Wong and Munro, 2014). A set of Golgi-localized Golgin tethers was artificially targeted to mitochondria of mammalian cells, after which their ability to redirect Golgi-bound carriers to the ectopic destination was monitored. By adapting this and other methodologies, the role and specificity of putative polysaccharide trafficking regulators can be investigated.

All these approaches offer the potential to deepen our spatiotemporal understanding and help model the highly choreographed trafficking events leading to cell wall deposition during both normal and stressful growth conditions.

AUTHOR CONTRIBUTIONS

GD, MR, and RS designed and wrote the manuscript.

FUNDING

This work was supported by NSF IOS 1258135, MCB 1818219 to GD, the USDA CA-D-PLS-2132-H, and funds from ABC PRB to GD.

ACKNOWLEDGMENTS

We apologize to our colleagues whose work we were not able to include in this review due to length limitations.

REFERENCES

- Anderson, C. T., Carroll, A., Akhmetova, L., and Somerville, C. (2010). Real-time imaging of cellulose reorientation during cell wall expansion in Arabidopsis roots. *Plant Physiol.* 152, 787–796. doi: 10.1104/pp.109.15.0128
- Anderson, C. T., Wallace, I. S., and Somerville, C. R. (2012). Metabolic click-labeling with a fucose analog reveals pectin delivery, architecture, and dynamics in Arabidopsis cell walls. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1329–1334. doi: 10.1073/pnas.1120429109
- Assaad, F. F., Huet, Y., Mayer, U., and Jurgens, G. (2001). The cytokinesis gene KEULE encodes a Sec1 protein that binds the syntaxin KNOLLE. *J. Cell Biol.* 152, 531–543. doi: 10.1083/jcb.152.3.531
- Atmodjo, M. A., Hao, Z. Y., and Mohnen, D. (2013). Evolving views of pectin biosynthesis. *Annu. Rev. Plant Biol.* 64, 747–749. doi: 10.1146/annurev-arplant-042811-105534
- Baluska, F., and Volkmann, D. (2002). Pictures in cell biology. Actin-driven polar growth of plant cells. *Trends Cell Biol.* 12:14. doi: 10.1016/S0962-8924(01)02208-5
- Bashline, L., Li, S., Anderson, C. T., Lei, L., and Gu, Y. (2013). The endocytosis of cellulose synthase in Arabidopsis is dependent on mu2, a clathrin-mediated endocytosis adaptin. *Plant Physiol.* 163, 150–160. doi: 10.1104/pp.113.22.1234
- Beeckman, T., Przemeck, G. K., Stamatiou, G., Lau, R., Terryn, N., De Rycke, R., et al. (2002). Genetic complexity of cellulose synthase a gene function in Arabidopsis embryogenesis. *Plant Physiol.* 130, 1883–1893. doi: 10.1104/pp.102.010603
- Bhuin, T., and Roy, J. K. (2014). Rab proteins: the key regulators of intracellular vesicle transport. *Exp. Cell Res.* 328, 1–19. doi: 10.1016/j.yexcr.2014.07.027
- Bohlenius, H., Morch, S. M., Godfrey, D., Nielsen, M. E., and Thordal-Christensen, H. (2010). The multivesicular body-localized GTPase ARFA1b/1c is important for callose deposition and ROR2 syntaxin-dependent preinvasive basal defense in barley. *Plant Cell* 22, 3831–3844. doi: 10.1105/tpc.110.078063
- Bombardier, J. P., and Munson, M. (2015). Three steps forward, two steps back: mechanistic insights into the assembly and disassembly of the SNARE complex. *Curr. Opin. Chem. Biol.* 29, 66–71. doi: 10.1016/j.cbpa.2015.10.003
- Boruc, J., and Van Damme, D. (2015). Endomembrane trafficking overarching cell plate formation. *Curr. Opin. Plant Biol.* 28, 92–98. doi: 10.1016/j.cbpa.2015.09.007
- Boutte, Y. (2018). Lipids at the crossroad: shaping biological membranes heterogeneity defines trafficking pathways. *PLoS Biol.* 16:e2005188. doi: 10.1371/journal.pbio.2005188
- Brownfield, L., Ford, K., Doblin, M. S., Newbigin, E., Read, S., and Bacic, A. (2007). Proteomic and biochemical evidence links the callose synthase in *Nicotiana glauca* pollen tubes to the product of the NaGSL1 gene. *Plant J.* 52, 147–156. doi: 10.1111/j.1365-3113.2007.03219.x
- Brownfield, L., Wilson, S., Newbigin, E., Bacic, A., and Read, S. (2008). Molecular control of the glucan synthase-like protein NaGSL1 and callose synthesis during growth of *Nicotiana glauca* pollen tubes. *Biochem. J.* 414, 43–52. doi: 10.1042/BJ20080693
- Burton, R. A., Gidley, M. J., and Fincher, G. B. (2010). Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat. Chem. Biol.* 6, 724–732. doi: 10.1038/nchembio.439
- Cannon, M. C., Terneus, K., Hall, Q., Tan, L., Wang, Y., Wegenhart, B. L., et al. (2008). Self-assembly of the plant cell wall requires an extensin scaffold. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2226–2231. doi: 10.1073/pnas.0711980105
- Chebli, Y., and Geitmann, A. (2017). Cellular growth in plants requires regulation of cell wall biochemistry. *Curr. Opin. Cell Biol.* 44, 28–35. doi: 10.1016/j.cb.2017.01.002
- Chen, H. W., Persson, S., Grebe, M., and McFarlane, H. E. (2018). Cellulose synthesis during cell plate assembly. *Physiol. Plant.* 164, 17–26. doi: 10.1111/ppl.12703

- Chen, X. Y., Liu, L., Lee, E., Han, X., Rim, Y., Chu, H., et al. (2009). The Arabidopsis callose synthase gene *GSL8* is required for cytokinesis and cell patterning. *Plant Physiol.* 150, 105–113. doi: 10.1104/pp.108.133918
- Chow, C. M., Neto, H., Foucart, C., and Moore, I. (2008). Rab-A2 and Rab-A3 GTPases define a trans-golgi endosomal membrane domain in Arabidopsis that contributes substantially to the cell plate. *Plant Cell* 20, 101–123. doi: 10.1105/tpc.107.052001
- Clausen, M. H., Willats, W. G., and Knox, J. P. (2003). Synthetic methyl hexagalacturonate hapten inhibitors of anti-homogalacturonan monoclonal antibodies LM7, JIM5 and JIM7. *Carbohydr. Res.* 338, 1797–1800. doi: 10.1016/S0008-6215(03)00272-6
- Collins, N. C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J. L., et al. (2003). SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* 425, 973–977. doi: 10.1038/nature02076
- Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6, 850–861. doi: 10.1038/nrm1746
- Cosgrove, D. J. (2016). Plant cell wall extensibility: connecting plant cell growth with cell wall structure, mechanics, and the action of wall-modifying enzymes. *J. Exp. Bot.* 67, 463–476. doi: 10.1093/jxb/erv511
- Crowell, E. F., Bischoff, V., Desprez, T., Rolland, A., Stierhof, Y. D., Schumacher, K., et al. (2009). Pausing of Golgi bodies on microtubules regulates secretion of cellulose synthase complexes in Arabidopsis. *Plant Cell* 21, 1141–1154. doi: 10.1105/tpc.108.065334
- Davis, D. J., Kang, B. H., Heringer, A. S., Wilkop, T. E., and Drakakaki, G. (2016). Unconventional protein secretion in plants. *Methods Mol. Biol.* 1459, 47–63. doi: 10.1007/978-1-4939-3804-9_3
- De Caroli, M., Lenucci, M. S., Di Sansebastiano, G. P., Dalessandro, G., De Lorenzo, G., and Piro, G. (2011). Protein trafficking to the cell wall occurs through mechanisms distinguishable from default sorting in tobacco. *Plant J.* 65, 295–308. doi: 10.1111/j.1365-313X.2010.04421.x
- De Souza, A. P., Alvim Kamei, C. L., Torres, A. F., Pattathil, S., Hahn, M. G., Trindade, L. M., et al. (2015). How cell wall complexity influences saccharification efficiency in *Miscanthus sinensis*. *J. Exp. Bot.* 66, 4351–4365. doi: 10.1093/jxb/erv183
- Dettmer, J., Hong-Hermesdorf, A., Stierhof, Y. D., and Schumacher, K. (2006). Vacuolar H⁺-ATPase activity is required for endocytic and secretory trafficking in Arabidopsis. *Plant Cell* 18, 715–730. doi: 10.1105/tpc.105.037978
- Dhonukshe, P., Baluska, F., Schlicht, M., Hlavacka, A., Samaj, J., Friml, J., et al. (2006). Endocytosis of cell surface material mediates cell plate formation during plant cytokinesis. *Dev. Cell* 10, 137–150. doi: 10.1016/j.devcel.2005.11.015
- Di Rubbo, S., Irani, N. G., Kim, S. Y., Xu, Z. Y., Gadeyne, A., Dejonghe, W., et al. (2013). The clathrin adaptor complex AP-2 mediates endocytosis of brassinosteroid insensitive1 in Arabidopsis. *Plant Cell* 25, 2986–2997. doi: 10.1105/tpc.113.114058
- Drakakaki, G. (2015). Polysaccharide deposition during cytokinesis: challenges and future perspectives. *Plant Sci.* 236, 177–184. doi: 10.1016/j.plantsci.2015.03.018
- Drakakaki, G., Van De Ven, W., Pan, S., Miao, Y., Wang, J., Keinath, N. F., et al. (2012). Isolation and proteomic analysis of the SYP61 compartment reveal its role in exocytic trafficking in Arabidopsis. *Cell Res.* 22, 413–424. doi: 10.1038/cr.2011.129
- Ebine, K., Fujimoto, M., Okatani, Y., Nishiyama, T., Goh, T., Ito, E., et al. (2011). A membrane trafficking pathway regulated by the plant-specific RAB GTPase ARA6. *Nat. Cell Biol.* 13, 853–859. doi: 10.1038/ncb2270
- Ebine, K., and Ueda, T. (2015). Roles of membrane trafficking in plant cell wall dynamics. *Front. Plant Sci.* 6:878. doi: 10.3389/fpls.2015.00878
- El Kasmi, F., Krause, C., Hiller, U., Stierhof, Y. D., Mayer, U., Conner, L., et al. (2013). SNARE complexes of different composition jointly mediate membrane fusion in Arabidopsis cytokinesis. *Mol. Biol. Cell* 24, 1593–1601. doi: 10.1091/mbc.E13-02-0074
- Ellinger, D., Naumann, M., Falter, C., Zwikowics, C., Jamrow, T., Manisseri, C., et al. (2013). Elevated early callose deposition results in complete penetration resistance to powdery mildew in Arabidopsis. *Plant Physiol.* 161, 1433–1444. doi: 10.1104/pp.112.211011
- Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010). Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol.* 153, 403–419. doi: 10.1104/pp.110.156000
- Fernandez-Suarez, M., and Ting, A. Y. (2008). Fluorescent probes for super-resolution imaging in living cells. *Nat. Rev. Mol. Cell Biol.* 9, 929–943. doi: 10.1038/nrm2531
- Ford, K. L., Chin, T., Srivastava, V., Zeng, W., Doblin, M. S., Bulone, V., et al. (2016). Comparative "Golgi" proteome study of *Lolium multiflorum* and *Populus trichocarpa*. *Proteomes* 4:E23. doi: 10.3390/proteomes4030023
- Frankova, L., and Fry, S. C. (2013). Biochemistry and physiological roles of enzymes that 'cut and paste' plant cell-wall polysaccharides. *J. Exp. Bot.* 64, 3519–3550. doi: 10.1093/jxb/ert201
- Fruholz, S., Fassler, F., Kolukisaoglu, U., and Pimpl, P. (2018). Nanobody-triggered lockdown of VSRs reveals ligand reloading in the Golgi. *Nat. Commun.* 9:643. doi: 10.1038/s41467-018-02909-6
- Gadeyne, A., Sanchez-Rodriguez, C., Vanneste, S., Di Rubbo, S., Zauber, H., Vanneste, K., et al. (2014). The TPLATE adaptor complex drives clathrin-mediated endocytosis in plants. *Cell* 156, 691–704. doi: 10.1016/j.cell.2014.01.039
- Geelen, D., Leyman, B., Batoko, H., Di Sansebastiano, G. P., Moore, I., and Blatt, M. R. (2002). The abscisic acid-related SNARE homolog NtSyr1 contributes to secretion and growth: evidence from competition with its cytosolic domain. *Plant Cell* 14, 387–406. doi: 10.1105/tpc.010328
- Gendre, D., McFarlane, H. E., Johnson, E., Mouille, G., Sjodin, A., Oh, J., et al. (2013). Trans-Golgi network localized ECHIDNA/Ypt interacting protein complex is required for the secretion of cell wall polysaccharides in Arabidopsis. *Plant Cell* 25, 2633–2646. doi: 10.1105/tpc.113.112482
- Gendre, D., Oh, J., Boutte, Y., Best, J. G., Samuels, L., Nilsson, R., et al. (2011). Conserved Arabidopsis ECHIDNA protein mediates trans-Golgi-network trafficking and cell elongation. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8048–8053. doi: 10.1073/pnas.1018371108
- Griffiths, G., and Simons, K. (1986). The trans Golgi network: sorting at the exit site of the Golgi complex. *Science* 234, 438–443. doi: 10.1126/science.2945253
- Groen, A. J., Sancho-Andres, G., Breckels, L. M., Gatto, L., Aniento, F., and Lilley, K. S. (2014). Identification of trans-golgi network proteins in *Arabidopsis thaliana* root tissue. *J. Proteome Res.* 13, 763–776. doi: 10.1021/pr4008464
- Gu, F., Bringham, M., Combs, J. R., Yang, J., Bergmann, D. C., and Nielsen, E. (2016). Arabidopsis CSLD5 functions in cell plate formation in a cell cycle-dependent manner. *Plant Cell* 28, 1722–1737. doi: 10.1105/tpc.16.00203
- Gunl, M., Gille, S., and Pauly, M. (2010). OLIGO mass profiling (OLIMP) of extracellular polysaccharides. *J. Vis. Exp.* 40:2046. doi: 10.3791/2046
- Gunl, M., Kraemer, F., and Pauly, M. (2011a). Oligosaccharide mass profiling (OLIMP) of cell wall polysaccharides by MALDI-TOF/MS. *Methods Mol. Biol.* 715, 43–54. doi: 10.1007/978-1-61779-008-9_3
- Gunl, M., Neumetzler, L., Kraemer, F., De Souza, A., Schultink, A., Pena, M., et al. (2011b). AXYS8 encodes an alpha-fucosidase, underscoring the importance of apoplastic metabolism on the fine structure of Arabidopsis cell wall polysaccharides. *Plant Cell* 23, 4025–4040. doi: 10.1105/tpc.111.089193
- Guseman, J. M., Lee, J. S., Bogenschutz, N. L., Peterson, K. M., Virata, R. E., Xie, B., et al. (2010). Dysregulation of cell-to-cell connectivity and stomatal patterning by loss-of-function mutation in Arabidopsis *CHORUS* (*GLUCAN SYNTHASE-LIKE 8*). *Development* 137, 1731–1741. doi: 10.1242/dev.049197
- Gutierrez, R., Lindeboom, J. J., Paredes, A. R., Emons, A. M., and Ehrhardt, D. W. (2009). Arabidopsis cortical microtubules position cellulose synthase delivery to the plasma membrane and interact with cellulose synthase trafficking compartments. *Nat. Cell Biol.* 11, 797–806. doi: 10.1038/ncb1886
- Hachez, C., Laloux, T., Reinhardt, H., Cavez, D., Degand, H., Grefen, C., et al. (2014). Arabidopsis SNAREs SYP61 and SYP121 coordinate the trafficking of plasma membrane aquaporin PIP2;7 to modulate the cell membrane water permeability. *Plant Cell* 26, 3132–3147. doi: 10.1105/tpc.114.127159
- Hall, Q., and Cannon, M. C. (2002). The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in Arabidopsis. *Plant Cell* 14, 1161–1172. doi: 10.1105/tpc.010477
- Haraszti, R. A., Didiot, M. C., Sapp, E., Leszyk, J., Shaffer, S. A., Rockwell, H. E., et al. (2016). High-resolution proteomic and lipidomic analysis of exosomes and microvesicles from different cell sources. *J. Extracell. Vesicles* 5:32570. doi: 10.3402/jev.v5.32570
- Heard, W., Sklenar, J., Tome, D. F., Robatzek, S., and Jones, A. M. (2015). Identification of regulatory and cargo proteins of endosomal and secretory

- pathways in *Arabidopsis thaliana* by proteomic dissection. *Mol. Cell. Proteomics* 14, 1796–1813. doi: 10.1074/mcp.M115.050286
- Heese, M., Gansel, X., Sticher, L., Wick, P., Grebe, M., Granier, F., et al. (2001). Functional characterization of the KNOLLE-interacting t-SNARE AtSNAP33 and its role in plant cytokinesis. *J. Cell Biol.* 155, 239–249. doi: 10.1083/jcb.200107126
- Herce, H. D., Schumacher, D., Schneider, A. F. L., Ludwig, A. K., Mann, F. A., Fillies, M., et al. (2017). Cell-permeable nanobodies for targeted immunolabelling and antigen manipulation in living cells. *Nat. Chem.* 9, 762–771. doi: 10.1038/nchem.2811
- Hofmann, N. R. (2017). Basal versus nonbasal polarity: different endomembrane trafficking pathways establish different patterns. *Plant Cell* 29:1. doi: 10.1105/tpc.17.00065
- Hoogenboom, J., Berghuis, N., Cramer, D., Geurts, R., Zuilhof, H., and Wennekes, T. (2016). Direct imaging of glycans in *Arabidopsis* roots via click labeling of metabolically incorporated azido-monosaccharides. *BMC Plant Biol.* 16:220. doi: 10.1186/s12870-016-0907-0
- Johnson, K. L., Cassin, A. M., Lonsdale, A., Bacic, A., Doblin, M. S., and Schultz, C. J. (2017a). Pipeline to identify hydroxyproline-rich glycoproteins. *Plant Physiol.* 174, 886–903. doi: 10.1104/pp.17.00294
- Johnson, K. L., Cassin, A. M., Lonsdale, A., Wong, G. K., Soltis, D. E., Miles, N. W., et al. (2017b). Insights into the evolution of hydroxyproline-rich glycoproteins from 1000 plant transcriptomes. *Plant Physiol.* 174, 904–921. doi: 10.1104/pp.17.00295
- Jurgens, G. (2005). Cytokinesis in higher plants. *Annu. Rev. Plant Biol.* 56, 281–299. doi: 10.1146/annurev.arplant.55.031903.141636
- Kang, B. H., Nielsen, E., Preuss, M. L., Mastronarde, D., and Staehelin, L. A. (2011). Electron tomography of RabA4b- and PI-4K β 1-labeled trans Golgi network compartments in *Arabidopsis*. *Traffic* 12, 313–329. doi: 10.1111/j.1600-0854.2010.01146.x
- Karnahl, M., Park, M., Krause, C., Hiller, U., Mayer, U., Stierhof, Y. D., et al. (2018). Functional diversification of *Arabidopsis* SEC1-related SM proteins in cytokinetic and secretory membrane fusion. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6309–6314. doi: 10.1073/pnas.1722611115
- Keegstra, K. (2010). Plant cell walls. *Plant Physiol.* 154, 483–486. doi: 10.1104/pp.110.161240
- Kim, S. J., and Bassham, D. C. (2011). TNO1 is involved in salt tolerance and vacuolar trafficking in *Arabidopsis*. *Plant Physiol.* 156, 514–526. doi: 10.1104/pp.110.168963
- Kim, S. J., and Brandizzi, F. (2016). The plant secretory pathway for the trafficking of cell wall polysaccharides and glycoproteins. *Glycobiology* 26, 940–949. doi: 10.1093/glycob/cwv044
- Kim, S. J., Zemelis-Durfee, S., Jensen, J. K., Wilkerson, C. G., Keegstra, K., and Brandizzi, F. (2018). In the grass species *Brachypodium distachyon*, the production of mixed-linkage (1,3;1,4)-beta-glucan (MLG) occurs in the Golgi apparatus. *Plant J.* 93, 1062–1075. doi: 10.1111/tpj.13830
- Komis, G., Novak, D., Ovecka, M., Samajova, O., and Samaj, J. (2018). Advances in imaging plant cell dynamics. *Plant Physiol.* 176, 80–93. doi: 10.1104/pp.17.00962
- Kong, Y., Peña, M. J., Renna, L., Avci, U., Pattathil, S., Tuomivaara, S. T., et al. (2015). Galactose-depleted xyloglucan is dysfunctional and leads to dwarfism in *Arabidopsis*. *Plant Physiol.* 167, 1296–1306. doi: 10.1104/pp.114.255943
- Kong, Z., Ioki, M., Braybrook, S., Li, S., Ye, Z.-H., Julie Lee, Y.-R., et al. (2015). Kinesin-4 functions in vesicular transport on cortical microtubules and regulates cell wall mechanics during cell elongation in plants. *Mol. Plant* 8, 1011–1023. doi: 10.1016/j.molp.2015.01.004
- Kulich, I., Cole, R., Drdova, E., Cvrckova, F., Soukup, A., Fowler, J., et al. (2010). *Arabidopsis* exocyst subunits SEC8 and EXO70A1 and exocyst interactor ROH1 are involved in the localized deposition of seed coat pectin. *New Phytol.* 188, 615–625. doi: 10.1111/j.1469-8137.2010.03372.x
- Kulich, I., Vojtikova, Z., Sabol, P., Ortmannova, J., Nedela, V., Tihlarikova, E., et al. (2018). Exocyst Subunit EXO70H4 has a specific role in callose synthase secretion and silica accumulation. *Plant Physiol.* 176, 2040–2051. doi: 10.1104/pp.17.01693
- Lampugnani, E. R., Khan, G. A., Somssich, M., and Persson, S. (2018). Building a plant cell wall at a glance. *J. Cell Sci.* 131:jcs.207373. doi: 10.1242/jcs.207373
- Lauber, M. H., Waizenegger, I., Steinmann, T., Schwarz, H., Mayer, U., Hwang, I., et al. (1997). The *Arabidopsis* KNOLLE protein is a cytokinesis-specific syntaxin. *J. Cell Biol.* 139, 1485–1493. doi: 10.1083/jcb.139.6.1485
- Law, A. H., Chow, C. M., and Jiang, L. (2012). Secretory carrier membrane proteins. *Protoplasma* 249, 269–283. doi: 10.1007/s00709-011-0295-0
- Lee, Y. R., and Liu, B. (2013). The rise and fall of the phragmoplast microtubule array. *Curr. Opin. Plant Biol.* 16, 757–763. doi: 10.1016/j.pbi.2013.10.008
- Lei, L., Singh, A., Bashline, L., Li, S., Yingling, Y. G., and Gu, Y. (2015). CELLULOSE SYNTHASE INTERACTIVE1 is required for fast recycling of cellulose synthase complexes to the plasma membrane in *Arabidopsis*. *Plant Cell* 27, 2926–2940. doi: 10.1105/tpc.15.00442
- Leucci, M. R., Di Sansebastiano, G. P., Gigante, M., Dalessandro, G., and Piro, G. (2007). Secretion marker proteins and cell-wall polysaccharides move through different secretory pathways. *Planta* 225, 1001–1017. doi: 10.1007/s00425-006-0407-9
- Li, R., Rodriguez-Furlan, C., Wang, J., Van De Ven, W., Gao, T., Raikhel, N. V., et al. (2017). Different Endomembrane trafficking pathways establish apical and basal polarities. *Plant Cell* 29, 90–108. doi: 10.1105/tpc.16.00524
- Liu, L., Guo, Z., Tieu, Q., Castle, A., and Castle, D. (2002). Role of secretory carrier membrane protein SCAMP2 in granule exocytosis. *Mol. Biol. Cell* 13, 4266–4278. doi: 10.1091/mbc.e02-03-0136
- Liu, L., Liao, H., Castle, A., Zhang, J., Casanova, J., Szabo, G., et al. (2005). SCAMP2 interacts with Arf6 and phospholipase D1 and links their function to exocytotic fusion pore formation in PC12 cells. *Mol. Biol. Cell* 16, 4463–4472. doi: 10.1091/mbc.e05-03-0231
- Lunn, D., Gaddipati, S. R., Tucker, G. A., and Lycett, G. W. (2013a). Null mutants of individual RABA genes impact the proportion of different cell wall components in stem tissue of *Arabidopsis thaliana*. *PLoS One* 8:e75724. doi: 10.1371/journal.pone.0075724
- Lunn, D., Phan, T. D., Tucker, G. A., and Lycett, G. W. (2013b). Cell wall composition of tomato fruit changes during development and inhibition of vesicle trafficking is associated with reduced pectin levels and reduced softening. *Plant Physiol. Biochem.* 66, 91–97. doi: 10.1016/j.plaphy.2013.02.005
- Luo, Y., Scholl, S., Doering, A., Zhang, Y., Irani, N. G., Di Rubbo, S., et al. (2015). V-ATPase activity in the TGN/EE is required for exocytosis and recycling in *Arabidopsis*. *Nat. Plants* 1:15094. doi: 10.1038/nplants.2015.94
- Lynch, M. A., and Staehelin, L. A. (1992). Domain-specific and cell type-specific localization of two types of cell wall matrix polysaccharides in the clover root tip. *J. Cell Biol.* 118, 467–479. doi: 10.1083/jcb.118.2.467
- McCann, M. C., Wells, B., and Roberts, K. (1992). Complexity in the spatial localization and length distribution of plant cell-wall matrix polysaccharides. *J. Microscopy Oxford* 166, 123–136. doi: 10.1111/j.1365-2818.1992.tb01511.x
- McFarlane, H. E., Doring, A., and Persson, S. (2014). The cell biology of cellulose synthesis. *Annu. Rev. Plant Biol.* 65, 69–94. doi: 10.1146/annurev-arplant-050213-040240
- McFarlane, H. E., Young, R. E., Wasteneys, G. O., and Samuels, A. L. (2008). Cortical microtubules mark the mucilage secretion domain of the plasma membrane in *Arabidopsis* seed coat cells. *Planta* 227, 1363–1375. doi: 10.1007/s00425-008-0708-2
- McMichael, C. M., and Bednarek, S. Y. (2013). Cytoskeletal and membrane dynamics during higher plant cytokinesis. *New Phytol.* 197, 1039–1057. doi: 10.1111/nph.12122
- Meents, M. J., Watanabe, Y., and Samuels, A. L. (2018). The cell biology of secondary cell wall biosynthesis. *Ann. Bot.* 121, 1107–1125. doi: 10.1093/aob/mcy005
- Miart, F., Desprez, T., Biot, E., Morin, H., Belcram, K., Hofte, H., et al. (2014). Spatio-temporal analysis of cellulose synthesis during cell plate formation in *Arabidopsis*. *Plant J.* 77, 71–84. doi: 10.1111/tpj.12362
- Mishin, A. S., Belousov, V. V., Solntsev, K. M., and Lukyanov, K. A. (2015). Novel uses of fluorescent proteins. *Curr. Opin. Chem. Biol.* 27, 1–9. doi: 10.1016/j.cbpa.2015.05.002
- Moller, I., Marcus, S. E., Haeger, A., Verhertbruggen, Y., Verhoef, R., Schols, H., et al. (2008). High-throughput screening of monoclonal antibodies against plant cell wall glycans by hierarchical clustering of their carbohydrate microarray binding profiles. *Glycoconj. J.* 25, 37–48. doi: 10.1007/s10719-007-9059-7
- Moore, P. J., Darvill, A. G., Albersheim, P., and Staehelin, L. A. (1986). Immunogold localization of xyloglucan and rhamnogalacturonan I in the cell

- walls of suspension-cultured sycamore cells. *Plant Physiol.* 82, 787–794. doi: 10.1104/pp.82.3.787
- Moore, P. J., and Staehelin, L. A. (1988). Immunogold localization of the cell-wall-matrix polysaccharides rhamnogalacturonan I and xyloglucan during cell expansion and cytokinesis in *Trifolium pratense* L.; implication for secretory pathways. *Planta* 174, 433–445. doi: 10.1007/BF00634471
- Moore, P. J., Swords, K. M., Lynch, M. A., and Staehelin, L. A. (1991). Spatial organization of the assembly pathways of glycoproteins and complex polysaccharides in the Golgi apparatus of plants. *J. Cell Biol.* 112, 589–602. doi: 10.1083/jcb.112.4.589
- Mravec, J., Kracun, S. K., Rydahl, M. G., Westereng, B., Miart, F., Clausen, M. H., et al. (2014). Tracking developmentally regulated post-synthetic processing of homogalacturonan and chitin using reciprocal oligosaccharide probes. *Development* 141, 4841–4850. doi: 10.1242/dev.113365
- Nguema-Ona, E., Vire-Gibouin, M., Gotte, M., Plancot, B., Lerouge, P., Bardor, M., et al. (2014). Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. *Front. Plant Sci.* 5:499. doi: 10.3389/fpls.2014.00499
- Nikolovski, N., Rubtsov, D., Segura, M. P., Miles, G. P., Stevens, T. J., Dunkley, T. P., et al. (2012). Putative glycosyltransferases and other plant golgi apparatus proteins are revealed by LOPIT proteomics. *Plant Physiol.* 160, 1037–1051. doi: 10.1104/pp.112.204263
- Obel, N., Erben, V., Schwarz, T., Kuhnelt, S., Fodor, A., and Pauly, M. (2009). Microanalysis of plant cell wall polysaccharides. *Mol. Plant* 2, 922–932. doi: 10.1093/mp/ssp046
- Otegui, M. S., Mastrorade, D. N., Kang, B. H., Bednarek, S. Y., and Staehelin, L. A. (2001). Three-dimensional analysis of syncytial-type cell plates during endosperm cellularization visualized by high resolution electron tomography. *Plant Cell* 13, 2033–2051. doi: 10.1105/tpc.13.9.2033
- Otegui, M. S., and Staehelin, L. A. (2004). Electron tomographic analysis of post-meiotic cytokinesis during pollen development in *Arabidopsis thaliana*. *Planta* 218, 501–515. doi: 10.1007/s00425-003-1125-1
- Paredez, A. R., Somerville, C. R., and Ehrhardt, D. W. (2006). Visualization of cellulose synthase demonstrates functional association with microtubules. *Science* 312, 1491–1495. doi: 10.1126/science.1126551
- Park, S., Szumlanski, A. L., Gu, F., Guo, F., and Nielsen, E. (2011). A role for CSLD3 during cell-wall synthesis in apical plasma membranes of tip-growing root-hair cells. *Nat. Cell Biol.* 13, 973–980. doi: 10.1038/ncb2294
- Parsons, H. T., and Lilley, K. S. (2018). Mass spectrometry approaches to study plant endomembrane trafficking. *Semin. Cell Dev. Biol.* 80, 123–132. doi: 10.1016/j.semcdb.2017.10.014
- Pattathil, S., Avci, U., Baldwin, D., Swennes, A. G., McGill, J. A., Popper, Z., et al. (2010). A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol.* 153, 514–525. doi: 10.1104/pp.109.151985
- Pattathil, S., Avci, U., Miller, J. S., and Hahn, M. G. (2012). Immunological approaches to plant cell wall and biomass characterization: glycome profiling. *Methods Mol. Biol.* 908, 61–72. doi: 10.1007/978-1-61779-956-3_6
- Pauly, M., and Keegstra, K. (2016). Biosynthesis of the plant cell wall matrix Polysaccharide Xyloglucan. *Annu. Rev. Plant Biol.* 67, 235–259. doi: 10.1146/annurev-arplant-043015-112222
- Poulsen, C. P., Dilokpimol, A., and Geshi, N. (2015). Arabinogalactan biosynthesis: implication of AtGALT29A enzyme activity regulated by phosphorylation and co-localized enzymes for nucleotide sugar metabolism in the compartments outside of the Golgi apparatus. *Plant Signal. Behav.* 10:e984524. doi: 10.4161/15592324.2014.984524
- Poulsen, C. P., Dilokpimol, A., Mouille, G., Burrow, M., and Geshi, N. (2014). Arabinogalactan glycosyltransferases target to a unique subcellular compartment that may function in unconventional secretion in plants. *Traffic* 15, 1219–1234. doi: 10.1111/tra.12203
- Ravikumar, R., Steiner, A., and Assaad, F. F. (2017). Multisubunit tethering complexes in higher plants. *Curr. Opin. Plant Biol.* 40, 97–105. doi: 10.1016/j.pbi.2017.08.009
- Rosquete, M. R., Davis, D. J., and Drakakaki, G. (2018). The plant trans-golgi network: not just a matter of distinction. *Plant Physiol.* 176, 187–198. doi: 10.1104/pp.17.01239
- Roth, J., Taatjes, D. J., Lucocq, J. M., Weinstein, J., and Paulson, J. C. (1985). Demonstration of an extensive trans-tubular network continuous with the Golgi apparatus stack that may function in glycosylation. *Cell* 43, 287–295. doi: 10.1016/0092-8674(85)90034-0
- Ruprecht, C., Bartetzko, M. P., Senf, D., Dallabernadina, P., Boos, I., Andersen, M. C. F., et al. (2017). A synthetic glycan microarray enables epitope mapping of plant cell wall glycan-directed antibodies. *Plant Physiol.* 175, 1094–1104. doi: 10.1104/pp.17.00737
- Rybak, K., Steiner, A., Synek, L., Klaeger, S., Kulich, I., Facher, E., et al. (2014). Plant cytokinesis is orchestrated by the sequential action of the TRAPP II and exocyst tethering complexes. *Dev. Cell* 29, 607–620. doi: 10.1016/j.devcel.2014.04.029
- Sampedro, J., Gianzo, C., Iglesias, N., Guitian, E., Revilla, G., and Zarra, I. (2012). AtBGAL10 is the main xyloglucan beta-galactosidase in Arabidopsis, and its absence results in unusual xyloglucan subunits and growth defects. *Plant Physiol.* 158, 1146–1157. doi: 10.1104/pp.111.192195
- Samuels, A. L., Giddings, T. H. Jr., and Staehelin, L. A. (1995). Cytokinesis in tobacco BY-2 and root tip cells: a new model of cell plate formation in higher plants. *J. Cell Biol.* 130, 1345–1357. doi: 10.1083/jcb.130.6.1345
- Sanchez-Rodriguez, C., Shi, Y., Kesten, C., Zhang, D., Sancho-Andres, G., Ivakov, A., et al. (2018). The cellulose synthases are cargo of the TPLATE adaptor complex. *Mol. Plant* 11, 346–349. doi: 10.1016/j.molp.2017.11.012
- Sanderfoot, A. A., Assaad, F. F., and Raikhel, N. V. (2000). The Arabidopsis genome. An abundance of soluble N-ethylmaleimide-sensitive factor adaptor protein receptors. *Plant Physiol.* 124, 1558–1569. doi: 10.1104/pp.124.4.1558
- Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315
- Segui-Simarro, J. M., Austin, J. R., White, E. A., and Staehelin, L. A. (2004). Electron tomographic analysis of somatic cell plate formation in meristematic cells of Arabidopsis preserved by high-pressure freezing. *Plant Cell* 16, 836–856. doi: 10.1105/tpc.017749
- Showalter, A. M. (1993). Structure and function of plant cell wall proteins. *Plant Cell* 5, 9–23. doi: 10.1105/tpc.5.1.9
- Showalter, A. M., and Basu, D. (2016a). Extensin and arabinogalactan-protein biosynthesis: glycosyltransferases, research challenges, and biosensors. *Front. Plant Sci.* 7:814. doi: 10.3389/fpls.2016.00814
- Showalter, A. M., and Basu, D. (2016b). Glycosylation of arabinogalactan-proteins essential for development in Arabidopsis. *Commun. Integr. Biol.* 9:e1177687. doi: 10.1080/19420889.2016.1177687
- Smertenko, A., Assaad, F., Baluska, F., Bezanilla, M., Buschmann, H., Drakakaki, G., et al. (2017). Plant cytokinesis: terminology for structures and processes. *Trends Cell Biol.* 27, 885–894. doi: 10.1016/j.tcb.2017.08.008
- Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., et al. (2004). Toward a systems approach to understanding plant cell walls. *Science* 306, 2206–2211. doi: 10.1126/science.1102765
- Staehelin, L. A., and Hepler, P. K. (1996). Cytokinesis in higher plants. *Cell* 84, 821–824. doi: 10.1016/S0092-8674(00)81060-0
- Tamura, K., Shimada, T., Kondo, M., Nishimura, M., and Hara-Nishimura, I. (2005). KATAMARI/MURUS3 is a novel Golgi membrane protein that is required for endomembrane organization in Arabidopsis. *Plant Cell* 17, 1764–1776. doi: 10.1105/tpc.105.031930
- Tan, L., Showalter, A. M., Egelund, J., Hernandez-Sanchez, A., Doblin, M. S., and Bacic, A. (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. *Front. Plant Sci.* 3:140. doi: 10.3389/fpls.2012.00140
- Thiele, K., Wanner, G., Kindzierski, V., Jurgens, G., Mayer, U., Pachel, F., et al. (2009). The timely deposition of callose is essential for cytokinesis in Arabidopsis. *Plant J.* 58, 13–26. doi: 10.1111/j.1365-313X.2008.03760.x
- Toller, A., Brownfield, L., Neu, C., Twell, D., and Schulze-Lefert, P. (2008). Dual function of Arabidopsis glucan synthase-like genes GSL8 and GSL10 in male gametophyte development and plant growth. *Plant J.* 54, 911–923. doi: 10.1111/j.1365-313X.2008.03462.x
- Toyooka, K., Goto, Y., Asatsuma, S., Koizumi, M., Mitsui, T., and Matsuoka, K. (2009). A mobile secretory vesicle cluster involved in mass transport from the Golgi to the plant cell exterior. *Plant Cell* 21, 1212–1229. doi: 10.1105/tpc.108.058933
- Toyooka, K., and Kang, B. H. (2014). Reconstructing plant cells in 3D by serial section electron tomography. *Methods Mol. Biol.* 1080, 159–170. doi: 10.1007/978-1-62703-643-6_13

- Ueda, T., Sato, M. H., and Uemura, T. (2012). "The Role of RAB GTPases and SNARE proteins in plant endocytosis and post-golgi trafficking," in *Endocytosis in Plants*, ed. J. Šamaj (Berlin: Springer), 201–216.
- van de Meene, A. M., Doblin, M. S., and Bacic, A. (2017). The plant secretory pathway seen through the lens of the cell wall. *Protoplasma* 254, 75–94. doi: 10.1007/s00709-016-0952-4
- van Oostende-Triplet, C., Guillet, D., Triplet, T., Pandzic, E., Wiseman, P. W., and Geitmann, A. (2017). Vesicle dynamics during plant cell cytokinesis reveals distinct developmental phases. *Plant Physiol.* 174, 1544–1558. doi: 10.1104/pp.17.00343
- Vernoud, V., Horton, A. C., Yang, Z., and Nielsen, E. (2013). Analysis of the small GTPase gene superfamily of Arabidopsis. *Plant Physiol.* 131, 1191–1208. doi: 10.1104/pp.013052
- Viotti, C., Bubeck, J., Stierhof, Y. D., Krebs, M., Langhans, M., Van Den Berg, W., et al. (2010). Endocytic and secretory traffic in Arabidopsis merge in the trans-Golgi network/early endosome, an independent and highly dynamic organelle. *Plant Cell* 22, 1344–1357. doi: 10.1105/tpc.109.072637
- Voiniciuc, C., Pauly, M., and Usadel, B. (2018). Monitoring polysaccharide dynamics in the plant cell wall. *Plant Physiol.* 176, 2590–2600. doi: 10.1104/pp.17.01776
- Vukasinovic, N., and Zarsky, V. (2016). Tethering complexes in the arabidopsis endomembrane system. *Front. Cell Dev. Biol.* 4:46. doi: 10.3389/fcell.2016.00046
- Wallace, M. L., Anderson, S. J., Mazumdar, S., Kong, L., and Mulsant, B. H. (2012). Incorporating temporal features of repeatedly measured covariates into tree-structured survival models. *Biom. J.* 54, 181–196. doi: 10.1002/bimj.201100013
- Wang, B., McClosky, D. D., Anderson, C. T., and Chen, G. (2016). Synthesis of a suite of click-compatible sugar analogs for probing carbohydrate metabolism. *Carbohydr. Res.* 433, 54–62. doi: 10.1016/j.carres.2016.07.012
- Wang, H., Zhuang, X., Wang, X., Law, A. H., Zhao, T., Du, S., et al. (2016). A distinct pathway for polar exocytosis in plant cell wall formation. *Plant Physiol.* 172, 1003–1018.
- Wang, J., Ding, Y., Hillmer, S., Miao, Y., Lo, S. W., Wang, X., et al. (2010). EXPO, an exocyst-positive organelle distinct from multivesicular endosomes and autophagosomes, mediates cytosol to cell wall exocytosis in Arabidopsis and tobacco cells. *Plant Cell* 22, 4009–4030. doi: 10.1105/tpc.110.080697
- Wang, P., Chen, X., Goldbeck, C., Chung, E., and Kang, B. H. (2017). A distinct class of vesicles derived from the trans-Golgi mediates secretion of xylogalacturonan in the root border cell. *Plant J.* 92, 596–610. doi: 10.1111/tjp.13704
- Wang, X., Chung, K. P., Lin, W., and Jiang, L. (2017). Protein secretion in plants: conventional and unconventional pathways and new techniques. *J. Exp. Bot.* 69, 21–37. doi: 10.1093/jxb/erx262
- Wattelet-Boyer, V., Brocard, L., Jonsson, K., Esnay, N., Joubes, J., Domergue, F., et al. (2016). Enrichment of hydroxylated C24- and C26-acyl-chain sphingolipids mediates PIN2 apical sorting at trans-Golgi network subdomains. *Nat. Commun.* 7:12788. doi: 10.1038/ncomms12788
- Wong, M., and Munro, S. (2014). Membrane trafficking. The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins. *Science* 346:1256898. doi: 10.1126/science.1256898
- Woollard, A. A., and Moore, I. (2008). The functions of Rab GTPases in plant membrane traffic. *Curr. Opin. Plant Biol.* 11, 610–619. doi: 10.1016/j.pbi.2008.09.010
- Worden, N., Park, E., and Drakakaki, G. (2012). Trans-Golgi network-an intersection of trafficking cell wall components(f). *J. Integr. Plant Biol.* 54, 875–886. doi: 10.1111/j.1744-7909.2012.01179.x
- Worden, N., Wilkop, T. E., Esteve, V. E., Jeannotte, R., Lathe, R., Vernhettes, S., et al. (2015). CESA TRAFFICKING INHIBITOR inhibits cellulose deposition and interferes with the trafficking of cellulose synthase complexes and their associated proteins KORRIGAN1 and POM2/CELLULOSE SYNTHASE INTERACTIVE PROTEIN1. *Plant Physiol.* 167, 381–393. doi: 10.1104/pp.114.249003
- Young, R. E., McFarlane, H. E., Hahn, M. G., Western, T. L., Haughn, G. W., and Samuels, A. L. (2008). Analysis of the Golgi apparatus in Arabidopsis seed coat cells during polarized secretion of pectin-rich mucilage. *Plant Cell* 20, 1623–1638. doi: 10.1105/tpc.108.058842
- Yu, M., and Zhao, J. (2012). The cytological changes of tobacco zygote and proembryo cells induced by beta-glucosyl Yariv reagent suggest the involvement of arabinogalactan proteins in cell division and cell plate formation. *BMC Plant Biol.* 12:126. doi: 10.1186/1471-2229-12-126
- Zeng, W., Ford, K. L., Bacic, A., and Heazlewood, J. L. (2018). N-linked Glycan Micro-heterogeneity in Glycoproteins of Arabidopsis. *Mol. Cell. Proteomics* 17, 413–421. doi: 10.1074/mcp.RA117.000165
- Zhang, G. F., and Staehelin, L. A. (1992). Functional compartmentation of the golgi apparatus of plant cells: immunocytochemical analysis of high-pressure frozen- and freeze-substituted sycamore maple suspension culture cells. *Plant Physiol.* 99, 1070–1083. doi: 10.1104/pp.99.3.1070
- Zhang, J., Hill, D. R., and Sylvester, A. W. (2007). Diversification of the RAB guanosine triphosphatase family in dicots and monocots. *J. Integr. Plant Biol.* 49, 1129–1141. doi: 10.1111/j.1672-9072.2007.00520.x
- Zhang, L., Zhang, H., Liu, P., Hao, H., Jin, J. B., and Lin, J. (2011). Arabidopsis R-SNARE proteins VAMP721 and VAMP722 are required for cell plate formation. *PLoS One* 6:e26129. doi: 10.1371/journal.pone.0026129
- Zhen, Y., and Stenmark, H. (2015). Cellular functions of Rab GTPases at a glance. *J. Cell Sci.* 128, 3171–3176. doi: 10.1242/jcs.166074
- Zheng, H., Bednarek, S. Y., Sanderfoot, A. A., Alonso, J., Ecker, J. R., and Raikhel, N. V. (2002). NPSN11 is a cell plate-associated SNARE protein that interacts with the syntaxin KNOLLE. *Plant Physiol.* 129, 530–539. doi: 10.1104/pp.003970
- Zheng, Y., Wang, X., Chen, Y., Wagner, E., and Cosgrove, D. J. (2018). Xyloglucan in the primary cell wall: assessment by FESEM, selective enzyme digestions and nanogold affinity tags. *Plant J.* 93, 211–226. doi: 10.1111/tjp.13778
- Zhu, C., Ganguly, A., Baskin, T. I., McClosky, D. D., Anderson, C. T., Foster, C., et al. (2015). The fragile Fiber1 kinesin contributes to cortical microtubule-mediated trafficking of cell wall components. *Plant Physiol.* 167, 780–792. doi: 10.1104/pp.114.251462
- Zhu, X., Li, S., Pan, S., Xin, X., and Gu, Y. (2018). CSI1, PATROL1, and exocyst complex cooperate in delivery of cellulose synthase complexes to the plasma membrane. *Proc. Natl. Acad. Sci. U.S.A.* 115, E3578–E3587. doi: 10.1073/pnas.1800182115
- Zhu, Y., and Chen, X. (2017). Expanding the scope of metabolic glycan labeling in *Arabidopsis thaliana*. *Chembiochem* 18, 1286–1296. doi: 10.1002/cbic.201700069
- Zuo, J., Niu, Q. W., Nishizawa, N., Wu, Y., Kost, B., and Chua, N. H. (2000). KORRIGAN, an Arabidopsis endo-1,4-beta-glucanase, localizes to the cell plate by polarized targeting and is essential for cytokinesis. *Plant Cell* 12, 1137–1152. doi: 10.1105/tpc.12.7.1137

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Sinclair, Rosquete and Drakakaki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

Edited by:

Laura Elizabeth Bartley,
The University of Oklahoma,
United States

Reviewed by:

Rui Shi,
North Carolina State University,
United States
Jenny C. Mortimer,
Lawrence Berkeley National
Laboratory (LBNL), United States

*Correspondence:

Udaya C. Kalluri
kalluriudayc@ornl.gov

† Present address:

Raghuram Badmi,
Division of Biotechnology and Plant
Health, Norsk Institutt for Bioøkonomi,
Ås, Norway
Raja S. Payyavula,
Eurofins Lancaster Laboratories,
Richmond, VA, United States
Hao-Bo Guo,
College of Engineering and Computer
Science, SimCenter, University
of Tennessee at Chattanooga,
Chattanooga, TN, United States
Robert W. Sykes,
Los Alamos National Laboratory,
Los Alamos, NM, United States

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 29 June 2018

Accepted: 26 October 2018

Published: 05 December 2018

Citation:

Badmi R, Payyavula RS, Bali G,
Guo H-B, Jawdy SS, Gunter LE,
Yang X, Winkler KA, Collins C,
Rottmann WH, Yee K, Rodriguez M Jr,
Sykes RW, Decker SR, Davis MF,
Ragauskas AJ, Tuskan GA and
Kalluri UC (2018) A New
Calmodulin-Binding Protein
Expresses in the Context
of Secondary Cell Wall Biosynthesis
and Impacts Biomass Properties
in *Populus*. *Front. Plant Sci.* 9:1669.
doi: 10.3389/fpls.2018.01669

A New Calmodulin-Binding Protein Expresses in the Context of Secondary Cell Wall Biosynthesis and Impacts Biomass Properties in *Populus*

Raghuram Badmi^{1,2†}, Raja S. Payyavula^{1,2†}, Garima Bali^{1,3}, Hao-Bo Guo^{4†}, Sara S. Jawdy^{1,2}, Lee E. Gunter^{1,2}, Xiaohan Yang^{1,2}, Kimberly A. Winkler⁵, Cassandra Collins⁵, William H. Rottmann⁵, Kelsey Yee^{1,2}, Miguel Rodriguez Jr.^{1,2}, Robert W. Sykes^{1,6†}, Stephen R. Decker^{1,6}, Mark F. Davis^{1,6}, Arthur J. Ragauskas^{1,7}, Gerald A. Tuskan^{1,2} and Udaya C. Kalluri^{1,2*}

¹ BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ² The Center for Bioenergy Innovation and Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ³ Georgia Institute of Technology, Atlanta, GA, United States, ⁴ Department of Biochemistry and Cellular and Molecular Biology, The University of Tennessee, Knoxville, TN, United States, ⁵ ArborGen Inc., Ridgeville, SC, United States, ⁶ National Renewable Energy Laboratory, Golden, CO, United States, ⁷ Department of Chemical and Biomolecular Engineering, The University of Tennessee, Knoxville, TN, United States

A greater understanding of biosynthesis, signaling and regulatory pathways involved in determining stem growth and secondary cell wall chemistry is important for enabling pathway engineering and genetic optimization of biomass properties. The present study describes a new functional role of *PdIQD10*, a *Populus* gene belonging to the IQ67-Domain1 family of *IQD* genes, in impacting biomass formation and chemistry. Expression studies showed that *PdIQD10* has enhanced expression in developing xylem and tension-stressed tissues in *Populus deltoides*. Molecular dynamics simulation and yeast two-hybrid interaction experiments suggest interactions with two calmodulin proteins, CaM247 and CaM014, supporting the sequence-predicted functional role of the *PdIQD10* as a calmodulin-binding protein. *PdIQD10* was found to interact with specific *Populus* isoforms of the Kinesin Light Chain protein family, shown previously to function as microtubule-guided, cargo binding and delivery proteins in *Arabidopsis*. Subcellular localization studies showed that *PdIQD10* localizes in the nucleus and plasma membrane regions. Promoter-binding assays suggest that a known master transcriptional regulator of secondary cell wall biosynthesis (*PdWND1B*) may be upstream of an *HD-ZIP III* gene that is in turn upstream of *PdIQD10* gene in the transcriptional network. RNAi-mediated downregulation of *PdIQD10* expression resulted in plants with altered biomass properties including higher cellulose, wall glucose content and greater biomass quantity. These results present evidence in support of a new functional role for an *IQD* gene family member, *PdIQD10*, in secondary cell wall biosynthesis and biomass formation in *Populus*.

Keywords: secondary cell wall, cellulose, signaling, kinesin, calcium-calmodulin, biomass, *IQD*, *Populus*

INTRODUCTION

Plant cell walls play essential structural and functional roles as strength-conferring and signal-responsive organelles made up of polysaccharides (cellulose, hemicellulose and pectin); polyphenols; lignin and polypeptides (wall glycoproteins and wall associated proteins). A greater understanding of the biosynthesis, signaling and regulatory pathways involved in the coordinated deposition and remodeling of plant cell walls is of great interest to fundamental plant science as well as applied biomass improvement research fields (Demura and Ye, 2010; Mizrachi et al., 2012). Improvements in the precision of pathway engineering efforts and acceleration in breeding approaches will need complementation of the existing knowledgebase of known cell wall pathway genes with functional characterization of genes that co-express with known marker cell wall pathway genes (Ruprecht et al., 2011; Sibout et al., 2017).

The primary characteristics that render a given unit of lignocellulosic biomass more amenable to conversion to biofuel include higher content and greater accessibility to the wall glucose or cellulose, reduced lignin content and cross-linkages among cellulose, lignin and hemicellulose (Klemm et al., 2005; McCann and Carpita, 2008; Mansfield, 2009; Scheller and Ulvskov, 2010; Fu et al., 2011; Ding et al., 2012; Cragg et al., 2015; Busse-Wicher et al., 2016; Gall et al., 2017). Toward the goal of identifying high potential candidate genes involved in signaling and regulation of biosynthesis of cell walls with high substrate content, we studied overlaps in expression data collected from developmental (xylem development) and physiological response (tension stress) phases when more wood cells, thicker cell walls and walls with higher cellulose content are formed. These include three studies; firstly, a tension stress response profiling study to characterize transcriptome response to bending/leaning in stems of *Populus*, while undergoing enhanced xylem cell proliferation and cellulose production in new cell wall layers composed of over 90% cellulose (Abraham et al., 2012). Secondly, a xylem proteomics study to profile a tissue type where enhanced cellulose and secondary cell wall production occurs (Kalluri et al., 2009). Lastly, a co-expression network analysis study to identify tightly co-expressed genes with a *Populus* secondary *CesA* gene (Yang et al., 2011). *PdIQD10* (Potri.001G375700) gene, predicted to

code for a calmodulin-binding protein, was identified as highly upregulated during phases of enhanced cellulose biosynthesis.

Calcium (Ca^{2+})-related signaling pathways constitute a major cellular signaling mechanism in response to a stress or developmental trigger and are prevalent among all eukaryotes (Clapham, 2007; Dodd et al., 2010). Ca^{2+} ion levels serve as important secondary messengers by inducing intracellular dose-dependent signals that are transduced or decoded via Ca^{2+} sensor proteins. Ca^{2+} -dependent protein kinases (CDPKs), CaMs and CMLs, and CBLs are the three major Ca^{2+} sensor proteins in plants. CDPKs have an intrinsic kinase domain that can directly transduce the signal to the target proteins upon sensing the Ca^{2+} signal; whereas, CaMs/CMLs and CBLs trigger a conformational change in their structure upon Ca^{2+} perception and interact with their target proteins to transduce the Ca^{2+} signal. CBLs interact specifically with CIPKs to transduce the Ca^{2+} signals for various intracellular processes or responses. CaMs are a large class of Ca^{2+} sensor proteins with 7 *CaM* and 50 *CML* genes encoded in the *Arabidopsis* genome (Abel et al., 2013; Bürstenbinder et al., 2013). CaMs and CMLs are known to interact with wide array of proteins with varied functions such as metabolic enzymes, transcriptional regulators, protein kinases, cytoskeletal proteins and ion transporters (Snedden and Fromm, 2001). In *Populus*, members of the Ca^{2+} -calmodulin module are known to play important roles in induction of freeze tolerance (Lin et al., 2004) and salt stress (Chang et al., 2006) responses.

Recent studies have described a new plant-specific class of calmodulin interacting proteins with conserved IQ-67 domains, referenced after the isoleucine and glutamine (IQ)- amino acid rich region and the central domain of 67 conserved amino acid residues (Abel et al., 2005; Levy et al., 2005; Bürstenbinder et al., 2013). These conserved IQD-67 domain containing proteins are referred to as IQDs and belong to a structurally conserved large gene family of 33 predicted members in *Arabidopsis* (Abel et al., 2005). Although, previously identified calmodulin-binding proteins were Ca^{2+} -dependent, further studies revealed the occurrence of Ca^{2+} -independent calmodulin interacting proteins. One such example is of intestinal brush border myosin I, which interacts with Ca^{2+} -free form of calmodulin called 'apocalmodulin' (Bahler and Rhoads, 2002). Members of IQD family also appear to interact with CaMs in Ca^{2+} -dependent and Ca^{2+} -independent manner (Bahler and Rhoads, 2002; Abel et al., 2005). The genome of *Populus trichocarpa* encodes for 40 IQD genes (Hui et al., 2014).

Co-expression studies in *Arabidopsis* have previously reported the sequence ortholog of *PdIQD10*, *AtIQD10* (AT3G15050), as co-expressed with *AtCesAs* as well as *AtSND1*, *AtNST1*, and *AtMYB103*, a set of transcription factors involved in secondary cell wall biosynthesis (Endler and Persson, 2011; Xu et al., 2013). However, no functional evidence for involvement in either biosynthesis or signaling pathways related to secondary cell wall formation have been found. While these co-expression studies suggest a potential role of an IQD gene in the context of cell wall biosynthesis, functional evidence at molecular, cellular or plant level has not been reported.

Abbreviations: CaMs, calmodulins; CBLs, calcineurin B-like proteins; cDNA, complementary DNA; CDPK, Ca^{2+} -dependent protein kinases; *CesA*, cellulose synthase; Chi, chitinase-like protein; CHUP, chloroplast unusual positioning; CIPKs, CBL-interacting protein kinases; CMLs, Calmodulin-like proteins; CSC, cellulose synthase complex; CSII, cellulose synthase-interactive protein I; CTAB, Cetyl trimethylammonium bromide; DICE, defect in cell elongation; DOE, Department of Energy; DUF, Domain of Unknown Function; GAUT, Galacturonosyltransferase; GUS, β -glucuronidase; HD-ZIP III, class III homeodomain-leucine zipper protein; HYP, hypothetical protein with unknown function; IQD, IQ-domain containing protein; IRX10, beta-1,4-xylosyltransferase (irregular xylem phenotype); KLCR or KLC, kinesin light chain-related protein; KOR, Korrigan; Lac, laccase; MAP, Microtubule-associated protein; MEGA, Molecular Evolutionary Genetics Analysis; NEK, NIMA-related kinases; NIMA, Never in mitosis A; ONPG, ortho-nitrophenyl- β -galactoside; RIC4, Rac-interactive binding; SCPL, Serine carboxypeptidase; SND, secondary wall-associated NAC domain protein; UC, plastocyanin-like domain (PCLD)-containing protein (UC-like); UGPase, UDP-glucose pyrophosphorylase; UTR, Un-translated Regions; WND, wood-associated NAC domain transcription factor.

Functional evidence for a distinct *Arabidopsis* IQD gene family member, *AtIQD1*, has been presented in the context of glucosinolate metabolism and defense response as well as cell growth and microtubule organization (Levy et al., 2005; Abel et al., 2013; Bürstenbinder et al., 2013, 2017b). *AtIQD1* was shown to interact with KLCR1 or KLCs and localize to microtubules as well as to nucleus (Bürstenbinder et al., 2013). It is proposed that the cargo transport function of kinesin is activated over microtubule cross-link or sliding function in the presence of KLCs (Wong and Rice, 2010). *AtIQD1* was hypothesized to function as a molecular scaffold in cargo transport across microtubule tracks of the cell (Hirokawa et al., 2009; Verhey et al., 2011; Bürstenbinder et al., 2013). The current understanding of the cellular and plant functional context of IQD proteins is limited and primarily derived from a research study by Bürstenbinder et al. (2013). Considering that the Ca^{2+} /calmodulin signaling system is integral in mediating a diverse range of plant processes (Bürstenbinder et al., 2017a), individual members of the large IQD gene family may be differentiated to express and function in the context of distinct developmental and physiological phases.

Here, we present evidence in support of a new functional role for an IQD gene family member, *PdIQD10*, in secondary cell wall biosynthesis and biomass formation in *Populus deltoides*.

MATERIALS AND METHODS

Phylogenetic and Sequence Analysis

Protein sequences of *Populus* IQD and CaM (Table 1) homologs were retrieved from *Phytozome* v9.1: *Populus trichocarpa* v3.0, and the NCBI database. Phylogenetic analysis was performed in the MEGA 7.0.25 (Molecular Evolutionary Genetics Analysis) program using the Maximum Likelihood method (Payyavula et al., 2014; Kumar et al., 2016). Bootstrap values were calculated from 1000 independent bootstrap runs. Protein sequence alignment was performed using Clustal W.

Construct Development and Plant Transformation

The *PdIQD10* RNAi construct was developed by PCR-amplifying a 156 bp nucleotide sequence overlapping the 3' coding and UTR regions (Supplementary Figure S1), cloning into a Gateway entry vector and then binary vector via LR Clonase recombination were transformed into wild-type *P. deltoides* 'WV94' according to previously published methods for *Agrobacterium*-based transformation of *Populus* (Meilan and Ma, 2006; Kumar et al., 2016). For subcellular localization, the full length protein-coding sequence corresponding to *PdIQD10* (Potri.001G375700) was amplified from the *P. deltoides* xylem cDNA library (primers are listed in Table 2) using Q5 High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA, United States), and cloned in a pENTR vector (Invitrogen, Carlsbad, CA, United States). After sequence confirmation, the coding region fragments were recombined into one of the Gateway binary vectors pGWB405, pGWB444 and pGWB454 (Waadt and Kudla, 2008; Nakagawa et al., 2009) using LR

TABLE 1 | List of calmodulin and calmodulin-like proteins employed in the study.

Sl. No	<i>Populus</i> Potri ID	Gene names in this paper	Cloned and tested	<i>Arabidopsis</i> orthologs
1	Potri.016G024700.2	CaM247	Yes	AT3G43810.1
2	Potri.001G117900.1	CaM1179	Yes	AT1G12310.1
3	Potri.T124900.1	CaM1249	Yes	AT3G07490.1
4	Potri.004G035100.1	CaM351	Yes	AT5G57580
5	Potri.005G128100.1	CaM1281	Yes	AT3G50770.1
6	Potri.008G134300.1	CaM1343	Yes	AT1G24620.1
7	Potri.017G029700.1	CaM297	Yes	AT3G07490.1
8	Potri.014G093900.1	CaM939	Yes	AT4G01150.1
9	Potri.007G031900.1	CaM319	Yes	AT3G50770.1
10	Potri.006G043700.1	CaM437	Yes	AT3G10190.1
11	Potri.004G122900.1	CaM1229	Yes	AT5G39670.1
12	Potri.001G314900.1	CaM3149	Yes	AT3G24110.1
13	Potri.013G040300.1	CaM403	Yes	AT3G22930.1
14	Potri.002G047300.1	CaM473	Yes	AT3G22930.1
15	Potri.002G239100.1	CaM2391	Yes	AT3G07490.1
16	Potri.015G032600.1	CaM326	Yes	AT5G37780.1
17	Potri.005G052800.1	CaM528	Yes	AT4G14640.1
18	Potri.012G041000.1	CaM410	Yes	AT5G37780.1
19	Potri.009G021500.1	CaM215	Yes	AT2G27030.3
20	Potri.003G095700.1	CaM957	Yes	AT1G32250.1
21	Potri.015G052600.1	CaM526	Yes	AT1G18530.1
22	Potri.002G001400.1	CaM014	Yes	
23	Potri.006G065900.1	CaM659	Yes	
24	Potri.005G259900.1	CaM2599	Yes	AT5G42380.1
25	Potri.012G048200.1	CaM482	Yes	AT1G18210.1
26	Potri.006G026700.1	CaM267	Yes	
27	Potri.006G043900.1	CaM439	No	AT5G04020.1
28	Potri.016G040900.1	CaM409	No	AT3G10190.1
29	Potri.001G375400.1	CaM3754	No	AT1G53210.1
30	Potri.017G054900.1	CaM549	No	AT3G24110.1
31	Potri.001G332900.1	CaM3329	No	AT1G76650.1
32	Potri.002G239200.1	CaM2392	No	AT3G25440.1
33	Potri.007G128600.1	CaM1286	No	AT2G43290.1
34	Potri.001G222200.1	CaM2222	No	AT3G56800.1
35	Potri.003G116800.1	CaM1168	No	AT1G24620.1
36	Potri.007G042700.1	CaM427	No	

clonase (Invitrogen), and the plasmid DNA from a single colony each was used to transform *Agrobacterium*. Tobacco infiltration and protein localization were performed as described previously (Waadt and Kudla, 2008). *Agrobacterium* harboring binary constructs of *PdIQD10*, *CaM014* and *CaM247* were cultured overnight in LB media. After a brief centrifugation, the supernatant was removed and the pellet was dissolved in 10 mM $MgCl_2$ and OD at A_{600} was adjusted to 0.5. The culture was infiltrated into 4-week old tobacco leaves. After 48–72 h, roughly 4 mm² leaf sections were cut and fixed in 3.7% formaldehyde, 50 mM NaH_2PO_4 and 0.2% Triton X-100 for 30 min, then rinsed with phosphate-buffered saline (PBS) and stained in DAPI (4,6'-diamidino-2-phenylindole, 1.5 μ g ml⁻¹ in PBS) for 30 min. For protoplast transfection experiments, *CaM247* and *CaM014* were cloned into CD3-1654 vector

TABLE 2 | List of primers used in the study.

SI No.	Primer name	Primer sequence (5'-3')	Purpose
1	IQD10-RNAiF	CACCCCGGGGCTGTGGATTGTCAAGCTG	RNAi construct
2	IQD10-RNAiR	TCTAGATTACACGCTGCATGCCA	RNAi construct
3	IQD10-IF-BDF	CATGGAGGCCGAATTCATGGTTGTCTCTTCTCTGGG	Yeast two-hybrid
4	IQD10-IF-BDR	GCAGGTCGACGGATCCACTTGAATCAGCAGATGAGAC	
5	KLC400-IF-ADF	GGAGGCCAGTGAATTCATGTTGGGCCTAGTCTCTGC	
6	KLC400-IF-ADR	TCATCTGCAGCTCGAGCGAGCTAGAATCTAGGAGATTTAC	
7	KLC3200-IF-ADF	GGAGGCCAGTGAATTCATGCCAGGTCTTGTCTCTG	
8	KLC3200-IF-ADR	TCATCTGCAGCTCGAGAATTCATAAACCCCACTTC	
9	KLC4700-ADF	GGAGGCCAGTGAATTCATGCCGGGACTAGCCATGG	
10	KLC4700-ADR	TCATCTGCAGCTCGAGCGAGACCTCAAGATCATCG	
11	KLC7800-ADF	GGAGGCCAGTGAATTCATGCCTGGTCTTGTCTCTG	
12	KLC7800-ADR	TCATCTGCAGCTCGAGAATTCATAAACCCCACTTC	
13	MAP599-ADF	GGAGGCCAGTGAATTCATGGAGAAAGCACATACAAAATC	
14	MAP599-ADR	TCATCTGCAGCTCGAGCTTAACAGGCTTCCAGGCATG	
15	DICE1-ADF	GGAGGCCAGTGAATTCATGCTGTACCGACGGGACCTG	
16	DICE1-ADR	TCATCTGCAGCTCGAGCCTGTCTAATTTACATGTGC	
17	Chi1416-ADF	GGAGGCCAGTGAATTCATGGAGGCCAATGGTGTTTTTC	
18	Chi1416-ADR	TCATCTGCAGCTCGAGAGAAGATGATGCAGAGGCAGAG	
21	CHUP1-ADF	GGAGGCCAGTGAATTCATGCGATTCCAATCTCCTG	
22	CHUP1-ADR	TCATCTGCAGCTCGAGTTGTTTGAGACTACCCATGC	
23	RIC4-ADF	GGAGGCCAGTGAATTCATGCATGCATGCCCTTTTCAG	
24	RIC4-ADR	TCATCTGCAGCTCGAGAGAGAGCCTGGAATGATCGACC	
25	HYP2316-ADF	GGAGGCCAGTGAATTCATGGGAAGTGAGCGGTGCAATA	
26	HYP2316-ADR	TCATCTGCAGCTCGAGCATGGGCAACCGGAGAAGTGAC	
27	UC2-ADF	GGAGGCCAGTGAATTCATGGCCGGTTTAATTTCAAGATC	
28	UC2-ADR	TCATCTGCAGCTCGAGGAAAACCATCACAAAAACGGCG	
29	Lac17-ADF	GGAGGCCAGTGAATTCATGGGTGCTTCATTTCTTCCATC	
30	Lac17-ADR	TCATCTGCAGCTCGAGACACTTGGGAAGATCGGCTGG	
31	PtiCesA7-A-ADF	GGAGGCCAGTGAATTCATGGAAGCCAGTGCTGGACTTG	
32	PtiCesA7-A-ADR	TCATCTGCAGCTCGAGACAGTTGATTCCACATTGCTTG	
33	MAP1733-ADF	GGAGGCCAGTGAATTCATGGGTTCGTTCCAAATTGG	
34	MAP1733-ADR	TCATCTGCAGCTCGAGGGAGATAACCGGCCAGAAAAGG	
35	PdGAUT12.1-ADF	GGAGGCCAGTGAATTCATGCAGCTTCATATATCGCCGAG	
36	PdGAUT12.1-ADR	TCATCTGCAGCTCGAGAGATGGCCTAATATGACAGCC	
37	IRX10-ADF	GGAGGCCAGTGAATTCATGAGGACATGCTTGTGGG	
38	IRX10-ADR	TCATCTGCAGCTCGAGCCAAGTTTCAGGTCCCCAACTG	
39	NEK5-ADF	GGAGGCCAGTGAATTCATGGAGGCAGATACCGGTGAAG	
40	NEK5-ADR	TCATCTGCAGCTCGAGGGTCCCTCCATTAATTTTTTTTAGAC	
41	NEK6-ADF	GGAGGCCAGTGAATTCATGGAGGCAGATGCTGCTGAAG	
42	NEK6-ADR	TCATCTGCAGCTCGAGGGTCCCTTCATTGAGTTTTTGC	
43	PdSCPL14-ADF	GGAGGCCAGTGAATTCATGGATTTGGTACCCAAGGTG	
44	PdSCPL14-ADR	TCATCTGCAGCTCGAGTCTGTTTACAAGGAAGAGATTC	
45	PdSCPL41-ADF	GGAGGCCAGTGAATTCATGGATTTGGTACCTAAGGTG	
46	PdSCPL41-ADR	TCATCTGCAGCTCGAGTCTCACAGAAGGCAGGATTC	
47	PdWND1B-ADF	GGAGGCCAGTGAATTCATGCCTGAGGATATGATGAATC	
48	PdWND1B-ADR	TCATCTGCAGCTCGAGTACCGATAAGTGGCATAATGG	
49	MAP2698-ADF	GGAGGCCAGTGAATTCATGGACCCTGCGAGTATACTTG	
50	MAP2698-ADR	TCATCTGCAGCTCGAGACTTCGTCTTTCTTCTCTGTTTC	
51	KOR1-ADF	GGAGGCCAGTGAATTCATGAGCATGTATGGTAGAGATC	
52	KOR1-ADR	TCATCTGCAGCTCGAGTGGTCTCCAAGGTGCTGGAGGTG	
53	UGPase-ADF	GGAGGCCAGTGAATTCATGGCTACTGATACGGCCAAG	
54	UGPase-ADR	TCATCTGCAGCTCGAGGAGTCTCCGGGCCATTG	
55	CaM247-IF-ADF	GGAGGCCAGTGAATTCATGGCGGATCAATTGACGGATG	

(Continued)

TABLE 2 | Continued

SI No.	Primer name	Primer sequence (5'-3')	Purpose
56	CaM247-IF-ADR	TCATCTGCAGCTCGAGCTTAGCCATCATAACCTTGAC	
57	CaM1179IF-ADF	GGAGGCCAGTGAATTCATGGGCAAGGATCTGAGCGACG	
58	CaM1179IF-ADR	TCATCTGCAGCTCGAGCTTCGTAACCATTCTGGCAATAAAATC	
59	CaM-1249-ADF	GGAGGCCAGTGAATTCATGCCAACTATTTTGCTTAGG	
60	CaM-1249-ADR	TCATCTGCAGCTCGAGACCCCAAAGCACTAAGCCACCACCC	
61	CaM-351-ADF	GGAGGCCAGTGAATTCATGCAAAACCGGTATGTTGAGAG	
62	CaM-351-ADR	TCATCTGCAGCTCGAGAGAATCATCTAACTCTATGAGCTGTGCTC	
63	CaM-1281-ADF	GGAGGCCAGTGAATTCATGGCAACTGATAGAGTTTC	
64	CaM-1281-ADR	TCATCTGCAGCTCGAGGGCCATCATTTGATTAAACTCATGAAAATC	
65	CaM-1343-ADF	GGAGGCCAGTGAATTCATGGGGTTTAAATGCCTTTTC	
66	CaM-1343-ADR	TCATCTGCAGCTCGAGACCCCTAGTTCCTCTTAGAGTATC	
67	CaM-297-ADF	GGAGGCCAGTGAATTCATGCCAACTATTTTGCTTAGG	
68	CaM-297-ADR	TCATCTGCAGCTCGAGACCCCAAAGCACTAAGCCACCACCC	
69	CaM939-ADF	GGAGGCCAGTGAATTCATGGCATCGACGGTGATTGC	
70	CaM939-ADR	TCATCTGCAGCTCGAGGCTCGTATTGCCTAATATATATTTTC	
71	CaM319-ADF	GGAGGCCAGTGAATTCATGGAACTGAAAGAGCTTC	
72	CaM319-ADR	TCATCTGCAGCTCGAGAGCCATCATTTGATTAAACTC	
73	CaM437-ADF	GGAGGCCAGTGAATTCATGAAATTCATCAAAAACTC	
74	CaM437-ADR	TCATCTGCAGCTCGAGTCTCCGCAGCTCCATCATAAC	
75	CaM1229-ADF	GGAGGCCAGTGAATTCATGCACATCATTATAATTTA	
76	CaM1229-ADR	TCATCTGCAGCTCGAGACAAAACTGTTCTCCATGAAC	
77	CaM3149-ADF	GGAGGCCAGTGAATTCATGAAAGATTCACTTGGAAC	
78	CaM3149-ADR	TCATCTGCAGCTCGAGAGTTCCTTGCTCACAGGAATATC	
79	CaM403-ADF	GGAGGCCAGTGAATTCATGGTGGATGTGTTTACAGAG	
80	CaM403-ADR	TCATCTGCAGCTCGAGAATTGCCAGCATCATTTCTC	
81	CaM473-ADF	GGAGGCCAGTGAATTCATGGCAGATGCACTGGCTGG	
82	CaM473-ADR	TCATCTGCAGCTCGAGGAAAGCCATCGCCATCATCC	
83	CaM2391-ADF	GGAGGCCAGTGAATTCATGGATCCGGCAGAGCTACG	
84	CaM2391-ADR	TCATCTGCAGCTCGAGACTTGAACCTAGGGCAGCAAATC	
85	CaM326-ADF	GGAGGCCAGTGAATTCATGACGGAGCAGCTAACAGAG	
86	CaM326-ADR	TCATCTGCAGCTCGAGCTTGCCAGCATCATCCTCAC	
87	CaM528-ADF	GGAGGCCAGTGAATTCATGGTAGATGTGTTGACGGAAG	
88	CaM528-ADR	TCATCTGCAGCTCGAGAGCAGCCAGCATTATTCTCAC	
89	CaM410-ADF	GGAGGCCAGTGAATTCATGTCTGAGCAGCTAACGGAG	
90	CaM410-ADR	TCATCTGCAGCTCGAGCTTGCCAAACATCATCCTCAC	
91	CaM215-ADF	GGAGGCCAGTGAATTCATGGCCGATCAGCTGACCG	
92	CaM215-ADR	TCATCTGCAGCTCGAGTAACACATTACAGCATTTTTTC	
93	CaM957-ADF	GGAGGCCAGTGAATTCATGAGCAACAAGCAACCAGC	
94	CaM957-ADR	TCATCTGCAGCTCGAGAACCCAAGAATTATCAAAAG	
95	CaM526-ADF	GGAGGCCAGTGAATTCATGACAGAGGCAGCCCTACATG	
96	CaM526-ADR	TCATCTGCAGCTCGAGAGTGATGCCTAAGAACTCC	
97	CaM014-ADF	GGAGGCCAGTGAATTCATGAAATCACACTCGGTTTC	
98	CaM014-ADR	TCATCTGCAGCTCGAGGCGCATCATGACAGCAAACTC	
99	CaM659-ADF	GGAGGCCAGTGAATTCATGATGATGATCATTACCATC	
100	CaM659-ADR	TCATCTGCAGCTCGAGGAAGACAGAAACAAGCTTGC	
101	CaM2599-ADF	GGAGGCCAGTGAATTCATGAAGTCACACTCAGTTTC	
102	CaM2599-ADR	TCATCTGCAGCTCGAGGCGCATCATGACAGAAAACCTC	
103	CaM482-ADF	GGAGGCCAGTGAATTCATGGCAACCCCGAACGCCAAAAC	
104	CaM482-ADR	TCATCTGCAGCTCGAGGGAACCCATATTAGCGGCCATC	
105	CaM267-ADF	GGAGGCCAGTGAATTCATGGCGGATCAATTGACTGATG	
106	CaM267-ADR	TCATCTGCAGCTCGAGCTTAGCCATCATAACCTTGAC	
108	IQdom-BDF	CATGGAGGCCGAATTCGAATATATAGCTGCTGTCAGG	
109	IQdom-BDR	GCAGGTGCACGGATCCCCTCACGGCGCCCTTAAGACGG	

(Continued)

TABLE 2 | Continued

SI No.	Primer name	Primer sequence (5'-3')	Purpose
110	C014-pEF	CACCATGAAATCACACTCGGTTTCATC	pENTR gateway cloning
111	C014-pER	GCGCATCATGACAGCAAACCTC	
112	C014-pER w stop	CTAGCGCATCATGACAGCAAACCTC	qPCR primers
113	UBCc-F	CTGAAGAAGGAGATGACAGCAC	
114	UBCc-R	GCATCCCTTCAACACAGTTTCAC	
115	IQD10-RTF	TGGAGAGCCAACCCTACCCAATA	
116	IQD10-RTR	GCAGGTTTGGTTGGAAGTGTCGTT	GST-fusion recombinant proteins
117	18S-F	AATTGTTGGTCTTCAACGAGGAA	
118	18S-R	AAAGGGCAGGGACGTAGTCAA	
119	PdWND1B-pX6p1F	GAAGGTCGTGGGATCCATGCCTGAGGATATGATGAATC	
120	PdWND1B-pX6p1R	GTCGACCCGGGAATTCACCGATAAGTGGCATAATGG	Promoter amplicons for EMSA assay
121	PdHB3-6p1F	GAAGGTCGTGGGATCCATGGCGCTTTCTATGCACAG	
122	PdHB3-6p1R	GTCGACCCGGGAATTCACAAAGGACCAAGTTAATGAAC	
123	ProMYB002-F	ACCTCTCTCATTTTCCCTGCG	
124	ProMYB002-R	TCCCTGTCACTAGAAAGGTGATCT	
125	ProHB3-F	GCCTGCCTCTCATTATTCTCTAC	
126	ProHB3-R	CACCTAAAGAAAGAACTAAAACCTTG	
127	ProIQD1-F	AAGTTAAACGGTTGAAGATGG	
128	ProIQD1-R	GGTCAAAGAGAGAAGCAACAC	
129	ProCaM014-F	CCGGAAATGCCGAGAAGGACTCG	
130	ProCaM014-R	AAAGGAAATGAGGGTGTGGTCATGGC	
131	ProAct-F	ACCTACTTCGTTTGGTCATTGTTA	
132	ProAct-R	CAAATACAACATAGTTCTCTCCAC	

obtained from the Arabidopsis Biological Resource Center. *Populus* protoplasts were isolated and transfected as described in Guo et al. (2012). FM-64 (#T13320) was used as plasma membrane marker and mCherry-VirD2NLS was used as nuclear marker. Fluorescence visualization and imaging was performed on a Zeiss LSM710 confocal laser scanning microscope (Carl Zeiss Microscopy, Thornwood, NY, United States) equipped with a Plan-Apochromat 63x/1.40 oil immersion objective. To increase the accessibility of images obtained from these subcellular localization experiments (Wong, 2011), the yellow color channel was converted to magenta uniformly across all images in the CMYK color spectrum. The original RGB color scheme images are also provided in Supplemental Files.

Plant Growth and Sampling

Transgenic and empty vector transformed control plants were acclimated from the tissue culture to Ray Leach tubes containing equal parts Fafard 52:perlite:vermiculite. After 2 months, the plants were moved to bigger pots (6 l) and propagated in a greenhouse maintained at 25°C with 16 h day length under drip irrigation and fertilized weekly with 200 ppm N. At the time of harvest (7-month old plants), plant height was measured from shoot tip to stem base, and diameter was measured at ~5 cm from the soil line. In our preliminary study, the basal, lignified 10 cm stem portion was harvested, debarked, air-dried, and used for carbohydrate composition, cellulose, lignin, S:G ratio, and sugar analyses. Initial studies were performed on 18 transgenic lines (plus 9 control lines) and additional studies were performed

on three to four selected lines. Plants for additional studies were generated from fresh internodal stem cuttings. The tissues collected were young leaves (leaf plastochron index, LPI-0 and 1), mature leaves (LPI-6) and stems (internode portion between LPI 6 and 8), which were frozen in liquid nitrogen and stored at -80°C until used.

For tension stress response study, *Populus deltoides* plants were grown erect as control or with bending stress to generate tension wood on the outer bent side and opposite wood on the inner stem side as follows. The six plants per treatment were grown from stem cuttings rooted and grown in the greenhouse for ~40 days prior to start of the experiment. Plants were tied in a bent position to induce tension wood formation and the control plants were left straight but were also tied to stakes to simulate the same thigmotactic set-up as the plants under tension, for a 2-week period of time. Xylem tissue was collected from freshly harvested samples, flash-frozen and processed as previously described (Kalluri et al., 2009).

RNA Extraction and Gene Expression Studies

RNA from the ground and frozen stem samples was extracted using a Plant RNA extraction kit (Sigma, St. Louis, MO, United States), with modifications. Briefly, 100 mg of frozen sample was extracted with a 850 µl CTAB buffer maintained at 65°C followed by chloroform:isoamylalcohol (24:1 v/v). After passing the supernatant through a filtration column, the eluent was diluted with 500 µl of 95% EtOH and passed through a binding column. Further steps, including on-column DNase

digestion, were followed as per the manufacturer protocol. cDNA was synthesized from 1.5 µg of RNA using oligo dT primers and RevertAid Reverse Transcriptase (ThermoFisher). Quantitative reverse transcriptase PCR (qRT-PCR) was performed in a 384 well plate using cDNA (3 ng), gene specific primers (250 nM, list provided in **Table 2**) and iTaq Universal SYBR Green Supermix (1X, Bio Rad). Gene expression was calculated by a ΔC_T or $\Delta\Delta C_T$ method using the expression of housekeeping genes 18S ribosomal RNA and Ubiquitin-conjugating enzyme E2 for template normalization (Gene accession numbers and primer sequence information can be found in Payyavula et al. (2014)).

Cellulose and Lignin Properties

The cellulose content in an air-dried stem sample was estimated using the anthrone method (Updegraff, 1969) as well as wet chemistry-based HPLC analysis. For the semi-quantitative anthrone assay, stem sample (25 mg) was first digested with 500 µl of acetic-nitric acid reagent (100 ml of 80% acetic acid mixed with 10 ml of nitric acid) at 98°C for 30 min. After cooling, the sample was centrifuged. The supernatant was discarded and the sample was washed with water. After a brief centrifugation, the water was removed and the pellet was digested with 67% (v/v) sulfuric acid for 1 h at room temperature. An aliquot of the mix was diluted (1:10) with water. In a PCR tube, 10 µl of diluted reaction mix, 40 µl of water and 100 µl of freshly prepared anthrone reagent (0.5 mg anthrone ml⁻¹ of cold concentrated sulfuric acid) were added and heated for 10 min at 96°C. The samples were cooled and the absorbance (A₆₃₀) was measured. The cellulose content was then estimated based on the absorbance of glucose standards. Holocellulose and α-cellulose samples were prepared and employed in gel permeation chromatography (GPC) and for a ¹³C-CPMAS NMR analysis of cellulose using established protocols.

For quantitative wet chemistry assay, roughly 25 mg of air-dried stem sample was weighed in a 2 ml tube and twice extracted at 85°C to a total of 2 ml of ethanol (80%). To eliminate pigments that interfere with sugar analysis, the supernatant was collected in a new 2 ml tube and re-extracted with 50 mg of activated charcoal (Sigma). A 1-ml aliquot of the pigment-free extract was incubated overnight in a heating block maintained at 50°C. The resulting pellet was dissolved in 120 µl of water and a 10 µl aliquot was used in sucrose and glucose estimation using kits (Sigma). Starch from the pellet was digested by 1 U of α-amylase (from *Aspergillus oryzae*, Sigma) and amyloglucosidase (from *Aspergillus niger*, Sigma). After starch removal, the pellet was dried overnight at 95°C and used to estimate structural sugars. Roughly 5 mg of sample was weighed in a 2-ml tube and digested with 50 µl of 75% v/v H₂SO₄ for 60 min. The reaction was diluted by adding 1.4 ml water. The tubes were sealed using lid-locks and autoclaved for 60 min in a liquid cycle. After cooling, the sample was neutralized with CaCO₃ and sugar composition was estimated with high performance liquid chromatography (HPLC, LaChrom Elite® system, Hitachi High Technologies America, Inc.), as described previously.

Lignin content and the syringyl-to-guaiacyl ratio (S/G ratio) were determined based on pyrolysis molecular beam mass spectrometry (MBMS) of dried and ground stem biomass samples as described previously (Mielenz et al., 2009; Kalluri et al., 2016).

In silico Identification of the PdlQD10 Interacting Partners

Putative interacting partners of PdlQD10 (**Table 3**) were identified using various *in-silico* databases like String (using POPTR_0001s38470 as query)¹, ATTED-II² and phytozome co-expression database³. For the *Arabidopsis* genes identified using ATTED-II, the closest homolog of *Populus* was chosen as a putative interacting partner. The genes GAUT12.1 (Biswal et al., 2015; Potri.001G416800.1), IRX10 (Porth et al., 2018; Potri.001G068100.1), PdSCPL14 and PdSCPL41 (Zhu et al., 2018; Potri.004G037800.1; Potri.011G046600.1) were named according to the respective *Populus* literature. The genes DICE1 (Le et al., 2018; Potri.006G269800.1), DUF1218 (Wilson-Sánchez et al., 2017; Potri.004G235000.1), CHUP1 (Oikawa et al., 2008; Potri.001G279000.1), RIC4 (Gu and Nielsen, 2013; Potri.002G233400.1), UC2 (Xu et al., 2017; Potri.002G101300.1), Lac17 (Voxeur et al., 2017; Potri.001G401300.1), NEK5 and NEK6 (Vigneault et al., 2007; Potri.016G051900.1; Potri.006G056300.1) were named according to the Arabidopsis convention reflecting their closest homologs. The names of three Microtubule Associated Proteins - MAP599, MAP1733 and MAP2698 were based on (Quentin et al., 2016) with the suffix numbers corresponding to the last four digits of respective gene IDs.

Calmodulin (CaM) and calmodulin-like (CML) genes encoded by the *Populus* genome were identified by repeated BLAST searches in the Phytozome database using *Arabidopsis* CaM/CML protein sequences as query. The identified proteins were named as they appear in **Table 1** in which the numbers following the CaM symbol are derived from their respective Potri IDs. The amino acid sequences of *Arabidopsis* and *Populus* CaM/CML proteins were used to build a phylogenetic tree using Maximum Likelihood method of the MEGA 7.0.25 software.

Yeast Two-Hybrid Assay

The coding sequences of the 21 putative interactors and 26 calmodulin/calmodulin-like CaM genes were cloned into pGADT7 (Clontech) vector and PdlQD10 into pGBKT7 (Clontech) vector using In-Fusion® Advantage PCR Cloning Kit (Clontech). The primers used for cloning are listed in **Table 2**. The generated constructs were transformed into yeast Y2H Gold (Clontech) competent cells using Fast™ Yeast Transformation (G-Biosciences Cat. #GZ-1) kit. Transformed yeast cells were plated on Yeast Synthetic Drop-out (SD) media lacking Leucine and Tryptophan amino acids (Sigma-Aldrich #Y0750) to select the positive transformants for both plasmids. At least two clones were individually tested either on SD media lacking histidine,

¹<http://string-db.org/>

²<http://atted.jp/>

³<https://phytozome.jgi.doe.gov/pz/portal.html#>

TABLE 3 | List of putative interacting partners tested.

Sl. No	Potri ID	Gene names	Putative function (At orthologs)	Prediction database
1	Potri.T059900.1	MAP599	Microtubule-associated protein	ATTED co-expression
2	Potri.014G108300.1	DICE1	DEFECT IN CELL ELONGATION1	
3	Potri.010G141600.1	Chi1416	Chitinase-like protein 2	
4	Potri.004G235000.1	DUF1218	Transmembrane protein with unknown function	
5	Potri.001G279000.1	CHUP1	Chloroplast unusual positioning1	
6	Potri.002G233400.1	RIC4	Cdc42- and Rac-interactive binding (CRIB) motif-containing protein	
7	Potri.001G231600.1	HYP2316	Similar to hypothetical protein with unknown function	
8	Potri.002G101300.1	UC2	Plastocyanin-like domain-containing protein	
9	Potri.001G401300.1	Lac17	Laccase-like	Phytozome co-expression
10	Potri.006G181900.1	PtiCesA7-A	Similar to cellulose synthase	
11	Potri.003G173300.1	MAP1733	Microtubule-associated protein	
12	Potri.001G416800.1	PdGAUT12.1	Galacturonosyltransferase 12-related	
13	Potri.001G068100.1	IRX10	beta-1,4-xylosyltransferase	STRING
14	Potri.016G051900.1	NEK5	Serine/Threonine-protein kinase	
15	Potri.006G056300.1	NEK6	Serine/Threonine-protein kinase	
16	Potri.004G037800.1	PdSCPL14	Serine carboxypeptidase	
17	Potri.011G046600.1	PdSCPL41	Serine carboxypeptidase	Genes with cell-wall related functions
18	Potri.001G448400.1	PdWND1B	NAC-domain transcription factor	
19	Potri.006G269800.1	MAP2698	Microtubule-associated protein	
20	Potri.001G078900.1	KOR1	Korrigan	
21	Potri.017G144700.1	UGPase	UDP-glucose pyrophosphorylase	

leucine, and tryptophan (Sigma-Aldrich #Y2146) or on SD media lacking histidine, leucine, tryptophan and adenine (Sigma-Aldrich #Y2021) in three different transformation experiments to determine the interaction result.

β-galactosidase Assay

The strength of the observed interactions was assessed by quantifying the activity of β-galactosidase using ONPG substrate as described in the Yeast Protocols Handbook (Clontech). Briefly, overnight yeast cultures in SD selection medium are inoculated in YPD medium to grow until mid-log phase (OD₆₀₀ of 1 ml = 0.5–0.8). Yeast cells were harvested, washed once in Z-buffer (40 mM Na₂HPO₄·7H₂O, 60 mM NaH₂PO₄·7H₂O, 10 mM KCl, 1 mM MgSO₄·7H₂O, pH 7.0) and subjected to repeated freeze-thaw cycles to break open yeast cells. The lysate is then supplemented with Z-buffer, β-mercaptoethanol and ONPG (final ~100 μg) and incubated at 30°C for 24 h. The reactions were stopped by adding Na₂CO₃ and the absorbance was measured at OD₄₂₀. β-galactosidase activity was calculated by using the equation: β-galactosidase units = 1,000 × OD₄₂₀/(t × V × OD₆₀₀). The assay was performed three times using yeast transformants from three independent transformation experiments.

Structural Modeling/Molecular Dynamics Simulation (MD)

Structural models of PdIQD10 (Potri.001G375700.1), six CaMs (Potri.016G024700.2, Potri.002G001400.1, Potri.001G222200.1, Potri.006G026700.1, Potri.009G021500.1, Potri.012G041000.1) and an PdIQD10-domain (Potri.001G375700.1) models were built using the iterative threading assembly refinement (I-TASSER, version 5.0) (Roy et al., 2010) protein structure modeling toolkit. A 200-ns molecular dynamics (MD) simulation was performed on the complex formed by the PdIQD10-domain and calmodulin. For the MD simulation a water box with at least 15 Å to the edge of the protein was used, and sodium/chloride ions were added to balance the net charge of the whole system. The software NAMD (Phillips et al., 2005) was used for the MD simulation. The CHARMM protein force field (Best et al., 2012) and TIP3P water model (Jorgensen et al., 1983) were adopted in all MD simulations. A time step of 2-fs was applied with the SHAKE algorithm to fix the bonds involving hydrogen atoms. In the MD simulation, after a 50,000 steps energy minimization, the temperature of the system was gradually heated to 300 K with a rate of 0.001 K per time step. The MD simulations were performed under an NPT ensemble with the system pressure of 1 atm and temperature of 300 K maintained by the Langevin

piston controls. Cutoff of switching between 9 and 11 Å was applied for the non-bonded interactions, and particle mesh Ewald summation with a grid spacing of 1.35 Å were applied for long range electrostatic interactions, respectively.

In vitro Promoter Binding Experiments

Full-length *PdHb3* (Potri.011G098300) and *PdWND1B* (Potri.001G448400) were cloned in pGEX-6P-1vector and GST-fused recombinant proteins were isolated and used for the assays. Promoter regions representing 250 bp upstream of transcriptional start sites in genomic sequences of *PdHb3* (Potri.011G098300), *PdIQD10* (Potri.001G375700) and *PdCaM014* (Potri.002G001400) and *PdMYB002* (Potri.001G258700) genes were used for the assays. Electrophoretic Mobility Shift assays (EMSA) were performed using Thermo Scientific LightShift™ Chemiluminescent EMSA Kit according to manufacturer's instructions. Briefly, ~2 picomoles of the amplified promoter regions were biotin labeled using Pierce™ Biotin 3' End DNA Labeling Kit. Biotin labeled fragments (~100 femtomoles) were incubated with 200–300 nanograms of the GST-fusion recombinant proteins in the reaction buffer (1X Binding buffer (20148A), 2.5% glycerol (20148F), 5 mM MgCl₂ (20148I), 50 ng/μl poly dI:dC (20148E), 0.05% NP-40 (20148G)) for 20 min at room temperature. For competition assays, ~200 fold-excess of the respective unlabeled DNA fragments were used. The reactions were separated on 6% DNA Retardation Gels (EC6365BOX), transferred on to nylon membrane and crosslinked at 120 mJ/cm² using ULTRA-LUM UVC 515 Ultraviolet Multilinker for 60 s. The membrane was then followed with the detection procedure of the kit and the chemiluminescence was detected using BioRad ChemiDoc™ XRS+ System.

Transcriptional Activator Assay

The coding sequences (CDS) of *PdIQD10*, and four *KLC* homologs; *KLC400*, *KLC3200*, *KLC4700*, and *KLC7800* were in-frame cloned in Gal4 binding domain (GD) effector vector (Wang et al., 2007). For the trans activator assays, the GD-fusion constructs were co-transfected with *Gal4:GUS* reporter construct into *Populus* 717 protoplasts (Guo et al., 2012). Empty GD effector vector was co-transfected with reporter vectors for the control experiments. The transfected protoplasts were incubated in dark for 16–20 h and GUS activity was quantitatively measured. All the protoplast transfections were included with equal amounts of 35S:Luciferase reporter construct and Luciferase activity was used for normalization of GUS activity. The quantification of GUS and Luciferase were performed as below.

For, quantitative measurements of β-glucuronidase and Luciferase, transfected protoplasts were lysed using 1X Cell Culture Lysis Reagent (Promega Cat. # E1531) followed by incubation on ice for 5 min. Cell-debris was separated by centrifugation at 2000 rpm for 3 min and the supernatant was used for the assays. For GUS activities, equal amounts of cell lysate was incubated with 1X solution of 4-methylumbelliferyl β-D-glucuronide (MUG) in the reaction buffer [10 mM Tris (pH

8.0), and 2 mM MgCl₂] at 37°C for 1 h. The reactions were stopped by adding 0.2 M Na₂CO₃ and fluorescence was measured at 460 nm when excited at 355 nm. For luciferase activities, cell lysate was mixed with Luciferase Assay Reagent (Promega Cat. # E1500) and the luminescence was measured. Both fluorescence and luminescence were measured using BioTek™ Synergy™ 2 Multi-Mode Microplate Reader.

RESULTS

PdIQD10 Gene Belongs to the IQ 67-Domain Containing IQD Family and Shows Enhanced Expression in the Context of Secondary Cell Wall Biosynthesis

PdIQD10 (Potri.001G375700), was identified from our previously undertaken studies including; tension stress response characterization, proteomics of developing xylem, and co-expression network analysis of *Populus* and *Arabidopsis* stem tissues (Kalluri et al., 2009; Yang et al., 2011; Abraham et al., 2012). Gene-specific qRT-PCR assays undertaken to support the findings of these three studies confirmed the enhanced expression of *PdIQD10* gene in tension-stressed xylem and secondary walled cells (Figure 1A). Among the native tissues profiled including leaf, stem and root tissues, *PdIQD10* showed a significantly higher expression in the xylem tissue (Figure 1B). *PdIQD10* expression data from the recent transcriptome study by Shi et al. (2017) is in agreement with our findings showing higher *PdIQD10* expression in xylem and fiber libraries (Supplementary Figure S2). Sequence analysis suggests that *PdIQD10*, codes for a predicted calmodulin-binding protein belonging to the family IQ67-domain containing proteins IQD (referenced after the isoleucine and glutamine (IQ)- amino acid rich region and the central domain of 67 conserved amino acid residues). A BLASTP search of the Phytozome database (*Populus* genome v. 3.0) using 33 *Arabidopsis* IQD protein sequences as query identified a total of 42 *PtIQDs*. The phylogenetic tree was constructed with predicted *PtIQD* and *AtIQD* protein sequences using a Maximum Likelihood algorithm in MEGA7.0.25 with 1000 bootstrap replicates (Figure 2). The present manuscript describes the functional characterization of a corresponding *P. deltoides* ortholog, *PdIQD10*, of *PtIQD10* gene (Potri.001G375700 in the reference *P. trichocarpa* genome).

PdIQD10 Is a Calmodulin-Binding Protein

The calmodulin binding ability of targeted *Arabidopsis* IQ67-domain-containing proteins has been previously demonstrated (Levy et al., 2005; Abel et al., 2013; Bürstenbinder et al., 2013). To test the predicted ability of *PdIQD10* to bind a calmodulin protein, calmodulin/calmodulin-like genes encoded by *Populus* genome were identified using Phytozome⁴ database. Table 1 lists 36 calmodulin and calmodulin-like *CaM* genes their inclusion in the yeast two-hybrid assay and their corresponding *Arabidopsis*

⁴<https://phytozome.jgi.doe.gov>

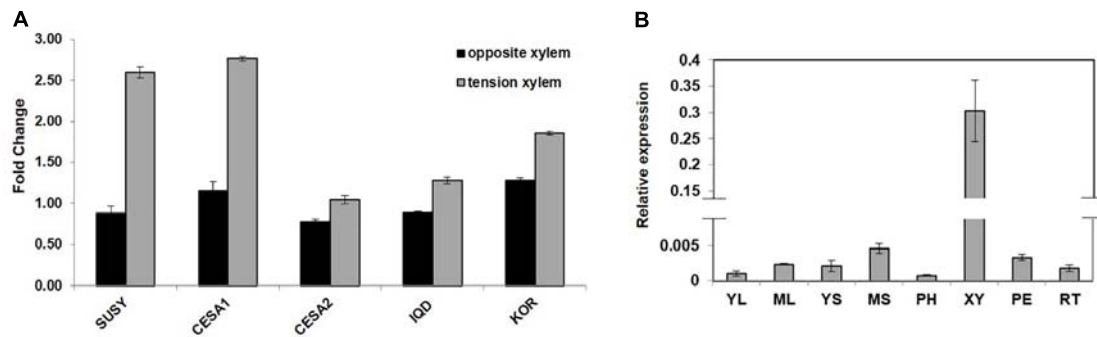


FIGURE 1 | (A) Expression profile of cell wall related genes in tension stressed and opposite xylem tissue. qRT-PCR data showing the relative transcript levels of the chosen genes involved in cell-wall formation. Known cellulose pathway marker genes such as sucrose synthase (*SUSY*; *Potri.006G136700*) cellulose synthases (*CESA1*; *Potri.004G059600*, *PdCESA8-B*; *CESA2*; *Potri.018G103900*, *PdCESA7-B*), and *KOR* (*Potri.001G078900*) display higher expression in the tension stressed xylem tissue along with *IQD* (*PdlQD10*; *Potri.001G375700*) gene as compared to the opposite xylem. **(B)** Transcript profiling of *PdlQD10* in various tissues of *Populus*. The expression levels of *PdlQD10* was determined by qRT-PCR method using *Populus* cDNA libraries of Young Leaf (YL), Mature Leaf (ML), Young Stem (YS), Mature Stem (MS), Phloem (PH), Xylem (XY), Petiole (PE) and Root (RT). As shown in the figure, the expression of *PdlQD10* transcript is highest in xylem tissue followed by mature stem. Standard deviation was calculated across biological replicate libraries ($n = 3$). A break in Y-axis denotes discontinuity in scale. The corresponding figure with continuous Y-axis scale is provided as **Supplementary Figure S9**. Relative expression was based on changes in critical threshold (R_C) values relative to housekeeping genes.

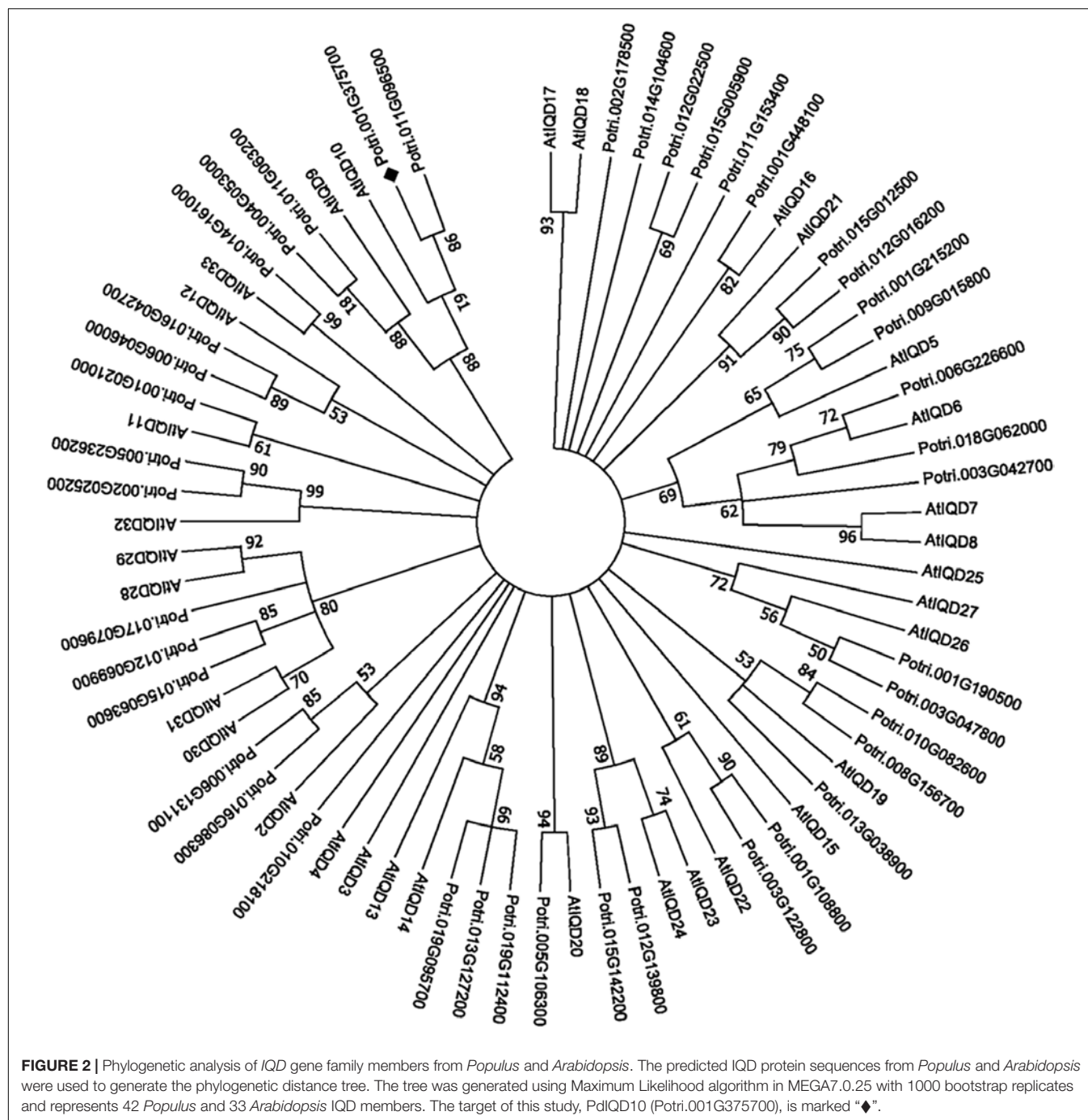
homologs. *PdlQD10* was cloned into pGBKT7 vector and was used as a bait to identify interactions with the full-length *CaM* cDNAs cloned in pGADT7 vector in targeted one-to-one yeast two-hybrid assays. Among a total of 26 interactions screened, *CaM014* and *CaM2599*, induced the growth of yeast auxotroph colonies on minimal media lacking histidine, leucine and tryptophan indicating weak positive interactions (**Figure 3A**). The quantification of interaction strengths by β -galactosidase activity add support to the weak interaction inferred from slow colony formation. Furthermore, although *CaM014* and *CaM2599* share 79.2% identity in their protein sequences (**Supplementary Figure S3**), interaction with *CaM014* was found to be stronger relative to that with *CaM2599*. *In vitro* studies undertaken by Bürstenbinder et al. (2013) suggest that IQ67-domain is the calmodulin binding region of IQD proteins. To investigate this prediction and the importance of the region flanking IQ67 domain for interaction with CaMs, IQ67-domain of *PdlQD10* (*PdlQD10*-domain) was cloned into pGBKT7 vector and used as a bait with CaMs. Contrary to theoretical expectation that IQ67-domain-version can interact with most CaMs, specificity in interaction was observed. *CaM247* and *CaM351* were found to interact with IQ67-domain of *PdlQD10* whereas *CaM014* and *CaM2599* were found not to interact (**Figure 3B**). The observed specificity of IQ67 domain-only version is in contrast with that observed for the full-length version of the IQD, may suggest that IQ67-domain functions differently than the *PdlQD10* full length protein. It is interesting to learn that the full-length *PdlQD10* and *PdlQD10*-domain have non-overlapping calmodulin interacting partners. One reason for the discrepancy may be the differences involved in physiologies of yeast and that of the plant. Alternatively, the longer linker may translate into a 3-D structural conformational change that determines specificity. The sequence alignments of *CaM247*, *CaM014* and *CaM2599* protein sequences showed higher similarity between *CaM014* and *CaM2599*, than with *CaM247*, which is shorter at its N-terminal

(**Supplementary Figure S4**). This observation is also consistent in phylogenetic tree of *CaM*/*CaM*-like proteins where *CaM014* and *CaM2599* are in a single clade (**Supplementary Figure S5**).

A complementary computational approach based on molecular dynamics (MD) simulations was undertaken to probe the potential interaction sites between IQD and *CaM* proteins. MD simulations showed that the predicted transient interaction of IQD-domain only with *CaM247* is stabilized in the presence of salt (**Figure 3C**). Specifically, the flip-and-flop of *CaM* around the IQD domain makes many of the interactions between the two molecules transient (**Figure 3C**). However, there are several strong electrostatic interactions (i.e., salt bridges) that persist throughout the MD simulations, including specific amino acid positions within *CaM247* and IQD-domain only; E15-R29, E12-R29, E85-R7, D81-R7, and E121-K16. In addition to salt-bridge interactions, MD simulations showed the potential for a hybrid duplex structure interaction between full-length *PdlQD10* (R82 to L93) and *CaM014* (S8 to P19) (**Figure 3D**). Salt-bridge interactions between the basic residues (Arg and the backbone of N-terminus) of *PdlQD10* and acidic residues (Asp and Glu) of *CaM014*, are noted as; M1-E138, R17-D34, R87-E56, R87-D52, R82-E160 and R76-E92.

***PdlQD10*, *CaM014* and *CaM247* Co-localize in Similar Subcellular Compartments**

Subcellular protein localization experiments were undertaken to understand the possible proximity and biological relevance of the observed interactions and its impacts on biomass properties of *Populus*. Fluorescent signal was observed under a confocal microscope, 72 h after an *Agrobacterium* clone carrying full length *PdlQD10* in pGWB405 vector was infiltrated in 4–6-week-old leaves of *Nicotiana benthamiana* (*N. benthamiana*). The signal from GFP tagged-*PdlQD10* was detected in the



cytoplasm and plasma membrane (Figure 4A). However, repeated localization experiments revealed that PdIQD10 also localizes to the nucleus (Figure 4B). CaM247 and CaM014 were cloned in-frame in pGWB454 vector expressing mRFP fusion tag. The fluorescence from both CaMs was also detected from nucleus and the plasma membrane further supporting their associated functions with PdIQD10 (Figure 4C and Supplementary Figure S6). PdIQD10-GFP co-localizes with CaM247-mRFP and CaM014-mRFP in the nucleus and in the plasma membrane supporting spatio-temporal co-localization

patterns of the signaling molecules (Figure 4D). Similar protoplast assays under plasmolysis conditions will further confirm the localization in plasma membrane.

PdIQD10 Is Potentially Regulated by Secondary Cell Wall Transcription Factors

In order to understand the functional context of expression of *PdIQD10* gene in secondary wall forming cells, we undertook

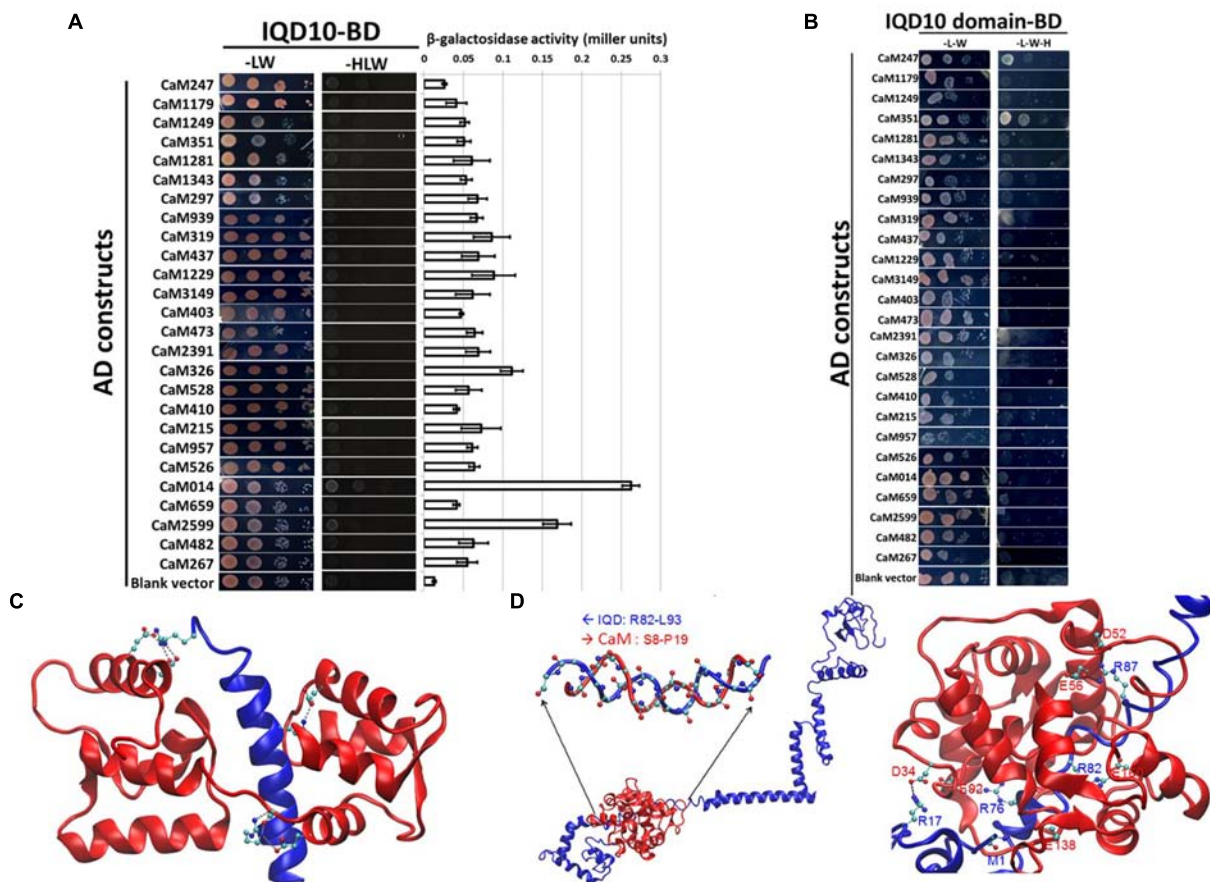


FIGURE 3 | Yeast two-hybrid interaction analysis of PdIQD10 with CaMs. **(A)** Yeast two-hybrid interaction analysis of full-length PdIQD10 with 26 calmodulins encoded by *Populus* genome. This figure shows the positive interaction of PdIQD10 with CaM014 and CaM2599. Bar graph on the right represents the average β -galactosidase activities assayed from three independent yeast colonies to determine the strength of the interactions. **(B)** Yeast two-hybrid interaction analysis of IQ67-domain of PdIQD10 with different calmodulins. The IQ67-domain only version of PdIQD10 interacts with CaM247 and CaM351 but not with CaM014 and CaM2599 indicating the difference in functionalities and specificities between full-length PdIQD10 and IQ-67 domain only version. **(C)** A cartoon representation of MD simulation showing IQD-domain only and CaM247 protein interactions. The flip-and-flop of Calmodulin (CaM, in Red) around the IQD domain (in Blue) makes many of the interactions between the two molecules transient. Strong electrostatic interactions (i.e., salt bridges) persist. **(D)** A cartoon representation of the complex of full-length IQD (blue) and CaM014 (red); (left) the inset shows a hybrid duplex between R82 to L93 of IQD and S8 to P19 of CaM mainly formed by dipole-dipole interactions between the backbone of the regions of both proteins; (right) Salt-bridge interactions between the basic residues (Arg and the backbone of N-terminus) of IQD and acidic residues (Asp and Glu) of CaM with basic residues.

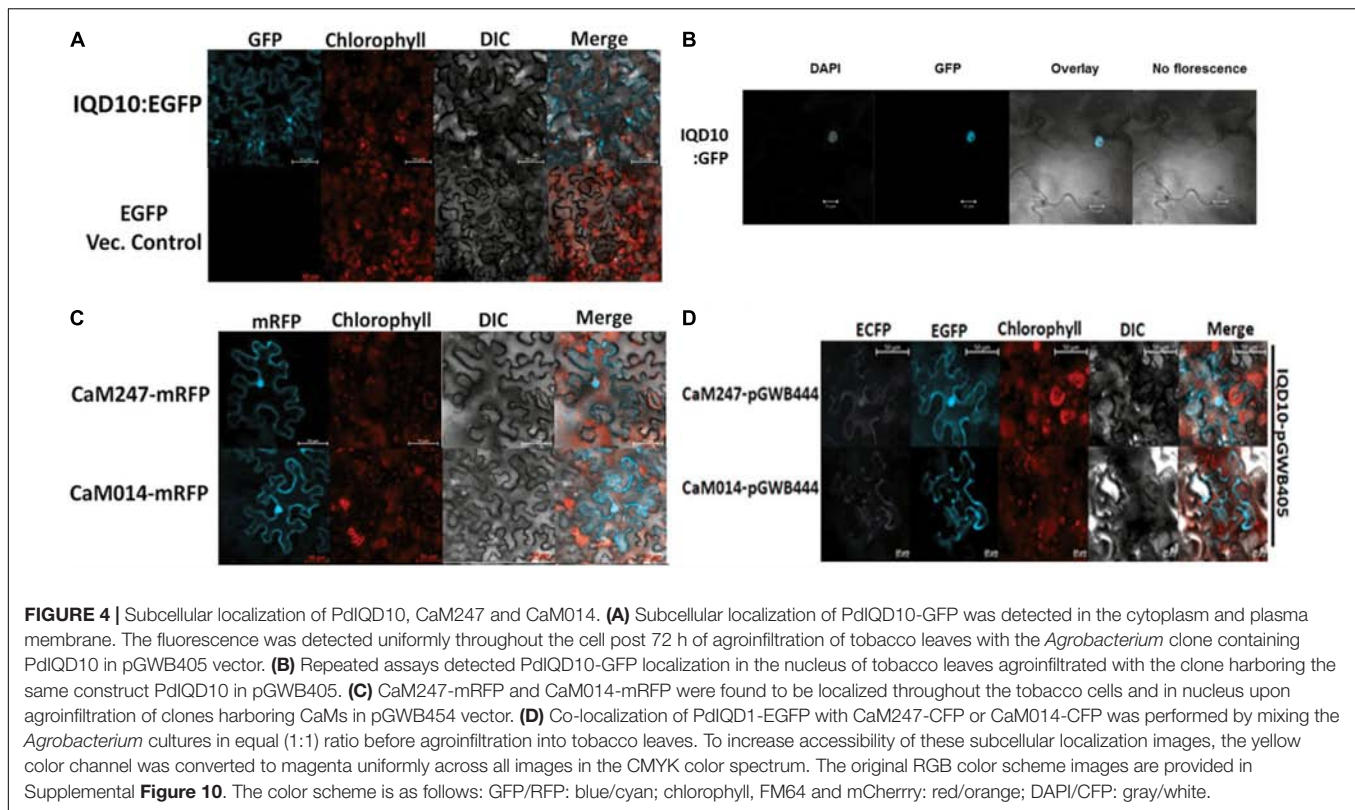
promoter binding assays using the secondary cell wall pathway transcription factors, PdWND1B, a *Populus* ortholog of the known master regulator of secondary cell wall biosynthesis in *Arabidopsis*, SND1 NAC domain transcription factor (Zhong et al., 2006; Zhao et al., 2014), and PdHB3, belonging to the HD-ZIP III family of transcription factors with known roles in stem development (Du et al., 2011; Robischon et al., 2011; Zhu et al., 2013). Electrophoretic Mobility Shift assays (EMSA) using GST-tagged fusion protein revealed that PdWND1B binds to the promoter of PdHB3 (Figure 5A). Binding of PdWND1B to the PdIQD10 promoter was observed to be weak compared to the binding to known positive control of MYB002 promoter (Lin et al., 2013) (Figure 5B). PdHB3 was observed to have moderate binding affinity to the promoters of PdIQD10 and its interacting partner CaM014 and the binding is competed out upon inclusion of unlabeled promoters (Figures 5C–E).

These experiments suggest that the *Populus* SND1 ortholog, PdWND1B, is regulating a transcriptional network that includes direct regulation of PdHB3, which in turn may regulate PdIQD10 and CaM014.

PdIQD10 Interacts With KLCR Proteins

Potential new pathway players and interactor proteins can be represented among tightly co-expressing gene sets. In order to identify new protein interacting partners for PdIQD10, we queried co-expression and protein interaction databases for identifying potential interactors in the context of stem development and cell wall biosynthesis. Selected top hits that were identified from *Populus* STRING⁵, *Arabidopsis* ATTED

⁵<http://version10.string-db.org/cgi/input.pl>



II⁶ and *Populus* Phytozome co-expression databases were employed in yeast two-hybrid assays (**Table 3**). The first set of proteins included sequence homologs of KLCR or KLC, a motor protein involved in unidirectional cargo transport along the microtubular network, which has previously been shown to interact with *Arabidopsis* IQD1 in a yeast two-hybrid cDNA library screen (Bürstenbinder et al., 2013). To test the functional conservation of PdIQD10 in *Populus*, we studied the interaction between PdIQD10 and four xylem-expressing KLCR-1 or KLC isoforms in *Populus*, viz., KLCR400 (Potri.014G100400), KLCR3200 (Potri.003G143200), KLCR7800 (Potri.001G087800) and KLCR4700 (Potri.008G094700) using yeast two-hybrid assays (**Figure 6**). These results suggest that three of the four KLCR proteins tested, KLCR400, KLCR3200 and KLCR7800, interact strongly with PdIQD10 (**Figure 6**). Domain structures of KLCR proteins shows a distinct primary protein structure for the fourth KLCR isoform, KLCR4700, which has an extended N-terminal region and a discontinuous tetratricopeptide repeat (TPRs) region at the N-terminal (**Supplementary Figure S7**). The C-terminal regions of these four KLCR proteins are similar in their domain architectures. It is plausible that interactions of PdIQD10 with specific KLCR proteins may be determined by variations in the N-terminal domain structure. Observed interaction of PdIQD10 with three out of four tested KLCRs indicates two things: first, PdIQD10 may interact with specific KLCRs and second, PdIQD10 might act as a cargo or a cargo associated (molecular scaffold) protein similar to the

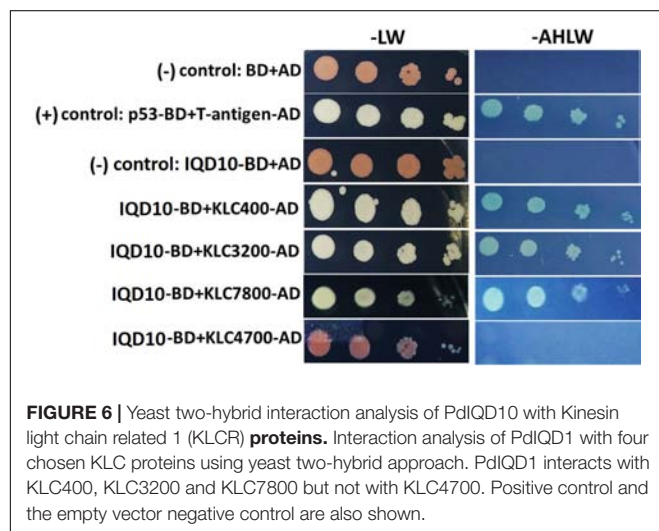
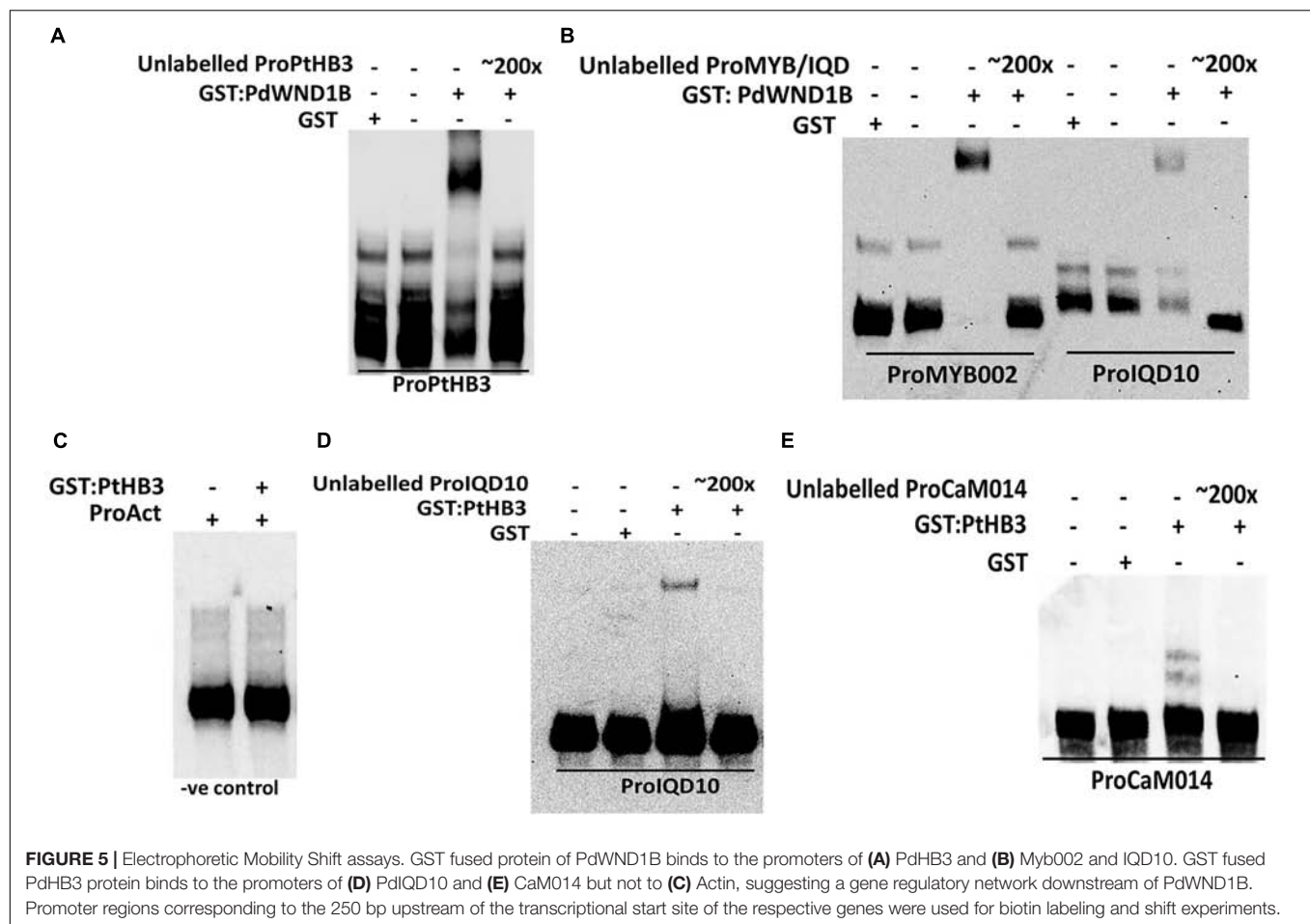
AtIQD1. Furthermore, the localization of PdIQD10 to the plasma membrane is in agreement with the understanding that cargos along the microtubules are generally transported by kinesins to the cell-periphery (Hirokawa et al., 2009; Verhey et al., 2011).

Additional potential interactor proteins were picked from publicly available co-expression databases. Selected proteins included in yeast two hybrid assays fell into the known or putative functional categories of cell wall formation (such as, cellulose synthase: PtiCesA7-A, Galacturonosyltransferase: GAUT12.1, glycosyltransferase family 47 member involved in xylan backbone synthesis; IRX10), cytoskeletal rearrangement (Microtubule associated proteins: MAP599, MAP1733, MAP2698), vascularization (NIMA related kinases: NEK5, NEK6) and/or transcriptional regulation of secondary cell wall formation (NAC domain transcription factor, PdWND1B), which are aligned with the functional context of PdIQD10 proposed here. Based on measured β -galactosidase activity, NEK6, SCPL14 and SCPL41 (Serine carboxypeptidase-like), appear to display weak interactions with PdIQD10 while assays with other proteins showed no interaction with PdIQD10 (**Supplementary Figure S8**).

PdIQD10 RNAi-Downregulated Transgenic Lines Show Altered Biomass Properties

In order to evaluate the functional role of PdIQD10 in stem development and chemistry, transgenic *Populus deltoides*

⁶<http://atted.jp/>



plants with stable, RNAi-mediated, downregulation of the *PdIQD10* gene were generated (**Figure 7B**). Preliminary growth assessment of 6-month-old greenhouse-grown *PdIQD10* RNAi downregulated lines relative to empty vector controls and a previously reported comparator *PdKOR* RNAi line with reduced

growth phenotype (**Figure 7A**) (Kalluri et al., 2016) showed that *PdIQD10* lines may have increased growth relative to controls. For a deep dive into phenotypic characterization of *PdIQD10* RNAi plants, three biological replicates each of three independent transformation lines of *PdIQD10* and empty vector constructs was undertaken. These assessments showed that *PdIQD10* RNAi lines displayed greater plant height, stem diameter and stem density (**Figures 7C–E**) as compared to the empty vector control.

Sugar composition analysis of cell wall showed higher glucose level but no significant consistent change in galactose, xylose, and arabinose content in independent RNAi lines relative to the control (**Figure 8**). Altered glucose level in *PdIQD10* RNAi lines relative to control suggest the potential functional significance of PdIQD10 in wall cellulose and hemicellulose composition and secondary cell wall biosynthesis pathways. Wet chemistry-based quantification of cellulose content in dried stem samples showed a higher percentage of cellulose in the *PdIQD10* RNAi samples relative to control (**Figure 8**). NMR techniques showed that a higher cellulose crystallinity, while gel permeation chromatography analysis showed a lower degree of polymerization for cellulose in the RNAi samples (**Figure 8**). MBMS analysis of lignin content and S/G ratios suggests that the impact of *PdIQD10* RNAi on lignin content was not significant (**Figure 8**).

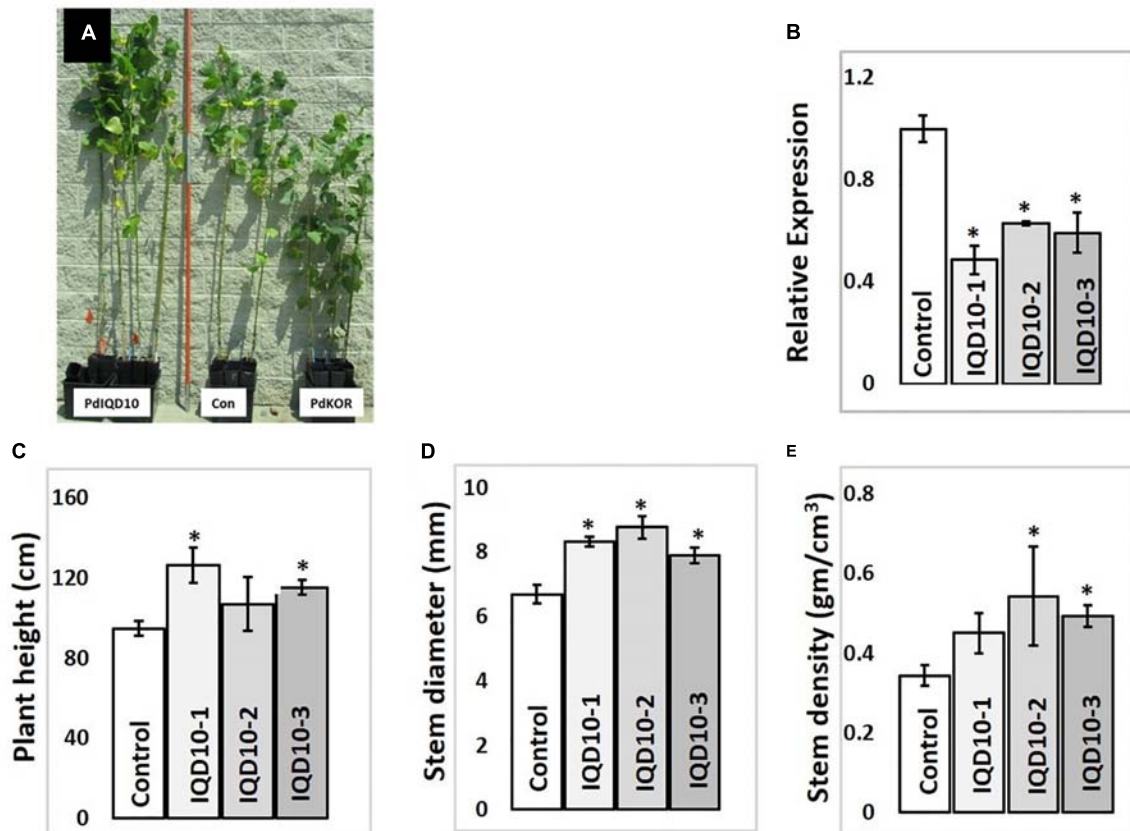


FIGURE 7 | Growth characterization of *PdIQD10* RNAi lines **(A)** 6-month -old greenhouse grown transgenic plants representing *IQD* RNAi (*PdIQD10-1*), empty vector control and a comparator reduced biomass line *PdKOR* RNAi (Kalluri et al., 2016). **(B)** qRT-PCR validation of RNAi-mediated downregulation of *PdIQD10* gene in three independent RNAi lines. **(C)** Measurements of average plant height, **(D)** stem diameter and **(E)** stem density of *PdIQD10* RNAi lines in comparison with that of vector control lines. Data represent means \pm SE ($n = 3-5$). *Indicates statistically significant, $p \leq 0.05$ based on Student's *t*-tests.

PdIQD Interacting KLC Proteins Show Transactivator Function

In addition to proposed cellular roles of KLC proteins in plants in signal and cargo transport along the microtubular network, a cellular functional role in transcriptional activation has also been reported in literature (Li et al., 2011, 2012). These studies reported that BC12/GDD1 (Gibberellin-Deficient Dwarf1), a rice Kinesin-Like Protein that is bound to microtubules in an ATP-dependent manner also binds to the promoter of ent-kaurene oxidase (KO2) – an enzyme in gibberellic acid (GA) biosynthesis). Protoplast assays revealed that GDD1 has transcriptional activation activity and that T-DNA insertion line *gdd1* has reduced accumulation of GA. Our Yeast two-hybrid experiments designed to test such a transactivator ability for *Populus* KLCs showed that three out of four KLC protein isoforms were able to autoactivate the transcription of reporter genes *HIS3*, *ADE2* and *MEL1* without the interacting partner protein (Figure 9A). To rule out the possibility that this might be the result of targeted nuclear localization in yeast two-hybrid system and to test these results *in planta*, we made use of *Populus* protoplast transient system. The Gal4

DNA binding domain (GD) fused KLCs were co-transfected with *Gal4:GUS* reporter construct to test if KLCs induce GUS transcript activation. GD-KLCs were able to activate the transcription of *Gal4* promoter-fused GUS reporter suggesting the transcriptional activation functions of KLCs (Figure 9B). GD-KLC7800 transformation induced highest GUS activity indicating its relatively stronger transactivation ability relative to GD-KLC3200, GD-KLC400, and GD-KLC4700, displaying lower and differential transactivation strengths.

Yeast two-hybrid analysis to dissect the interacting domains of *PdIQD10* with KLCs provided additional insights. *PdIQD10* protein spanning from the initial start codon to the end of *PdIQD10*-domain is referred to here as IQD10a, from start of the *PdIQD10*-domain up till the stop codon as IQD10b and from the end of *PdIQD10*-domain up to the stop codon as IQD10c (Figure 10A). Yeast transformants carrying Binding Domain (BD) fused IQD10c was able to autoactivate the transcription of *HIS3* reporter thus inducing Yeast growth on minimal media lacking Histidine, Leucine and Tryptophan (Figure 10A) without the interacting partner KLC400. This observation indicated two mechanisms: (i) IQD10c might have transactivation functions and (ii) The N-terminal *PdIQD10*-domain fragment might be

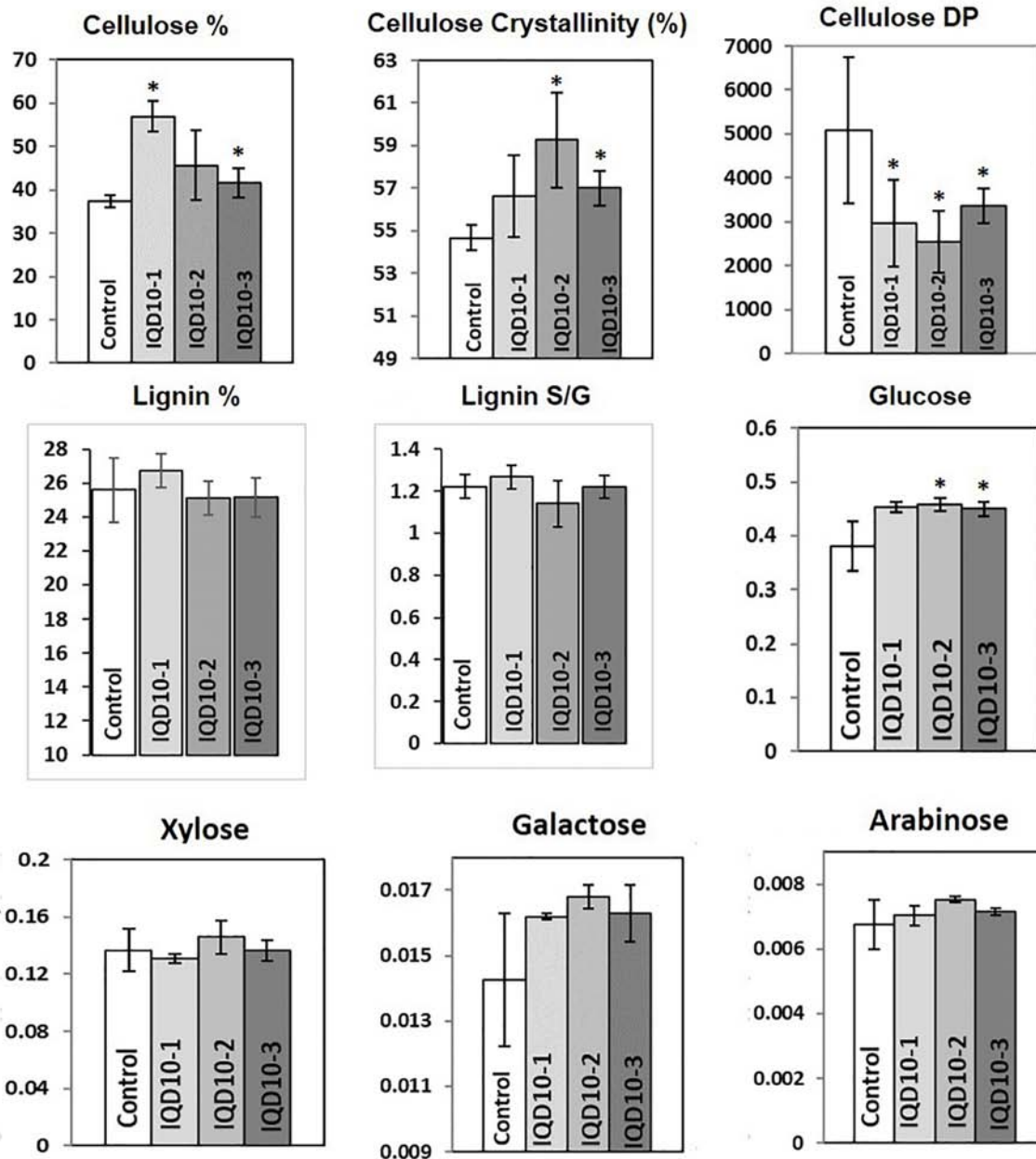


FIGURE 8 | Cell wall characteristics of *PdIQD10* RNAi lines. Wet chemistry analysis of cellulose content, NMR analysis of cellulose crystallinity, GPC analysis of cellulose DP, MBMS analysis of lignin content and S/G ratio, and analysis of sugar composition of cell walls in control and three independent *PdIQD10* RNAi lines. Data represent means \pm SD ($n = 3-7$). *Indicates statistically significant, $p \leq 0.05$ based on Student's *t*-tests.

involved in suppressing the transactivation activity of IQD10c as the full-length BD fused PdIQD10 does not show autoactivation in Yeast (Figures 3A, 10 and Supplementary Figure S8). Based on these observations, transactivation properties of PdIQD10 full length and IQD10c proteins were tested with or without KLC protein co-transfection in *Populus* protoplasts.

The protoplasts transfected with empty GD vector and GD fused PdIQD10 displayed weak GUS activity to the same extent suggesting full length PdIQD10 has no transactivation

activity (Figure 10B). However, transfection of GD-fused IQD10c induced GUS activity more than twice that of the GD control and full length PdIQD10 indicating GD-IQD10c may be able to activate the *Gal4* promoter. Furthermore, co-transfections with KLC7800 and KLC3200 showed induction of PdIQD10 transactivation activity to a similar level as IQD10c, which is higher relative to control (Figure 10B). These results suggest a potential activation effect of KLC proteins on PdIQD10.

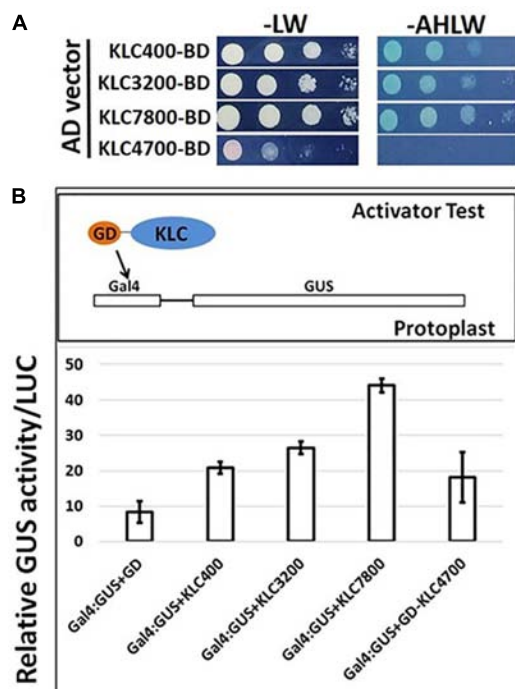


FIGURE 9 | Transactivation analysis of KLC proteins. **(A)** Yeast two-hybrid assays with KLCs in BD vector co-transfected with empty AD vectors and grown on minimal media lacking adenine, histidine, leucine and tryptophan. **(B)** Protoplasts co-transfected with GD-fused KLCs along with *Gal4*:GUS reporter construct display GUS activity.

DISCUSSION

The present study provides evidence in support of a functional role for a new calmodulin-binding protein member in the contexts of secondary cell wall biosynthesis based on

expression and molecular biology studies, and in the context of biomass formation based on transgenic *Populus* plant characterization.

PdIQD10 Is Preferentially Expressed in the Context of Secondary Cell Wall Formation and Is Potentially Regulated by Secondary Cell Wall Transcription Factors

The transcript accumulation of *PdIQD10* was significantly higher in tension-stressed, secondary wall-enriched, xylem tissue along with other known cell wall marker genes such as sucrose synthase (*SUSY*), cellulose synthases (*CESAs*), and *KOR* (*KORRIGAN*) (**Figure 1A**). The qRT-PCR assay of *PdIQD10* gene expression showed highest levels in xylem tissue (> 100-fold) relative to other tissue/organ libraries (**Figure 1B**).

Promoter binding assays support a functional context for *PdIQD10* in the secondary cell wall biosynthesis transcriptional regulatory network where the *Populus* homolog of the known master regulator SND1, PdWND1B, binds to *PdHb3* promoter and *PdHb3* protein in turn binds to *PdIQD10* promoter.

PdIQD10 Is Potentially a Component of Multi-Protein Signaling Complex

The members of the IQD family are known for their interaction with the calmodulin and calmodulin-like proteins (Bahler and Rhoads, 2002; Abel et al., 2013; Bürstenbinder et al., 2013, 2017b). We have shown that the full-length *PdIQD10* interacts weakly with CaM014 and CaM2599 out of 26 CaMs in yeast-two-hybrid assays. The *PdIQD10*-domain interacts with CaM247 and CaM351 but not CaM014 and CaM2599. In agreement with *Arabidopsis* IQD1 studies (Bürstenbinder et al., 2013), *PdIQD10* was found to interact with KLC proteins (**Figure 6**). Subcellular localization experiments reveal that *PdIQD10*,

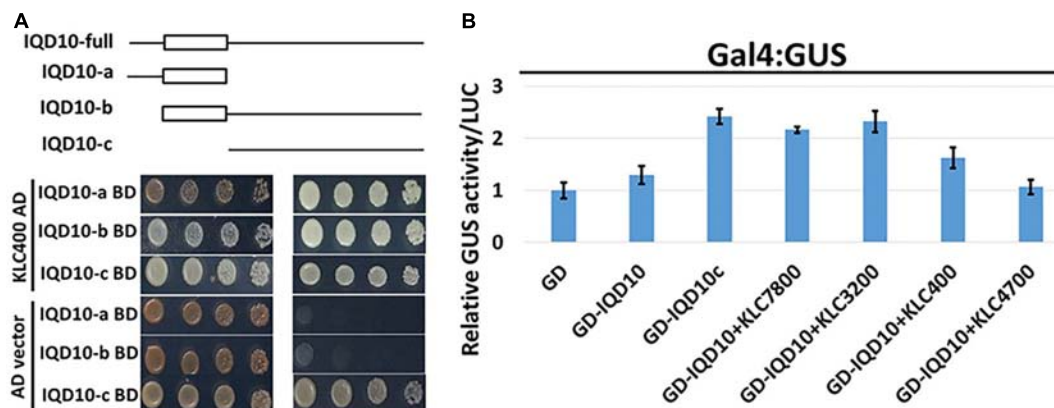
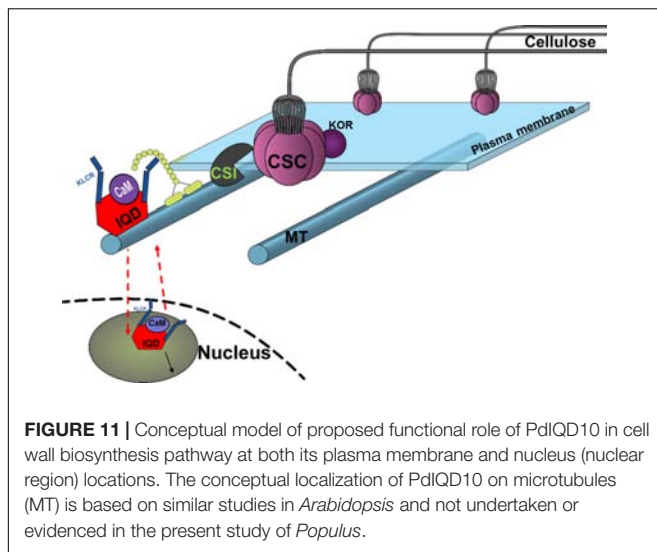


FIGURE 10 | Domain split analysis of *PdIQD10* using transactivation assay. **(A)** Yeast two-hybrid analysis of *PdIQD10* domains with KLC400 indicates the possible interaction of KLC400 with IQD10a and IQD10b fragments. Also, the empty vector AD control transfection with split *PdIQD10* domains in BD shows autoactivation indicating its transactivation functions. **(B)** Protoplasts co-transfected with *Gal4*:GUS reporter construct with versions of *PdIQD10* in GD-fused constructs, in combination or not with KLCs, show increase in GUS activities indicating increased transactivation activity of *PdIQD10* upon interaction with KLCs.



CaM014 and CaM2599 localize to the plasma membrane and the nucleus (**Figures 4A–D**). These observations indicate the possible complex formation between PdIQD10, CaMs and the kinesin-light chain proteins, similar to the model proposed for *Arabidopsis* IQD1 protein complex (Bürstenbinder et al., 2013) (**Figure 11**).

The putative PdIQD10 interacting partner proteins selected from various databases (**Table 3**) either have known cell wall related functions or represent potential members in cell wall or stem development pathways. Yeast two hybrid assays showed weak or no interactions with selected co-expressing factors such as NEK5, NEK6, PtiCesA7-A, IRX10, GAUT12.1, SCPL14, and SCPL41 (**Supplementary Figure S8**). The weak interactions may be due to the physiology of the yeast and the lack of the functional scaffolding protein for two proteins to interact or a measure of no functional competence for biological interaction between targeted proteins. CesA complexes (CSCs) are known to move in plasma membrane guided by the underlying cortical microtubular network and directionally deposit cellulose microfibrils in the cell wall (Gu et al., 2010; Endler and Persson, 2011). Studies from Lei et al. provided evidence that CSI interact with CesAs and regulate their function in a microtubule dependent or independent manner (Lei et al., 2013) as well as their turnover or recycling from plasma membrane (Lei et al., 2015). Golgi complex and vesicle trafficking network are integral to the cellulose biosynthesis process with CSCs known to assemble in the Golgi complex. More recently, a Golgi-localized protein, STELLO, has been identified as a key regulator of cellulose biosynthesis via its functional role in assembly of CSCs in Golgi and its trafficking to the plasma membrane (Zhang et al., 2016). These recent studies suggest there are additional signaling and regulatory factors involved in the coordinated recruitment of CSCs to microtubules, and their integration into and turnover in plasma membrane. Functional characterization of proteins co-expressing and/or interacting with factors implicated in cell wall pathways would lead to new insights. Based on the findings in the present study, we hypothesize that cell wall

biosynthesis components or signals regulating the bioprocess may be carried as potential cargo in PdIQD10-KLC mediated directional transport along the microtubules (Bürstenbinder et al., 2013) (**Figure 11**).

PdIQD10 Is a Potential New Signaling Factor in the Secondary Cell Wall and Stem Growth Bioprocesses

Our protein interaction studies support the role of predicted calmodulin-binding *PdIQD10* gene as a CaM-binding protein. CaMs and Ca^{2+} -signaling encompasses a core signaling mechanism to mediate signal transduction from plasma membrane, cytosolic and nuclear compartments. Our subcellular localization assays show that PdIQD10 localizes both in the nucleus and plasma membrane as well as bind to other known xylem expressing, signaling complex proteins. Considering IQD's basic cellular role in Ca^{2+} /CaM signaling and the strong expression of this particular IQD, *PdIQD10*, in the context of cells undergoing secondary wall biosynthesis, along with phenotypic ramifications on secondary cell wall chemistry and biomass formation, we propose that PdIQD10 is a potential new signaling factor in the secondary cell wall and stem growth bioprocesses. Follow-on studies are needed to clarify the functional role as general or specific to aspects of cell wall formation.

The nuclear localization and transactivation ability of KLC proteins and their influence on the transactivation activity of PdIQD10 suggest that the complex of PdIQD10-KLCs may potentially regulate genes involved in secondary cell wall biosynthesis or stem developmental processes. Further studies are needed to clarify the additional protein composition of the complex and genes regulated by such a complex and their cascading mechanistic influence on *Populus* biomass properties.

RNAi lines interestingly displayed increased biomass formation as reflected by their increased plant height, stem diameter and stem density. RNAi plant stem samples also showed a changes in cellulose properties and wall sugar composition relative to control stems. The observation of increased stem growth and cellulose content via knockdown of *PdIQD10* is interesting, which leads to the new hypothesis that this gene, in part, partakes in mediating the tight feedback-loop in regulation of cellulose content and biosynthesis in secondary walls in *Populus* stems. The known co-expression of *PdIQD10* with other secondary cell wall polysaccharide pathway genes such as *PdIRX10*, beta-1,4-xylosyltransferase, a homolog of *AtIRX10* implicated in hemicellulose biosynthesis, combined with the observation that *PdIQD10* RNAi lines have altered wall glucose levels relative to control, suggest a potential role of *PdIQD10* gene in secondary cell wall (cellulose and hemicellulose) biosynthesis pathways. Our yeast-two-hybrid assays with PdIQD10 and selected proteins identified to be co-expressing from public databases showed weak or no interactions (**Supplementary Figure S8**). Given the frequency of false positives and false negatives in yeast-two hybrid assays in general, observations here will need alternate lines of evidence such as *in vivo* protein-protein interactions,

protein pull-down assays or computational interactome predictions.

CONCLUSION

The present study provides evidence in support of a new functional context for an IQD family member in secondary cell wall biosynthesis and biomass formation. Specifically, the study shows that *PdIQD10* gene codes for a calmodulin-binding protein, expresses under tissue contexts of higher secondary cell wall formation, potentially regulated by secondary cell wall transcription factor, has subcellular localization contexts of plasma membrane and nucleus, and forms protein complexes with CaM and KLC proteins. Stable knockdown of gene expression in transgenic *PdIQD10* RNAi *Populus* plants impacted the properties of resultant stem biomass. *PdIQD10* RNAi plants displayed enhanced growth, accompanied by quantitatively modest, yet significant, concomitant changes in cellulose content and crystallinity and wall sugar composition relative to the control.

In the future, quantitative analysis of cell wall pathway and interactors proteins in *PdIQD10* RNAi and control lines and pulldown assays with tagged-IQD overexpression lines will aid in clarifying *PdIQD10*'s involvement in signaling and response for biosynthesis of a specific wall component, cellulose, wall remodeling and integrity sensing, and/or stem developmental programs. The significant phenotypic effect on cellulose properties observed here will require clarification of a direct impact on cellulose biosynthesis pathway by imaging anomalies in movement or turnover of GFP-tagged CesAs in *IQD10*-modified plants. Experimental studies designed to probe the binding potential of *PdIQD10* to microtubule proteins in the presence or absence of calcium ions or calcium chelators would further inform its molecular activity as a signaling protein. While the present study is a first characterization of a single isoform from the large IQD protein family in *Populus*, the full functional repertoire of the large IQD gene family in *Populus* is still unknown. Last but not the least, garnering functional genomics evidence for additional genes of unknown function strongly co-expressing with xylem or secondary cell wall development marker genes will greatly expand the critical knowledge base needed to understand and optimize plant cell wall properties.

AUTHOR'S NOTE

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE

Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

AUTHOR CONTRIBUTIONS

RB designed and conducted experiments to understand molecular function including protein-interaction assays and wrote manuscript. RP designed and conducted transgenic plant phenotyping, expression analysis and wrote manuscript. SJ and LG carried out plant growth and phenotyping. XY designed the RNAi construct, H-BG conducted MD simulations. KW, CC, and WR generated transgenic plants in tissue culture and stool beds. KY and MR conducted wall chemistry assays. RS, SD, and MD carried out MBMS and biomass sugar assays. GB and AR conducted NMR and GPC analysis of cellulose. GT and UK conceived the study and wrote the manuscript.

FUNDING

The work was supported by U.S DOE BioEnergy Science Center and the Center for Bioenergy Innovation. The BioEnergy Science Center and the Center for Bioenergy Innovation are US Department of Energy Bioenergy Research Centers, supported by the Office of Biological and Environmental Research in the Department of Energy's Office of Science. The funding body has no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

ACKNOWLEDGMENTS

The authors wish to thank Zack Moore for assistance with plant care and propagation of materials employed in the greenhouse experiments. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DEAC05-00OR22725.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.01669/full#supplementary-material>

FIGURE S1 | Construct design of the *PdIQD10* RNAi construct.

FIGURE S2 | Expression analysis of *PdIQD10* in LCM-derived stem tissue and xylem cell types of *Populus* using database published by Shi et al. (2017).

FIGURE S3 | Multiple sequence alignment of CaM014 and CaM2599 protein sequences.

FIGURE S4 | Multiple sequence alignment of CaM247, CaM014, and CaM2599 protein sequences showing the shorter N-terminal region of CaM247.

FIGURE S5 | Phylogenetic tree of calmodulin/calmodulin-like (CaM) gene family members of *Populus* and *Arabidopsis*. The tree was generated using Maximum Likelihood algorithm in MEGA7.0.25 with 1000 bootstrap replicates and

represents 36 *Populus* and 25 *Arabidopsis* calmodulin/calmodulin-like protein sequences.

FIGURE S6 | Subcellular localization of CaM247 and CaM014 in *Populus* protoplasts. To increase accessibility of these subcellular localization images, the yellow color channel was converted to magenta uniformly across all images in the CMYK color spectrum. The original RGB color scheme images are provided in **Supplementary Figure S10**. The color scheme is as follows: GFP/RFP: blue/cyan; chlorophyll, FM64 and mCherry: red/orange.

FIGURE S7 | Domain Architecture of the four *Populus* KLC proteins.

REFERENCES

- Abel, S., Bürstenbinder, K., and Müller, J. (2013). The emerging function of IQD proteins as scaffolds in cellular signaling and trafficking. *Plant Signal. Behav.* 8:e24369. doi: 10.4161/psb.24369
- Abel, S., Savchenko, T., and Levy, M. (2005). Genome-wide comparative analysis of the IQD gene families in *Arabidopsis thaliana* and *Oryza sativa*. *BMC Evol. Biol.* 5:72. doi: 10.1186/1471-2148-5-72
- Abraham, P., Adams, R., Giannone, R. J., Kalluri, U., Ranjan, P., Erickson, B., et al. (2012). Defining the boundaries and characterizing the landscape of functional genome expression in vascular tissues of *Populus* using shotgun proteomics. *J. Proteome Res.* 11, 449–460. doi: 10.1021/pr200851y
- Bahler, M., and Rhoads, A. (2002). Calmodulin signaling via the IQ motif. *FEBS Lett.* 513, 107–113. doi: 10.1016/S0014-5793(01)03239-2
- Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., et al. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J. Chem. Theory Comput.* 8, 3257.
- Biswal, A. K., Hao, Z., Pattathil, S., Yang, X., Winkler, K., Collins, C., et al. (2015). Downregulation of GAUT12 in *Populus deltoides* by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. *Biotechnol. Biofuels* 8:41. doi: 10.1186/s13068-015-0218-y
- Bürstenbinder, K., Mitra, D., and Quegwer, J. (2017a). Functions of IQD proteins as hubs in cellular calcium and auxin signaling: a toolbox for shape formation and tissue-specification in plants? *Plant Signal. Behav.* 12:e1331198. doi: 10.1080/15592324.2017.1331198
- Bürstenbinder, K., Möller, B., Plötner, R., Stamm, G., Hause, G., Mitra, D., et al. (2017b). The IQD family of calmodulin-binding proteins links calmodulin signaling to microtubules, membrane microdomains, and the nucleus. *Plant Physiol.* 173, 1692–1708. doi: 10.1104/pp.16.01743
- Bürstenbinder, K., Savchenko, T., Müller, J., Adamson, A. W., Stamm, G., Kwong, R., et al. (2013). Arabidopsis calmodulin-binding protein IQ67-domain 1 localizes to microtubules and interacts with kinesin light chain-related protein-1. *J. Biol. Chem.* 288, 1871–1882. doi: 10.1074/jbc.M112.396200
- Busse-Wicher, M., Grantham, N. J., Lyczakowski, J. J., Nikolovski, N., and Dupree, P. (2016). Xylan decoration patterns and the plant secondary cell wall molecular architecture. *Biochem. Soc. Trans.* 44, 74–78. doi: 10.1042/BST20150183
- Chang, Y., Chen, S. L., Yin, W. L., Wang, R. G., Liu, Y. F., Shi, Y., et al. (2006). Growth, gas exchange, abscisic acid, and calmodulin response to salt stress in three poplars. *J. Integr. Plant Biol.* 48, 286–293. doi: 10.1111/j.1744-7909.2006.00194.x
- Clapham, D. E. (2007). Calcium signaling. *Cell* 131, 1047–1058. doi: 10.1016/j.cell.2007.11.028
- Cragg, S. M., Beckham, G. T., Bruce, N. C., Bugg, T. D., Distel, D. L., Dupree, P., et al. (2015). Lignocellulose degradation mechanisms across the tree of life. *Curr. Opin. Chem. Biol.* 29, 108–119. doi: 10.1016/j.cbpa.2015.10.018
- Demura, T., and Ye, Z. H. (2010). Regulation of plant biomass production. *Curr. Opin. Plant Biol.* 13, 299–304. doi: 10.1016/j.cbpa.2010.03.002
- Ding, S. Y., Liu, Y. S., Zeng, Y. N., Himmel, M. E., Baker, J. O., and Bayer, E. A. (2012). How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? *Science* 338, 1055–1060. doi: 10.1126/science.1227491
- Dodd, A. N., Kudla, J., and Sanders, D. (2010). The language of calcium signaling. *Annu. Rev. Plant Biol.* 61, 593–620. doi: 10.1146/annurev-arplant-070109-104628
- Du, J., Miura, E., Robischon, M., Martinez, C., and Groover, A. (2011). The *Populus* Class III HD ZIP transcription factor POPCORONA affects cell differentiation during secondary growth of woody stems. *PLoS One* 6:e17458. doi: 10.1371/journal.pone.0017458
- Endler, A., and Persson, S. (2011). Cellulose synthases and synthesis in *Arabidopsis*. *Mol. Plant* 4, 199–211. doi: 10.1093/mp/ssq079
- Fu, C., Mielenz, J. R., Xiao, X., Ge, Y., Hamilton, C. Y., Rodriguez, M., et al. (2011). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3803–3808. doi: 10.1073/pnas.1100310108
- Gall, D. L., Ralph, J., Donohue, T. J., and Noguera, D. R. (2017). Biochemical transformation of lignin for deriving valued commodities from lignocellulose. *Curr. Opin. Biotechnol.* 45, 120–126. doi: 10.1016/j.copbio.2017.02.015
- Gu, F., and Nielsen, E. (2013). Targeting and regulation of cell wall synthesis during tip growth in plants. *J. Integr. Plant Biol.* 55, 835–846. doi: 10.1111/jipb.12077
- Gu, Y., Kaplinsky, N., Bringmann, M., Cobb, A., Carroll, A., Sampathkumar, A., et al. (2010). Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. *Proc. Natl. Acad. Sci. U S A* 107, 12866–12871. doi: 10.1073/pnas.1007092107
- Guo, J., Morrell-Falvey, J. L., Labbé, J. L., Muchero, W., Kalluri, U. C., Tuskan, G. A., et al. (2012). Highly efficient isolation of populus mesophyll protoplasts and its application in transient expression assays. *PLoS One* 7:e44908. doi: 10.1371/journal.pone.0044908
- Hirokawa, N., Noda, Y., Tanaka, Y., and Niwa, S. (2009). Kinesin superfamily motor proteins and intracellular transport. *Nat. Rev. Mol. Cell Biol.* 10, 682–696. doi: 10.1038/nrm2774
- Hui, M., Lin, F., Zhu, C., Xue, C., Zhao, H., and Yan, X. (2014). Genome-wide identification and expression analysis of the IQD gene family in *Populus trichocarpa*. *Plant Sci.* 229, 96–110. doi: 10.1016/j.plantsci.2014.08.017
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi: 10.1063/1.445869
- Kalluri, U. C., Hurst, G. B., Lankford, P. K., Ranjan, P., and Pelletier, D. A. (2009). Shotgun proteome profile of *Populus* developing xylem. *Proteomics* 9, 4871–4880. doi: 10.1002/pmic.200800854
- Kalluri, U. C., Payyavula, R. S., Labbé, J. L., Engle, N., Bali, G., Jawdy, S. S., et al. (2016). Down-regulation of KORRIGAN-Like Endo- β -1,4-glucanase genes impacts carbon partitioning, mycorrhizal colonization and biomass production in *Populus*. *Front. Plant Sci.* 7:1455. doi: 10.3389/fpls.2016.01455
- Klemm, D., Heublein, B., Fink, H., and Bohn, A. (2005). Cellulose: fascinating biopolymer and sustainable raw material. *Angew. Chem. Int. Ed. Engl.* 44, 3358–3393. doi: 10.1002/anie.200460587
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Le, P. Y., Jeon, H. W., Kim, M. H., Park, E. J., Lee, H., Hwang, I., et al. (2018). Gain-of-function mutation of AtDICE1, encoding a putative endoplasmic reticulum-localized membrane protein, causes defects in anisotropic cell elongation by disturbing cell wall integrity in *Arabidopsis*. *Ann. Bot.* 122, 151–164. doi: 10.1093/aob/mcy049

- Lei, L., Li, S., Du, J., Bashline, L., and Gu, Y. (2013). Cellulose synthase INTERACTIVE3 regulates cellulose biosynthesis in both a microtubule-dependent and microtubule-independent manner in Arabidopsis. *Plant Cell* 25, 4912–4923. doi: 10.1105/tpc.113.116715
- Lei, L., Singh, A., Bashline, L., Li, S., Yingling, Y. G., and Gu, Y. (2015). Cellulose synthase interactive1 is required for fast recycling of cellulose synthase complexes to the plasma membrane in Arabidopsis. *Plant Cell* 27, 2926–2940. doi: 10.1105/tpc.15.00442
- Levy, M., Wang, Q., Kaspi, R., Parrella, M. P., and Abel, S. (2005). Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. *Plant J.* 43, 79–96. doi: 10.1111/j.1365-313X.2005.02435.x
- Li, J., Jiang, J., Qian, Q., Xu, Y., Zhang, C., Xiao, J., et al. (2011). Mutation of rice BC12/GDD1, which encodes a kinesin-like protein that binds to a GA biosynthesis gene promoter, leads to dwarfism with impaired cell elongation. *Plant Cell* 23, 628–640. doi: 10.1105/tpc.110.081901
- Li, J., Xu, Y., and Chong, K. (2012). The novel functions of kinesin motor proteins in plants. *Protoplasma* 249, S95–S100. doi: 10.1007/s00709-011-0357-3
- Lin, S. Z., Zhang, Z. Y., Lin, Y. Z., Zhang, Q., and Guo, H. (2004). The role of calcium and calmodulin in freezing-induced freezing resistance of *Populus tomentosa* cuttings. *J. Plant Physiol. Mol. Biol.* 30, 59–68.
- Lin, Y. C., Li, W., Sun, Y. H., Kumari, S., Wei, H., Li, Q., et al. (2013). SND1 transcription factor-directed quantitative functional hierarchical genetic regulatory network in wood formation in *Populus trichocarpa*. *Plant Cell* 25, 4324–4341. doi: 10.1105/tpc.113.117697
- Mansfield, S. D. (2009). Solutions for dissolution–engineering cell walls for deconstruction. *Curr. Opin. Biotechnol.* 20, 286–294. doi: 10.1016/j.copbio.2009.05.001
- McCann, M. C., and Carpita, N. C. (2008). Designing the deconstruction of plant cell walls. *Curr. Opin. Plant Biol.* 11, 314–320. doi: 10.1016/j.pbi.2008.04.001
- Meilan, R., and Ma, C. (2006). “Poplar (*Populus* spp.),” in *Methods in Molecular Biology, Agrobacterium Protocols*, Vol. 34, ed. K. Wang (Totowa, NJ: Humana Press), 143–151. doi: 10.1385/1-59745-131-2:143
- Mielenz, J. R., Sykes, R., Yung, M., Novaes, E., Kirst, M., et al. (2009). “High-throughput screening of plant cell-wall composition using pyrolysis molecular beam mass spectroscopy,” in *Biofuels: Methods and Protocols, Methods in Molecular Biology*, Vol. 581, ed. J. R. Mielenz (New York, NY: Humana Press), 169–183.
- Mizrachi, E., Mansfield, S. D., and Myburg, A. A. (2012). Cellulose factories: advancing bioenergy production from forest trees. *New Phytol.* 194, 54–62. doi: 10.1111/j.1469-8137.2011.03971.x
- Nakagawa, T., Ishiguro, S., and Kimura, T. (2009). Gateway vectors for plant transformation. *Plant Biotechnol.* 26, 275–284. doi: 10.5511/plantbiotechnology.26.275
- Oikawa, K., Yamasato, A., Kong, S. G., Kasahara, M., Nakai, M., Takahashi, F., et al. (2008). Chloroplast outer envelope protein CHUP1 is essential for chloroplast anchorage to the plasma membrane and chloroplast movement. *Plant Physiol.* 148, 829–842. doi: 10.1104/pp.108.123075
- Payyavula, R. S., Tschaplinski, T. J., Jawdy, S. S., Sykes, R. W., Tuskan, G. A., and Kalluri, U. C. (2014). Metabolic profiling reveals altered sugar and secondary metabolism in response to UGPase overexpression in *Populus*. *BMC Plant Biol.* 14:265. doi: 10.1186/s12870-014-0265-8
- Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., et al. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802. doi: 10.1002/jcc.20289
- Porth, I., Maghuly, F., El-Kassaby, Y. A., and Mansfield, S. (2018). Localization of gene expression, tissue specificity of *Populus xylosyltransferase* genes by isolation and functional characterization of their promoters. *Plant Cell Tissue Organ Cult.* 134, 503–508. doi: 10.1007/s11240-018-1426-5
- Quentin, M., Baures, I., Hoeffle, C., Caillaud, M. C., Allasia, V., Panabieres, F., et al. (2016). The Arabidopsis microtubule associated protein MAP65-3 supports infection by filamentous biotrophic pathogens by down-regulating salicylic acid-dependent defenses. *J. Exp. Bot.* 67, 1731–1743. doi: 10.1093/jxb/erv564
- Robischon, M., Du, J., Miura, E., and Groover, A. (2011). The *Populus* class III HD ZIP, popREVOLUTA, influences cambium initiation and patterning of woody stems. *Plant Physiol.* 155, 1214–1225. doi: 10.1104/pp.110.167007
- Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5
- Ruprecht, C., Mutwil, M., Saxe, F., Eder, M., Nikoloski, Z., and Persson, S. (2011). Large-scale co-expression approach to dissect secondary cell wall formation across plant species. *Front. Plant Sci.* 2:23. doi: 10.3389/fpls.2011.00023
- Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315
- Shi, R., Wang, J. P., Lin, Y. C., Li, Q., Sun, Y. H., Chen, H., et al. (2017). Tissue and cell-type co-expression networks of transcription factors and wood component genes in *Populus trichocarpa*. *Planta* 245, 927–938. doi: 10.1007/s00425-016-2640-1
- Sibout, R., Proost, S., Hansen, B. O., Vaid, N., Giorgi, F. M., Ho-Yue-Kuang, S., et al. (2017). Expression atlas and comparative coexpression network analyses reveal important genes involved in the formation of lignified cell wall in *Brachypodium distachyon*. *New Phytol.* 215, 1009–1025. doi: 10.1111/nph.14635
- Snedden, W. A., and Fromm, H. (2001). Calmodulin as a versatile calcium signal transducer in plants. *New Phytol.* 151, 35–66. doi: 10.1046/j.1469-8137.2001.00154.x
- Updegraff, D. M. (1969). Semimicro determination of cellulose in biological materials. *Anal. Biochem.* 32, 420–424. doi: 10.1016/S0003-2697(69)80009-6
- Verhey, K. J., Kaul, N., and Soppina, V. (2011). Kinesin assembly and movement in cells. *Annu. Rev. Biophys.* 40, 267–288. doi: 10.1146/annurev-biophys-042910-155310
- Vigneault, F., Lachance, D., Cloutier, M., Pelletier, G., Levasseur, C., and Seguin, A. (2007). Members of the plant NIMA-related kinases are involved in organ development and vascularization in poplar, Arabidopsis and rice. *Plant J.* 51, 575–588. doi: 10.1111/j.1365-313X.2007.03161.x
- Voxeur, A., Soubigou-Taconnat, L., Legée, F., Sakai, K., Antelme, S., Durand-Tardif, M., et al. (2017). Altered lignification in mur1-1 a mutant deficient in GDP-L-fucose synthesis with reduced RG-II cross linking. *PLoS One* 12:e0184820. doi: 10.1371/journal.pone.0184820
- Waadt, R., and Kudla, J. (2008). In planta visualization of protein interactions using bimolecular fluorescence complementation (BiFC). *CSH Protoc.* 2008:pdb.prot4995. doi: 10.1101/pdb.prot4995
- Wang, S., Chang, Y., Guo, J., and Chen, J.-G. (2007). Arabidopsis Ovate Family Protein 1 is a transcriptional repressor that suppresses cell elongation. *Plant J.* 50, 858–872. doi: 10.1111/j.1365-313X.2007.03096.x
- Wilson-Sánchez, D., Martínez-López, S., Navarro-Cartagena, S., Jover-Gil, S., and Micol, J. L. (2017). Members of the DEAL subfamily of the DUF1218 gene family are required for bilateral symmetry but not for dorsoventrality in Arabidopsis leaves. *New Phytol.* 217, 1307–1321. doi: 10.1111/nph.14898
- Wong, B. (2011). Points of view: color blindness. *Nat. Methods* 8:441. doi: 10.1038/nmeth.1618
- Wong, Y. L., and Rice, S. E. (2010). Kinesin's light chains inhibit the head- and microtubule-binding activity of its tail. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11781–11786. doi: 10.1073/pnas.1005854107
- Xu, L., Wang, X. J., Wang, T., and Li, L. B. (2017). Genome-wide identification, classification, and expression analysis of the phytoecanin gene family in *Phalaenopsis equestris*. *Biol. Plant* 61, 445–452. doi: 10.1007/s10535-017-0716-9
- Xu, P., Kong, Y., Li, X., and Li, L. (2013). Identification of molecular processes needed for vascular formation through transcriptome analysis of different vascular systems. *BMC Genomics* 14:217. doi: 10.1186/1471-2164-14-217
- Yang, X., Ye, C. Y., Bisaria, A., Tuskan, G. A., and Kalluri, U. C. (2011). Identification of candidate genes in Arabidopsis and Populus cell wall biosynthesis using text-mining, co-expression network analysis and comparative genomics. *Plant Sci.* 181, 675–687. doi: 10.1016/j.plantsci.2011.01.020
- Zhang, Y., Nikolovski, N., Sorieul, M., Vellosillo, T., McFarlane, H. E., Dupree, R., et al. (2016). Golgi-localized STELLA proteins regulate the assembly and trafficking of cellulose synthase complexes in Arabidopsis. *Nat. Commun.* 7:11656. doi: 10.1038/ncomms11656
- Zhao, Y. J., Sun, J. Y., Xu, P., Zhang, R., and Li, L. G. (2014). Intron-mediated alternative splicing of WOOD-ASSOCIATED NAC TRANSCRIPTION FACTOR1B regulates cell wall thickening during fiber development in Populus species. *Plant Physiol.* 164, 765–776. doi: 10.1104/pp.113.231134

- Zhong, R., Demura, T., and Ye, Z. H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*. *Plant Cell* 18, 3158–3170. doi: 10.1105/tpc.106.047399
- Zhu, D., Chu, W., Wang, Y., Yan, H., Chen, Z., and Xiang, Y. (2018). Genome-wide identification, classification and expression analysis of the serine carboxypeptidase-like protein family in poplar. *Physiol. Plant.* 162, 333–352. doi: 10.1111/ppl.12642
- Zhu, Y., Song, D., Sun, J., Wang, X., and Li, L. (2013). PtrHB7, a class III HD-Zip gene, plays a critical role in regulation of vascular cambium differentiation in *Populus*. *Mol. Plant* 6, 1331–1343. doi: 10.1093/mp/sss164

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Badmi, Payyavula, Bali, Guo, Jawdy, Gunter, Yang, Winkeler, Collins, Rottmann, Yee, Rodriguez, Sykes, Decker, Davis, Ragauskas, Tuskan and Kalluri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Progress and Opportunities in the Characterization of Cellulose – An Important Regulator of Cell Wall Growth and Mechanics

Sintu Rongpipi¹, Dan Ye¹, Enrique D. Gomez^{1,2,3*} and Esther W. Gomez^{1,4*}

¹ Department of Chemical Engineering, The Pennsylvania State University, University Park, PA, United States, ² Department of Materials Science and Engineering, The Pennsylvania State University, University Park, PA, United States, ³ Materials Research Institute, The Pennsylvania State University, University Park, PA, United States, ⁴ Department of Biomedical Engineering, The Pennsylvania State University, University Park, PA, United States

OPEN ACCESS

Edited by:

Laura Elizabeth Bartley,
The University of Oklahoma,
United States

Reviewed by:

Doriano Lamba,
Italian National Research Council, Italy
Yunqiao Pu,
Oak Ridge National Laboratory (DOE),
United States

*Correspondence:

Enrique D. Gomez
edg12@psu.edu
Esther W. Gomez
ewg10@psu.edu

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 29 June 2018

Accepted: 06 December 2018

Published: 01 March 2019

Citation:

Rongpipi S, Ye D, Gomez ED and
Gomez EW (2019) Progress
and Opportunities
in the Characterization of Cellulose –
An Important Regulator of Cell Wall
Growth and Mechanics.
Front. Plant Sci. 9:1894.
doi: 10.3389/fpls.2018.01894

The plant cell wall is a dynamic network of several biopolymers and structural proteins including cellulose, pectin, hemicellulose and lignin. Cellulose is one of the main load bearing components of this complex, heterogeneous structure, and in this way, is an important regulator of cell wall growth and mechanics. Glucan chains of cellulose aggregate via hydrogen bonds and van der Waals forces to form long thread-like crystalline structures called cellulose microfibrils. The shape, size, and crystallinity of these microfibrils are important structural parameters that influence mechanical properties of the cell wall and these parameters are likely important determinants of cell wall digestibility for biofuel conversion. Cellulose–cellulose and cellulose–matrix interactions also contribute to the regulation of the mechanics and growth of the cell wall. As a consequence, much emphasis has been placed on extracting valuable structural details about cell wall components from several techniques, either individually or in combination, including diffraction/scattering, microscopy, and spectroscopy. In this review, we describe efforts to characterize the organization of cellulose in plant cell walls. X-ray scattering reveals the size and orientation of microfibrils; diffraction reveals unit lattice parameters and crystallinity. The presence of different cell wall components, their physical and chemical states, and their alignment and orientation have been identified by Infrared, Raman, Nuclear Magnetic Resonance, and Sum Frequency Generation spectroscopy. Direct visualization of cell wall components, their network-like structure, and interactions between different components has also been made possible through a host of microscopic imaging techniques including scanning electron microscopy, transmission electron microscopy, and atomic force microscopy. This review highlights advantages and limitations of different analytical techniques for characterizing cellulose structure and its interaction with other wall polymers. We also delineate emerging opportunities for future developments of structural characterization tools and multi-modal analyses of cellulose and plant cell walls. Ultimately, elucidation of the structure of plant cell walls across multiple length scales will be imperative for establishing structure-property relationships to link cell wall structure to control of growth and mechanics.

Keywords: cellulose microfibrils, cellulose allomorphs, cellulose crystallinity, X-ray diffraction, X-ray scattering, vibrational spectroscopy, nuclear magnetic resonance spectroscopy, atomic force microscopy

INTRODUCTION

The plant cell wall is a complex, heterogeneous network of several polymers and structural proteins. It provides mechanical strength and plays key roles in plant growth, cell differentiation, intercellular communication, water movement, and defense (Cosgrove, 2005). Most higher plants contain both primary and secondary cell walls. The primary cell wall is a thin, flexible, and highly hydrated structure that surrounds the growing cell, while secondary cell wall is a stronger and more rigid structure that starts to deposit when the cell ceases to grow. These cell wall types differ in function, rheological and mechanical properties, and in the arrangement, mobility, and structure of matrix polymers (Cosgrove and Jarvis, 2012). Primary walls are comprised of mainly cellulose, pectin, and xyloglucans with lesser amounts of arabinoxylans and structural proteins. Hydration of the pectin matrix facilitates the slippage and separation of cellulose microfibrils during expansive growth. The strength and rigidity of secondary walls come from a more oriented arrangement of cellulose microfibrils and the presence of lignin. Secondary cell walls are composed mainly of cellulose, lignin, xylans, and glucomannans, and are also less hydrated when compared to primary walls (Cosgrove and Jarvis, 2012).

Cellulose is the primary structural component responsible for much of the mechanical strength of the cell wall. The distribution and orientation of cellulose microfibrils within the cell wall contribute to the control of cell growth. The alignment of microfibrils provides the cell with mechanical anisotropy that enables preferential expansion in one direction (Jordan and Dumais, 2010). In addition to its biological significance, cellulose is an important raw material for textiles, paper, construction materials, and many industrially important chemical derivatives. It is also the most abundant carbohydrate on earth, and is a promising source for renewable energy.

The chemical structure of cellulose consists of linear chains of glucose units linked by β -1,4-glycosidic bonds. Glucan chains of cellulose aggregate via hydrogen bonds and van der Waals forces to form a long thread-like crystalline structure called a cellulose microfibril (Harris et al., 2010). Important structural properties of cellulose include crystallite shape and size and crystallinity. Many different analytical techniques have been employed to study the structure and assembly of cellulose microfibrils in cell walls, yet a comprehensive understanding over multiple length scales remains elusive.

Structural characterization approaches currently used to examine plant cell walls are based on four broad categories of techniques: diffraction/scattering, spectroscopy, microscopy, and physicochemical assays. **Figure 1** highlights these structural characterization tools and the length scales at which they can reveal information about cell wall structure. Solid state ^{13}C nuclear magnetic resonance (NMR) studies led to the discovery of two cellulose allomorphs (VanderHart and Atalla, 1984). The crystal structures of cellulose I α and I β were then determined with the help of X-ray, electron, and neutron diffraction studies (Sugiyama et al., 1991b; Abe et al., 1997; Nishiyama et al., 2002, 2003). Further details about structural differences between these two forms were described by Raman

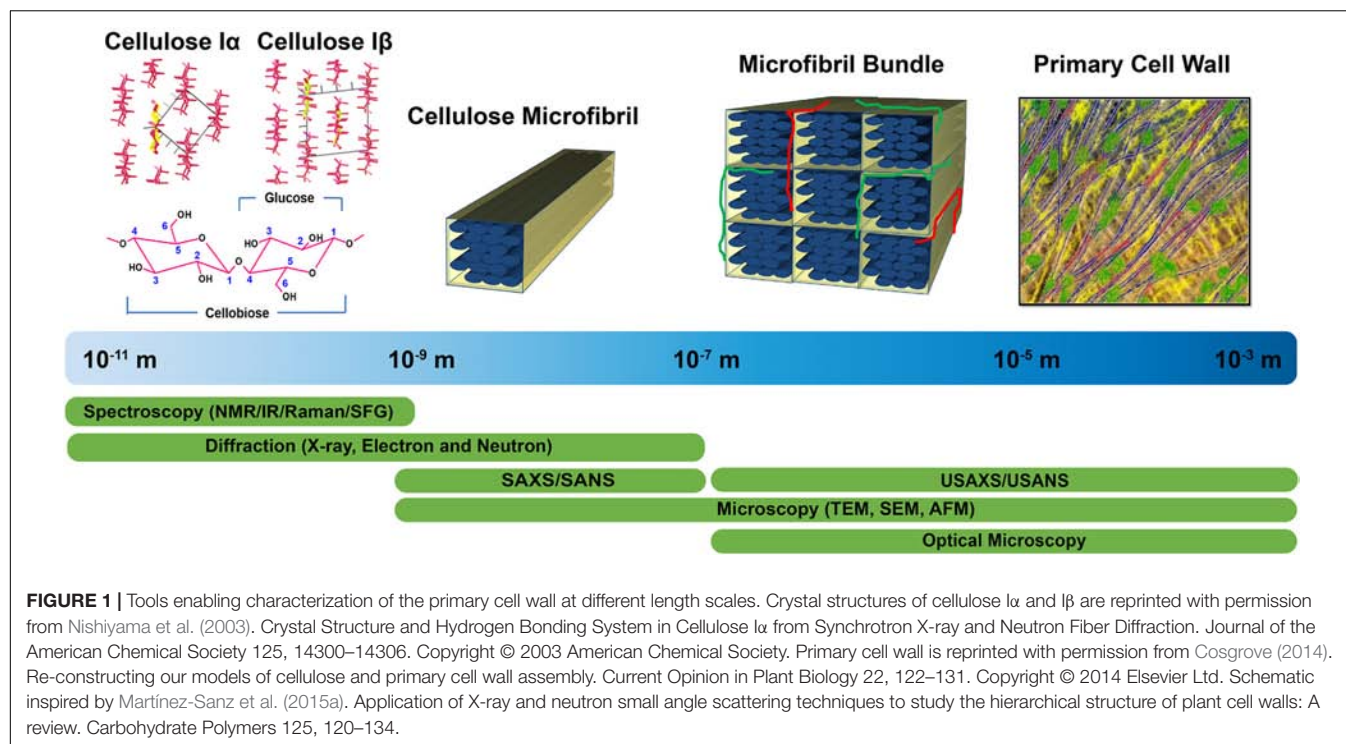
and Fourier-transform infrared (FTIR or IR) spectroscopy, which indicated that glucan chains have similar conformations but differ in hydrogen bonding patterns (Atalla and VanderHart, 1999). The selective detection of cellulose allomorphs is also possible through an emerging spectroscopic technique called sum frequency generation (SFG) spectroscopy (Kim et al., 2013). Beyond the crystal structure, X-ray diffraction (XRD), NMR, and IR and Raman spectroscopy are widely used to estimate the amount of crystalline cellulose present (degree of crystallinity) in plant cell walls. Crystallinity is also determined by some physico-chemical methods, such as the Updegraff method, iodine adsorption, sorption of water vapor, and enthalpy of wetting.

The supramolecular structure of the primary cell wall has been widely characterized by microscopic techniques. Many structural parameters such as crystallite size as well as fibril dimensions, cross-section, and spacing have been directly visualized (Cox and Juniper, 1973; Davies and Harris, 2003; Ding et al., 2014). Electron microscopy has been most widely used to image the fibrillar features of cellulose, but can nevertheless introduce artifacts during sample preparation. Therefore, other microscopic techniques, including scanning probe microscopy, fluorescence microscopy, confocal microscopy, and polarized light microscopy (Abe et al., 1997; Thomson et al., 2007; Choong et al., 2016), are now being explored to visualize the cell wall in its native state with minimal sample preparation.

Complementary to microscopy, the dimensions and packing of cellulose microfibrils are also examined by scattering and spectroscopic techniques (Fernandes et al., 2011; Newman et al., 2013; Zhang et al., 2016). Due to the minimal sample preparation required, scattering is ideal for characterizing the cell wall in its native state. Scattering approaches also offer the benefit of enabling investigation of a large size range, thus allowing for the arrangement of individual microfibrils as well as the aggregates of microfibrils to be examined.

Altogether, the combination of various techniques to characterize the organization of cell wall components opens the door to the examination of interactions between cellulose and other cell wall polysaccharides, potentially revealing various aspects of cell wall assembly (Martínez-Sanz et al., 2015a). For example, a combination of different imaging techniques such as atomic force microscopy (AFM), transmission electron microscopy (TEM), field emission scanning electron microscopy (FESEM), and confocal microscopy has been used to examine alteration in cellulose microfibril arrangement in the primary cell walls of the *Arabidopsis* xxt1 xxt2 double mutant that lacks detectable xyloglucan (Xiao et al., 2016). The study revealed that cellulose microfibrils are highly aligned in xyloglucan mutants as compared to those in wild type, suggesting that xyloglucan functions as a spacer between cellulose microfibrils in the primary cell wall.

This review summarizes techniques that are used for the characterization of structure and interactions of cellulose in plant cell walls, particularly cellulose crystallinity, microfibril size, and spatial organization along with cellulose–cellulose and cellulose–matrix interactions. We discuss both established and emerging techniques used for the molecular and microstructural characterization of cellulose structure, and highlight the strengths



and limitations of each technique. In addition, the review introduces several characterization techniques that are presently not widely used for studying plant cell walls, but given their capabilities, might prove to be powerful tools to reveal new information regarding structure and organization.

CRYSTALLINE STRUCTURE OF NATIVE CELLULOSE AND ITS ALLOMORPHS

Six polymorphic forms of cellulose (Cellulose I, II, III_I, III_{II}, IV_I, and IV_{II}) that are interconvertible have been identified (O'Sullivan, 1997). Natural cellulose is found in the form of cellulose I, which has two allomorphs – cellulose I α and cellulose I β (VanderHart and Atalla, 1984; Sugiyama et al., 1991a). Cellulose I α is the dominant form in primitive organisms like bacteria and algae while Cellulose I β is dominant in higher plants. The existence of these two forms was established by spectroscopic techniques while their lattice structures were revealed by diffraction techniques. Both techniques are widely used to identify the two forms of cellulose in plant cell walls and they are also used to quantify the relative abundances of the cellulose forms. This section highlights studies that revealed the cellulose unit cell parameters by diffraction techniques, and also discusses methods for identifying the two different forms (cellulose I α and I β) most commonly found in nature.

Revealing the Unit Cell Parameters of Cellulose

The unit cell parameters of the two allomorphs of native cellulose were established through X-ray, electron, and neutron

diffraction techniques. These techniques work on the principle of Bragg's law to determine the d -spacing of atomic planes using electromagnetic waves. Thus, although diffraction data is often represented as intensity versus scattering angle θ , it is useful to represent it as a function of scattering vector q instead to normalize for the radiation wavelength λ ($q = 4 \pi \sin(\theta/2)/\lambda$). Diffraction techniques are used for two main purposes: (i) determination of the three-dimensional structure of molecules and thus their crystallographic form, and (ii) assessment of the degree of crystallinity. Due to the weak diffraction from primary cell walls, the majority of studies on the unit cell parameters have focused on cellulose from algae, bacteria, and secondary cell walls. We briefly discuss these findings in this section, but also emphasize available data on primary cell walls.

The first X-ray diffraction (XRD) patterns of cellulose fibers were collected from wood, hemp, and bamboo in 1913 (Nishikawa and Ono, 1913). The quantification of cellulose crystal parameters began with data derived from XRD of plant fibers including Ramie, hemp, flax, spruce, and cotton (Sponsler, 1928). The lattice parameters of cellulose from different sources like algae, bacteria, and plants are well summarized (O'Sullivan, 1997).

Neutron diffraction (Beg et al., 1974; Ahmed et al., 1976) and electron diffraction (Honjo and Watanabe, 1958) studies have provided complementary structural information about cellulose I, enabling improvement of structural models developed from XRD data. Specifically, synchrotron X-ray techniques and neutron diffraction have enabled near atomic resolution. High-resolution synchrotron 2D data from oriented fibers of *Halocynthia*, which is nearly pure cellulose I β , is shown in Figure 2 (Nishiyama et al., 2002). The data

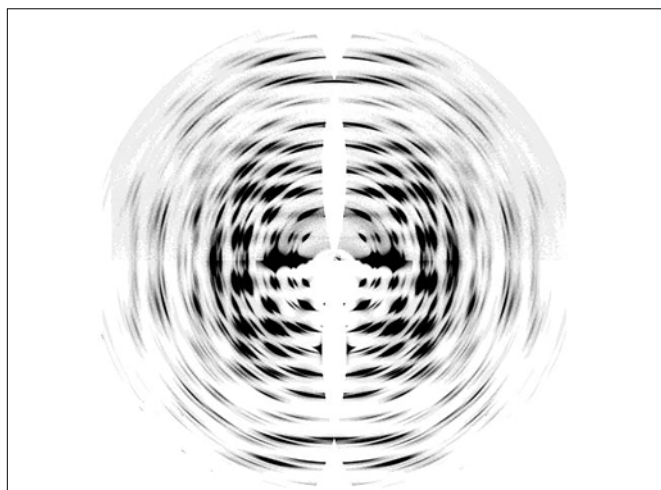


FIGURE 2 | Synchrotron XRD 2D pattern of Halocynthia cellulose I β . Reprinted with permission from Nishiyama et al. (2002). Crystal Structure and Hydrogen-Bonding System in Cellulose I β from Synchrotron X-ray and Neutron Fiber Diffraction. *Journal of the American Chemical Society* 124, 9074–9082. Copyright © 2002, American Chemical Society.

have a resolution better than 1 Å with more than 300 unique reflections. The high resolution of this data was important to determine atomic coordinates in the unit cell of cellulose I β .

Synchrotron X-ray experiments can provide accurate locations for carbon and oxygen atoms, but cannot do so for hydrogen atoms due to their small X-ray scattering cross-sections. Neutron diffraction of intra-crystalline deuterated cellulose samples has revealed important information about the intermolecular hydrogen bond network in cellulose I α and I β (Nishiyama et al., 2002, 2003). These experiments reveal that no inter-sheet hydrogen bonds exist in crystals of cellulose I α and I β , and the sheets are held together by hydrophobic interactions and weak C-H \cdots O bonds. The hydrogen bonds O3-H \cdots O5 could be visualized through Fourier difference maps calculated from neutron diffraction data. These maps give information about missing atoms in the crystal structure by subtracting the calculated structure factors from observed ones. These studies also showed that within each cellulose sheet the intramolecular hydrogen bond at O3 is well organized while the intermolecular hydrogen bond for O2 and O6 is disordered over two possible networks. Furthermore, the relative occurrence of these networks differs in the two cellulose allomorphs. Also, the bond length and bond angle of the intrachain O3-H \cdots O5 hydrogen bonds alternate between two different geometries in cellulose I α and I β . While the alternating geometry of the bond is along the same chain in I α , it is between two distinct chains in I β .

Electron diffraction has made significant contributions in differentiating between the structures of the two crystalline phases of native cellulose, and established that cellulose I α and I β have different lattice systems (Sugiyama et al., 1991a,b). Cellulose I α has a triclinic lattice with one chain per unit cell and cellulose

I β has a monoclinic lattice with two chains per unit cell, as shown in **Figure 3**. This technique has the advantage of producing intense diffraction patterns from a very small amount of sample, but the patterns can only be observed for a very short time for an organic substance like cellulose due to radiation damage caused by the electron beam.

High resolution synchrotron X-ray experiments have also been used to determine precise lattice parameters and the compositional ratio of cellulose I α and I β in native cellulose from different sources including algae, bacteria, and plants (Wada et al., 1997). XRD peaks were deconvoluted using six types of profile functions such as Gaussian, Lorentzian, intermediate Lorentzian, modified Lorentzian, pseudo-Voigt, and Pearson VII. The pseudo-Voigt profile gave the best fit and was used to determine lattice spacings as shown in **Table 1**. The relative content of cellulose I α was also determined based on the assumption that the first two equatorial reflections in the XRD pattern of *Valonia* cellulose are composites of cellulose I α (100) and cellulose I β (1 $\bar{1}$ 0), and of cellulose I α (110) and cellulose I β (010) reflections. The two reflections were thus deconvoluted into four independent reflections using pseudo-voigt functions. The cellulose I α content y_α was then estimated as:

$$y_\alpha = \frac{J_{I\alpha 100} + J_{I\alpha 010}}{J_{I\alpha 100} + J_{I\beta 1\bar{1}0} + J_{I\alpha 010} + J_{I\beta 110}} \quad (1)$$

where $J_{I\alpha XXX}$ denotes the integrated intensities J from I α and I β reflections. The cellulose I α fraction was found to be 0.65 for *Valonia* cellulose, which was nearly equal to the value of 0.64 reported for *Valonia* cellulose from ^{13}C NMR (Yamamoto and Horn, 1994).

X-ray diffraction is perhaps more widely used to study cell walls than other techniques because of multiple reasons, including less sensitivity of the sample to radiation damage, easier sample preparation, and easier data acquisition when compared to electron diffraction, and the ability to examine samples without the need of deuteration when compared to neutron diffraction. Nevertheless, because large single crystals of cellulose are not readily available, XRD studies are typically performed using protocols for powder diffraction, and the final results can depend on the model assumptions. Also, one of the limitations of diffraction techniques is that their results are averaged over space and time. These techniques cannot provide a dynamic visualization of the cellulose structure that is required to explain some of its properties. The complementary use of various spectroscopy techniques, such as NMR, IR, Raman and, more recently, neutron spectroscopy, have been beneficial to elucidating cellulose structure. A recent report on inelastic neutron scattering of cellulose explored the dynamics of hydrogen bond networks (Araujo et al., 2018). The effects of increasing water content in kraft cellulose was observed in the inelastic neutron scattering bands that are assigned to the hydroxymethyl group. Formation of ice microcrystals due to shock-freezing led to partial disruption of the hydrogen-bond network, which could be concluded from shifts of the OH vibrational mode observed in the spectra.

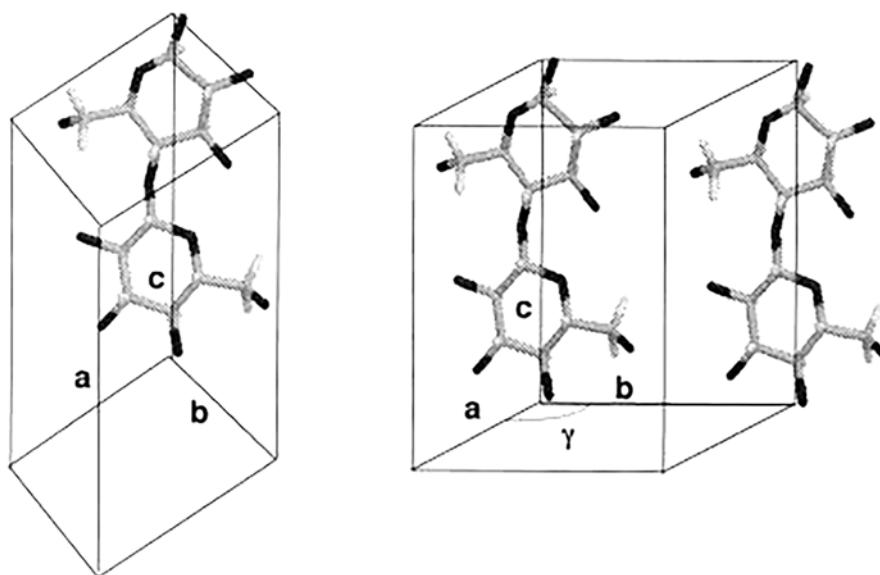


FIGURE 3 | Chain packing in unit cell of cellulose I α (left) and cellulose I β (right). Reproduced from Koyama et al. (1997). Parallel-Up Structure Evidences the Molecular Directionality during Biosynthesis of Bacterial Cellulose PNAS 94 (17), 9091–9095. Copyright © 1997, The National Academy of Sciences of the United States.

TABLE 1 | d -spacings of native cellulose calculated from synchrotron-based X-ray diffraction studies (Wada et al., 1997).

	d_1 , nm [composite of triclinic (100) and monoclinic (1 $\bar{1}$ 0) reflections]	d_2 , nm [composite of triclinic (010) and monoclinic (110) reflections]	d_3 , nm [composite of triclinic (110) and monoclinic (200) reflections]
Valonia	0.610	0.531	0.392
Cladophora	0.611	0.531	0.392
Chaetomorpha	0.608	0.530	0.391
Bacterial cellulose	0.614	0.530	0.394
Halocynthia	0.601	0.535	0.390
Cotton	0.601	0.536	0.393
Ramie	0.597	0.534	0.394
Kouzo	0.596	0.534	0.393

Identifying Allomorphs of Native Cellulose

The early crystallographic data of native cellulose from different sources were inconsistent with each other with respect to chain packing (French et al., 1987), and the assumption of twofold screw symmetry ($P2_1$ space group) was inconsistent with reflections observed in electron diffraction (Atalla, 1987). Additionally, the findings from applying new spectroscopic techniques to cellulose could not be rationalized on the basis of the then existing crystallographic models. The inconsistencies were resolved through solid state (SS) ^{13}C NMR spectral studies that led to the conclusion that native cellulose (cellulose I) is composed of two crystalline forms: cellulose I α and I β (Atalla and Vanderhart, 1984). The two allomorphs are identified

in plant cell walls, through spectroscopic and diffraction techniques as discussed in the following section. **Figure 4** shows the XRD pattern and spectra obtained from NMR, SFG, IR, and Raman spectroscopy for different forms of cellulose. These techniques present spectra with distinct features for each of the allomorphs and can be used to estimate the relative contents of the forms of cellulose in a sample.

NMR spectroscopy provides qualitative and quantitative information about atoms in a sample and their chemical environments. The technique can distinguish between chemically equivalent carbons located at magnetically non-equivalent sites. The application of Cross-Polarization Magic Angle Spinning (CP/MAS) ^{13}C NMR to study cellulose revealed that cellulose I α and I β can be differentiated in the NMR spectra based on the multiplicity of the C4 resonance peak near 88–90 ppm. Cellulose I α has a second peak in the down-field region while cellulose I β has it in the up-field region. The relative abundance of the allomorphs is calculated by deconvolution of the resonance peaks in the C4 region (Yamamoto and Horii, 1993). **Figure 4B** shows the NMR spectra of cellulose I β in comparison to other forms of cellulose. Cellulose I, II, and III can be distinguished on the basis of the chemical shifts of the C6 resonance peak; they have signals at 65.5–66.2, 63.5–64.1, and 62.1–62.8 ppm, respectively (Isogai et al., 1989).

IR and Raman spectroscopy are vibrational spectroscopic techniques that can provide complementary information on chemical functionality, molecular conformation, and hydrogen bonding. IR spectroscopy requires a dipole change while Raman requires a polarizability change as a molecule rotates or vibrates. One key advantage of Raman over IR spectroscopy for the study of hydrated cell walls is that water appears as broad absorption bands in IR spectra, while water bands have weak

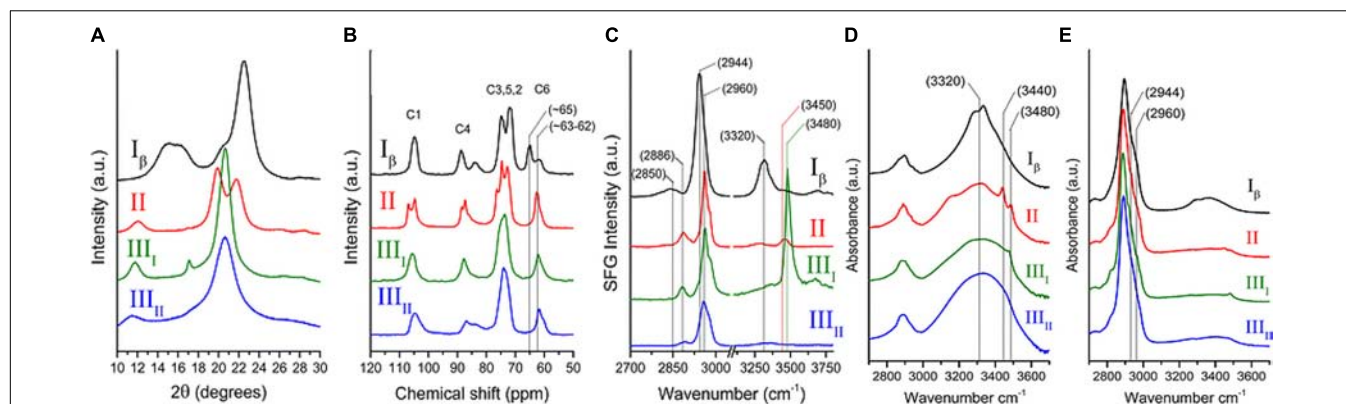


FIGURE 4 | Techniques used to characterize cellulose polymorphism: **(A)** XRD, **(B)** NMR, **(C)** SFG, **(D)** IR, and **(E)** Raman. Reprinted by permission from RightsLink Permissions: Springer Nature. Cellulose. Cellulose polymorphism study with sum-frequency-generation (SFG) vibration spectroscopy: identification of exocyclic CH₂OH conformation and chain orientation, Lee, C. M., Mittal, A., Barnette, A. L., Kafle, K., Park, Y. B., Shin, H., Johnson, D. K., Park, S., and Kim, S. H., Copyright © 2013.

intensities in Raman spectra. Moreover, changes in the refractive index of the material can cause variations in IR background but not in Raman, because excitation frequencies are far from absorption bands (Agarwal, 2014). When comparing the IR and Raman spectra of cellulose I α and I β , differences are observed in the OH-stretching region (3200–3600 cm⁻¹). In IR spectra, cellulose I α has peaks at 3240 and 750 cm⁻¹ while cellulose I β has peaks at 3270 and 710 cm⁻¹ (Sugiyama et al., 1991a). These findings suggest that the two phases have similar chain conformations, but differ in hydrogen bonding patterns and dihedral angles at the glycosidic linkages. Line shape analyses of these characteristic peaks can be carried out to determine the mass fractions of cellulose I α and I β in various cellulose samples (Yamamoto et al., 1996). **Figures 4D,E** compare the IR and Raman spectra of cellulose I β with the spectra obtained for cellulose II, III_I, and III_{II}. The main differences in the spectra are seen for the region above 3000 cm⁻¹. Cellulose I β has a distinct peak at about 3320 cm⁻¹, cellulose II has two peaks at about 3450 and 3480 cm⁻¹, cellulose III_I has one peak at about 3480 cm⁻¹, while cellulose III_{II} has no distinct sharp peak in this region.

Sum frequency generation (SFG) vibrational spectroscopy is a non-linear optical spectroscopy tool that is sensitive to non-centrosymmetric crystalline materials. As discussed in the Crystallinity of Cellulose, Spectroscopic Techniques Section, SFG is sensitive to structural ordering over an optical coherence length that enables it to characterize the structural hierarchy of cellulose microfibrils in the cell wall (Kim et al., 2013). NMR, IR, and Raman spectroscopy are widely used to study the conformation of purified cellulose, but their application is limited when it comes to native cellulose or lignocellulosic biomass, where spectral interference from other cell wall components cannot be avoided. The non-centrosymmetric requirement of SFG negates the interferences from SFG-inactive groups and thus enables the identification of exocyclic CH₂OH conformation and chain orientation of forms of cellulose as shown in **Figure 4C** (Lee et al., 2013). Similar to IR and Raman spectroscopy, SFG also

exhibits characteristic peaks for cellulose I α at 3240 cm⁻¹ and for cellulose I β at 3270 cm⁻¹ (Lee et al., 2015b).

CRYSTALLINITY OF CELLULOSE

Crystallinity is the ratio of crystalline to crystalline plus amorphous content by volume, and as such is a measure of structural order. Crystallinity affects mechanical properties such as strength and stiffness of cellulose and cellulose-derived materials. Higher cellulose crystallinity results in increased Young's modulus, tensile strength, density, and hardness (Lionetto et al., 2012). It is also an important parameter in many micromechanical models for wood (Bergander and Salmén, 2002; Hofstetter et al., 2005). Furthermore, the relative level of crystalline versus amorphous material within cellulose can influence the accessibility and reactivity of a given cellulose substrate to enzymes for biomass conversion. Given the importance of this metric, the crystallinity of cellulose has been estimated by many techniques, including XRD, IR and Raman spectroscopy, SS-NMR, SFG spectroscopy, Differential Scanning Calorimetry (DSC), and a variety of physicochemical assays. The measured crystallinity of cellulose can vary significantly depending on the technique and analysis approach used, with variations of up to 30–40% in reported values for cellulose-based materials (Thygesen et al., 2005; Park et al., 2010; Kljun et al., 2011; Agarwal et al., 2013; Karimi and Taherzadeh, 2016). **Table 2** summarizes the crystallinity of cellulose derived from different sources as determined by XRD and NMR (Park et al., 2010). The lack of consensus reflects the challenges in measuring the degree of order in plant cell walls and the limitations of the aforementioned techniques, which we discuss below.

Physicochemical Methods

The Updegraff method is a commonly used chemical method for determining the amount of crystalline cellulose in a sample (Updegraff, 1969). This method

TABLE 2 | Crystallinities of cellulose from different sources determined by XRD and NMR analysis methods.

Cellulose sample	XRD*			NMR**
	Peak height	Peak deconvolution	Amorphous subtraction	C4 peak separation
Bacterial microcrystalline cellulose	95.2	73.1	82.4	73.8
Avicel PH-101	91.7	60.6	77.7	56.7
SigmaCell 50	91.2	61.3	79.4	56.1
SigmaCell 20	84.8	64.2	67.0	52.6
JT Baker cellulose	85.5	61.5	69.1	49.1
Fluka cellulose	82.9	52.9	61.6	48.6
SolkaFloc cellulose	78.3	56.8	57.2	43.9
Sigma α -cellulose	78.0	55.9	54.4	41.5

*XRD: X-ray Diffraction; **NMR: Nuclear Magnetic Resonance Spectroscopy. All values are means (Park et al., 2010).

involves extraction of lignin, hemicellulose, and xylosans with an acetic acid/nitric acid reagent, leaving behind crystalline cellulose. Cellulose is then dissolved in 67% H₂SO₄, and the amount of crystalline cellulose can be determined after treatment with an anthrone reagent to enable colorimetric analysis (Scott and Melvin, 1953; Kumar and Turner, 2015).

In principle, cellulose crystallinity should be related to accessibility. The moisture sorption of cellulose takes place primarily by hydrogen bonding of water to accessible hydroxyls in less ordered regions at the surfaces of elementary fibrils and their random fibrillar aggregations at relative humidities lower than 50–60%. Thus, moisture regain of cellulose is a more direct measure of cellulose accessibility to reactants, rather than crystallinity. It is common practice to relate accessibility to crystallinity through the following equation (Howson, 1949):

$$A = \sigma X + (100 - X) \quad (2)$$

where A is the percentage of accessible cellulose in the sample, σ is the fraction of accessible cellulose on the surface of crystalline regions, and X is the percentage of crystalline cellulose in the sample.

The determination of accessibility of glucan chains based on deuterium exchange is based on the assumption that accessible OH groups in amorphous regions of cellulose readily exchange their hydrogen atoms for deuterium while the OH groups in crystalline regions exchange more slowly. Accordingly, the reaction curve for exchange reactions has two separate regions: an initial rapid rate region followed by a slow rate regime (Frilette et al., 1948), and the crystallinity has been related to accessibility similarly as shown in equation 2.

Because iodine is reported to be adsorbed in the amorphous regions of cellulose, measurements of iodine adsorption have

also been used to determine crystallinity (Hessler and Power, 1954). The amount of iodine adsorption per gram of cellulose has been linked to the fraction of amorphous cellulose within a sample. The crystallinity was then estimated by subtracting the amorphous fraction from 100.

A recent report has attempted to calculate the absolute degree of crystallinity of cellulose based on sorption of water vapor and enthalpy of wetting (Ioelovich, 2016). The crystallinity x of cellulose is calculated from sorption of water using the following equation that is derived from the sigmoidal isotherm of sorption of water vapor by semi-crystalline cellulose:

$$x = 1 - 2A(1 - 2.61 \ln \phi) \quad (3)$$

where A is the relative amount of water in cellulose by mass and ϕ is the relative vapor pressure at a constant temperature of 25°C. Under the assumption that water molecules interact with amorphous domains of cellulose and this interaction is accompanied by release of heat, the enthalpy of wetting is directly proportional to the amount of amorphous cellulose content. Then the crystallinity can also be determined by:

$$x = 1 - \frac{\Delta H}{\Delta H_{\text{am}}} \quad (4)$$

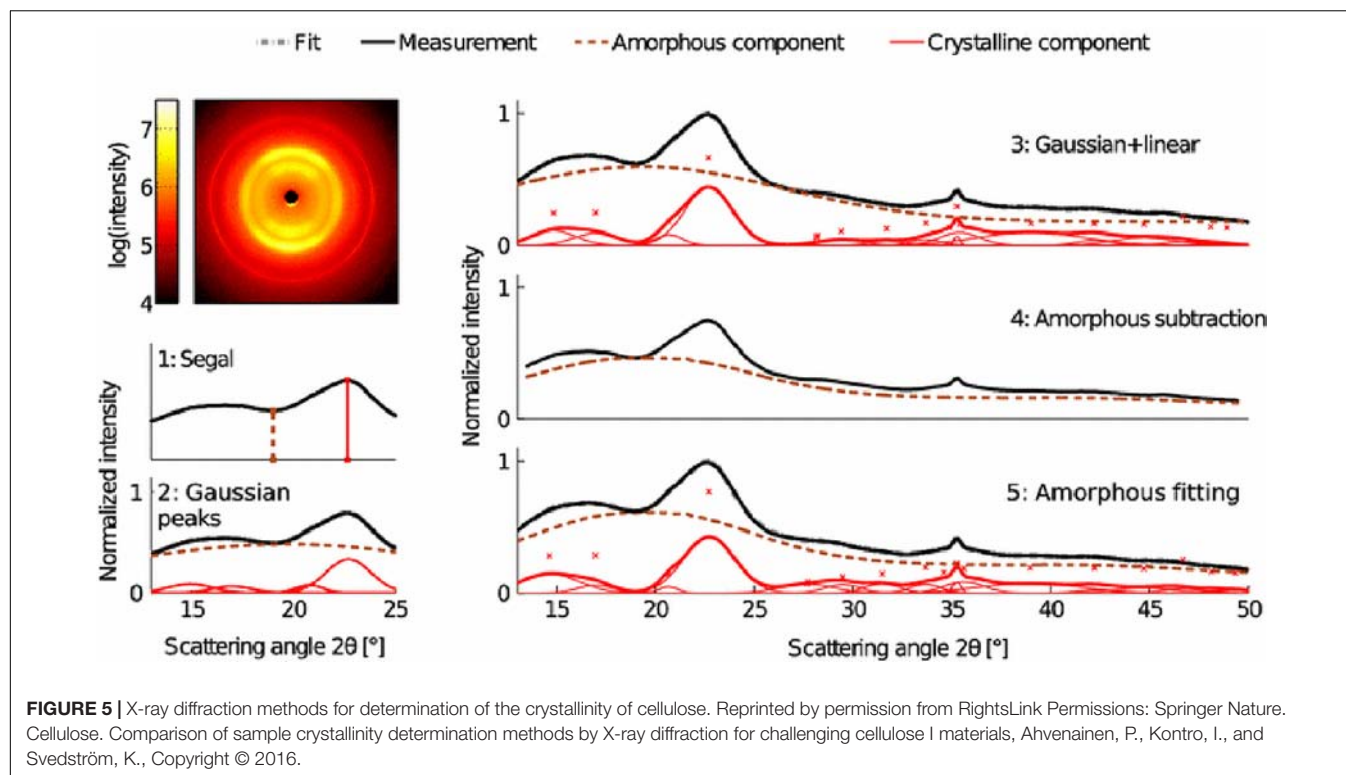
where ΔH_{am} is the enthalpy of wetting of purely amorphous cellulose. A value of $\Delta H_{\text{am}} = -167.5$ J/g has been reported and used to estimate crystallinity (Ioelovich, 2016). The crystallinity of microcrystalline cellulose samples was found to range from 0.72 to 0.75, as determined from the enthalpy of wetting and water sorption methods.

When compared to the crystallinity found from XRD measurements, physicochemical methods typically report a higher value of crystallinity. One possible origin of the discrepancy is the compositional and structural heterogeneity of cell walls, in particular of primary cell walls, that might complicate access to non-crystalline components. This would invalidate the assumption of a direct relationship between crystallinity and the physical and chemical properties investigated by these methods.

X-Ray Diffraction

X-ray diffraction is the most widely used technique for determining the crystallinity of cellulose due to its established reliability and accuracy, and minimal sample preparation requirements. XRD gives a measure of crystallinity as the mass fraction of crystalline cellulose within the entire sample (Ahvenainen et al., 2016). As shown in **Figure 5**, three methods are widely used for estimation of crystallinity from XRD, including: (i) the peak height or Segal method; (ii) peak deconvolution of crystalline and amorphous peaks; and (iii) the amorphous subtraction or Ruland-Vonk method. These approaches are discussed extensively in various reviews (Park et al., 2010; Kim et al., 2013; Ju et al., 2015; Karimi and Taherzadeh, 2016) and are described briefly below.

The peak height method, also called the Segal method (Segal et al., 1959), is the most widely used analysis approach



to characterize the crystallinity of cellulosic samples. The crystallinity x is calculated by:

$$x = \frac{I_{200} - I_{AM}}{I_{200}} \quad (5)$$

where I_{200} is the height of the (200) peak and I_{AM} is the height of the minimum between the (200) and (110) peaks. This method is not very accurate as the exact amount of the crystalline fraction is proportional to the peak area rather than to the peak height. Also, the underlying assumption of equation 5 is that scattering intensities from amorphous and crystalline content are equivalent per unit volume, which actually depends on the details of the structure factor of each of these phases. As a consequence, the crystallinity obtained using this method is dependent on crystallite size and cellulose allomorph (Ju et al., 2015).

The second method is based on peak deconvolution of crystalline and amorphous peaks. In XRD data, crystalline cellulose is represented by several intense peaks at (110), (102), (200), and (004) for cellulose I β and a single broad peak for the amorphous phase. Gaussian, Lorentzian, and Voigt functions are commonly used for peak fitting and the ratio of the area of the crystalline peaks to the total area is defined as the crystallinity. The accuracy of this method depends on selecting the correct peaks that correspond to the actual diffraction contributed by each fraction.

In the third method, also called the amorphous subtraction or Ruland–Vonk method (Ruland, 1961), the crystallinity is defined as the ratio of an area above an amorphous profile to the total area. The amorphous profile is obtained either from a polynomial function or a pattern measured from experimentally

prepared material believed to be entirely amorphous, such as ball-milled cellulose, regenerated cellulose, xylan, or lignin powder. In this method, a scaling factor is applied to the amorphous spectrum so that after subtraction from the original spectrum, no negative signal occurs in the residual spectrum. Often, the scaled amorphous background touches the diffractogram somewhere in the low q (low 2θ) region where the intensity is most poorly determined due to the fine adjustment of slits and the effects of axial divergence, so the method is sensitive to instrumental inaccuracies. It can also be difficult to compare samples of different origin. In addition, it can be challenging to compare results from different studies due to the variability in the amorphous standard used.

The crystallinity obtained from XRD can depend on crystallite size and preferred orientation of crystallites. The use of area-based fitting methods can better avoid the effects of crystallite size than peak height-based methods. The effects of preferred orientations can be mitigated by use of 2D Rietveld refinement, which includes the contribution of all diffraction peaks and two-dimensional diffraction data. Both 1D and 2D Rietveld refinement of XRD data are reported to accurately determine the degree of crystallinity (Thygesen et al., 2005; De Figueiredo and Ferreira, 2014; Driemeier, 2014). Because 2D Rietveld analysis is done on 2D diffraction data, it takes into account the preferred orientation and thus is considered more accurate for textured samples (Ahvenainen et al., 2016).

Additional approaches to estimate the crystallinity of cellulose from XRD data have also been developed, including the Hermans–Weidinger method (Gusev, 1978) and the Debye method (Thygesen et al., 2005), although these approaches

are less widely used in comparison to the abovementioned analyses. The Hermans–Weidinger method was developed for the determination of polymer crystallinity based on the proportionality of X-ray scattering intensities of crystalline and amorphous parts of a polymer. The proportionality is expressed as:

$$\frac{x_1}{x_2} = \frac{I_{c1}}{I_{c2}} \quad (6)$$

where x_1 is the degree of crystallinity and I_{c1} is the scattering intensity from the crystalline region. Crystallinity of a sample (labeled 1 in equation 6) can be determined only when a sample of known crystallinity (labeled 2) is available. The Debye method is similar to the Rietveld refinement method with the difference being that it requires simulation and fitting of the diffractogram to the experimental data to determine the quality of the fit (Thygesen et al., 2005). This approach has an advantage over the Rietveld method as the crystallite dimensions are included explicitly in the simulations and not fitted by analytical peak profile functions. This enables the Debye method to give the most reliable estimate of the crystalline part of the diffraction pattern, but it is less commonly used due to the heavy computing efforts required.

A robust estimate of the crystallinity from XRD measurements requires consideration of various approaches for data analysis. Even then, the limitations highlighted above preclude confidence in absolute values, although relative values for the crystallinity can reveal trends in samples that differ minimally (e.g., within the same species). Often, the term “crystallinity index” is used for crystallinities obtained from XRD to emphasize the challenges with comparing these values to those extracted from other techniques.

Spectroscopic Techniques

The intra- and inter-molecular hydrogen bonds found in crystalline cellulose can be analyzed using IR spectroscopy. The absorption band between 1420 and 1430 cm^{-1} (A_{1430}) is assigned to a symmetric CH_2 bending vibration, known as the “crystallinity band,” and the band appearing between 893 and 898 cm^{-1} (A_{898}) is assigned to C–O–C stretching at β -(1→4)-glycosidic linkages, known as the “amorphous band” (Nelson and O’Connor, 1964). Two terms related to crystallinity of cellulose have been defined, namely Lateral Order Index (LOI) and Total Crystallinity Index (TCI). LOI, also called the empirical crystallinity index, is the ratio of the intensities of A_{1430} to A_{898} and is sensitive to the amount of crystalline versus amorphous regions in cellulose. A lower LOI indicates a more amorphous structure (O’Connor et al., 1958). TCI is the ratio of the absorption band at 1372 to 2900 (Nelson and O’Connor, 1964; Poletto et al., 2014). The band at 1372 cm^{-1} is assigned to C–H bending and is reported to be affected by the amorphous content of a cellulose sample while the band at 2900 cm^{-1} is assigned to C–H and CH_2 stretching and is reported to be unaffected by changes in crystallinity. Taking the ratio of intensities of these bands as TCI enables the crystallinity index to be insensitive to sources of variation other than changes in crystallinity. IR spectroscopy is routinely used to characterize

woody biomass meant for biofuel conversion (Amiri and Karimi, 2015; Noori and Karimi, 2016).

Different peak ratios in Raman spectra have been reported in literature as a measure of crystallinity. The relative intensity ratios of the Raman bands 1481 and 1462 cm^{-1} in cellulose I (Schenzel et al., 2005) and that of 380 and 1096 cm^{-1} bands (Agarwal et al., 2010) are both reported as measures of the crystallinity. Unfortunately, both IR and Raman spectroscopy face challenges when characterizing the crystallinity present in primary cell walls due to the interference of signals from other wall components.

In the ^{13}C NMR spectra of cellulose, the peak at 89 ppm is assigned to C4 in crystalline cellulose and the peak at 84 ppm to amorphous cellulose (Atalla and VanderHart, 1999). The crystallinity from NMR spectra is defined as the integral area of the C4 peak from 87 to 93 ppm divided by the total integral area assigned to the C4 peaks (from 80 to 93 ppm). This method has been used to determine the degree of crystallinity in wood (Newman and Hemmingson, 1990; Newman et al., 1993) and to study the effect of crystallinity on enzymatic degradation of cellulose (Mansfield and Meder, 2003). It has also been applied to estimate crystallinity in primary cell walls of cellulose synthase mutants of *Arabidopsis thaliana* (Harris et al., 2012).

As introduced earlier, the non-centrosymmetric requirement of SFG allows selective detection of cellulose in plant cell walls and characterization of its structural properties. SFG has also been used to determine the amount of crystalline cellulose in secondary cell wall samples, which was estimated by applying a calibration curve from Avicel to the intensity of the CH_2 SFG peak of cellulose at 2945 cm^{-1} (Barnette et al., 2012). The limitations of this technique lie in the assumption of 100% crystalline Avicel, the assumption of the same signals from Avicel cellulose and from the biological systems under study, and the neglect of the effect of crystal size. Perhaps as a consequence, the technique has not yet been reported for crystallinity studies of primary cell walls.

CELLULOSE MICROFIBRIL SIZE AND ORGANIZATION

Direct visualization of the cell wall through light microscopy shows the existence of cellulose in a bundled fibrillar structure. High resolution electron microscopy reveals microfibrils that are aggregated, such that individual microfibrils (sometimes termed elementary fibrils) have cross-sections of 2–4 nm and lengths of 100 nm or more (Kraissig, 1992). Complete understanding of this fibrillar network requires the characterization of structural parameters, including fibril length, lateral size and shape, as well as the spatial arrangement of microfibrils. These parameters have a strong influence on the mechanical and physicochemical properties of cellulose and its derivatives. The following section discusses the characterization of the abovementioned parameters through different techniques such as microscopy, diffraction/scattering, spectroscopy, and chemical methods. We cover examples from studies of bacterial cellulose, primary cell walls, and secondary cell walls.

Size and Shape of Cellulose Microfibrils

Perhaps the simplest approach to estimate the dimensions of microfibrils relies on physicochemical methods. Under the assumption that the microfibril length is equal to the chain length, the length is estimated from the degree of polymerization (DP) of residual cellulose that remains after an initial drastic drop upon dissolution in dilute acid. This degree of polymerization is called the leveling off DP, and the crystallite length is estimated as the product of the leveling off DP and length of one monomer unit. The DP of cellulose has also been determined through light scattering, osmotic pressure, and gel permeation chromatography (Levi and Sellen, 1967; Holt et al., 1973). The crystallite width is calculated by observing the reactivity of cellulose toward dilute mineral acid and deuterium oxide. Under the hypothesis that both acid hydrolysis and deuteration take place in the amorphous regions, but only deuteration takes place on the surface, the number of molecules per side of a rectangular cross-section is calculated and multiplied by the average of the (101) and ($\overline{101}$) spacings for cellulose I. For example, values for the crystallite width are 31 Å for cotton and 33 Å for Ramie, with crystallite lengths of about 100 nm for both (Scallan, 1971). As discussed below, these crystallite widths are consistent with measurements from electron microscopy and other techniques.

Various approaches have attempted to directly image the size and shape of microfibrils (Table 3). The use of electron microscopy along with techniques like metal shadowing (Bayley et al., 1957; Beer and Setterfield, 1958), negative staining (Heyn, 1966; Revol, 1982; Manley, 2003), and diffraction contrast imaging (Bourret et al., 1972; Revol, 1982) have revealed valuable structural information about cellulose from several sources including valonia, jute, cotton, and ramie fibers. Based on the findings from X-ray diffraction/scattering and electron microscopy of cellulose materials following different chemical treatments, two descriptions of microfibrils developed. One hypothesis stated that each microfibril has a single crystalline core whose size is almost the same as a microfibril, while an alternative hypothesis stated that each microfibril was composed of elementary microfibrils of 35 Å width (Nieduszynski and Preston, 1970). The former hypothesis was supported with studies on bacterial cellulose, where apparent crystallite lateral dimensions are much larger than 35 Å, and not necessarily in its multiples. Cellulose crystallites from *Chaetomorpha melagonium* and *Acetobacter xylinum* were found to measure between 100 and 200 Å when studied through X-ray diffraction and electron microscopy (Colvin, 1963; Nieduszynski and Preston, 1970). Further work based on high resolution imaging techniques was crucial to resolve these conflicting descriptions of cellulose organization, as described below.

Lattice imaging of native cellulose from ramie fibers and different algal and bacterial sources was made possible with high resolution electron microscopy in combination with negative staining, metal shadowing, and diffraction contrast imaging (Sugiyama et al., 1985; Kuga and Brown, 1987a,b). These studies established that each microfibril corresponds to a single crystalline entity. Negative staining of sections of cellulose from cotton, ramie, and jute fibers revealed lateral dimensions between

TABLE 3 | Microfibril diameter from different sources of cellulose obtained through the use of different analytical characterization techniques.

Source of cellulose	Microfibril diameter (nm)	Techniques*
<i>Arabidopsis thaliana</i>	5.8 ± 0.17	AFM (Davies and Harris, 2003)
Celery collenchyma	2.4–3.6	NMR, SAXS, WAXS (Kennedy et al., 2007a)
	2.9–3.0	SANS, WAXS (Thomas et al., 2013)
	2.6–3.0	SAXS (Kennedy et al., 2007b)
Celery parenchyma	6.0–25.0	AFM (Thimm et al., 2000)
Cotton	2.5–4.0	TEM (Heyn, 1966)
	4.9–6.1	TEM (Nieduszynski and Preston, 1970)
	5.5	SAXS (Heyn, 1955)
Flax fibers	1.0–5.0	SAXS (Astley and Donald, 2001)
	2.8	SAXS (Heyn, 1955)
Jute	2.8	TEM (Heyn, 1966)
	2.8	SAXS (Heyn, 1955)
Maize	3.2–5.3	AFM (Ding and Himmel, 2006)
	2.5–3.5	WAXS, NMR (Rondeau-Mouro et al., 2003)
Mung bean	2.5–3.2	WAXS, NMR (Newman et al., 2013)
Oak wood	2.9–3.1	WAXS, SAXS (Svedström et al., 2012)
Onion	8.0–10.0	NMR (Ha et al., 1998)
	4.4 ± 0.13	AFM (Davies and Harris, 2003)
Quince	2.0	NMR (Ha et al., 1998)
Ramie	3.6–4.8	TEM (Heyn, 1966)
	5.9	TEM (Nieduszynski and Preston, 1970)
	4.3	SAXS (Heyn, 1955)
Spruce wood	2.5	TEM, WAXS, SAXS (Jakob et al., 1995)
	2.9	WAXS (Andersson et al., 2000)
	3.1–3.2	SANS, WAXS (Fernandes et al., 2011)
	2.9–3.1	WAXS (Peura et al., 2007)
Sugi wood	2.4–2.6	SAXS (Suzuki and Kamiyama, 2004)
Tunicin	3.4–7.6	TEM (Nieduszynski and Preston, 1970)
<i>Valonia ventricosa</i>	18.0	TEM (Revol, 1982)
	10.0–20.0	WAXS (Caulfield, 1971)
	3.0	WAXS, NMR, IR (Horikawa et al., 2009)

*AFM, atomic force microscopy; NMR, nuclear magnetic resonance spectroscopy; SAXS, small angle X-ray scattering; SANS, small angle neutron scattering; WAXS, wide angle X-ray scattering (synonymous with XRD); TEM, transmission electron microscopy; IR, infrared spectroscopy.

25 and 40 Å (Heyn, 1966). As shown in Figure 6, transmission electron microscopy (TEM) with negative staining has also been used to demonstrate individual cellulose microfibrils that result from various alkaline treatments of vascular bundles of banana rachis (Zuluaga et al., 2009). Using electron diffraction and dark field electron microscopy, cellulose crystallites from

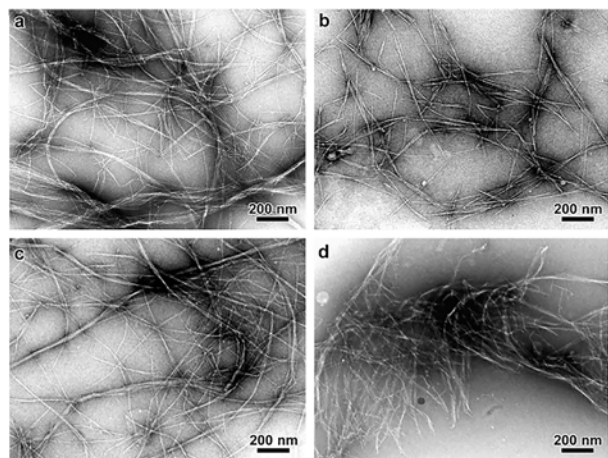


FIGURE 6 | Transmission Electron Microscopy micrographs comparing the morphology of cellulose microfibrils isolated by different chemical treatments. **(a)** Peroxide alkaline, **(b)** peroxide-alkaline-hydrochloric acid, **(c)** 5 wt% potassium hydroxide, and **(d)** 18 wt% potassium hydroxide. The combination of peroxide alkaline and hydrochloric acid or the application of a high concentration (18 wt%) potassium hydroxide solution leads to shorter microfibrils, suggesting these treatments can cause microfibril scission. Reprinted from Carbohydrate Polymers, 76, Zuluaga, R., Putaux, J. L., Cruz, J., Vélez, J., Mondragon, I., Gañán, P. Cellulose microfibrils from banana rachis: Effect of alkaline treatments on structural and morphological features, 51–59, Copyright © 2009, with permission from Elsevier.

algae (*Valonia ventricosa*) were found to be above 1000 Å in length and 140 to 180 Å in width (Bourret et al., 1972). Thus, although the “elementary” unit appears to be a microfibril of a few nanometers, dimensions of cellulose crystallites appear to vary depending on the source. In a similar way, no agreement has been reached on the cross-sectional shape of cellulose found from imaging. The cross-section of valonia microfibrils was found to be almost square-shaped with an average size of 180–200 Å (Revol, 1982; Sugiyama et al., 1985) while the cross-section of tunicate cellulose was found to be parallelogram shaped (Helbert et al., 1998a,b). Even though valuable information has been obtained about cellulose microfibrils from electron microscopy, the sample preparation that generally requires drying could introduce artifacts through modifications in the physical structure of native cellulose, such as collapse and aggregation of microfibrils. This has limited the study of microfibril shape and diameter in primary cell walls through TEM.

As an alternative to electron microscopy, scanning force microscopy (SFM), also termed atomic force microscopy (AFM), and optical microscopy are techniques that can visualize cellulose microfibrils with spatial resolution ranging from the micrometer to the sub-nanometer scale in biologically relevant environments. AFM techniques reveal the surface topology by measuring the interaction between a fine physical probe and the surface of the sample. Imaging contrast is based on variations of the sample topology, modulus, or interaction with the probe. AFM can record the surface topography and properties at the nanoscale by scanning a sample under a sharp stylus or tip, which is often made from silicon or silicon nitride. The stylus is attached to

a cantilever, which is deflected as the stylus interacts with the surface. Images are produced by measuring the deflection of the cantilever as the sample is scanned. Alternatively, atomic force microscopes can be operated in constant-force mode in which a feedback system is used to keep the deflection constant (Prater et al., 1990). AFM enables direct characterization of sample surfaces with high spatial resolution (0.1–100 nm) and minimal sample preparation; thus, AFM is ideal for characterizing the structure of cell walls, as many features can be detected within this resolution range (Yarbrough et al., 2009). Samples need not be fixed, stained, dried, or metal coated as in the case of electron microscopy. Even if a pectin layer is present, the tip can probe through this soft layer to reveal the microfibril structure underneath in primary cell walls (Zhang et al., 2014, 2016).

The earliest cellulose-containing biological samples studied using AFM were dried cells of archae-bacterium *Halobacterium halobium* (Butt et al., 1990); later studies focused on bacterial polysaccharides (Gunning et al., 1995) and cellulose from root hair cell wall of *Zea mays* and *Raphanus sativus* (van der Wel et al., 1996). AFM has also been used to visualize cellulose microfibrils in hydrated primary cell walls from apple, water chestnut, potato, and carrot (Kirby et al., 1996). These measurements supported the polylaminate description of cell wall structure. Furthermore, the effect of hydration on the diameter of cellulose microfibrils in celery parenchymal cell walls was studied using AFM. It was found that the measured diameters depend on the water content of the samples and also on the procedure of dehydration, with diameters ranging from 15.2 ± 0.4 nm before dehydration to 25.1 ± 0.8 nm after dehydration (Thimm et al., 2000). Nevertheless, as the tip scans across the surface, it can lead to broadening of lateral features due to the width of the tip itself, leading to differences in measured microfibril diameters from AFM in comparison to other techniques. Measuring the height of microfibrils (in the z-direction) resolves this problem, as was done to find the dimensions of cellulose microfibrils from partially hydrated cell walls of onions and *A. thaliana* (Davies and Harris, 2003). Microfibrils were 4–6 nm in diameter and contain a single cellulose crystallite, 2–3 nm wide, which is surrounded by non-cellulosic polysaccharides. It was also found that removal of pectin from the cell wall improved the accuracy of measurements. AFM studies of maize parenchyma cell wall indicated microfibril dimensions similar to that found in onion and *A. thaliana* as discussed above, although the authors proposed a 36-chain model for each microfibril (Ding and Himmel, 2006). AFM has also been used to compare cellulose microfibrils in different scales of onion (Kafle et al., 2014; Tittmann and Xi, 2014; Zhang et al., 2014). These studies showed that the microfibrils are more ordered in older scales than in younger scales. Altogether, previous work has demonstrated AFM as a powerful tool for imaging of the cell wall in physiological environments.

Scanning Electron Microscopy (SEM) is an alternative approach to image the surface of plant cell walls. Sample preparation for SEM is simpler than for TEM, because electron-transparent samples and heavy metal staining are not required. SEM allows imaging cell walls directly and has been used to observe the cell wall structure of both primary (Crow

and Murphy, 2000; Carpita et al., 2001) and secondary cell walls (Awano et al., 2002; Kim et al., 2012). Measurements of microfibril dimensions are consistent with estimates derived from AFM (Zheng et al., 2018). Nevertheless, SEM usually requires dehydration or critical-point drying, removal of the top pectin layer (if present), as well as deposition of a conductive coating to prevent charging, which may cause artifacts. As a consequence, the technique is often used to complement other microscopic and spectroscopic techniques. For example, SEM has been used along with IR spectroscopy to study the cell wall architecture of Maize coleoptile (Carpita et al., 2001), and with AFM to study different plant tissues like cucumber hypocotyls, *A. thaliana*, and onions (Marga et al., 2005; Xiao and Anderson, 2016; Zhang et al., 2016).

In addition to estimates from real-space images, estimates of microfibril dimensions have been obtained from reciprocal space techniques. These approaches have the advantage of averaging structural features over large areas. Line broadening in X-ray diffraction (XRD, or wide-angle X-ray scattering, WAXS) is directly related to the coherence length t , as given by the Scherrer formula:

$$t = \frac{k \lambda}{\beta \cos \theta} \quad (7)$$

where λ is the X-ray wavelength, θ is the Bragg angle, k is a shape factor that is often 0.89, and β is the angular half width of the line profile. The coherence length is equivalent to the crystal size if fluctuations or defects in the crystal lattice are not cumulative, such that deviations from ideal average lattice positions do not disrupt the long-range order of the lattice. Under this assumption, early applications of this approach measured the cellulose crystallite size for valonia, tunicin, cotton, ramie, *Acetobacter xylinum*, and *Chaetomorpha melagonium* (Nieduszynski and Preston, 1970; Caulfield, 1971). Line broadening of the equatorial reflections (200) and (110/1 $\bar{1}$ 0) give the lateral dimension while the meridional reflection (004) gives the longitudinal dimension. The reported crystal widths from XRD (100–200 Å) significantly exceed values reported for microfibril diameters from electron microscopy (35 Å) and other techniques (see Table 3). One possible explanation is that microfibrils aggregate and strong interactions maintain lattice coherence, thereby leading to apparent larger crystal dimensions from X-ray experiments.

Analyses of XRD data have also attempted to resolve diffraction peaks into Gaussian and Cauchy profiles (Hindeleh and Johnson, 1972, 1974). The obtained crystallite sizes did not support the existence of elementary microfibrils. The results, however, depend a lot on the details of the model adopted for peak fitting, such as the type of fitting function and background subtraction. Other factors like crystal morphology, distortions, and size distribution also affect the results.

In addition to X-ray diffraction, small-angle scattering techniques have also been employed to examine the dimensions of microfibrils. These techniques involve analysis of the intensity of radiation scattered from the sample as a function of the scattering vector q . Focusing on small scattering angles can reveal the size and shape of objects, such as the diameter of

rod-like microfibrils. Diameters of highly oriented fibrils were obtained from Small Angle X-ray Scattering (SAXS) of ramie, cotton, jute, flax, and cordura using Guinier plots for cylindrical particles (Heyn, 1955). The sizes obtained for jute, ramie, and cotton were in agreement with coherence lengths (crystal sizes) previously obtained from XRD and with diameters obtained from electron microscopy with negative staining (Heyn, 1966). Nevertheless, the weak spatial organization of primary cell walls make interpretation of SAXS profiles challenging; yet, SAXS has successfully been used to examine the size and arrangement of cellulose fibrils in secondary cell walls of spruce wood (*Picea abies*). An almost constant diameter of 2.5 nm with a standard deviation as small as 0.14 nm was found for measurements from 10 different trees (Jakob et al., 1994). This microfibril diameter was in good agreement with that obtained from TEM, which reported the diameter to be 2.4 nm but with a standard deviation as high as 1.3 nm. Other work has demonstrated good agreement between SAXS profiles and Fourier transforms of TEM micrographs (Jakob et al., 1995).

An advantage of SAXS is the ability to perform experiments under moist environments; for example, hydration-dependent structural changes of cellulose microfibrils in spruce wood have been examined (Jakob et al., 1996). The packing density and fibril center-to-center distance was estimated, and it was found that the structure of the cell wall was independent of hydration if the moisture content was above the saturation point of fibrils. Comparable measurements were not possible for moisture content below the saturation point, as the scattering from pores could not be neglected. Similarly, SAXS has been used to study the effect of hydration on cellulose from different sources including *Acetobacter xylinus*, flax, sugi wood, and celery collenchyma (Astley and Donald, 2001; Astley et al., 2001; Suzuki and Kamiyama, 2004; Kennedy et al., 2007b). Such studies are mostly on secondary cell walls as in the case of flax or wood. Celery collenchyma offers a convenient experimental platform for studying hydrated primary cell walls through scattering as they have unusually well oriented microfibrils. It has been reported that hydration increases the mean microfibril spacing from 3.8 nm in dry cell walls to 5.4 nm in hydrated cell walls of celery collenchyma (Kennedy et al., 2007b).

The low scattering contrast between cellulose and other cell wall polymers makes the analysis of X-ray scattering patterns difficult. Small Angle Neutron Scattering (SANS) provides an advantage over SAXS in this context. Because hydrogen scatters much more strongly than deuterium, neutron scattering contrast can be enhanced by replacing H₂O with D₂O, or by deuterating components of the cell wall. A SANS study of primary cell walls in celery collenchyma characterized the microfibril diameter and shape (Thomas et al., 2013). The diameter was found to be about 2.9–3.0 nm and this value corresponds to 24 chains in a microfibril with a rectangular cross-section. These results of microfibril diameter and cross-section were similar to the findings of a SANS study of secondary cell wall in spruce wood; nevertheless, the presence of extensive disorder in primary cell walls prevented a conclusive result (Fernandes et al., 2011).

A challenge with scattering approaches is that, in principle, multiple structures can lead to the same scattering profiles.

Thus, complementary data is crucial to develop structural models capable of explaining scattering data. This is especially true for primary cell walls, which exhibit poorly ordered packing, and as a consequence, scattering data from these tissues is more challenging to interpret. As such, the application of spectroscopic techniques, such as SS-NMR and IR, to primary cell walls is important to complement scattering and microscopy.

One early report that combined spectroscopy with imaging investigated onion and quince cell walls with fibril diameters established by electron microscopy of 8–10 nm and 2 nm, respectively (Ha et al., 1998). The authors proposed that six microfibrils aggregate in onion, such that each elementary fibril is about 2 nm; a strongly charged hemicellulose coating in quince is proposed to keep these microfibrils isolated. Two independent approaches were adopted for measuring the crystallite diameter, by calculating the proportion of surface to interior chains and through spin-diffusion experiments to measure the distance between surface and interior chains. Altogether, the two methods suggest that fibrils from onion and quince have similar crystallite diameters of approximately 2 nm.

The lateral dimensions of cellulose crystallites from 10 different sources were estimated using ^{13}C NMR signal strengths (Newman, 1999). Signals at 89 and 85 ppm were assigned to C4 in the interiors and on the surfaces of crystallites, respectively. Lateral dimensions were estimated from the relative signal areas under an assumption of a square microfibril cross-section. When compared with XRD results of the same samples, lateral dimensions obtained from NMR were found to be 10% higher, and this deviation was attributed to different molecular conformations of surface and interior chains that lead to broadening of XRD peaks. Using the same aforementioned peak assignment of surface and interior chains, NMR was also used to study the microfibril diameter of celery collenchyma and the results compared with that obtained from XRD and SAXS (Kennedy et al., 2007a). Assuming a constant microfibril diameter and circular model for its cross-section, the microfibril radius is calculated as:

$$\frac{A_I}{A} = \frac{(R - S)^2}{R^2} \quad (8)$$

where A_I/A is the relative area of signals from interior chains, R is the radius, and S is the thickness of the surface monolayer of chains calculated from cellulose I β lattice parameters as previously reported (Nishiyama et al., 2003). If no structural difference between surface and interior chains is assumed, the size of microfibrils obtained from NMR is in agreement with XRD results. Thus, NMR measurements can reconcile with the entire range of SAXS measurements depending on the different rotational orientation of surface chains that is assumed.

In addition to NMR, IR spectroscopy has been used to extract estimates of the microfibril size in higher plants, algae, and tunicates (Horikawa et al., 2009). This approach is based on an initial deuteration of OH groups in the entire crystalline region followed by re-hydrogenation at 25°C during which deuterated (OD) groups on the surface become re-hydrogenated (OH). Microfibril dimensions were then estimated from the absorbances (A) of OD and OH

groups. Defining R as an empirical parameter that is the ratio of the OD absorbance (A_{OD}) to the total absorbance by $R = A_{OD}/(A_{OH} + A_{OD})$ can then enable comparison with other measures of the microfibril diameter. Indeed, R was found to be highly correlated to the full width at half maximum of the (200) peak in XRD. Microfibrils were proposed to be flat based on the behavior of the re-hydrogenation process under heat treatment, which was consistent with observations by electron microscopy.

More recently, detailed studies on the cross-sectional shapes of cellulose crystallites and the number of chains in each microfibril have been attempted through spectroscopic techniques. These methods also provide valuable insights into aggregation and twinning of microfibrils, as well as conformational and packing disorder. SS-NMR and IR were used in combination with SANS and XRD to study the microfibril structure of spruce wood (Fernandes et al., 2011). The results of these studies favored a 24-chain model with a rectangular microfibril cross-section and the presence of twisting and disorder that increases toward the surface. Another study on celery collenchyma used NMR and IR of deuterated samples in combination with XRD, SANS, and WANS (neutron diffraction) (Thomas et al., 2013). This study suggests a 24-chain model with eight hydrogen bonded sheets of three chains and also the possibility of an 18-chain model if the presence of a hemicellulose chain is included. It also proposed the presence of high disorder in conformation, packing, and hydrogen bonding. Simulations of XRD profiles were compared with synchrotron XRD data and NMR results to predict the number of chains in microfibrils (Newman et al., 2013). The number of chains in a microfibril was estimated using the crystallinity x estimated from NMR spectra (Newman, 1999). The uncertainties involved in the estimation of k (shape factor) and x made it difficult to make a precise estimate, and a possibility of 17–22 chains was suggested. The study ruled out a 36-chain model on the basis of predicted peaks that did not match with the experimental diffractogram. Good fits were obtained for 24- and 18-chain models, with an even better fit for the 18-chain model with mixed cross-sectional shapes and the presence of occasional twinning.

Furthermore, studies of the cellulose synthase complex suggest a rosette that is a hexamer composed of trimers (Hill et al., 2014; Nixon et al., 2016; Vandavasi et al., 2016), which would be consistent with an 18-chain model. Using this as a starting point, a detailed study that combines X-ray diffraction and NMR data with predictions from computer simulations established a 5-layer cross-section with a 34443 chain arrangement as most probable (Kubicki et al., 2018). The ability to compare predicted and measured ^{13}C NMR shifts and diffraction spectra was able to rule out a 6×3 arrangement as highly unlikely, although a 6-layer 234432 cross-section is only slightly less likely than the 34443 configuration.

Cellulose Microfibril Angle

In contrast to the dispersed cellulose orientation of primary cell walls, cellulose microfibrils in woods are wound around the cell in a helical manner whose pitch is defined by the microfibril angle (MFA), which is described as the angle that the microfibrils

make with the long axis of the cell (Barnett and Bonham, 2007). Traditionally, the MFA has been used to describe the orientation of cellulose microfibrils in the S2 layer of secondary walls in woods because cellulose makes up the greatest proportion of the wall thickness and most affects the macroscopic physical properties (Senft and Bendetsen, 1985). The S2 MFA has a significant influence on tensile strength, stiffness, and shrinkage in wood (Cave, 1968). Both the longitudinal tensile strength and stiffness of wood have been shown to be markedly affected by MFAs; as the MFA increases, tensile strength and stiffness quickly decrease (Altaner and Jarvis, 2008). The MFA is also an important determinant of quality of wood products. It has a major effect on the stability of wood on drying and subsequent manufacturing processes (Zobel, 1961).

The techniques for measuring MFAs can be grouped into four categories: (1) Polarized light microscopy, (2) direct visualization through microscopy after physical or chemical treatment such as iodine staining, (3) XRD and SAXS, and (4) Near IR (NIR) spectroscopy. A detailed review of these techniques and their comparison is available (Donaldson, 2008), and a brief summary of results from various techniques is shown in Table 4.

Extracting MFAs from polarized light microscopy involves rotating cellulose fibers relative to the fiber long axis until the maximum extinction position (MEP) is reached, which occurs when the bright cell wall becomes dark (Preston, 1934; Page,

1969). The difference between the fiber axis and MEP gives an estimate of an average MFA. A disadvantage of this technique is that it requires samples consisting of a single cell wall, otherwise the orientation of microfibrils in opposing cell walls in front and back walls will inhibit accurate determination of the MEP (El-Hosseiny and Page, 1973).

Brightfield microscopy and confocal microscopy have been used to measure MFAs in iodine stained samples (Bailey and Vestal, 1937; Senft and Bendetsen, 1985; Donaldson and Frankland, 2004). This method involves precipitation of iodine crystals within the cell wall and hence, it is limited by the fact that not all wood samples react well with iodine; thus, iodine does not always uniformly disperse in all the cells. Because iodine sublimates fast, the measurements have to be taken rapidly. Higher accuracy measurements of MFAs were facilitated through high contrast images taken with confocal reflectance microscopy (Donaldson and Frankland, 2004) or electron microscopy (Wardrop and Preston, 1947; Frei et al., 1957; Dunning, 1968).

X-ray diffraction is perhaps the most commonly used method for determination of MFAs. Typically, MFA is obtained from XRD through the azimuthal distribution of the cellulose (200) equatorial reflection (Cave, 1968; Nelmes and Preston, 1968; Yamamoto et al., 1993). This method assumes that the cellulose crystals do not have a preferred orientation around the microfibril axis. SAXS can also provide MFA in a similar manner as XRD without this assumption (Jakob et al., 1994; Reiterer et al., 1998). SAXS has been used to estimate MFA in primary cell walls of single celled alga *Chara corallina* and multicellular hypocotyl of *A. thaliana* (Saxe et al., 2014). The work shows a bimodal MFA distribution such that the bulk of the microfibrils are oriented either transversely or longitudinally with broad scattering. The highly oriented microfibrils in secondary walls give an anisotropic SAXS pattern and the azimuthal intensity distribution of the resulting streaks is used to extract information on the distribution of MFA. This method has been adopted for wood cells in *Picea abies* (Jakob et al., 1994; Reiterer et al., 1998). These studies found that stiffer parts of trees have lower MFA when compared to the more flexible parts that have higher MFA, thereby supporting the correlation between cellulose MFA and mechanical properties of the cell wall.

Near IR spectroscopy has also been used to predict MFA by examining wood surfaces on the radial-longitudinal face (Jones et al., 2005; Schimleck et al., 2005). The method uses XRD data for calibration, and thus becomes inaccurate for higher angles because XRD data are less precise at high angles due to a reduced signal-to-noise ratio for the (200) reflection of the diffraction pattern (Schimleck et al., 2005).

Spatial Organization of Cellulose Microfibrils

Because cellulose microfibrils are the structural units of primary cell walls, the spatial arrangement of these microfibrils, including their bundling and packing, strongly impacts cell wall mechanics and growth. Traditionally, the mesoscale arrangement of microfibrils was studied largely by electron microscopy. The technique provided many valuable insights about the

TABLE 4 | Microfibril angle from different sources of cellulose obtained through the use of different characterization techniques.

Source of cellulose	Microfibril angle (°)	Techniques*
<i>Picea abies</i>	≤5 (earlywood), 20 (latewood)	SAXS (Jakob et al., 1994)
	8 (earlywood), 9 (latewood)	XRD (Sahlberg et al., 1997)
<i>Picea excelsa</i>	32–35 (normal wood)	XRD (Kantola and Seitsonen, 1961)
	39–43 (compression wood)	
	18 (normal wood)	SAXS (Kantola and Kähkönen, 1963)
	25–45 (compression wood)	
Cedar (branch)	39–57	PLM (Preston, 1934)
Japanese larch	37–79	PLM (Preston, 1934)
Abies nobilis	23–69	PLM (Preston, 1934)
Virginia pine	20	PLM (Mark, 1967)
Loblolly pine	4–25 (latewood)	SM (Hiller, 1964)
	19.22–34.06	NIR (Jones et al., 2005)
Slash pine	10–40 (latewood)	SM (Hiller, 1964)
Douglas fir	20 (early & normal wood)	XRD (El-osta et al., 1973)
	7–30	PLM (Erickson and Arima, 1974)
<i>Pinus radiata</i>	10.7–41.6	NIR (Schimleck et al., 2002)
	12–27	PLM (Boyd and Foster, 1974)

*PLM; polarized light microscopy; SM; staining methods; NIR; near IR spectroscopy; XRD; X-ray diffraction; SAXS; small angle X-ray scattering.

microstructure in cell walls, such as the development of network-like morphologies in growing cells of maize and oats coleoptiles (Mühlethaler, 1950). Microfibrils form a loosely reticulated network in a newly deposited cell wall, and gradually stiffen the wall with the addition of new microfibrils. Electron microscopy has also been used to study the cell wall architecture of near native onion primary cell walls at high resolution through shadowed replicas of rapidly frozen, deep-etched specimens (McCann et al., 1990). This study suggests hemicelluloses form the cross-links between cellulose microfibrils, and indicated a lamellate model for cellulose organization; microfibrils are co-aligned within each “lamellae,” multiple lamellae (ca. 100) are stacked on top of each other, but the net orientation of each lamellae is not necessarily correlated to other lamellae. Various aspects of this model were challenged by further work on native tissues, as described below.

Although limited to the structure near the surface, SEM and AFM provide an opportunity to image the spatial arrangement of microfibrils in primary cell walls. SEM has been demonstrated as a powerful tool to examine microfibril organization and will be discussed in more detail in the next section in the context of examining the interaction between cell wall components; AFM provides a relatively unique capability of imaging cell walls in their native state. For example, detailed observations of the primary cell walls of onion and *Arabidopsis* have elucidated multiple aspects of the cellulose network structure. Contrary to reports based on electron microscopy (McCann et al., 1990), high resolution images of microfibrils in their native state for onion did not support the hypothesis of microfibrils cross-linked by hemicellulose. Instead, AFM images show microfibril bundles with single microfibrils emerging in and out to form a reticulated network (Zhang et al., 2014, 2016). **Figure 7** shows a montage of high resolution AFM images of onion where the alignment of microfibrils and extensive microfibril bundling is visible. Often, multiple layers are visible, such that the relative orientation of the layers can be examined. The studies suggest a crossed polylamellate wall structure instead of a helicoidal arrangement.

As a complementary technique to AFM, fluorescence microscopy can characterize cellulose microfibrils with high sensitivity and selectivity to chosen markers despite low spatial resolution (~200 nm). Xyloglucan binding proteins, galactan-binding proteins, or antibodies have been used with fluorescent labels for visualizing the distribution of hemicellulosic components in cell walls (Hayashi and MacLachlan, 1984; Brunecky et al., 2008; Sandquist et al., 2010). Nevertheless, the large size of these proteins restricts penetration into interstitial spaces and nano-sized pores within the cell wall structure. The search for smaller probes led to the discovery of Carbohydrate Binding Modules (CBM) as suitable molecular probes for high-resolution fluorescence microscopy because of their compact size and specificity toward targeted substrates. According to their substrate specificity, CBMs are classified as Types A, B, and C, where Type A binds to the surface of crystalline polysaccharides, B binds internally to glycan chains, and C binds to termini of glycan chains (Gilbert et al., 2013). Fluorescence microscopy with CBMs as molecular probes has been used to investigate the structure of cellulosic material both in native and treated samples (Porter Stephanie et al., 2007; Kawakubo et al., 2009; Široký et al.,

2016). In addition, confocal microscopy with the fluorescent dye Pontamine Fast Scarlet 4B (S4B), a stain that shows higher specificity for cellulose than for other cell wall components, has been used to study the cell wall architecture and dynamics of cellulose microfibrils in growing cell walls of *A. thaliana* root cells (Anderson et al., 2010). Confocal fluorescence microscopy images from this study supported the passive reorientation theory of cellulose microfibrils, which states that newly deposited cellulose microfibrils are transversely oriented to the longitudinal axis and the microfibrils reorient during expansion. **Figure 8** shows confocal images of cellulose orientation in different cell wall layers using the S4B stain. As a function of time, the cellulose microfibrils reorient from approximately 47–30° with respect to the long axis of the epidermal cells.

Scattering methods again provide a complementary approach to microscopy. Ultra-small angle (USAXS) and very small-angle X-ray scattering (VSAXS) are being used with SAXS to study the hierarchical structure of cellulose. USAXS can probe length scales from 1 to 10 μm , thus enabling the study of microfibril bundles or aggregates, while VSAXS can probe length scales intermediate between that of SAXS and USAXS. The scattering patterns of untreated and pre-treated maize using these techniques reveal the presence of structures with sizes in between microfibrils of 30 nm diameter (likely microfibril aggregates) and 140 nm bundles (Inouye et al., 2014; Zhang et al., 2015). Yet, details regarding the origin of these scattering features remain elusive.

Another approach to examine the spatial arrangement of cell walls is based on SFG. The non-centrosymmetry and phase matching requirements and the coherence length on the order of hundreds of nanometers lead to signatures of the spatial organization of crystalline cellulose dispersed in amorphous matrices. In particular, the overall SFG intensity, the alkyl peak shape, and the alkyl/hydroxyl intensity ratio have been shown to depend on the mesoscale assembly of cellulose, such as the lateral packing and net directionality of microfibrils (Lee et al., 2014). Recent work shows that SFG can detect the difference in arrangement of cellulose microfibrils between primary and secondary cell walls (Lee et al., 2014, 2015a). On the basis of the CH/OH relative intensity in SFG, it was suggested that over the SFG coherence length, primary cell walls have a lower degree of antiparallel orientation of cellulose microfibrils (Lee et al., 2014). Furthermore, control samples with uniaxially aligned cellulose crystals in amorphous matrices were examined to identify spectral signatures corresponding to the distance between microfibrils, and these signatures are supported with predictions of the spectra. The work on these model systems suggests that the CH/OH intensity ratio in SFG spectra decreases non-linearly as the intercrystallite distance increases (Makarem et al., 2017). In addition, because SFG can be performed on hydrated samples, the effect of drying has been examined. Reversible changes in the SFG spectra with dehydration and rehydration were attributed to the presence of local strains due to drying (Huang et al., 2018). The consequence of such strains could be to perturb the packing of cellulose, thereby affecting the width and position of diffraction peaks. Further work is needed to determine the consequences of drying, and to ensure that

X-ray and electron beam techniques that rely on dry samples yield reliable and biologically relevant structural information.

The aforementioned techniques provide valuable insights into the arrangement of cellulose microfibrils in cell walls. Nevertheless, relating cell wall structure with cell growth and mechanics requires an understanding of the interaction of cellulose with other matrix polysaccharides. The different approaches and techniques focused in this area are discussed in the following section.

INTERACTION OF CELLULOSE MICROFIBRILS WITH OTHER MATRIX POLYSACCHARIDES

Cell wall properties are dependent upon the combined structure, chemistry, and mechanical properties of the constituents (Chebli and Geitmann, 2017). Cellulose–cellulose and cellulose–matrix interactions influence the strength and extensibility of cell walls, thus contributing to the regulation of cell growth. The major non-cellulosic polymers in primary walls are different from those in secondary walls (Cosgrove and Jarvis, 2012). Xyloglucans and pectin are dominant in primary walls, and the current structural model of the primary wall depicts a cellulose–hemicellulose network embedded in a pectin matrix. These constituents form the crucial load bearing components. In secondary walls of coniferous wood, cellulose microfibrils form aggregates with adjacent microfibrils directly attached to each other over part of their length, and most of the hemicellulose and lignin lie out of these aggregates, with glucomannans more closely associated with the microfibrils (Fernandes et al., 2011). These structural models were derived from chemical analysis, biochemical studies, and electron and optical microscopies (Carpita and Gibeau, 1993; Sarkar et al., 2009). New approaches to examine the interaction of cellulose and matrix polysaccharides involve scattering, spectroscopy, and microscopic techniques, such as AFM and FESEM. The following section discusses the application of these techniques to investigate interactions between cellulose and matrix polysaccharides.

The heterogeneity of the cell wall composition complicates the application of characterization techniques to the whole cell wall. Methods to isolate interactions of specific wall components can

be roughly classified in one of two ways: (i) top-down approaches and (ii) bottom-up approaches (Martínez-Sanz et al., 2015a). The top-down approach involves investigating the effects of removal of non-cellulosic components on the structure of the cell wall, while the bottom-up approach involves the incorporation of additives into the culture media of cellulose-producing bacteria to mimic the assembly process taking place during plant cell wall biosynthesis. Bottom-up approaches are limited in relevance to primary cell walls given that a detailed description of cell wall assembly is currently not available; nevertheless, such studies are potentially informative as we learn more about cell wall structure and assembly and we thus briefly discuss them here.

In the top-down approach, non-cellulosic components of the cell wall can be removed by techniques including enzymatic hydrolysis and acid hydrolysis (Pingali et al., 2010), and by treatment with base (Jungnikl et al., 2007), steam (Pingali et al., 2014), or ionic liquids (Cheng et al., 2011). The effects of enzymatic hydrolysis on the structure of the cellulose network have been widely studied by SAXS and SANS (Kent et al., 2010; Penttilä et al., 2010, 2013). These studies suggest that hydrolytic digestion proceeds from the outer surface and very often cannot penetrate into the substrate interior without agitation of the sample. In addition, SEM and TEM have been extensively used to follow structural changes in the cell wall after biomass pre-treatment (Sant'Anna and de Souza, 2012). SEM is the method of choice to describe anatomical features and degradation at cellular- and nano-resolution of biomass surfaces, while TEM is combined with techniques including ultra-thin sectioning, rapid-freezing followed by deep etching, ultrastructural cytochemistry, immunogold, and electron tomography to investigate ultrastructural changes in the cell wall. In a recent study, FESEM was used to investigate fiber bundling, organization, and the spatial location and conformation of xyloglucans in onion cell walls (Zheng et al., 2018). FESEM imaging was combined with digestions by substrate-specific endoglucanases and labeling with nanogold affinity tags for cellulose and xyloglucan (Figure 9). The study provided evidence of coverage of cellulose surfaces by xyloglucan to some extent, but distinct xyloglucan structures could not be imaged. In particular, a lack of evidence for xyloglucan tethered to multiple microfibrils suggests xyloglucan does not serve as load-bearing links between microfibrils.

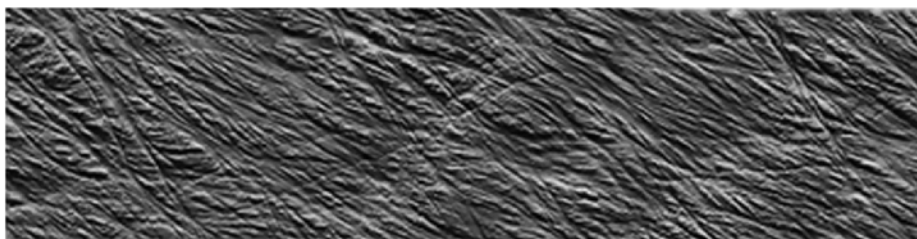


FIGURE 7 | Atomic Force Microscopy micrograph of cellulose microfibrils merging in and out of microfibril bundles. Reprinted with permission from Zhang et al. (2016). Spatial organization of cellulose microfibrils and matrix polysaccharides in primary plant cell walls as imaged by multichannel atomic force microscopy. The Plant Journal 85, 179–192. Copyright © 2016, John Wiley and Sons.

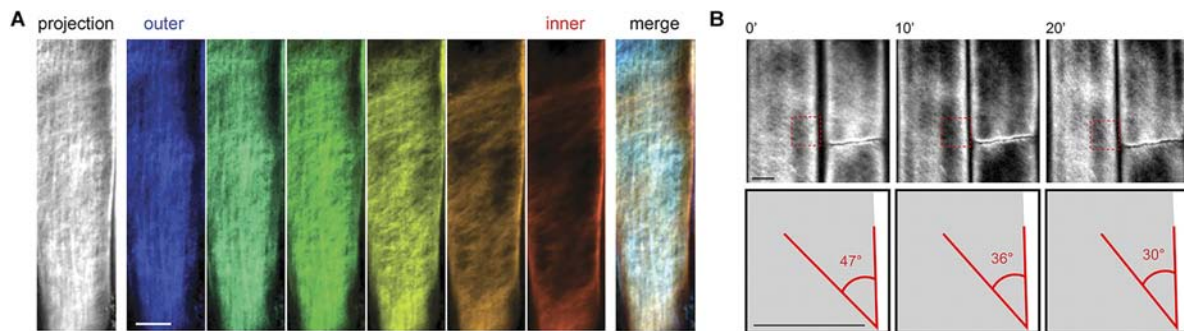


FIGURE 8 | Confocal microscopy images showing **(A)** cellulose orientation in different cell wall layers using S4B staining, and **(B)** rotation of stained microfibrils over time in *Arabidopsis*. Reprinted with permission from Anderson et al. (2010). Real-Time Imaging of Cellulose Reorientation during Cell Wall Expansion in *Arabidopsis* Roots. *Plant Physiology* 152 (2), 787–796. www.plantphysiol.org. Copyright © 2010 American Society of Plant Biologists.

Atomic force microscopy has also been applied to study the effect of chemical extraction procedures on the structure of cellulose microfibrils (Davies and Harris, 2003; Kirby et al., 2006). In a study of the effect of thermochemical treatment on maize cell wall (Chundawat et al., 2011), the ability of the AFM tip to differentiate between hydrophobic and hydrophilic regions was used to reveal that the native cell wall is mostly hydrophobic. Nevertheless, after thermochemical treatment, hydrophilic regions were found. An increased surface roughness could also be measured by AFM.

Alternatively to extraction, mutants have been used to examine the effects of modifying cell wall compositions and reveal interactions between cellulose and matrix components. Xyloglucan deficient mutants of *A. thaliana* (*xxt1 xxt2*) show highly aligned cellulose microfibrils in AFM images of the cell wall (Xiao et al., 2016). This increase in local order suggests that xyloglucan mediates interactions between cellulose microfibrils, as a spacer molecule that promotes microfibril dispersion within the cell wall. Pectin mutants of *A. thaliana* (*PGX1^{AT}*) lead to shorter homogalacturonan, and ^{13}C solid-state NMR reveals perturbations to the pectin-rich matrix and pectin-cellulose interactions in the cell walls of these plants (Phyo et al., 2017). The overall larger growth of pectin mutants and ^{13}C NMR characterization suggests that the pectin matrix influences wall dynamics during cell growth. The ongoing studies of mutants will continue to reveal fundamental interactions between cell wall components.

Another approach relies on labeling of components to provide sensitivity to specific interactions. Multidimensional solid-state NMR (MAS SS-NMR) spectroscopy, coupled with ^{13}C labeling of whole plants, enabled study of the spatial arrangement of cell wall polysaccharides in near-native cell walls. The analyses of cross-peaks in two- and three-dimensional MAS SS-NMR of ^{13}C labeled *A. thaliana* suggests that cellulose forms a single network with pectin and xyloglucans (Wang and Hong, 2016). The technique also revealed the existence of pectin-cellulose close contacts in primary cell walls (Wang et al., 2012). ^{13}C SS-NMR of mung bean cell walls detected xyloglucans of different mobilities including rigid and partly rigid (Bootten et al., 2004); the study suggests that the partly rigid xyloglucans are predominant in

the cell wall. In addition, polarization transfer in SS-NMR has been used to study water-polysaccharide interactions in primary cell walls of *Arabidopsis*. Results on water-pectin and water-cellulose spin diffusion support the single network model of the primary cell wall (White et al., 2014). Furthermore, MAS NMR of *Arabidopsis* stems revealed that xylans are found in both two and threefold screw conformations (Simmons et al., 2016). The twofold conformation is required for xylans to bind onto cellulose microfibrils.

Bottom-up approaches use a cellulose-producing bacteria such as *Gluconacetobacter xylinus* as a model system for the study of cellulose-matrix polysaccharide interactions. Cell wall polysaccharides like hemicellulose and pectin are incorporated into the culture media of the bacteria and composite pellicles are produced. The bacterial cellulose composites can then be used to examine how matrix polymers affect cellulose crystallization and how cellulose interacts with matrix polysaccharides. For example, XRD, SAXS, and SANS have been used widely to study composite pellicles with cell wall polysaccharides including xylan, xyloglucan, arabinoxylan, mannan, and pectin (Astley et al., 2001; Gu and Catchmark, 2012; Martínez-Sanz et al., 2015b). SAXS and XRD studies revealed that addition of xyloglucan affects the cellulose microfibril packing and crystalline structure; in contrast, addition of arabinoxylan does not impact these features of the cellulose network (Martínez-Sanz et al., 2015b). Spectroscopy and microscopy can also be used to examine the cellulose network within composite pellicles; nevertheless, the pellicles are highly hydrated and have strong aggregation tendency, so structural artifacts may be introduced during the drying process that is required for analysis. Recent SANS studies demonstrated that controlled incorporation of deuterium into bacterial cellulose does not introduce any structural changes in bacterial cellulose (Bali et al., 2013; He et al., 2014). The deuterated bacterial cellulose will have applications in elastic and inelastic neutron scattering experiments for studying cellulose structure and dynamics and interactions with wall polysaccharides.

Similar to scattering techniques, spectroscopic techniques including IR, Raman, SFG, and NMR have also been used to investigate interaction among cell wall components through both top-down and bottom-up approaches. IR spectroscopy has been

used to study the changes in cellulose polymorphism on addition of xyloglucan, xylan, arabinogalactan, and pectin to bacterial cellulose (Tokoh et al., 1998, 2002; Gu and Catchmark, 2012). Addition of xylan and xyloglucan results in an increase in the levels of cellulose I β and a decrease in crystallinity. Xyloglucan has a larger impact on cellulose assembly than pectin as addition of xyloglucan decreases crystallinity and increases disorder in the cellulose structure, but addition of pectin has no effect (Gu and Catchmark, 2012).

OPPORTUNITIES IN STRUCTURAL CHARACTERIZATION OF PLANT CELL WALLS

The on-going development of instrumentation and techniques for the study of soft matter structure leads to new opportunities

in the structural characterization of cell walls. In this section, we highlight some emerging techniques based on diffraction/scattering, imaging, and spectroscopy that may facilitate the creation of new knowledge on cell wall structure and assembly.

Diffraction and Scattering

The high brilliance of synchrotron radiation sources has enabled the application of X-ray microbeam diffraction and scattering techniques to weakly scattering samples like polymers and biopolymers. X-ray diffraction and scattering techniques can provide average structural parameters, but not information on local structures. Beam sizes of about 1 μm and sub- μm sizes can provide abundant local information, such as the spatial heterogeneity of materials and the structural change at a local position. An advantage of scanning X-ray diffractometry when compared to transmission electron scattering experiments is

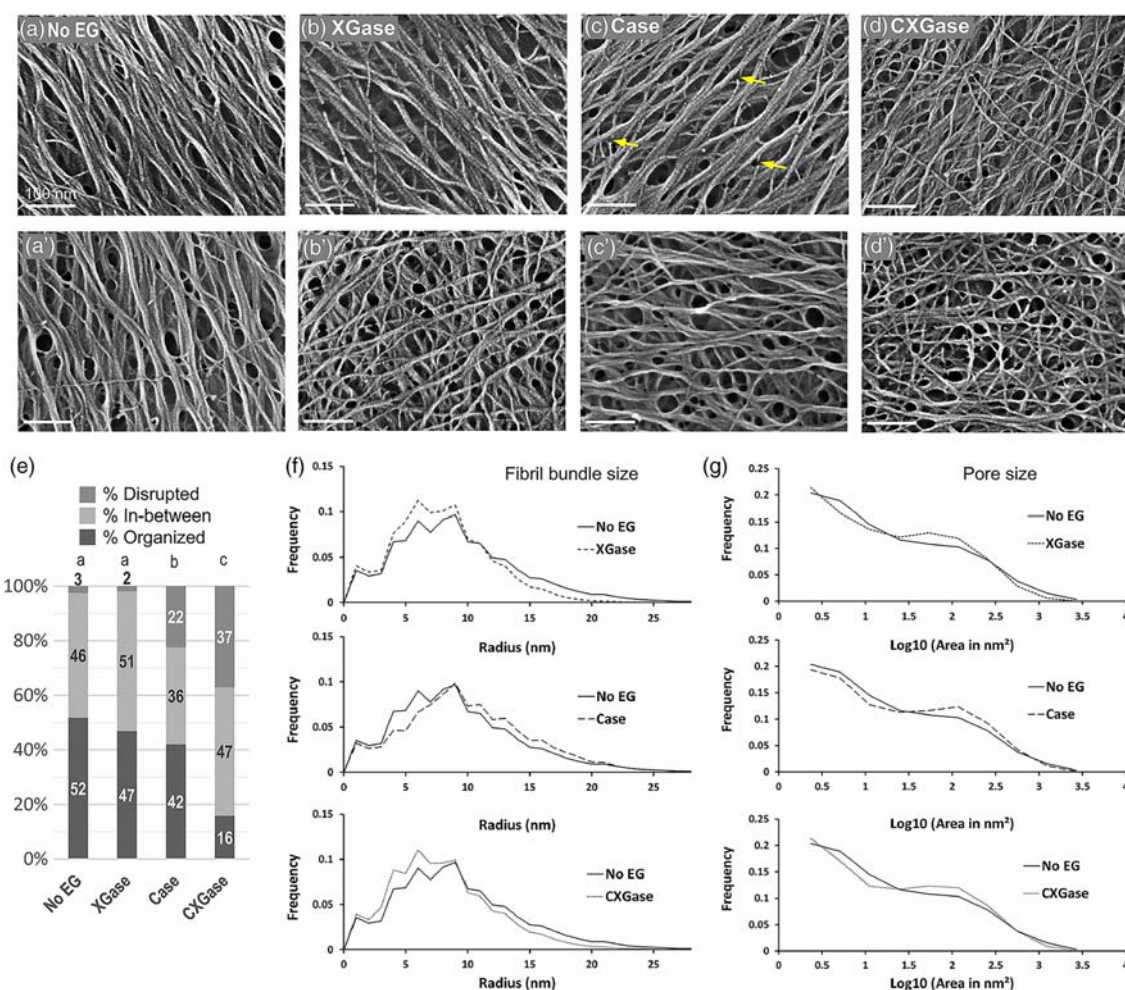


FIGURE 9 | Field emission scanning electron microscopy (FESEM) images of onion cell wall without enzyme treatment (a,a') and after different enzyme treatments (b–d,b'–d'). (e–g) Effect of treatments on wall properties. XGase is a family-12 glycosyl hydrolase (GH) that hydrolyzes xyloglucan but not unbranched glucan, Case is a family-12 GH that hydrolyzes non-crystalline cellulose only, and CXGase is a family-5 GH that hydrolyzes both non-crystalline cellulose and xyloglucan. No EG denotes no endoglucanase treatment. Reprinted with permission from Zheng et al. (2018). Xyloglucan in the primary cell wall: assessment by FESEM, selective enzyme digestions and nanogold affinity tags. *The Plant Journal* 93, 211–226. Copyright 2018, John Wiley and Sons.

the ability to examine single fibers without the necessity for sectioning (Riekkel, 2000).

The application of position-resolved synchrotron X-ray microdiffraction with beam size less than the thickness of a single cell wall enabled the imaging of the helical arrangement of cellulose microfibrils in cell walls of Norwegian spruce (Lichtenegger et al., 1999; Peura et al., 2005). X-ray microbeam diffraction has also been used to study the orientation, crystallite size, and crystallinity of cellulose microfibrils from various sources, including viscose rayon fibers (Müller et al., 2000), Japanese Cedar (Müller et al., 2002), and Norway spruce (Peura et al., 2007). For techniques aiming at analyzing small sample volumes, X-ray microdiffraction has a clear advantage over transmission electron microscopy/diffraction in terms of sample preparation and acquisition time. The application of SAXS with a beam size of a few micrometers (μ SAXS) revealed the strong alignment of cellulose microfibrils within single native flax fibers (Müller et al., 1998). Such position-resolved studies could potentially resolve the super-molecular structure of cellulose microfibrils.

Grazing Incidence Wide Angle X-ray Scattering (GIWAXS) is another synchrotron based technique (although it is becoming available in lab-scale instruments) that may be useful for primary plant cell walls. GIWAXS probes not only the surface but also beneath it. Because of its grazing incidence geometry, GIWAXS is a promising scattering technique for weakly scattering and fragile cell wall samples. The large beam footprint produces a better signal-to-noise ratio and also causes less radiation damage. GIWAXS with a 2D detector can reveal net orientation of crystals, called texturing (Baker et al., 2010 #690; Gomez et al., 2011 #1026; Rivnay et al., 2012 #3971). GIWAXS data from a cell wall sample can be used to estimate the degree of preferred orientation and crystallinity of cellulose crystals, which has not been previously demonstrated.

Resonant soft X-ray scattering (RSoXS) is a combination of conventional SAXS with soft X-ray spectroscopy that offers enhanced and tunable scattering contrast as well as elemental and chemical environment sensitivity (Virgili et al., 2007; Guo et al., 2013; Liu et al., 2016). Its large length scale accessibility, chemical sensitivity, and molecular bond orientation sensitivity makes RSoXS an attractive tool for studying different materials including biological assemblies. The different cell wall polysaccharides have similar electron density, so RSoXS could be useful in differentiating between them based on their chemical differences. Recent work has shown that RSoXS can reveal the structure of casein micelles and proteins by tuning to specific X-ray energies and thereby producing contrast between components (Ingham et al., 2015, 2016; Ye et al., 2018b). Furthermore, work on onion scales has demonstrated that tuning the X-ray energy to the Ca edge generates contrast between pectin and cellulose microfibrils, such that the spacing between microfibrils or microfibril bundles is revealed (Ye et al., 2018a). Thus, an opportunity exists to adopt a new chemically sensitive scattering technique for the study of plant cell walls.

In addition to X-ray scattering, there are opportunities for novel characterization approaches based on neutron scattering. Quasi-elastic neutron scattering (QENS) is sensitive

to reorganization of atoms and molecules on a pico-second to nano-second time scale over length scales of 1–500 Å. This broad spatial and temporal scale is ideal for studying complex biological systems as the scale is matched to atomic and molecular vibrational displacements, jump distances, and correlation lengths (Magazù and Migliardo, 2011). Because of the dependence of the relaxation times on the wave-vector, QENS can resolve spatial differences in the dynamics of water and biological macromolecules like proteins. The technique has also been applied to study water-cellulose dynamics in bacterial cellulose, which revealed the existence of two distinct populations of water in the bacterial cellulose system: surface water and water confined in the spaces between the microfibrils (O'Neill et al., 2017). Even though the nanoscale structure and composition of bacterial cellulose is markedly different from plant cell walls, the feasibility of the study presents the technique as a promising tool for the study of native plant cell walls as well.

Microscopy

Recent advances in optical, X-ray, and electron imaging tools provide new opportunities for the study of cell walls. Optical microscopes cannot distinguish between two objects separated by a lateral distance less than approximately half the wavelength of light used to image the specimen. This resolution limitation is referred to as the diffraction limit. The diffraction limit for optical microscopy is about 200–300 nm in the lateral direction and 500–700 nm in the axial (vertical) direction for confocal microscopy, which makes subcellular structures too small to be resolved in detail. This presents a problem when optical microscopy is used to investigate plant cell wall features of about a few nanometers in size. In such cases, the signal collected by optical microscopy represents an ensemble average of signals from different wall constituents. Super Resolution Fluorescence Microscopy (SRFM) refers to a host of techniques that overcome the resolution limitation caused by the diffraction limit in conventional fluorescence microscopy (Huang et al., 2009). With SRFM, three-dimensional imaging with an optical resolution of about 20 nm in the lateral direction and 40–50 nm in the axial dimension has been achieved. These techniques can employ non-linear optical effects to reduce the size of the excitation point spread function through Stimulated Emission-Depletion (STED) or Saturated Structured Illumination microscopy (SSIM). Furthermore, some techniques are also based on the localization of individual fluorescent molecules, such as stochastic optical reconstruction microscopy (STORM), photoactivated localization microscopy (PALM), and fluorescence photoactivation localization microscopy (FPALM). Recent advances have enabled 3D imaging (Huang et al., 2008), multicolor imaging (Bossi et al., 2008), and live cell imaging (Westphal et al., 2007) with SRFM.

Another approach to increase the spatial resolution beyond the diffraction barrier relies on combining near-field optical techniques with scanning probe microscopy. Near-field scanning optical microscopy (NSOM) obtains high optical and spatial resolution through the use of a tapered optical fiber with a sub-wavelength aperture of about 100 nm in diameter. Because these tips are made from optical fibers, they are fragile and

easily damaged, which can lead to artifacts. Due to these issues, NSOM with aperture-less probes with tip enhancement are being used to compartmentalize signals collected from the near field and the far field (Fragola et al., 2004). Such an ability holds promise for characterization of plant cell walls as signals from different cell wall components could be differentiated. Other apertureless tip-enhanced imaging techniques that may be able to chemically characterize plant cell walls on the cellular scale are Tip-enhanced Raman imaging, Near-field coherent anti-stokes Raman Scattering microscopy and two-photon excitation fluorescence (TPEF) spectroscopy. The capabilities of these techniques have been discussed in detail in a review (Yarbrough et al., 2009).

Total Internal Reflection Fluorescence Microscopy (TIRFM) is well suited for optical sectioning at cell-substrate regions with a thin region of fluorescence excitation (Axelrod, 2001; Mattheyses et al., 2010). The laser beam is incident on the glass-substrate interface at an angle beyond the critical angle. Due to the nature of the evanescent field, the excitation volume is large in the transverse dimension but highly confined in the axial dimension. This greatly reduces background fluorescence from out-of-focus planes and results in images with a very high signal-to-noise ratio. TIRFM has proven to be a powerful approach for examination of animal cells and for single-molecule experiments. It is particularly useful for analysis of dynamics of molecules and processes near the plasma membrane as it obscures the fluorescence from the bulk of the cell. Indeed, a recent study applied TIRFM to examine protein endocytosis in the plant plasma membrane (Johnson and Vert, 2017); however, application of TIRFM to the cell wall which lies adjacent to the plasma membrane has not yet been demonstrated. Furthermore, the use of multi-angle TIRFM opens the possibility of examining the distribution of proteins within plant samples in the axial direction (Fu et al., 2016).

High resolution can also be achieved using short-wavelength radiation. Scanning Transmission X-ray Microscopy (STXM) can generate microscopic images of a thin section of a specimen by raster-scanning a focused X-ray beam while the transmitted X-ray intensity is recorded as a function of the sample position. This technique falls under the category of 'spectro-microscopy' as X-ray absorption spectra can be obtained from microscopic features of the sectioned sample. The technique is based on synchrotron radiation and leverages X-ray absorption spectra that are characteristic of chemical states of atomic species or crystalline structures of materials (Warwick et al., 1998). Thus, STXM is useful for elemental identification and spatial mapping of heterogeneous materials (Ade and Hitchcock, 2008). The main advantages of STXM are minimal radiation damage (when compared to electron microscopy), ability to analyze hydrated samples, and ability to probe alignment of molecular orbitals due to polarization dependence. Various STXM based techniques like C-XANES and C-NEXAFS have been used to carry out chemical analysis of plant biomass (Cody, 2000; Mancosky et al., 2005; Cody et al., 2009). STXM based spectrotomography is also able to do morphological 3D visualization and quantitative chemical mapping in bacteria (Wang et al., 2011).

STXM based on soft X-ray spectromicroscopy is a powerful technique that holds promise for characterization of plant samples with advantages of high spatial resolution and chemical sensitivity similar to mid infrared spectromicroscopy. The major problem in characterizing cell wall samples is the heterogeneous matrix and the spectra obtained are often dominated by the component in highest concentration. The use of X-ray fluorescent probes can be used with high resolution STXM to overcome the limitations of molecular sensitivity. Using the combination of confocal laser microscopy with fluorescent probes and STXM could be a valuable approach for studying plant cell wall samples. The approach has been demonstrated successfully in microbial biofilms (Lawrence et al., 2003).

Advances in TEM, and in particular in cryogenic transmission electron microscopy (cryo-TEM), provides new opportunities that are also based on short-wavelength radiation. Recent advances in direct electron detectors and automatic image acquisition have significantly advanced structural biology, as these instrumental developments allow for better signal to noise and acquisition of large data sets that can be averaged to improve resolution. The process begins with vitrification, in which the sample solution is rapidly cooled and water molecules form an amorphous solid instead of crystallizing. Resolutions of approximately 3 Å or lower have been achieved by cryo-TEM (Bartesaghi et al., 2015; Dellisanti, 2015). The technique can analyze large and complex biological assemblies that are often difficult to crystallize for X-ray crystallography or are too large and complex for NMR. 3D images of samples can be reconstructed from tilted 2D images through cryogenic electron tomography (cryo-ET). Both cryo-TEM and cryo-ET hold promise for characterization of plant cell walls as they allow analysis of the preserved hydrated state. The use of cryo-TEM to study the cell wall organization of *Staphylococcus aureus* has been demonstrated (Matias and Beveridge, 2006). Cryo-ET has also been used for 3D visualization of cell wall ultrastructure at a resolution of about 2 nm without isolation of cell walls (Sarkar et al., 2014). The microfibril diameter within *Arabidopsis* cell walls found from the study was comparable to diameters measured from AFM. Nevertheless, the sample preparation required for this approach is lengthy and arduous when compared to the more commonly used imaging techniques like AFM and FESEM. Application of faster sample preparation protocols might contribute to more routine use of the technique.

The use of ionizing radiation, as in X-ray or electron microscopy, is limited by the damage caused from the beam. An alternative approach is Scanning Acoustic Microscopy (SAM), which makes use of acoustic waves to create images of microscopic objects. Unlike optical microscopy, SAM does not require any staining or fixation, so it can be used for imaging live cells. Also, it can non-invasively observe not only the surface but also the internal structure of the specimen with sub-micron resolution. In addition, SAM is capable of measuring mechanical properties like the loss factor and modulus of tissues (Maeva et al., 2009). The interactions between ultrasonic waves and matter determine the size of the receiving signal and thus create contrast; contrast is generated on the basis of different acoustic impedances of different materials and is also due to absorption

of acoustic waves in the material. Conventional SAM operates in the range of 20–200 MHz while High Frequency SAM (HF-SAM) operates in the 0.4–2 GHz range. HF-SAM has been used to study the hydrated primary cell wall of onion epidermis (Tittmann and Xi, 2014). In this study, SAM was able to detect that enzymatic removal of pectin influences the mechanical properties of primary cell wall. Thus, SAM presents potential as a powerful tool to study not only the structure and mechanics of the cell wall in its natural state but also the interactions between the different wall components through the top-down approach of enzymatic treatments.

As discussed earlier, microscopic techniques are widely used for direct visualization of plant cell walls. Nevertheless, only a few examples of quantitative image analysis have been reported. Typical image analysis includes determining particle sizes, area, length, porosity, and other useful measurements. The availability of open source and open architecture image processing software like ImageJ (Schneider et al., 2012) has contributed to the ability to readily quantify various parameters from microscopic images. For example, ImageJ has been used to process and analyze AFM images to quantify different cellulose microfibril parameters like width and orientation (Boudaoud et al., 2014; Kafle et al., 2014; Zhang et al., 2016, 2017). Several open source image analysis software packages including SOAX (Xu et al., 2015; Zhang et al., 2017) and FibreApp (Usov and Mezzenga, 2015) in addition to ImageJ offer immense opportunities to be used for quantitative analyses of microscopic images of cell walls.

Yet, quantitative analysis of microscopy images of the cell wall is still challenging as the structure is highly heterogeneous. It is even more difficult in primary cell walls due to the higher degree of disorder. A number of times, the arrangement of microfibrils have been reported to mimic a ‘liquid crystal’ like structure (Reis et al., 1991; Himmel et al., 2007). The molecules in such structures seem to have a certain degree of preferred orientation. The amount of ‘order’ in such states can be defined by an order parameter that describes ordering in liquid crystals (Nishiguchi et al., 2017). Currently available image analysis tools have capabilities that enable estimation of such order from microscopy images of cell walls.

Spectroscopy

The region of the electromagnetic spectrum from 0.1 to 10 THz ($3.3\text{--}333.6\text{ cm}^{-1}$) is described as the terahertz (THz) region. THz spectroscopy has the ability to distinguish between samples with good and poor long-range order and thus can probe the crystallinity of materials (McIntosh et al., 2012). For THz radiation, crystalline materials present well-defined absorption peaks while amorphous phases present featureless spectra. It can differentiate between different crystalline phases as well (Strachan et al., 2005). THz-time domain spectroscopy (TDS) has been applied to determine the degree of crystallinity of microcrystalline cellulose samples (Vieira and Pasquini, 2014). Because THz radiation is responsible for long-range periodic vibrations in crystals, the absorption bands can be directly related to the degree of crystallinity of the sample. As such, the technique can selectively detect crystalline cellulose and holds promise for characterization of cellulose in native cell wall samples.

Atomic force microscopy-based infrared (AFM-IR) spectroscopy combines the spatial resolution of AFM with the chemical analysis capability of IR spectroscopy. It was developed to overcome the diffraction barrier limitation of IR spectroscopy and the inability of AFM to discriminate materials on the basis of chemical composition. The nanometer scale spatial resolution of AFM-IR allows IR microspectroscopy to investigate many life science problems like subcellular imaging and spectroscopy of bacterial and mammalian cells. Extension of this technique to plant cell wall studies may reveal important information about the spatial distribution of various cell wall components.

CONCLUSION AND OUTLOOK

Describing the structure of cellulose has direct implications on understanding the anisotropic growth and mechanics of plants, designing efficient biofuel conversion, and developing biomass-based products. Nevertheless, complete elucidation of the structure of cellulose and its interaction with matrix components has not been possible due to the complexity and heterogeneity of the cell wall and its variability from species to species. Ambiguity in the interpretation of structural characterization data obtained from different plant sources could sometimes be explained by a complimentary technique. For example, in the case of crystal parameters of native cellulose, the ambiguity in XRD results were resolved by NMR spectroscopy, which established the existence of two cellulose allomorphs, cellulose Ia and Ib.

Ambiguities can also be seen among results quantifying certain properties measured through different techniques. For example, there is a large mismatch between estimates of the lateral dimension of cellulose crystallites obtained from XRD and from electron microscopy. Most often, this mismatch is attributed to artifacts induced by sample preparation for electron microscopy. Due to the structural complexity, a single technique cannot characterize a cellulose microfibril completely. Recent reports present cell wall characterization through combined application of complementary techniques like diffraction, scattering, spectroscopy, and microscopy. This combination of techniques is also applied for examining the interaction between cellulose and other cell wall polysaccharides by either studying cellulose microfibrils after sequential removal of other polysaccharides (top-down approach) or by studying the effect on cellulose after introducing additives to bacterial cellulose composites (bottom-up approach). Recent developments in SS-NMR have also enabled the study of interactions of cell wall components with each other and with water directly in native primary cell walls of plants.

The perpetual concern of introduction of artifacts during cell wall preparation has been reduced with the application of techniques like AFM and X-ray scattering that require minimal sample preparation. These approaches allow for the characterization of cell walls in near native states. The use of hybrid techniques like AFM-IR/Raman and advanced scattering techniques like RSoXS can provide chemical sensitivity along with high spatial resolution. In addition, use of advanced

microscopic techniques like cryogenic electron tomography (cryo-ET) can create 3D reconstructions of nearly native cell walls and use of image analysis tools can quantify aspects of the microstructure.

Despite tremendous progress to date, many aspects of cellulose structure, and cellulose–cellulose and cellulose–matrix interactions are not well understood. The relation between the nanostructure of the cell wall and its macroscopic properties remains elusive. The current model of the primary cell wall suggests the existence of ‘biomechanical hotspots,’ which are sites of close contacts between cellulose microfibrils mediated by xyloglucans (Cosgrove, 2014). These are proposed as the control sites for wall extension. Nevertheless, many questions regarding the creation, destruction, location, and functioning mechanism of these structures are yet to be answered. Furthermore, recent work shows a linear correlation between the FWHM of cellulose (200) diffraction peaks and *d*-spacing for different sources (Huang et al., 2018). Thus, the *d*-spacing of cellulose is inversely proportional to the crystallite size. This inverse proportionality might be from enhanced thermal fluctuations and higher para-crystalline disorder in smaller crystals; yet, further work is warranted to ascertain the origin of this empirical relationship. We predict that the application of emerging approaches and multi-modal analyses (combination of multiple techniques) will generate

new insights on the abovementioned topics and on other open questions regarding the regulation of cell wall growth and mechanics.

AUTHOR CONTRIBUTIONS

EWG, EDG, DY, and SR contributed to the conception and design of the review. All authors wrote the manuscript, contributed to manuscript revisions, and read and approved the submitted version.

FUNDING

This work was supported as part of the Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the United States Department of Energy, Office of Science, Basic Energy Sciences under award no. DE-SC0001090.

ACKNOWLEDGMENTS

The authors acknowledge Dan Cosgrove for educational discussions.

REFERENCES

- Abe, H., Funada, R., Ohtani, J., and Fukazawa, K. (1997). Changes in the arrangement of cellulose microfibrils associated with the cessation of cell expansion in tracheids. *Trees* 11, 328–332. doi: 10.1007/s004680050092
- Ade, H., and Hitchcock, A. P. (2008). NEXAFS microscopy and resonant scattering: composition and orientation probed in real and reciprocal space. *Polymer* 49, 643–675. doi: 10.1016/j.polymer.2007.10.030
- Agarwal, U. P. (2014). 1064 nm FT-Raman spectroscopy for investigations of plant cell walls and other biomass materials. *Front. Plant Sci.* 5:490. doi: 10.3389/fpls.2014.00490
- Agarwal, U. P., Reiner, R. R., and Ralph, S. A. (2013). Estimation of cellulose crystallinity of lignocelluloses using near-IR FT-Raman spectroscopy and comparison of the Raman and Segal-WAXS methods. *J. Agric. Food Chem.* 61, 103–113. doi: 10.1021/jf304465k
- Agarwal, U. P., Reiner, R. S., and Ralph, S. A. (2010). Cellulose I crystallinity determination using FT-Raman spectroscopy: univariate and multivariate methods. *Cellulose* 17, 721–733. doi: 10.1007/s10570-010-9420-z
- Ahmed, A. U., Ahmed, N., Aslam, J., Butt, N. M., Khan, Q. H., and Atta, M. A. (1976). Neutron diffraction studies of the unit cell of cellulose II. *J. Polym. Sci.* 14, 561–564. doi: 10.1002/pol.1976.130140909
- Ahvenainen, P., Kontro, I., and Svedström, K. (2016). Comparison of sample crystallinity determination methods by X-ray diffraction for challenging cellulose I materials. *Cellulose* 23, 1073–1086. doi: 10.1007/s10570-016-0881-6
- Altaner, C. M., and Jarvis, M. C. (2008). Modelling polymer interactions of the ‘molecular Velcro’ type in wood under mechanical stress. *J. Theor. Biol.* 253, 434–445. doi: 10.1016/j.jtbi.2008.03.010
- Amiri, H., and Karimi, K. (2015). Improvement of acetone, butanol, and ethanol production from woody biomass using organosolv pretreatment. *Bioprocess Biosyst. Eng.* 38, 1959–1972. doi: 10.1007/s00449-015-1437-0
- Anderson, C. T., Carroll, A., Akhmetova, L., and Somerville, C. (2010). Real-time imaging of cellulose reorientation during cell wall expansion in Arabidopsis roots. *Plant Physiol.* 152, 787–796. doi: 10.1104/pp.109.150128
- Andersson, S., Serimaa, R., Torkkeli, M., Paakkari, T., Saranpää, P., and Pesonen, E. (2000). Microfibril angle of Norway spruce [*Picea abies* (L.) Karst.] compression wood: comparison of measuring techniques. *J. Wood Sci.* 46, 343–349. doi: 10.1007/BF00776394
- Araujo, C., Freire, C. S. R., Nolasco, M. M., Ribeiro-Claro, P. J. A., Rudić, S., Silvestre, A. J. D., et al. (2018). Hydrogen bond dynamics of cellulose through inelastic neutron scattering spectroscopy. *Biomacromolecules* 19, 1305–1313. doi: 10.1021/acs.biomac.8b00110
- Astley, O. M., Chanliaud, E., Donald, A. M., and Gidley, M. J. (2001). Structure of Acetobacter cellulose composites in the hydrated state. *Int. J. Biol. Macromol.* 29, 193–202. doi: 10.1016/S0141-8130(01)00167-2
- Astley, O. M., and Donald, A. M. (2001). A small-angle X-ray scattering study of the effect of hydration on the microstructure of flax fibers. *Biomacromolecules* 2, 672–680. doi: 10.1021/bm005643l
- Atalla, R. H. (1987). “Structures of Cellulose,” in *The Structures of Cellulose*, ed. R. H. Atalla, (Washington, DC: American Chemical Society), 1–14. doi: 10.1021/bk-1987-0340
- Atalla, R. H., and Vanderhart, D. L. (1984). Native cellulose: a composite of two distinct crystalline forms. *Science* 223, 283–285. doi: 10.1126/science.223.4633.283
- Atalla, R. H., and VanderHart, D. L. (1999). The role of solid state ¹³C NMR spectroscopy in studies of the nature of native celluloses. *Solid State Nuclear Magn. Reson.* 15, 1–19. doi: 10.1016/S0926-2040(99)00042-9
- Awano, T., Takabe, K., and Fujita, M. (2002). Xylan deposition on secondary wall of *Fagus crenata* fiber. *Protoplasma* 219, 106–115. doi: 10.1007/s007090200011
- Axelrod, D. (2001). Total internal reflection fluorescence microscopy in cell biology. *Traffic* 2, 764–774. doi: 10.1034/j.1600-0854.2001.21104.x
- Bailey, I. W., and Vestal, M. R. (1937). The orientation of cellulose in the secondary wall of tracheary cells. *J. Arnold Arbor.* 18, 185–195. doi: 10.1104/pp.17.00885
- Baker, J. L., Jimison, L. H., Mannsfeld, S., Volkman, S., Yin, S., Subramanian, V., et al. (2010). Quantification of thin film crystallographic orientation using X-ray diffraction with an area detector. *Langmuir* 26, 9146–9151. doi: 10.1021/la904840q
- Bali, G., Foston, M. B., O’Neill, H. M., Evans, B. R., He, J., and Ragauskas, A. J. (2013). The effect of deuteration on the structure of bacterial cellulose. *Carbohydr. Res.* 374, 82–88. doi: 10.1016/j.carres.2013.04.009

- Barnett, J. R., and Bonham, V. A. (2007). Cellulose microfibril angle in the cell wall of wood fibres. *Biol. Rev.* 79, 461–472. doi: 10.1017/S1464793103006377
- Barnette, A. L., Lee, C., Bradley, L. C., Schreiner, E. P., Park, Y. B., Shin, H., et al. (2012). Quantification of crystalline cellulose in lignocellulosic biomass using sum frequency generation (SFG) vibration spectroscopy and comparison with other analytical methods. *Carbohydr. Polym.* 89, 802–809. doi: 10.1016/j.carbpol.2012.04.014
- Bartasaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J. L. S., et al. (2015). 2.2 Å resolution cryo-EM structure of β -galactosidase in complex with a cell-permeant inhibitor. *Science* 348, 1147–1151. doi: 10.1126/science.aab1576
- Bayley, S. T., Colvin, J. R., Cooper, F. P., and Martin-Smith, C. A. (1957). The structure of the primary epidermal cell wall of avena coleoptiles. *J. Biophys. Biochem. Cytol.* 3, 171–182. doi: 10.1083/jcb.3.2.171
- Beer, M., and Setterfield, G. (1958). Fine structure in thickened primary walls of collenchyma cells of celery petioles. *Am. J. Bot.* 45, 571–580. doi: 10.1002/j.1537-2197.1958.tb13169.x
- Beg, M. M., Aslam, J., Khan, Q. H., Butt, N. M., Rolandson, S., and Ahmed, A. U. (1974). Neutron diffraction studies of the unit cell of cellulose I. *J. Polym. Sci.* 12, 311–318. doi: 10.1002/pol.1974.130120602
- Bergander, A., and Salmén, L. (2002). Cell wall properties and their effects on the mechanical properties of fibers. *J. Mater. Sci.* 37, 151–156. doi: 10.1023/A:1013115925679
- Bootten, T. J., Harris, P. J., Melton, L. D., and Newman, R. H. (2004). Solid-state ^{13}C -NMR spectroscopy shows that the xyloglucans in the primary cell walls of mung bean (*Vigna radiata* L.) occur in different domains: a new model for xyloglucan–cellulose interactions in the cell wall. *J. Exp. Bot.* 55, 571–583. doi: 10.1093/jxb/erh065
- Bossi, M., Fölling, J., Belov, V. N., Boyarskiy, V. P., Medda, R., Egner, A., et al. (2008). Multicolor far-field fluorescence nanoscopy through isolated detection of distinct molecular species. *Nano Lett.* 8, 2463–2468. doi: 10.1021/nl801471d
- Boudaoud, A., Burian, A., Borowska-Wykręt, D., Uyttewaal, M., Wrzalik, R., Kwiatkowska, D., et al. (2014). FibrilTool, an ImageJ plug-in to quantify fibrillar structures in raw microscopy images. *Nat. Protoc.* 9, 457–463. doi: 10.1038/nprot.2014.024
- Bouret, A., Chanzy, H., and Lazaro, R. (1972). Crystallite features of valonia cellulose by electron diffraction and dark-field electron microscopy. *Biopolymers* 11, 893–898. doi: 10.1002/bip.1972.360110412
- Boyd, J. D., and Foster, R. C. (1974). Tracheid anatomy changes as responses to changing structural requirements of the tree. *Wood Sci. Technol.* 8, 91–105. doi: 10.1007/BF00351365
- Brunecky, R., Vinzant Todd, B., Porter Stephanie, E., Donohoe Bryon, S., Johnson David, K., and Himmel Michael, E. (2008). Redistribution of xylan in maize cell walls during dilute acid pretreatment. *Biotechnol. Bioeng.* 102, 1537–1543. doi: 10.1002/bit.22211
- Butt, H. J., Wolff, E. K., Gould, S. A. C., Dixon Northern, B., Peterson, C. M., and Hansma, P. K. (1990). Imaging cells with the atomic force microscope. *J. Struct. Biol.* 105, 54–61. doi: 10.1016/1047-8477(90)90098-W
- Carpita, N. C., and Gibeau, D. M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30. doi: 10.1111/j.1365-3113.1993.tb00007.x
- Carpita, N. C., Marianne, D., Findlay, K., Wells, B., Shoue, D. A., Gareth, C., et al. (2001). Cell wall architecture of the elongating maize coleoptile. *Plant Physiol.* 127, 551–565. doi: 10.1104/pp.010146
- Caulfield, D. F. (1971). Crystallite sizes in wet and dry *Valonia ventricosa* textile. *Res. J.* 41, 264–280.
- Cave, I. D. (1968). The anisotropic elasticity of the plant cell wall. *Wood Sci. Technol.* 2, 268–278. doi: 10.1007/BF00350273
- Chebly, Y., and Geitmann, A. (2017). Cellular growth in plants requires regulation of cell wall biochemistry. *Curr. Opin. Cell Biol.* 44, 28–35. doi: 10.1016/j.cob.2017.01.002
- Cheng, G., Varanasi, P., Li, C., Liu, H., Melnichenko, Y. B., Simmons, B. A., et al. (2011). Transition of cellulose crystalline structure and surface morphology of biomass as a function of ionic liquid pretreatment and its relation to enzymatic hydrolysis. *Biomacromolecules* 12, 933–941. doi: 10.1021/bm101240z
- Choong, F. X., Bäck, M., Steiner, S. E., Melican, K., Nilsson, K. P. R., Edlund, U., et al. (2016). Nondestructive, real-time determination and visualization of cellulose, hemicellulose and lignin by luminescent oligothiophenes. *Sci. Rep.* 6:35578. doi: 10.1038/srep35578
- Chundawat, S. P. S., Donohoe, B. S., da Costa Sousa, L., Elder, T., Agarwal, U. P., Lu, F., et al. (2011). Multi-scale visualization and characterization of lignocellulosic plant cell wall deconstruction during thermochemical pretreatment. *Energy Environ. Sci.* 4, 973–984. doi: 10.1039/c0ee00574f
- Cody, G. D. (2000). Probing chemistry within the membrane structure of wood with soft X-ray spectral microscopy. *AIP Conf. Proc.* 507, 307–312. doi: 10.1063/1.1291162
- Cody, G. D., Brandes, J., Jacobsen, C., and Wirick, S. (2009). Soft X-ray induced chemical modification of polysaccharides in vascular plant cell walls. *J. Electr. Spectrosc. Relat. Phenomena* 170, 57–64. doi: 10.1016/j.elspec.2008.09.007
- Colvin, J. R. (1963). The size of the cellulose microfibril. *J. Cell Biol.* 17, 105–109. doi: 10.1083/jcb.17.1.105
- Cosgrove, D., and Jarvis, M. (2012). Comparative structure and biomechanics of plant primary and secondary cell walls. *Front. Plant Sci.* 3:204. doi: 10.3389/fpls.2012.00204
- Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6, 850–861. doi: 10.1038/nrm1746
- Cosgrove, D. J. (2014). Re-constructing our models of cellulose and primary cell wall assembly. *Curr. Opin. Plant Biol.* 22, 122–131. doi: 10.1016/j.pbi.2014.11.001
- Cox, G., and Juniper, B. (1973). Electron microscopy of cellulose in entire tissue. *J. Microsc.* 97, 343–355. doi: 10.1111/j.1365-2818.1973.tb03789.x
- Crow, E., and Murphy, R. J. (2000). Microfibril orientation in differentiating and maturing fibre and parenchyma cell walls in culms of bamboo (*Phyllostachys viridi-glaucescens* (Carr.) Riv. & Riv.). *Bot. J. Linnean Soc.* 134, 339–359. doi: 10.1111/j.1095-8339.2000.tb02357.x
- Davies, L. M., and Harris, P. J. (2003). Atomic force microscopy of microfibrils in primary cell walls. *Planta* 217, 283–289.
- De Figueiredo, L. P., and Ferreira, F. F. (2014). The rietveld method as a tool to quantify the amorphous amount of microcrystalline cellulose. *J. Pharm. Sci.* 103, 1394–1399. doi: 10.1002/jps.23909
- Dellisanti, C. (2015). A barrier-breaking resolution. *Nat. Struct. Mol. Biol.* 22, 361–361. doi: 10.1038/nsmb.3025
- Ding, S. Y., and Himmel, M. E. (2006). The maize primary cell wall microfibril: a new model derived from direct visualization. *J. Agric. Food Chem.* 54, 597–606. doi: 10.1021/jf051851z
- Ding, S.-Y., Zhao, S., and Zeng, Y. (2014). Size, shape, and arrangement of native cellulose fibrils in maize cell walls. *Cellulose* 21, 863–871. doi: 10.1007/s10570-013-0147-5
- Donaldson, L. (2008). Microfibril angle: measurement, variation and relationship: a review. *IWA J.* 29, 387–396. doi: 10.1163/22941932-90000192
- Donaldson, L., and Frankland, A. (2004). Ultrastructure of iodine treated wood. *Holzforchung* 58, 219–225. doi: 10.1515/HF.2004.034
- Driemeier, C. (2014). Two-dimensional Rietveld analysis of celluloses from higher plants. *Cellulose* 21, 1065–1073. doi: 10.1007/s10570-013-9995-2
- Dunning, C. E. (1968). Cell wall morphology of longleaf pine latewood. *Wood Sci. Technol.* 1, 65–76.
- El-Hosseiny, F., and Page, D. H. (1973). The measurement of fibril angle of wood fibres using polarised light. *Wood Fibre* 5:7.
- El-osta, M. L. M., Kellogg, R. M., Foschi, R. O., and Butters, R. G. (1973). A direct x-ray technique for measuring microfibril angle. *Wood Fibre* 5, 118–128.
- Erickson, H. D., and Arima, T. (1974). Douglas-fir wood quality studies Part II: effects of age and stimulated growth on fibril angle and chemical constituents. *Wood Sci. Technol.* 8, 255–265. doi: 10.1007/BF00351859
- Fernandes, A. N., Thomas, L. H., Altaner, C. M., Callow, P., Forsyth, V. T., Apperley, D. C., et al. (2011). Nanostructure of cellulose microfibrils in spruce wood. *Proc. Natl. Acad. Sci. U.S.A.* 108, E1195–E1203. doi: 10.1073/pnas.1108942108
- Fragola, A., Aigouy, L., Mignotte, P. Y., Formanek, F., and De Wilde, Y. (2004). Apertureless scanning near-field fluorescence microscopy in liquids. *Ultramicroscopy* 101, 47–54. doi: 10.1016/j.ultramic.2004.04.002

- Frei, E. V. A., Preston, R. D., and Ripley, G. W. (1957). The fine structure of the walls of conifer tracheids: VI. Electron microscope investigations of sections. *J. Exp. Bot.* 8, 139–146. doi: 10.1093/jxb/8.1.139
- French, A. D., Roughhead, W. A., and Miller, D. P. (1987). “X-ray diffraction studies of ramie cellulose I,” in *The Structures of Cellulose*, ed. R. H. Atalla (Washington, DC: American Chemical Society), 15–37.
- Frilette, V. J., Hanle, J., and Mark, H. (1948). Rate of exchange of cellulose with heavy water. *J. Am. Chem. Soc.* 70, 1107–1113. doi: 10.1021/ja01183a071
- Fu, Y., Winter, P. W., Rojas, R., Wang, V., McAuliffe, M., and Patterson, G. H. (2016). Axial superresolution via multiangle TIRF microscopy with sequential imaging and photobleaching. *Proc. Natl. Acad. Sci. U.S.A.* 113, 4368–4373. doi: 10.1073/pnas.1516715113
- Gilbert, H. J., Knox, J. P., and Boraston, A. B. (2013). Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr. Opin. Struct. Biol.* 23, 669–677. doi: 10.1016/j.sbi.2013.05.005
- Gomez, E. D., Barteau, K. P., Wang, H., Toney, M. F., and Loo, Y.-L. (2011). Correlating the scattered intensities of P3HT and PCBM to the current densities of polymer solar cells. *Chem. Commun.* 47, 436–438. doi: 10.1039/c0cc02927k
- Gu, J., and Catchmark, J. M. (2012). Impact of hemicelluloses and pectin on sphere-like bacterial cellulose assembly. *Carbohydr. Polym.* 88, 547–557. doi: 10.1016/j.carbpol.2011.12.040
- Gunning, A. P., Kirby, A. R., Morris, V. J., Wells, B., and Brooker, B. E. (1995). Imaging bacterial polysaccharides by AFM. *Polym. Bull.* 34, 615–619. doi: 10.1007/BF00423359
- Guo, C., Kozub, D. R., Kesava, S. V., Wang, C., Hexemer, A., and Gomez, E. D. (2013). Signatures of multiphase formation in the active layer of organic solar cells from resonant soft X-ray scattering. *ACS Macro Lett.* 2, 185–189. doi: 10.1021/mz300547x
- Gusev, G. V. (1978). Hermans-Weidinger X-ray diffraction technique for determining polymer crystallinity and the use of the Ruland ratio. *Polym. Sci. U.S.S.R.* 20, 1295–1297. doi: 10.1016/0032-3950(78)90270-8
- Ha, M. A., Apperley, D. C., Evans, B. W., Huxham, I. M., Jardine, W. G., Viëtor, R. J., et al. (1998). Fine structure in cellulose microfibrils: NMR evidence from onion and quince. *Plant J.* 16, 183–190. doi: 10.1046/j.1365-313x.1998.00291.x
- Harris, D., Bulone, V., Ding, S.-Y., and DeBolt, S. (2010). Tools for cellulose analysis in plant cell walls. *Plant Physiol.* 153, 420–426. doi: 10.1104/pp.110.154203
- Harris, D. M., Corbin, K., Wang, T., Gutierrez, R., Bertolo, A. L., Petti, C., et al. (2012). Cellulose microfibril crystallinity is reduced by mutating C-terminal transmembrane region residues CESA1A903V and CESA3T942I of cellulose synthase. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4098–4103. doi: 10.1073/pnas.1200352109
- Hayashi, T., and MacLachlan, G. (1984). Pea xyloglucan and cellulose. *Plant Physiol.* 75, 596–604. doi: 10.1104/pp.75.3.596
- He, J., Pingali, S. V., Chundawat, S. P. S., Pack, A., Jones, A. D., Langan, P., et al. (2014). Controlled incorporation of deuterium into bacterial cellulose. *Cellulose* 21, 927–936. doi: 10.1016/bs.mie.2015.08.031
- Helbert, W., Nishiyama, Y., Okano, T., and Sugiyama, J. (1998a). Molecular imaging of *Halocynthia papillosa* cellulose. *J. Struct. Biol.* 124, 42–50.
- Helbert, W., Sugiyama, J., Kimura, S., and Itoh, T. (1998b). High-resolution electron microscopy on ultrathin sections of cellulose microfibrils generated by glomerulocytes in *Polyzoa vesiculiphora*. *Protoplasma* 203, 84–90. doi: 10.1007/BF01280590
- Hessler, L. E., and Power, R. E. (1954). The use of iodine adsorption as a measure of cellulose fiber crystallinity. *Textile Res. J.* 24, 822–827. doi: 10.1177/004051755402400906
- Heyn, A. N. J. (1955). Small particle X-ray scattering by fibers, size and shape of microcrystallites. *J. Appl. Phys.* 26, 519–526. doi: 10.1063/1.1722032
- Heyn, A. N. J. (1966). The microcrystalline structure of cellulose in cell walls of cotton, ramie, and jute fibers as revealed by negative staining of sections. *J. Cell Biol.* 29, 181–197. doi: 10.1083/jcb.29.2.181
- Hill, J. L., Hammudi, M. B., and Tien, M. (2014). The *Arabidopsis* cellulose synthase complex: a proposed hexamer of CESA trimers in an equimolar stoichiometry. *Plant Cell* 26, 4834–4842. doi: 10.1105/tpc.114.131193
- Hiller, C. H. (1964). Correlation of fibril angle with wall thickness of tracheids in summerwood of slash and loblolly pine. *Tappi* 47, 125–128.
- Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., et al. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315, 804–807. doi: 10.1126/science.1137016
- Hindeleh, A. M., and Johnson, D. J. (1972). Crystallinity and crystallite size measurement in cellulose fibres: 1. Ramie and Fortisan. *Polymer* 13, 423–430. doi: 10.1016/0032-3861(72)90107-3
- Hindeleh, A. M., and Johnson, D. J. (1974). Crystallinity and crystallite size measurement in cellulose fibres: 2. Viscose rayon. *Polymer* 15, 697–705. doi: 10.1016/0032-3861(74)90020-2
- Hofstetter, K., Hellmich, C., and Eberhardsteiner, J. (2005). Development and experimental validation of a continuum micromechanics model for the elasticity of wood. *Eur. J. Mech.* 24, 1030–1053. doi: 10.1016/j.euromechsol.2005.05.006
- Holt, C., Mackie, W., and Sellen, D. B. (1973). Degree of polymerization and polydispersity of native cellulose. *J. Polym. Sci.* 42, 1505–1512. doi: 10.1002/polc.5070420353
- Honjo, G., and Watanabe, M. (1958). Examination of cellulose fibre by the low-temperature specimen method of electron diffraction and electron microscopy. *Nature* 181, 326–328. doi: 10.1038/181326a0
- Horikawa, Y., Clair, B., and Sugiyama, J. (2009). Varietal difference in cellulose microfibril dimensions observed by infrared spectroscopy. *Cellulose* 16, 1–8. doi: 10.1007/s10570-008-9252-2
- Howsmen, J. A. (1949). Water sorption and the poly-phase structure of cellulose fibers. *Textile Res. J.* 19, 152–162. doi: 10.1177/004051754901900303
- Huang, B., Bates, M., and Zhuang, X. (2009). Super resolution fluorescence microscopy. *Annu. Rev. Biochem.* 78, 993–1016. doi: 10.1146/annurev.biochem.77.061906.092014
- Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* 319, 810–813. doi: 10.1126/science.1153529
- Huang, S., Makarem, M., Kiemle, S. N., Zheng, Y., He, X., Ye, D., et al. (2018). Dehydration-induced physical strains of cellulose microfibrils in plant cell walls. *Carbohydr. Polym.* 197, 337–348. doi: 10.1016/j.carbpol.2018.05.091
- Ingham, B., Erlangga, G. D., Smialowska, A., Kirby, N. M., Wang, C., Matia-Merino, L., et al. (2015). Solving the mystery of the internal structure of casein micelles. *Soft Matter* 11, 2723–2725. doi: 10.1039/c5sm00153f
- Ingham, B., Smialowska, A., Erlangga, G. D., Matia-Merino, L., Kirby, N. M., Wang, C., et al. (2016). Revisiting the interpretation of casein micelle SAXS data. *Soft Matter* 12, 6937–6953. doi: 10.1039/c6sm01091a
- Inouye, H., Zhang, Y., Yang, L., Venugopalan, N., Fischetti, R. F., Gleber, S. C., et al. (2014). Multiscale deconstruction of molecular architecture in corn stover. *Sci. Rep.* 4:3756. doi: 10.1038/srep03756
- Ioelovich, M. (2016). Physicochemical methods for determination of cellulose crystallinity. *ChemXpress* 9, 245–251.
- Isogai, A., Usuda, M., Kato, T., Uryu, T., and Atalla, R. H. (1989). Solid-state CP/MAS carbon-13 NMR study of cellulose polymorphs. *Macromolecules* 22, 3168–3172. doi: 10.1021/ma00197a045
- Jakob, H. F., Fengel, D., Tschegg, S. E., and Fratzl, P. (1995). The elementary cellulose fibril in picea abies: comparison of transmission electron microscopy, small-angle X-ray scattering, and wide-angle X-ray scattering results. *Macromolecules* 28, 8782–8787. doi: 10.1021/ma00130a010
- Jakob, H. F., Fratzl, P., and Tschegg, S. E. (1994). Size and arrangement of elementary cellulose fibrils in wood cells: a small-angle X-ray scattering study of *Picea abies*. *J. Struct. Biol.* 113, 13–22. doi: 10.1006/jsbi.1994.1028
- Jakob, H. F., Tschegg, S. E., and Fratzl, P. (1996). Hydration dependence of the wood-cell wall structure in *Picea abies*. A small-angle X-ray scattering study. *Macromolecules* 29, 8435–8440. doi: 10.1021/ma9605661
- Johnson, A., and Vert, G. (2017). Single event resolution of plant plasma membrane protein endocytosis by tirf microscopy. *Front. Plant Sci.* 8:612. doi: 10.3389/fpls.2017.00612
- Jones, P. D., Schimleck, L. R., Peter, G. F., Daniels, R. F., and Clark, A. (2005). Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites in Georgia. *Can. J. Forest Res.* 35, 85–92. doi: 10.1139/x04-160
- Jordan, B. M., and Dumais, J. (2010). *Biomechanics of Plant Cell Growth*. *Encyclopedia of Life Sciences*. Hoboken, NJ: Wiley.

- Ju, X., Bowden, M., Brown, E. E., and Zhang, X. (2015). An improved X-ray diffraction method for cellulose crystallinity measurement. *Carbohydr. Polym.* 123, 476–481. doi: 10.1016/j.carbpol.2014.12.071
- Jungnikl, K., Paris, O., Fratzl, P., and Burgert, I. (2007). The implication of chemical extraction treatments on the cell wall nanostructure of softwood. *Cellulose* 15:407. doi: 10.1007/s10570-007-9181-5
- Kafle, K., Xi, X., Lee, C. M., Tittmann, B. R., Cosgrove, D. J., Park, Y. B., et al. (2014). Cellulose microfibril orientation in onion (*Allium cepa* L.) epidermis studied by atomic force microscopy (AFM) and vibrational sum frequency generation (SFG) spectroscopy. *Cellulose* 21, 1075–1086. doi: 10.1007/s10570-013-0121-2
- Kantola, M., and Kähkönen, H. (1963). Small-angle X-ray investigation of the orientation of crystallites in Finnish coniferous and deciduous wood fibres. *Ann. Acad. Sci. Fenn.* 137:14.
- Kantola, M., and Seitonen, S. (1961). X-ray orientation investigations on finish conifers. *Ann. Acad. Sci. Fenn.* 80, 1–15.
- Karimi, K., and Taherzadeh, M. J. (2016). A critical review of analytical methods in pretreatment of lignocelluloses: composition, imaging, and crystallinity. *Bioresour. Technol.* 200, 1008–1018. doi: 10.1016/j.biortech.2015.11.022
- Kawakubo, T., Karita, S., Araki, Y., Watanabe, S., Oyadomari, M., Takada, R., et al. (2009). Analysis of exposed cellulose surfaces in pretreated wood biomass using carbohydrate-binding module (CBM)–cyan fluorescent protein (CFP). *Biotechnol. Bioeng.* 105, 499–508. doi: 10.1002/bit.22550
- Kennedy, C. J., Cameron, G. J., Šturcová, A., Apperley, D. C., Altaner, C., Wess, T. J., et al. (2007a). Microfibril diameter in celery collenchyma cellulose: X-ray scattering and NMR evidence. *Cellulose* 14, 235–246. doi: 10.1007/s10570-007-9116-1
- Kennedy, C. J., Šturcová, A., Jarvis, M. C., and Wess, T. J. (2007b). Hydration effects on spacing of primary-wall cellulose microfibrils: a small angle X-ray scattering study. *Cellulose* 14, 401–408. doi: 10.1007/s10570-007-9129-9
- Kent, M. S., Cheng, G., Murton, J. K., Carles, E. L., Dibble, D. C., Zendejas, F., et al. (2010). Study of enzymatic digestion of cellulose by small angle neutron scattering. *Biomacromolecules* 11, 357–368. doi: 10.1021/bm9008952
- Kim, J. S., Awano, T., Yoshinaga, A., and Takabe, K. (2012). Ultrastructure of the innermost surface of differentiating normal and compression wood tracheids as revealed by field emission scanning electron microscopy. *Planta* 235, 1209–1219. doi: 10.1007/s00425-011-1569-7
- Kim, S. H., Lee, C. M., and Kafle, K. (2013). Characterization of crystalline cellulose in biomass: basic principles, applications, and limitations of XRD, NMR, IR, Raman, and SFG. *Korean J. Chem. Eng.* 30, 2127–2141. doi: 10.1007/s11814-013-0162-0
- Kirby, A. R., Gunning, A. P., Waldron, K. W., Morris, V. J., and Ng, A. (1996). Visualization of plant cell walls by atomic force microscopy. *Biophys. J.* 70, 1138–1143. doi: 10.1016/S0006-3495(96)79708-4
- Kirby, A. R., Ng, A., Waldron, K. W., and Morris, V. J. (2006). AFM investigations of cellulose fibers in bintje potato (*Solanum tuberosum* L.) cell wall fragments. *Food Biophys.* 1, 163–167. doi: 10.1007/s11483-006-9013-4
- Kljun, A., Benians, T. A. S., Goubet, F., Meulewaeter, F., Knox, J. P., and Blackburn, R. S. (2011). Comparative analysis of crystallinity changes in cellulose I polymers using ATR-FTIR, X-ray diffraction, and carbohydrate-binding module probes. *Biomacromolecules* 12, 4121–4126. doi: 10.1021/bm201176m
- Koyama, M., Helbert, W., Imai, T., Sugiyama, J., and Henrissat, B. (1997). Parallel-structure evidences the molecular directionality during biosynthesis of bacterial cellulose. *Proc. Natl. Acad. Sci. U.S.A.* 94, 9091–9095. doi: 10.1073/pnas.94.17.9091
- Kraissig, H. A. (1992). “Fibrillar Structure,” in *Cellulose: Structure, Accessibility and Reactivity*, ed. M. B. Huglin (Yverdon -les-Bains: Gordon and Breach Science Publishers).
- Kubicki, J. D., Yang, H., Sawada, D., O'Neill, H., Oehme, D., and Cosgrove, D. (2018). The shape of native plant cellulose microfibrils. *Sci. Rep.* 8:13983. doi: 10.1038/s41598-018-32211-w
- Kuga, S., and Brown, R. M. Jr. (1987a). Lattice imaging of ramie cellulose. *Polym. Commun.* 28, 311–314.
- Kuga, S., and Brown, R. M. (1987b). Practical aspects of lattice imaging of cellulose. *J. Electr. Microsc. Techn.* 6, 349–356. doi: 10.1002/jemt.1060060405
- Kumar, M., and Turner, S. (2015). Protocol: a medium-throughput method for determination of cellulose content from single stem pieces of *Arabidopsis thaliana*. *Plant Methods* 11:46. doi: 10.1186/s13007-015-0090-6
- Lawrence, J. R., Swerhone, G. D. W., Leppard, G. G., Araki, T., Zhang, X., West, M. M., et al. (2003). Scanning transmission X-ray, laser scanning, and transmission electron microscopy mapping of the exopolymeric matrix of microbial biofilms. *Appl. Environ. Microbiol.* 69, 5543–5554. doi: 10.1128/AEM.69.9.5543-5554.2003
- Lee, C. M., Kafle, K., Belias, D. W., Park, Y. B., Glick, R. E., Haigler, C. H., et al. (2015a). Comprehensive analysis of cellulose content, crystallinity, and lateral packing in *Gossypium hirsutum* and *Gossypium barbadense* cotton fibers using sum frequency generation, infrared and Raman spectroscopy, and X-ray diffraction. *Cellulose* 22, 971–989. doi: 10.1007/s10570-014-0535-5
- Lee, C. M., Kubicki, J. D., Fan, B., Zhong, L., Jarvis, M. C., and Kim, S. H. (2015b). Hydrogen-bonding network and OH stretch vibration of cellulose: comparison of computational modeling with polarized IR and SFG Spectra. *J. Phys. Chem. B* 119, 15138–15149. doi: 10.1021/acs.jpcc.5b08015
- Lee, C. M., Kafle, K., Park, Y. B., and Kim, S. H. (2014). Probing crystal structure and mesoscale assembly of cellulose microfibrils in plant cell walls, tunicate tests, and bacterial films using vibrational Sum Frequency Generation (SFG) spectroscopy. *Phys. Chem. Chem. Phys.* 16, 10844–10853. doi: 10.1039/c4cp00515e
- Lee, C. M., Mittal, A., Barnette, A. L., Kafle, K., Park, Y. B., Shin, H., et al. (2013). Cellulose polymorphism study with sum-frequency-generation (SFG) vibration spectroscopy: identification of exocyclic CH₂OH conformation and chain orientation. *Cellulose* 20, 991–1000. doi: 10.1007/s10570-013-9917-3
- Levi, M. P., and Sellen, D. B. (1967). The degree-of-polymerization and polydispersity of cellulose in beechwood, and its radial variation across fibre cell-walls. *Carbohydr. Res.* 5, 351–355. doi: 10.1016/S0008-6215(00)80513-3
- Lichtenegger, H., Muller, M., Paris, O., Riekel, C., and Fratzl, P. (1999). Imaging of the helical arrangement of cellulose fibrils in wood by synchrotron X-ray microdiffraction. *J. Appl. Crystallogr.* 32, 1127–1133. doi: 10.1107/S0021889899010961
- Lionetto, F., Del Sole, R., Cannolella, D., Vasapollo, G., and Maffezzoli, A. (2012). Monitoring wood degradation during weathering by cellulose crystallinity. *Materials* 5, 1910–1922. doi: 10.3390/ma5101910
- Liu, F., Brady, M. A., and Wang, C. (2016). Resonant soft X-ray scattering for polymer materials. *Eur. Polym. J.* 81, 555–568. doi: 10.1016/j.eurpolymj.2016.04.014
- Maeva, E., Severin, F., Miyasaka, C., Tittmann, B. R., and Maev, R. G. (2009). Acoustic imaging of thick biological tissue. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 56, 1352–1358. doi: 10.1109/TUFFC.2009.1191
- Magazù, S., and Migliardo, F. (2011). *Dynamics of Biological Macromolecules by Neutron Scattering*. Emirate of Sharjah: Bentham Science Publishers.
- Makarem, M., Sawada, D., O'Neill, H. M., Lee, C. M., Kafle, K., Park, Y. B., et al. (2017). Dependence of Sum Frequency Generation (SFG) spectral features on the mesoscale arrangement of SFG-active crystalline domains interspersed in SFG-inactive matrix: a case study with cellulose in uniaxially aligned control samples and alkali-treated secondary cell walls of plants. *J. Phys. Chem. C* 121, 10249–10257. doi: 10.1021/acs.jpcc.7b02770
- Mancosky, D. G., Lucia, L. A., Nanko, H., Wirrick, S., Rudie, A. W., and Braun, R. (2005). Novel visualization studies of lignocellulosic oxidation chemistry by application of C-near edge X-ray absorption fine structure spectroscopy. *Cellulose* 12, 35–41. doi: 10.1023/B:CELL.0000049352.60007.76
- Manley, R. S. J. (2003). Molecular morphology of cellulose. *J. Polym. Sci. Part A-2 Polym. Phys.* 9, 1025–1059. doi: 10.1002/pol.1971.160090604
- Mansfield, S. D., and Meder, R. (2003). Cellulose hydrolysis – the role of monocomponent cellulases in crystalline cellulose degradation. *Cellulose* 10, 159–169. doi: 10.1023/A:1024022710366
- Marga, F., Grandbois, M., Cosgrove, D. J., and Baskin, T. I. (2005). Cell wall extension results in the coordinate separation of parallel microfibrils: evidence from scanning electron microscopy and atomic force microscopy. *Plant J.* 43, 181–190. doi: 10.1111/j.1365-3113.2005.02447.x
- Mark, R. E. (1967). *Cell Wall Mechanics of Tracheids*. New Haven, CT: Yale University Press.
- Martínez-Sanz, M., Gidley, M. J., and Gilbert, E. P. (2015a). Application of X-ray and neutron small angle scattering techniques to study the hierarchical structure of plant cell walls: a review. *Carbohydr. Polym.* 125, 120–134. doi: 10.1016/j.carbpol.2015.02.010

- Martínez-Sanz, M., Lopez-Sanchez, P., Gidley, M. J., and Gilbert, E. P. (2015b). Evidence for differential interaction mechanism of plant cell wall matrix polysaccharides in hierarchically-structured bacterial cellulose. *Cellulose* 22, 1541–1563. doi: 10.1007/s10570-015-0614-2
- Matias, V. R. F., and Beveridge, T. J. (2006). Native cell wall organization shown by cryo-electron microscopy confirms the existence of a periplasmic space in *Staphylococcus aureus*. *J. Bacteriol.* 188, 1011–1021. doi: 10.1128/JB.188.3.1011-1021.2006
- Mattheyses, A. L., Simon, S. M., and Rappoport, J. Z. (2010). Imaging with total internal reflection fluorescence microscopy for the cell biologist. *J. Cell Sci.* 123, 3621–3628. doi: 10.1242/jcs.056218
- McCann, M. C., Wells, B., and Roberts, K. (1990). Direct visualization of cross-links in the primary plant cell wall. *J. Cell Sci.* 96, 323–334.
- McIntosh, A. I., Yang, B., Goldup, S. M., Watkinson, M., and Donnan, R. S. (2012). Terahertz spectroscopy: a powerful new tool for the chemical sciences? *Chem Soc. Rev.* 41, 2072–2082. doi: 10.1039/c1cs15277g
- Mühlethaler, K. (1950). Electron microscopy of developing plant cell walls. *Biochim. Biophys. Acta* 5, 1–9. doi: 10.1016/0006-3002(50)90143-0
- Müller, M., Cziha, C., Vogl, G., Fratzl, P., Schober, H., and Riek, C. (1998). Direct observation of microfibril arrangement in a single native cellulose fiber by microbeam small-angle X-ray scattering. *Macromolecules* 31, 3953–3957. doi: 10.1021/ma980004c
- Müller, M., Hori, R., Itoh, T., and Sugiyama, J. (2002). X-ray microbeam and electron diffraction experiments on developing xylem cell walls. *Biomacromolecules* 3, 182–186. doi: 10.1021/bm015605h
- Müller, M., Riek, C., Vuong, R., and Chanzy, H. (2000). Skin/core microstructure in viscose rayon fibres analysed by X-ray microbeam and electron diffraction mapping. *Polymer* 41, 2627–2632. doi: 10.1016/S0032-3861(99)00433-4
- Nelmes, B. J., and Preston, R. D. (1968). Cellulose microfibril orientation in rubbery wood. *J. Exp. Bot.* 19, 519–525. doi: 10.1093/jxb/19.3.519
- Nelson, M. L., and O'Connor, R. T. (1964). Relation of certain infrared bands to cellulose crystallinity and crystal lattice type. Part II. A new infrared ratio for estimation of crystallinity in celluloses I and II. *J. Appl. Polym. Sci.* 8, 1325–1341. doi: 10.1002/app.1964.070080323
- Newman, R. H. (1999). Estimation of the lateral dimensions of cellulose crystallites using ¹³C NMR signal strengths. *Solid State Nuclear Magn. Reson.* 15, 21–29. doi: 10.1016/S0926-2040(99)00043-0
- Newman, R. H., and Hemmingson, J. A. (1990). Determination of the degree of cellulose crystallinity in wood by carbon-13 nuclear magnetic resonance spectroscopy. *Holzforschung* 44, 351–355. doi: 10.1515/hfsg.1990.44.5.351
- Newman, R. H., Hemmingson, J. A., and Suckling, I. D. (1993). Carbon-13 nuclear magnetic resonance studies of kraft pulping. *Holzforschung* 47, 234–238. doi: 10.1515/hfsg.1993.47.3.234
- Newman, R. H., Hill, S. J., and Harris, P. J. (2013). Wide-Angle X-Ray scattering and solid-state nuclear magnetic resonance data combined to test models for cellulose microfibrils in mung bean cell walls. *Plant Physiol.* 163, 1558–1567. doi: 10.1104/pp.113.228262
- Nieduszynski, I., and Preston, R. D. (1970). Crystallite size in natural cellulose. *Nature* 225, 273–274. doi: 10.1038/225273a0
- Nishiguchi, D., Nagai, K. H., Chaté, H., and Sano, M. (2017). Long-range nematic order and anomalous fluctuations in suspensions of swimming filamentous bacteria. *Phys. Rev. E* 95:020601. doi: 10.1103/PhysRevE.95.020601
- Nishikawa, S., and Ono, S. (1913). Transmission of X-rays through fibrous, lamellar and granular substances. *Proc. Tokyo Math. Phys. Soc.* 7, 131–138.
- Nishiyama, Y., Langan, P., and Chanzy, H. (2002). Crystal structure and hydrogen-bonding system in cellulose I β from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* 124, 9074–9082. doi: 10.1021/ja0257319
- Nishiyama, Y., Sugiyama, J., Chanzy, H., and Langan, P. (2003). Crystal structure and hydrogen bonding system in cellulose Ia from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* 125, 14300–14306. doi: 10.1021/ja037055w
- Nixon, B. T., Mansouri, K., Singh, A., Du, J., Davis, J. K., Lee, J.-G., et al. (2016). Comparative structural and computational analysis supports eighteen cellulose synthases in the plant cellulose synthesis complex. *Sci. Rep.* 6:28696. doi: 10.1038/srep28696
- Noori, M. S., and Karimi, K. (2016). Chemical and structural analysis of alkali pretreated pinewood for efficient ethanol production. *RSC Adv.* 6, 65683–65690. doi: 10.1039/C6RA11486E
- O'Connor, R. T., DuPré, E. F., and Mitcham, D. (1958). Applications of infrared absorption spectroscopy to investigations of cotton and modified cottons: part I: physical and crystalline modifications and oxidation. *Textile Res. J.* 28, 382–392. doi: 10.1177/004051755802800503
- O'Neill, H., Pingali, S. V., Petridis, L., He, J., Mamontov, E., Hong, L., et al. (2017). Dynamics of water bound to crystalline cellulose. *Sci. Rep.* 7:11840. doi: 10.1038/s41598-017-12035-w
- O'Sullivan, A. C. (1997). Cellulose: the structure slowly unravels. *Cellulose* 4, 173–207. doi: 10.1023/A:1018431705579
- Page, D. H. (1969). A method for determining the fibrillar angle in wood tracheids. *J. Microsc.* 90, 137–143. doi: 10.1111/j.1365-2818.1969.tb00701.x
- Park, S., Baker, J. O., Himmel, M. E., Parilla, P. A., and Johnson, D. K. (2010). Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol. Biofuels* 3:10. doi: 10.1186/1754-6834-3-10
- Penttilä, P. A., Várnai, A., Fernández, M., Kontro, I., Liljeström, V., Lindner, P., et al. (2013). Small-angle scattering study of structural changes in the microfibril network of nanocellulose during enzymatic hydrolysis. *Cellulose* 20, 1031–1040. doi: 10.1007/s10570-013-9899-1
- Penttilä, P. A., Várnai, A., Leppänen, K., Peura, M., Kallonen, A., Jääskeläinen, P., et al. (2010). Changes in submicrometer structure of enzymatically hydrolyzed microcrystalline cellulose. *Biomacromolecules* 11, 1111–1117. doi: 10.1021/bm1001119
- Peura, M., Müller, M., Serimaa, R., Vainio, U., Sarén, M.-P., Saranpää, P., et al. (2005). Structural studies of single wood cell walls by synchrotron X-ray microdiffraction and polarized light microscopy. *Nuclear Instrum. Methods Phys. Res. Sec. B* 238, 16–20. doi: 10.1016/j.nimb.2005.06.011
- Peura, M., Müller, M., Vainio, U., Sarén, M.-P., Saranpää, P., and Serimaa, R. (2007). X-ray microdiffraction reveals the orientation of cellulose microfibrils and the size of cellulose crystallites in single Norway spruce tracheids. *Trees* 22:49. doi: 10.1007/s00468-007-0168-5
- Phyo, P., Wang, T., Xiao, C., Anderson, C. T., and Hong, M. (2017). Effects of pectin molecular weight changes on the structure, dynamics, and polysaccharide interactions of primary cell walls of *Arabidopsis thaliana*: insights from solid-state NMR. *Biomacromolecules* 18, 2937–2950. doi: 10.1021/acs.biomac.7b00888
- Pingali, S. V., O'Neill, H. M., Nishiyama, Y., He, L., Melnichenko, Y. B., Urban, V., et al. (2014). Morphological changes in the cellulose and lignin components of biomass occur at different stages during steam pretreatment. *Cellulose* 21, 873–878. doi: 10.1007/s10570-013-0162-6
- Pingali, S. V., Urban, V. S., Heller, W. T., McGaughey, J., O'Neill, H., Foston, M., et al. (2010). Breakdown of cell wall nanostructure in dilute acid pretreated biomass. *Biomacromolecules* 11, 2329–2335. doi: 10.1021/bm100455h
- Poletto, M., Ornaghi, H. L., and Zattera, A. J. (2014). Native cellulose: structure, characterization and thermal properties. *Materials* 7, 6105–6119. doi: 10.3390/ma7096105
- Porter Stephanie, E., Donohoe Bryon, S., Beery Kyle, E., Xu, Q., Ding, S. Y., Vinzant Todd, B., et al. (2007). Microscopic analysis of corn fiber using starch- and cellulose-specific molecular probes. *Biotechnol. Bioeng.* 98, 123–131. doi: 10.1002/bit.21409
- Prater, C. B., Butt, H. J., and Hansma, P. K. (1990). Atomic force microscopy. *Nature* 345:839. doi: 10.1038/345839a0
- Preston, R. D. (1934). The organization of the cell wall of the conifer tracheid. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 224, 131–174.
- Reis, D., Vian, B., Chanzy, H., and Roland, J.-C. (1991). Liquid crystal-type assembly of native cellulose-glucuronoxylans extracted from plant cell wall. *Biol. Cell* 73, 173–178. doi: 10.1016/0248-4900(91)90100-2
- Reiterer, A., Jakob, H. F., Stanzl-Tschegg, S. E., and Fratzl, P. (1998). Spiral angle of elementary cellulose fibrils in cell walls of *Picea abies* determined by small-angle x-ray scattering. *Wood Sci. Technol.* 32, 335–345. doi: 10.1007/BF00702790
- Revol, J. F. (1982). On the cross-sectional shape of cellulose crystallites in *Valonia ventricosa*. *Carbohydr. Polym.* 2, 123–134. doi: 10.1016/0144-8617(82)90058-3

- Riekkel, C. (2000). New avenues in x-ray microbeam experiments. *Rep. Prog. Phys.* 63, 233–262. doi: 10.1088/0034-4885/63/3/201
- Rivnay, J., Mannsfeld, S. C. B., Miller, C. E., Salleo, A., and Toney, M. F. (2012). Quantitative determination of organic semiconductor microstructure from the molecular to device scale. *Chem. Rev.* 112, 5488–5519. doi: 10.1021/cr3001109
- Rondeau-Mouro, C., Bouchet, B., Pontoire, B., Robert, P., Mazoyer, J., and Buléon, A. (2003). Structural features and potential texturising properties of lemon and maize cellulose microfibrils. *Carbohydr. Polym.* 53, 241–252. doi: 10.1016/S0144-8617(03)00069-9
- Ruland, W. (1961). X-ray determination of crystallinity and diffuse disorder scattering. *Acta Crystallogr.* 14, 1180–1185. doi: 10.1107/S0365110X61003429
- Sahlberg, U., Salmén, L., and Oscarsson, A. (1997). The fibrillar orientation in the S2-layer of wood fibres as determined by X-ray diffraction analysis. *Wood Sci. Technol.* 31, 77–86. doi: 10.1007/BF00705923
- Sandquist, D., Filonova, L., Schantz, L., Ohlin, M., and Daniel, G. (2010). Microdistribution of xyloglucan in differentiating poplar cells. *Bioresources* 5, 796–807.
- Sant'Anna, C., and de Souza, W. L. (2012). "Microscopy as a tool to follow deconstruction of lignocellulosic biomass," in *Current Microscopy Contributions to Advanced in Science and Technology*, ed. A. Méndez-Vilas (Badajoz: Formatex), 639–645.
- Sarkar, P., Bosneaga, E., and Auer, M. (2009). Plant cell walls throughout evolution: towards a molecular understanding of their design principles. *J. Exp. Bot.* 60, 3615–3635. doi: 10.1093/jxb/erp245
- Sarkar, P., Bosneaga, E., Yap, E. G. Jr., Das, J., Tsai, W.-T., Cabal, A., et al. (2014). Electron tomography of cryo-immobilized plant tissue: a novel approach to studying 3D macromolecular architecture of mature plant cell walls in situ. *PLoS One* 9:e106928. doi: 10.1371/journal.pone.0106928
- Saxe, F., Eder, M., Benecke, G., Aichmayer, B., Fratzl, P., Burgert, I., et al. (2014). Measuring the distribution of cellulose microfibril angles in primary cell walls by small angle X-ray scattering. *Plant Methods* 10:25. doi: 10.1186/1746-4811-10-25
- Scallan, A. M. (1971). A quantitative picture of the fringed micellar model of cellulose. *Textile Res. J.* 41, 647–653. doi: 10.1177/004051757104100802
- Schenzel, K., Fischer, S., and Brendler, E. (2005). New method for determining the degree of cellulose I crystallinity by means of FT raman spectroscopy. *Cellulose* 12, 223–231. doi: 10.1007/s10570-004-3885-6
- Schimleck, L. R., Evans, R., Ilic, J., and Matheson, A. C. (2002). Estimation of wood stiffness of increment cores by near-infrared spectroscopy. *Can. J. Forest Res.* 32, 129–135. doi: 10.1139/x01-176
- Schimleck, L. R., Kube, P. D., Raymond, C. A., Michell, A. J., and French, J. (2005). Estimation of whole-tree kraft pulp yield of *Eucalyptus nitens* using near-infrared spectra collected from increment cores. *Can. J. Forest Res.* 35, 2797–2805. doi: 10.1139/x05-193
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089
- Scott, T. A., and Melvin, E. H. (1953). Determination of dextran with anthrone. *Anal. Chem.* 25, 1656–1661. doi: 10.1021/ac60083a023
- Segal, L., Creely, J. J., Martin, A. E., and Conrad, C. M. (1959). An empirical method for estimating the degree of crystallinity of native cellulose using the X-Ray diffractometer. *Textile Res. J.* 29, 786–794. doi: 10.1177/004051755902901003
- Senft, J. F., and Bendtsen, B. A. (1985). Measuring microfibrillar angles using light microscopy. *Wood Fibre Sci.* 17:4.
- Simmons, T. J., Mortimer, J. C., Bernardinelli, O. D., Pöppler, A.-C., Brown, S. P., deAzevedo, E. R., et al. (2016). Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* 7:13902. doi: 10.1038/ncomms13902
- Široký, J., Benians, T. A. S., Russell, S. J., Bechtold, T., Paul Knox, J., and Blackburn, R. S. (2012). Analysis of crystallinity changes in cellulose II polymers using carbohydrate-binding modules. *Carbohydr. Polym.* 89, 213–221. doi: 10.1016/j.carbpol.2012.02.073
- Sponsler, O. L. (1928). The molecular structure of the cell wall of fibers. A summary of X-ray investigations. *Am. J. Bot.* 15, 525–536. doi: 10.1002/j.1537-2197.1928.tb09447.x
- Strachan, C. J., Taday, P. F., Newnham, D. A., Gordon, K. C., Zeitler, J. A., Pepper, M., et al. (2005). Using terahertz pulsed spectroscopy to quantify pharmaceutical polymorphism and crystallinity. *J. Pharm. Sci.* 94, 837–846. doi: 10.1002/jps.20281
- Sugiyama, J., Harada, H., Fujiyoshi, Y., and Uyeda, N. (1985). Lattice images from ultrathin sections of cellulose microfibrils in the cell wall of *Valonia macrophysa* Kütz. *Planta* 166, 161–168. doi: 10.1007/BF00397343
- Sugiyama, J., Persson, J., and Chanzy, H. (1991a). Combined infrared and electron diffraction study of the polymorphism of native celluloses. *Macromolecules* 24, 2461–2466. doi: 10.1021/ma00009a050
- Sugiyama, J., Vuong, R., and Chanzy, H. (1991b). Electron diffraction study on the two crystalline phases occurring in native cellulose from an algal cell wall. *Macromolecules* 24, 4168–4175. doi: 10.1021/ma00014a033
- Suzuki, H., and Kamiyama, T. (2004). Structure of cellulose microfibrils and the hydration effect in *Cryptomeria japonica*: a small-angle X-ray scattering study. *J. Wood Sci.* 50, 351–357. doi: 10.1007/s10086-003-0567-1
- Svedström, K., Bjurhager, I., Kallonen, A., Peura, M., and Serimaa, R. (2012). "Structure of oak wood from the Swedish warship Vasa revealed by X-ray scattering and microtomography. *Holzforschung* 66, 355–363. doi: 10.1515/hf.2011.157
- Thimm, J. C., Burritt, D. J., Ducker, W. A., and Melton, L. D. (2000). Celery (*Apium graveolens* L.) parenchyma cell walls examined by atomic force microscopy: effect of dehydration on cellulose microfibrils. *Planta* 212, 25–32. doi: 10.1007/s004250000359
- Thomas, L. H., Forsyth, V. T., Šturcová, A., Kennedy, C. J., May, R. P., Altaner, C. M., et al. (2013). Structure of cellulose microfibrils in primary cell walls from collenchyma. *Plant Physiol.* 161, 465–476. doi: 10.1104/pp.112.206359
- Thomson, C. I., Lowe, R. M., and Ragauskas, A. J. (2007). Imaging cellulose fibre interfaces with fluorescence microscopy and resonance energy transfer. *Carbohydr. Polym.* 69, 799–804. doi: 10.1016/j.carbpol.2007.01.023
- Thygesen, A., Oddershede, J., Lilholt, H., Thomsen, A. B., and Ståhl, K. (2005). On the determination of crystallinity and cellulose content in plant fibres. *Cellulose* 12:563. doi: 10.1007/s10570-005-9001-8
- Tittmann, B., and Xi, X. (2014). "Imaging and quantitative data acquisition of biological cell walls with atomic force microscopy and scanning acoustic microscopy," in *Microscopy: Advances in Scientific Research and Education*, ed. A. Méndez-Vilas (Badajoz: Formatex Research Center), 161–172.
- Tokoh, C., Takabe, K., Fujita, M., and Saiki, H. (1998). Cellulose synthesized by *Acetobacter xylinum* in the presence of acetyl glucosaminan. *Cellulose* 5, 249–261. doi: 10.1023/A:1009211927183
- Tokoh, C., Takabe, K. J., and Fujita, M. (2002). Cellulose synthesized by *Acetobacter xylinum* in the presence of plant cell wall polysaccharides. *Cellulose* 9, 65–74. doi: 10.1023/A:1015827121927
- Updegraff, D. M. (1969). Semimicro determination of cellulose in biological materials. *Anal. Biochem.* 32, 420–424. doi: 10.1016/S0003-2697(69)80099-6
- Usov, I., and Mezzenga, R. (2015). FiberApp: an open-source software for tracking and analyzing polymers, filaments, biomacromolecules, and fibrous objects. *Macromolecules* 48, 1269–1280. doi: 10.1021/ma502264c
- van der Wel, N. N., Putman, C. A. J., van Noort, S. J. T., de Grooth, B. G., and Emons, A. M. C. (1996). Atomic force microscopy of pollen grains, cellulose microfibrils, and protoplasts. *Protoplasma* 194, 29–39. doi: 10.1007/BF01273165
- Vandavasi, V. G., Putnam, D. K., Zhang, Q., Petridis, L., Heller, W. T., Nixon, B. T., et al. (2016). A Structural Study of CESA1 catalytic domain of arabidopsis cellulose synthesis complex: evidence for CESA trimers. *Plant Physiol.* 170, 123–135. doi: 10.1104/pp.15.01356
- VanderHart, D. L., and Atalla, R. H. (1984). Studies of microstructure in native celluloses using solid-state carbon-13 NMR. *Macromolecules* 17, 1465–1472. doi: 10.1021/ma00138a009
- Vieira, F. S., and Pasquini, C. (2014). Determination of cellulose crystallinity by terahertz-time domain spectroscopy. *Anal. Chem.* 86, 3780–3786. doi: 10.1021/ac4035746
- Virgili, J. M., Tao, Y. F., Kortright, J. B., Balsara, N. P., and Segalman, R. A. (2007). Analysis of order formation in block copolymer thin films using resonant soft X-ray scattering. *Macromolecules* 40, 2092–2099. doi: 10.1021/ma061734k

- Wada, M., Okano, T., and Sugiyama, J. (1997). Synchrotron-radiated X-ray and neutron diffraction study of native cellulose. *Cellulose* 4, 221–232. doi: 10.1023/A:1018435806488
- Wang, J., Hitchcock, A. P., Karunakaran, C., Prange, A., Franz, B., Harkness, T., et al. (2011). 3D chemical and elemental imaging by STXM spectromicroscopy. *AIP Conf. Proc.* 1365, 215–218. doi: 10.1063/1.3625342
- Wang, T., and Hong, M. (2016). Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *J. Exp. Bot.* 67, 503–514. doi: 10.1093/jxb/erv416
- Wang, T., Zabolina, O., and Hong, M. (2012). Pectin–cellulose interactions in the arabidopsis primary cell wall from two-dimensional magic-angle-spinning solid-state nuclear magnetic resonance. *Biochemistry* 51, 9846–9856. doi: 10.1021/bi3015532
- Wardrop, A. B., and Preston, R. D. (1947). Organisation of the cell walls of tracheids and wood fibres. *Nature* 160:911. doi: 10.1038/160911a0
- Warwick, T., Franck, K., Kortright, J. B., Meigs, G., Moronne, M., Myneni, S., et al. (1998). A scanning transmission x-ray microscope for materials science spectromicroscopy at the advanced light source. *Rev. Sci. Instr.* 69, 2964–2973. doi: 10.1063/1.1149041
- Westphal, V., Lauterbach, M. A., Nicola, A. D., and Hell, S. W. (2007). Dynamic far-field fluorescence nanoscopy. *New J. Phys.* 9:435. doi: 10.1101/pdb.prot069260
- White, P. B., Wang, T., Park, Y. B., Cosgrove, D. J., and Hong, M. (2014). Water-polysaccharide interactions in the primary cell wall of arabidopsis thaliana from polarization transfer solid-state NMR. *J. Am. Chem. Soc.* 136, 10399–10409. doi: 10.1021/ja504108h
- Xiao, C., and Anderson, C. T. (2016). Interconnections between cell wall polymers, wall mechanics, and cortical microtubules: teasing out causes and consequences. *Plant Signal. Behav.* 11:e1215396. doi: 10.1080/15592324.2016.1215396
- Xiao, C., Zhang, T., Zheng, Y., Cosgrove, D. J., and Anderson, C. T. (2016). Xyloglucan deficiency disrupts microtubule stability and cellulose biosynthesis in Arabidopsis, altering cell growth and morphogenesis. *Plant Physiol.* 170, 234–249. doi: 10.1104/pp.15.01395
- Xu, T., Vavylonis, D., Tsai, F.-C., Koenderink, G. H., Nie, W., Yusuf, E., et al. (2015). SOAX: a software for quantification of 3D biopolymer networks. *Sci. Rep.* 5:9081. doi: 10.1038/srep09081
- Yamamoto, H., and Horii, F. (1993). CP/MAS carbon-13 NMR analysis of the crystal transformation induced for Valonia cellulose by annealing at high temperatures. *Macromolecules* 26, 1313–1317. doi: 10.1021/ma00058a020
- Yamamoto, H., Horii, F., and Hirai, A. (1996). In situ crystallization of bacterial cellulose II. Influences of different polymeric additives on the formation of celluloses I α and I β at the early stage of incubation. *Cellulose* 3, 229–242. doi: 10.1007/BF02228804
- Yamamoto, H., and Horn, F. (1994). In Situ crystallization of bacterial cellulose I. Influences of polymeric additives, stirring and temperature on the formation celluloses I α and I β as revealed by cross polarization/magic angle spinning (CP/MAS)13C NMR spectroscopy. *Cellulose* 1, 57–66. doi: 10.1007/BF00818798
- Yamamoto, H., Okuyama, T., and Yoshida, M. (1993). Method of determining the mean microfibril angle of wood over a wide range by the improved Cave's method. *Mokuzai Gakkaishi* 39, 375–381.
- Yarbrough, J. M., Himmel, M. E., and Ding, S.-Y. (2009). Plant cell wall characterization using scanning probe microscopy techniques. *Biotechnol. Biofuels* 2:17. doi: 10.1186/1754-6834-2-17
- Ye, D., Kiemle, S. N., Rongpipi, S., Wang, X., Wang, C., Cosgrove, D. J., et al. (2018a). Resonant soft X-ray scattering reveals cellulose microfibril spacing in plant primary cell walls. *Sci. Rep.* 8:12449. doi: 10.1038/s41598-018-31024-1
- Ye, D., Le, T. P., Kuei, B., Zhu, C., Zwart, P. H., Wang, C., et al. (2018b). Resonant soft X-Ray scattering provides protein structure with chemical specificity. *Structure* 26, 1513.e3–1521.e3. doi: 10.1016/j.str.2018.07.018
- Zhang, T., Mahgoudy-Louyeh, S., Tittmann, B., and Cosgrove, D. J. (2014). Visualization of the nanoscale pattern of recently-deposited cellulose microfibrils and matrix materials in never-dried primary walls of the onion epidermis. *Cellulose* 21, 853–862. doi: 10.1007/s10570-013-9996-1
- Zhang, T., Vavylonis, D., Durachko, D. M., and Cosgrove, D. J. (2017). Nanoscale movements of cellulose microfibrils in primary cell walls. *Nat. Plants* 3:17056. doi: 10.1038/nplants.2017.56
- Zhang, T., Zheng, Y., and Cosgrove, D. J. (2016). Spatial organization of cellulose microfibrils and matrix polysaccharides in primary plant cell walls as imaged by multichannel atomic force microscopy. *Plant J.* 85, 179–192. doi: 10.1111/tpj.13102
- Zhang, Y., Inouye, H., Yang, L., Himmel, M. E., Tucker, M., and Makowski, L. (2015). Breakdown of hierarchical architecture in cellulose during dilute acid pretreatments. *Cellulose* 22, 1495–1504. doi: 10.1007/s10570-015-0592-4
- Zheng, Y., Wang, X., Chen, Y., Wagner, E., and Cosgrove, D. J. (2018). Xyloglucan in the primary cell wall: assessment by FESEM, selective enzyme digestions and nanogold affinity tags. *Plant J.* 93, 211–226. doi: 10.1111/tpj.13778
- Zobel, B. J. (1961). Inheritance of wood properties in conifers. *Silvae Genet.* 10, 65–70.
- Zuluaga, R., Putaux, J.-L., Cruz, J., Vélez, J., Mondragon, I., and Gañán, P. (2009). Cellulose microfibrils from banana rachis: effect of alkaline treatments on structural and morphological features. *Carbohydr. Polym.* 76, 51–59. doi: 10.1016/j.carbpol.2008.09.024

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Rongpipi, Ye, Gomez and Gomez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Recent Advances in the Transcriptional Regulation of Secondary Cell Wall Biosynthesis in the Woody Plants

Jin Zhang^{1,2}, Meng Xie^{1,2,3}, Gerald A. Tuskan^{1,2}, Wellington Muchero^{1,2*} and Jin-Gui Chen^{1,2*}

¹ Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ² Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ³ Department of Plant Sciences, University of Tennessee, Knoxville, TN, United States

OPEN ACCESS

Edited by:

Laura Elizabeth Bartley,
University of Oklahoma, United States

Reviewed by:

Berit Ebert,
The University of Melbourne, Australia
Samuel P. Hazen,
University of Massachusetts Amherst,
United States

*Correspondence:

Wellington Muchero
mucherow@ornl.gov
Jin-Gui Chen
chenj@ornl.gov

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 18 June 2018

Accepted: 28 September 2018

Published: 23 October 2018

Citation:

Zhang J, Xie M, Tuskan GA,
Muchero W and Chen J-G (2018)
Recent Advances
in the Transcriptional Regulation
of Secondary Cell Wall Biosynthesis
in the Woody Plants.
Front. Plant Sci. 9:1535.
doi: 10.3389/fpls.2018.01535

Plant cell walls provide structural support for growth and serve as a barrier for pathogen attack. Plant cell walls are also a source of renewable biomass for conversion to biofuels and bioproducts. Understanding plant cell wall biosynthesis and its regulation is of critical importance for the genetic modification of plant feedstocks for cost-effective biofuels and bioproducts conversion and production. Great progress has been made in identifying enzymes involved in plant cell wall biosynthesis, and in *Arabidopsis* it is generally recognized that the regulation of genes encoding these enzymes is under a transcriptional regulatory network with coherent feedforward and feedback loops. However, less is known about the transcriptional regulation of plant secondary cell wall (SCW) biosynthesis in woody species despite of its high relevance to biofuels and bioproducts conversion and production. In this article, we synthesize recent progress on the transcriptional regulation of SCW biosynthesis in *Arabidopsis* and contrast to what is known in woody species. Furthermore, we evaluate progress in related emerging regulatory machineries targeting transcription factors in this complex regulatory network of SCW biosynthesis.

Keywords: woody plants, *Populus*, secondary cell wall, transcription factor, transcriptional regulation

INTRODUCTION

Trees are important natural sources of sustainable energy and have important ecological and economical values (Tuskan, 1998; Richmond, 2000; Ragauskas et al., 2014). The majority of biomass of trees resides in the wood of stems, branches and roots. Wood is the major product of secondary growth derived from a lateral meristem, i.e., the vascular cambium, which forms xylem inwards and phloem outwards (Zhang et al., 2015). Prior to forming specialized cell types, cells in xylem and phloem undergo cell expansion and primary cell wall biosynthesis. However, wood is primarily composed of secondary cell walls (SCW) (Sundell et al., 2017). As the most abundant plant biomass worldwide, wood and fibers are widely used for various industrial applications, such as energy, pulping and textiles. In xylem, all the cell type firstly undergo SCW thickening and lignification, after which vessel elements and fibers undergo programmed cell death (PCD) (Courtois-Moreau et al., 2009).

Secondary cell walls, composed of lignin, cellulose and hemicelluloses, play an important role in plant development and stress responses (Houston et al., 2016). The maturation of SCWs reinforces specialized cells such as fibers and vessels, allowing them form mechanical tissues to provide structural support and protection while enabling negative pressure gradients generated during transpiration (Zhong et al., 2010a). The formation of SCW is a complex process requiring coordination of several metabolic pathways. Understanding the regulatory mechanism controlling SCW formation is critical for providing molecular and genetic bases for industrial applications (Zhong et al., 2013).

To date, a regulatory network consisting of several different types of transcription factors (TFs) and controlling SCW formation in the model plant *Arabidopsis* has been constructed (Zhong et al., 2010a; Taylor-Teeples et al., 2015). Recently, Rao and Dixon (2018) compared the transcriptional regulation models of SCW biosynthesis in grasses and *Arabidopsis*, and showed that the regulatory network of SCW development in grasses is relatively conserved with divergences. Compared to the annual herbaceous *Arabidopsis* and grasses, perennial woody species display extreme secondary growth that undergo seasonal changes that are impacted by various environmental stresses. Wood formation in perennial woody species is a dynamic and continuous process, which includes cambial cell proliferation, xylem cell differentiation, SCW thickening and PCD (Zhang et al., 2014). A comprehensive transcriptional regulatory network controlling secondary cell wall formation in woody species is still lacking. This review synthesizes the current advances of SCW regulatory network in plants in general and aims to highlight the recent progresses in this area in woody species. We also discuss the direction for future research in woody species.

THE FIRST LAYER OF TRANSCRIPTION FACTORS IN THE REGULATORY NETWORK IN SCW FORMATION

NAC (NAM, ATAF, and CUC) TFs are plant-specific transcriptional regulators and are widely involved in various biological processes, including growth/development and stress responses (Olsen et al., 2005). During SCW formation, a group of closely-related NAC TFs function as master switches, which were named SECONDARY WALL NACs (SWNs). In the first layer of the SCW regulatory network, SWNs are comprised of two types of NACs: VASCULAR-RELATED NAC DOMAINS (VNDs; VND1-7) and NAC SECONDARY WALL THICKENING PROMOTING FACTOR (NST)/SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN (SND) (NST1-3) (**Figure 1**). SWNs can bind to a 19 bp secondary wall NAC binding element (SNBE) sequences, (T/A)NN(C/T)(T/C/G)TNNNNNNA(A/C)GN(A/C/T)(A/T), and directly activate the expression of downstream TFs in the second layer, as well as structural genes involved in SCW biosynthesis, cell wall modification, and PCD (Zhong et al., 2010c). In addition, a 11 bp tracheary element-regulating *cis*-element (TERE) [CT(T/C)NAA(A/C)GCN(A/T)] was identified through an *in vitro* tracheary element (TE) transdifferentiation

study and was shown to be essential for TE-specific expression mediated by VNDs (Pyo et al., 2007; Ohashi-Ito et al., 2010; Yamaguchi et al., 2011).

The function of NACs in SCW formation was first reported in *Zinnia elegans*, in which a NAC TF Z567 was found to be up-regulated during the transdifferentiation from mesophyll cell into TEs in an *in vitro* culture system (Demura et al., 2002). Subsequently, in *Arabidopsis* suspension cells, seven homologs of Z567 were shown to be up-regulated during xylem vessel cell differentiation, which were named VND1 through VND7 (Kubo et al., 2005). VNDs individually display specific expression patterns and functions. For example, VND1-5 are expressed in the vessels of stem, but not expressed in interfascicular fibers. Moreover, VND4 and VND5 are expressed in vessels of the root hypocotyl. Overexpressing VND1-5 can activate the expression of TFs and structural genes involved in SCW biosynthesis and PCD (Zhou et al., 2014). VND6 and VND7 are specifically expressed in vessels, directing metaxylem and protoxylem vessel differentiation, respectively (Kubo et al., 2005). Different with the function of VNDs in vessels, NST1 and NST3/SND1 are master regulators of SCW biosynthesis in fibers (Mitsuda et al., 2007). In *Arabidopsis nst1-1 nst3-1* double mutant, the SCW thickening was completely suppressed in interfascicular fibers and secondary xylem without affecting the cell formation (Mitsuda and Ohme-Takagi, 2008). Similar to the functional redundancy of NST1 and NST3 in stem, NST1 and NST2 function redundantly in SCW formation in anthers (Mitsuda et al., 2005).

These master regulators have relatively conserved functions across plant species, though copy numbers vary by species. In *Medicago truncatula*, only one member, *MtNST1*, was identified corresponding to the three sequence homologs NST1-3 in *Arabidopsis*. A loss-of-function mutant, *mtnst1*, results in reduced lignin and cell wall polysaccharide contents through regulating the expression of most lignin biosynthetic and cellulose and hemicellulose biosynthetic genes (Zhao et al., 2010). *Oryza sativa* secondary wall NAC domain protein 1 (OsSWN1), an ortholog of *Arabidopsis* NST3/SND1, also regulates SCW formation in rice (Zhong et al., 2011; Chai et al., 2015). In *Arabidopsis*, ectopic expression of OsSWN1 induced massive ectopic deposition of lignified SCW in leaf mesophyll cells and in the epidermis and cortical cells of the inflorescence stems (Zhong et al., 2011). When OsSWN1 was heterologously expressed driven by the *Arabidopsis* NST3 promoter in the *nst1 nst3* double mutant, the pendent stem phenotype and the SCW lignification of inflorescent fibers were effectively rescued (Zhong et al., 2011), suggesting that OsSWN1 is functionally equivalent to *Arabidopsis* NST3/SND1. Subsequently, Sakamoto et al. (2016) overexpressed OsSWN1 in poplar using *Arabidopsis* NST3 promoter. The transgenic poplars displayed thickened SCW in xylem cells and phloem fiber cells but not in xylem vessels. A follow-up study indicated that overexpression of OsSWN1 in *Populus* altered lignin structure, but not lignin content, due to an unbalance induction of lignin biosynthetic genes (Nuoendagula et al., 2017). This further confirms that the function of the master switches is conserved across different species, whether they are annual or perennial, herbaceous or woody. Consistent with this notion, the similar regulatory

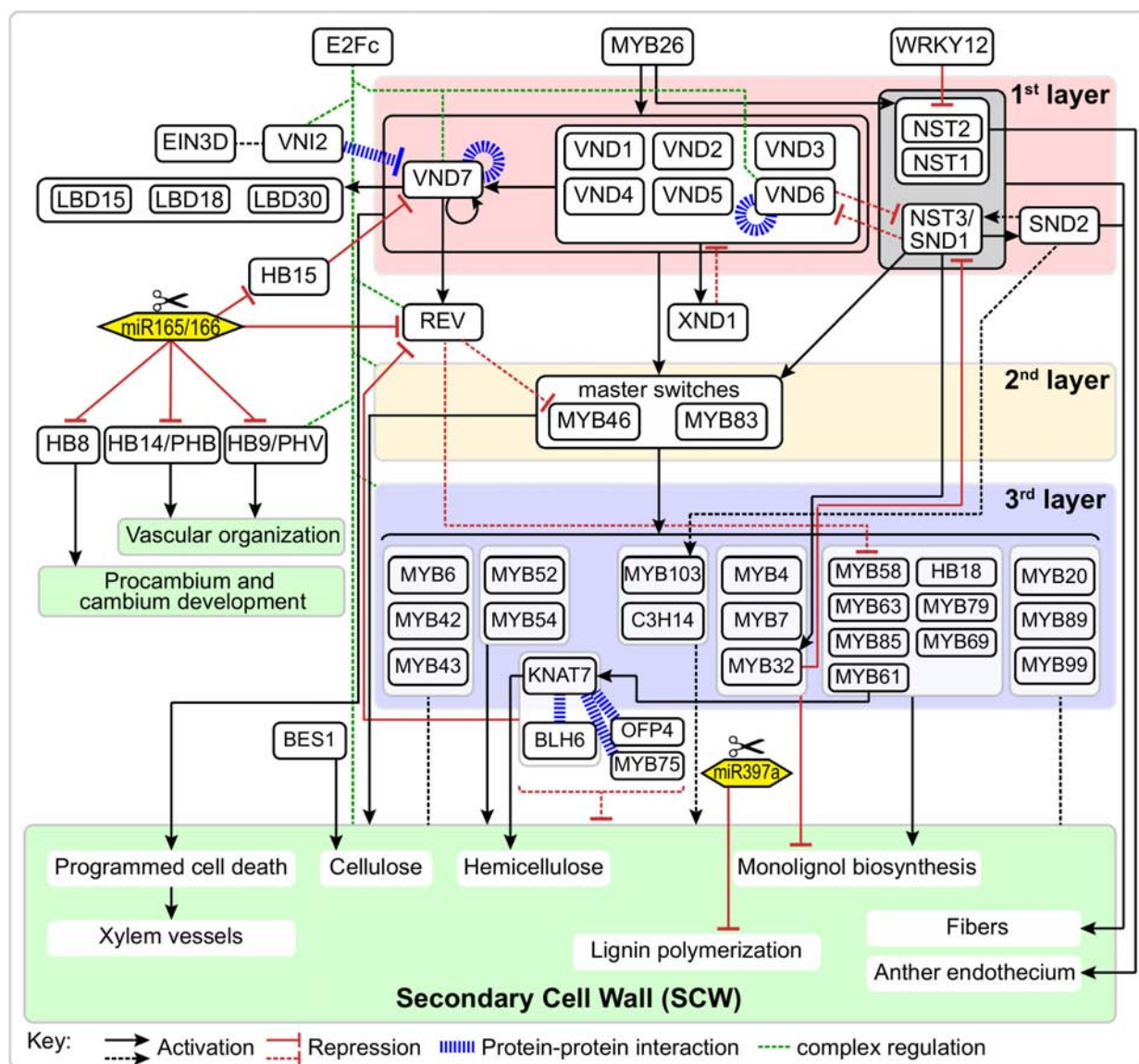


FIGURE 1 | Schematic representation of the regulatory network of secondary cell wall formation. Arrows and "T" at the ends of lines represent activation and repression, respectively. Blue bold dash lines represent protein-protein interactions. Solid and dotted lines represent direct or indirect regulation, respectively.

pathway is also observed in *Zea mays*, where *ZmNST3* and *ZmNST4* were specifically expressed in SCW-forming cells and functioned as master switches for SCW deposition through regulating the expression of *ZmMYB109/128/149* (Xiao et al., 2018).

In woody species, similar master switches have also been identified in *Populus* and *Eucalyptus*. In *Populus*, a group of wood-associated NAC domain TFs, *PtrWNDs*, were identified as master transcriptional switches in SCW biosynthesis. Ohtani et al. (2011) isolated 16 *Populus* NAC TFs and designated them as *PtVNS* (VND-, NST/SND- and SMB-related)/*PtrWND*. Among them, 12 members in NST and VND groups are expressed in developing xylem and phloem fibers, whereas only the VND group members are expressed in primary xylem

vessels. A homolog of *SND2* from *P. trichocarpa*, *PtSND2*, plays a similar role in the SCW biosynthesis. Chimeric repressor of *PtSND2* reduced the SCW thickness of xylem fibers and decreased lignin and cellulose contents in *Populus* (Wang et al., 2013). *PtrSND1-A2/PtrWND1B* (Potri.001G448400) was shown to be specifically expressed in secondary xylem fiber cells and suppression of *PtrWND1B* significantly inhibited fiber SCW thickening (Li Q. Z. et al., 2012; Zhao et al., 2014). Moreover, these master regulators function in gymnosperm trees. *Pinus pinaster* *PpNAC1* is a NST group TF, and it is a key regulator of phenylalanine biosynthesis through activating the expression of itself and *PpMYB4* (Table 1 and Figure 2) (Pascual et al., 2018). These results suggest that SWNs are ancestral master switches for the SCW formation, and that these master switches are

TABLE 1 | Summary of the transcription factors involved in secondary cell wall formation in woody species.

TF group	Species	TF	TF ortholog in <i>Arabidopsis</i>	Function	Reference
SWN	<i>Populus trichocarpa</i>	PtrSND1-A2 (PtrWND1B)	SND1	Positively regulate fiber cell wall thickening. Its splice variant, PtrSND1-A2 ^{IR} , is a dominant-negative regulator to suppress the transactivation of all PtrSND1 family members.	Li Q. Z. et al., 2012; Zhao et al., 2014
	<i>Populus trichocarpa</i>	PtrSND1-B1	SND1	Function as a master regulator to activate a hierarchical gene regulatory network during wood formation.	Lin et al., 2013
	<i>Eucalyptus grandis</i>	EgWND1	SND1	Transcriptional activator of SCW biosynthesis in wood.	Zhong et al., 2010a
	<i>Populus tomentosa</i> Carr.	PtoVNS11	SND1	Positively regulate lignin deposition and SCW thickening.	Yang et al., 2015
	<i>Populus trichocarpa</i>	PtSND2	SND2	Positively regulate fiber SCW thickening and lignin and cellulose biosynthesis.	Wang et al., 2013
	<i>Populus trichocarpa</i>	PtrWND2B	NST1/2	Activate SCW TFs and biosynthetic genes.	Zhong et al., 2010b
	<i>Populus trichocarpa</i>	PtrWND6B	VND6/7	Activate SCW TFs and biosynthetic genes.	Zhong et al., 2010b
	<i>Populus trichocarpa</i>	PtrVND6-C1	VND6	Its splice variant, PtrVND6-C1 ^{IR} , together with PtrSND1-A2 ^{IR} reciprocally cross-regulate the two TF families.	Lin et al., 2017
MYB	<i>Pinus pinaster</i>	PpNAC1	VND6	Positively regulate SCW formation.	Pascual et al., 2018
	<i>Populus trichocarpa</i>	PtrMYB3/20	MYB46/83	Activate the biosynthetic pathways of cellulose, xylan and lignin and are directly target of PtrWND2.	McCarthy et al., 2010
	<i>Eucalyptus gunnii</i>	EgMYB2	MYB46/83	Positively regulate SCW thickness and activates lignin biosynthetic genes.	Goicoechea et al., 2005
	<i>Pinus taeda</i>	PtMYB4	MYB46/83	Bind to AC elements and activate lignin biosynthetic genes.	Patzlaff et al., 2003a
	<i>Eriobotrya japonica</i>	EjMYB1	MYB58/63	Bind to AC elements and activate lignin biosynthetic genes.	Xu et al., 2014
	<i>Populus tomentosa</i> Carr.	PtoMYB216	MYB61	Positively regulate lignin biosynthetic pathway.	Tian et al., 2013
	<i>Populus tomentosa</i> Carr.	PtoMYB170	MYB61	Positively regulate lignin biosynthetic pathway and promote dark-induced stomatal closure.	Xu et al., 2017
	<i>Pinus taeda</i>	PtMYB8	MYB61	Positively regulate lignin biosynthetic and other cell wall-related genes.	Bomal et al., 2008
	<i>Picea glauca</i>	PgMYB8	MYB61	Positively regulate lignin biosynthesis.	Bedon et al., 2007
	<i>Populus deltoides</i>	PdMYB221	MYB4	Repressor. Negatively regulate SCW formation, including cellulose, xylose and lignin.	Tang et al., 2015
	<i>Populus tomentosa</i> Carr.	PtoMYB156	MYB4	Repressor. Repress phenylpropanoid biosynthesis and negatively regulate SCW formation.	Yang et al., 2017
	<i>Eucalyptus gunnii</i>	EgMYB1	MYB4	Repressor. Negatively regulate SCW formation.	Legay et al., 2007, 2010
	<i>Leucaena leucocephala</i>	LIMYB1	MYB4	Repressor. Negatively regulate lignin biosynthesis.	Omer et al., 2013
	<i>Eriobotrya japonica</i>	EjMYB2	MYB4	Repressor. EjMYB2 and EjMYB1 competitively interact with AC elements in the promoters of lignin biosynthetic genes.	Xu et al., 2014
	<i>Populus tremula</i> L. × <i>tremuloides</i> Michx.	PttMYB21a	MYB52	Repressor. Negatively regulate lignin biosynthesis.	Karpinska et al., 2004
	<i>Quercus suber</i>	QsMYB1	MYB68	Related to secondary growth and cork biosynthesis.	Almeida et al., 2013
	<i>Vitis vinifera</i> L.	VvMYB5a	MYB5	Involved in phenylpropanoid pathway.	Deluc et al., 2006
	<i>Pinus taeda</i>	PtMYB1	MYB42/43/20	Bind to AC elements and activate lignin biosynthetic genes.	Patzlaff et al., 2003b
	<i>Populus trichocarpa</i>	PtrMYB152	MYB43/20	Positively regulate SCW biosynthesis.	Wang et al., 2014

(Continued)

TABLE 1 | Continued

	<i>Populus tomentosa</i> Carr.	PtoMYB92	MYB42/85	Positively regulate lignin biosynthesis.	Li et al., 2015
	<i>Populus deltoides</i>	PdMYB10/128	MYB103	Positively regulate fiber SCW thickening and delay flowering.	Chai et al., 2014b
	<i>Populus deltoides</i>	PdMYB90/167	MYB52	Negatively regulate fiber and vessel SCW thickening and prompt flowering.	Chai et al., 2014b
	<i>Populus deltoides</i>	PdMYB92/125	MYB42	Repressor. Negatively regulate fiber and vessel SCW thickening and prompt flowering.	Chai et al., 2014b
WRKY	<i>Populus trichocarpa</i>	PtWRKY19	WRKY12	Repressor. Negatively regulate pith SCW formation.	Yang et al., 2016
	<i>Vitis vinifera</i> L.	VvWRKY2	WRKY3/4	Positively regulate lignin biosynthesis and affect S/G ratio.	Guillaumie et al., 2010
Others:					
HD Zip III	<i>Populus trichocarpa</i>	popREVOLUTA (PRE)	REV	Play fundamental roles in cambium initiation and patterning of secondary vascular tissues	Robischon et al., 2011
HD Zip III	<i>Populus tremula</i> × <i>alba</i>	POPCORONA (PCN)	HB15(CNA)	Involved in SCW lignification and regulate cell differentiation during secondary growth.	Du et al., 2011
HD Zip II	<i>Eucalyptus camaldulensis</i>	EcHB1	HAT22(ABIG1)	Negatively regulate lignin and hemicellulose content, increase fiber length and growth.	Sonoda et al., 2009
KNOX I	<i>Populus tremula</i> × <i>alba</i>	ARBORKNOX1 (ARK1)	STM	Affect internode elongation and secondary vascular cell types in stem, positively regulate lignin biosynthesis.	Groover et al., 2006
KNOX I	<i>Populus tremula</i> × <i>alba</i>	ARBORKNOX2 (ARK2)	BP	Negatively regulate SCW biosynthesis.	Du et al., 2009
KNOX II	<i>Populus trichocarpa</i> × <i>deltoides</i> , <i>P. balsamifera</i>	PtRKAT7	KNAT7	Negatively regulate SCW biosynthesis.	Li E. Y. et al., 2012
LBD	<i>Populus tremula</i> × <i>alba</i>	PtaLBD1	LBD1/11	Positively regulate ray cell development and phloem differentiation.	Yordanov et al., 2010
CCCH	<i>Populus deltoides</i>	PdC3H17 and PdC3H18	C3H14	Positively regulate SCW formation and are direct targets of PdMYB3 and PdMYB21.	Chai et al., 2014a
MADS-box	<i>Populus tremuloides</i>	PTM5	SOC1	Expressed in differentiating vascular cambium and xylem tissues.	Cseke et al., 2003
ERF	<i>Populus simonii</i> ? × ? <i>nigra</i>	PsnSHN2	SHN2	Positively regulate cellulose and hemicellulose biosynthesis, but negatively regulate lignin biosynthesis.	Liu et al., 2017
EIN	<i>Populus tremula</i>	EIN3D	EIN3	Possibly act upstream or together with VIN2 during wood formation.	Seyfferth et al., 2018

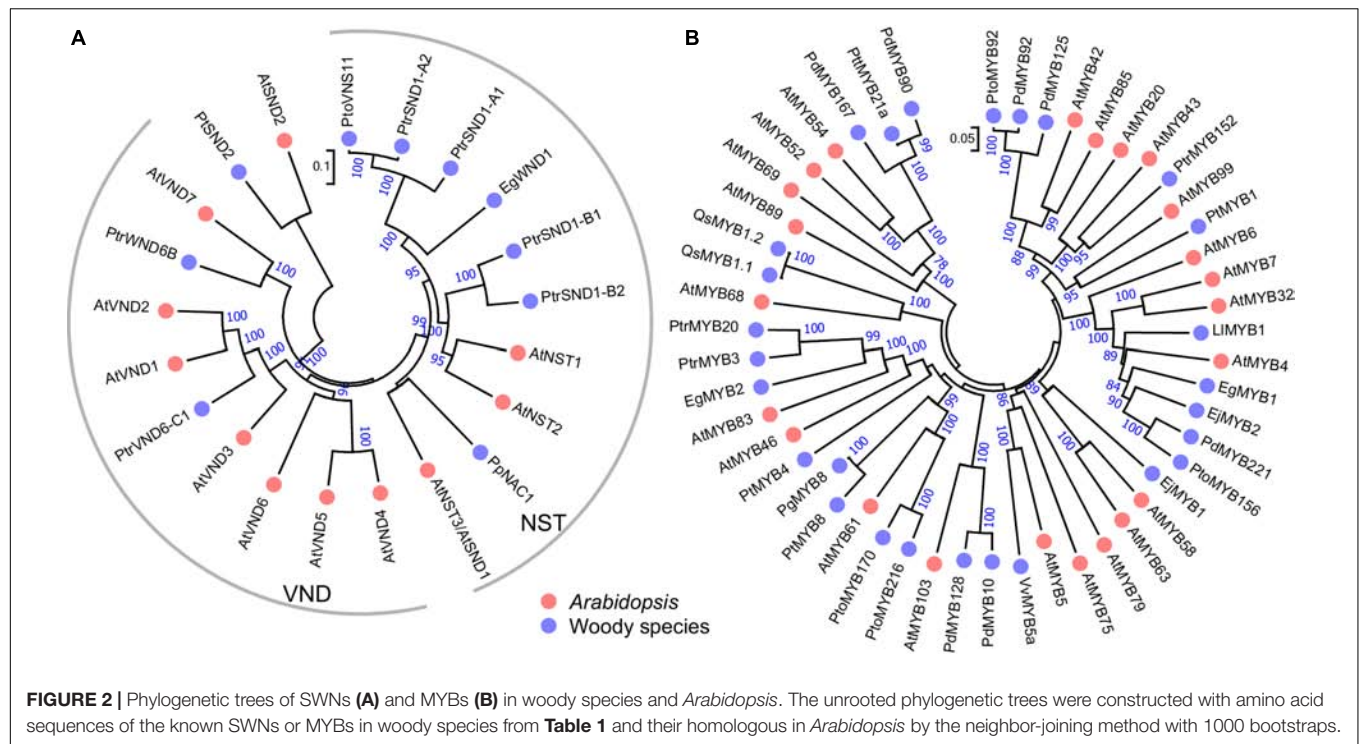
functionally conserved across different plant species, including woody species.

REGULATORS ASSOCIATED WITH THE FIRST LAYER OF TRANSCRIPTION FACTORS

In the first layer of SCW regulatory network, several TFs are involved in regulation or interaction with the master switches (**Figure 1**). For instance, VND-INTERACTING2 (VNI2) interacts with VND7 and VND1-5. Here, VNI2 functions as a transcriptional repressor to limit the expression of VND7-regulated vessel-specific genes (Yamaguchi et al., 2010). XYLEM NAC DOMAIN1 (XND1) is up-regulated in xylem, and it can negatively regulate xylem vessel differentiation (Zhao et al., 2008). A recent study indicates the function of XND1 in xylem differentiation depends on its C-terminal region containing

linear motifs (KII-acidic, LXCXE, E2F^{TD}-like and LXCXE-mimic) which can interact with the cell cycle and differentiation regulator RETINOBLASTOMA-RELATED (RBR) (Zhao et al., 2017). By using enhanced yeast one hybrid assays, Taylor-Teeple et al. (2015) identified E2Fc as a key upstream regulator of VND6, VND7 and other SCW biosynthetic genes. E2Fc is a known negative regulator of endoreduplication (del Pozo et al., 2002, 2006), but it can also act as a transcriptional activator (Kosugi and Ohashi, 2002; Heckmann et al., 2011). Prior to terminal differentiation, the elongating xylem cells likely undergo endoreduplication before SCW deposition via E2Fc-mediated activation or repression of VND7 in a concentration-dependent manner (Taylor-Teeple et al., 2015).

Other regulators associated with master switches in the first layer of SCW regulatory network include Homeobox HD-Zip class III (HD-Zip III), a small TF family that consists of five members in the *Arabidopsis* genome, i.e., REVOLUTA/INTERFASCICULAR FIBERLESS1 (REV/IFL1), PHABULOSA (PHB), PHAVOLUTA (PHV), HB8, and HB15



(CORONA). The HD-Zip III genes are negatively regulated by highly conserved miRNAs (Floyd et al., 2006) and all five HD-Zip III TFs are necessary for xylem cell specification and SCW synthesis. In *Populus*, *popREVOLUTA* (ortholog of *REV*) plays fundamental roles in the initiation of the cambium and in the regulation of the patterning of secondary vascular tissues (Robischon et al., 2011). The promoters of *REV* and *PHB* can be bound and regulated by VND7 (Taylor-Teeple et al., 2015). HB15 is necessary for repressing SCW biosynthesis in pith and disruption of the expression of *HB15* causes ectopic lignification in pith cells. An ortholog in *Populus*, *POPCORONA*, is involved in SCW lignification and regulates cell differentiation during secondary vascular growth (Du et al., 2011). Noticeably, the expression of *WRKY12* is up-regulated in *athb15* mutant (Du et al., 2015). As a negative regulator, *WRKY12* can directly bind to the *NST2* promoter to repress its expression, thus repressing the SCW thickening in pith cells (Wang et al., 2010). Finally, three homologous LOB domain TFs (LBD15, LBD18, and LBD30) are expressed in differentiating TEs and enhance the transcription of *VND7* in a positive feedback loop (Soyano et al., 2008; Ohashi-Ito et al., 2018).

In addition to the transcriptional regulation, the post-translational modifications play important roles in regulating the master switches in the first layer of SCW regulatory network. A study using tobacco BY-2 cells expressing VND7-YFP together with the treatment of proteasome inhibitor MG-132 showed that VND7 is also regulated by proteolysis (Yamaguchi et al., 2008). Recently, Kawabe et al. (2018) identified a recessive mutant with inhibited ectopic xylem cell differentiation in *35S::VND7-VP16-GR* lines and found this mutant is caused by a single amino

acid substitution (E36K) in S-nitrosogluthathione reductase (GSNOR1). GSNOR was first reported as a glutathione-dependent formaldehyde dehydrogenase and regulates the turnover of S-nitrosogluthathione, a natural nitric oxide donor. VND7 can be S-nitrosylated at Cys264 and Cys320, which are located near the transactivation domain. The *in vivo* S-nitrosylation of VND7 mediated by GSNOR1 affects VND7-downstream signaling events and thereby leading to deficient xylem vessel differentiation (Kawabe et al., 2018). Collectively, these regulators work with the first layer master switches to regulate their transcription or protein activity by providing post-translational modifications. These provide an additional layer of regulation at the top level to regulate SCW biosynthesis, which may possibly involve the integration of developmental or environmental signals since many of these regulators play roles in these processes (Jin et al., 2000; Preston et al., 2004; Romano et al., 2012).

THE SECOND LAYER OF TRANSCRIPTION FACTORS IN THE REGULATORY NETWORK IN SCW FORMATION

A series of additional TFs make up the second layer of regulation of the expression of SCW biosynthetic genes and other downstream genes. The master switches in the second layer are MYB46 and MYB83 (Figure 1), which are directly regulated by SND1 and its close homologs (NST1, NST2,

VND6 and VND7) (Zhong et al., 2007; McCarthy et al., 2009). *MYB46* and *MYB83* are functional redundant and are specifically expressed in fibers and vessels where SCW thickening occurs. Overexpression of *MYB46* or *MYB83* enhanced the biosynthetic pathways of lignin, cellulose and xylan, and resulted in ectopic deposition of SCW; whereas RNAi or dominant repression of *MYB46* and *MYB83* reduced SCW thickening of fibers and vessels (Zhong et al., 2007; McCarthy et al., 2009).

MYB46 and *MYB83* can regulate other SCW-related TFs or directly regulate the SCW structural genes. Based on results from the estrogen-inducible direct activation system, several downstream TFs, including *MYB43*, *MYB52*, *MYB54*, *MYB58*, *MYB63* and *KANT7*, have been identified as direct targets of *MYB46/83*. A 7-bp sequence ACC(A/T)A(A/C)(T/C) has been designated as the secondary wall MYB-responsive element (SMRE) (Zhong and Ye, 2012), similar to binding sequences of AC element [ACC(T/A)ACC] (Fornale et al., 2010) and P1 [CC(T/A)ACC] (Grote et al., 1994). Another 8-bp sequence [(T/C)ACC(A/T)A(A/C)(T/C)] has also been identified as *MYB46* specific binding sequence, namely *MYB46*-responsive *cis*-regulatory element (M46RE) (Kim et al., 2012; Ko et al., 2014). In addition, *MYB46/83* can directly regulate SCW structural genes. For example, *MYB46* directly regulates all three SCW-associated cellulose synthase genes (*CesA4*, *CesA7* and *CesA8*) (Kim et al., 2013) and a mannan synthase *CSLA9* (Kim et al., 2014b). Noticeably, the promoters of these genes contain multiple M46REs. A genome-wide screen of promoter sequences indicates the xylan biosynthetic genes (*IRX8*, *IRX9*, *IRX10*, *IRX14*, *IRX15* and *IRX15-L*) (Jensen et al., 2011; Kim et al., 2014a), lignin biosynthesis-related laccase (*LAC4/IRX12*, *LAC10* and *LAC11*) (Zhao et al., 2013), cytoskeleton-related genes (*Myosin5*, *microtubule-associated protein*), and homologous of *IRX15/15-L* (*DUF579s*) also contain multiple M46REs in their promoter regions (Jensen et al., 2011).

Similar to the master switches in the first layer of SCW regulatory network, the function of *MYB46* and *MYB83* is also highly conserved in woody species. For instance, *PtMYB3* and *PtMYB20* from *Populus*, *EgMYB2* from *Eucalyptus*, and *PtMYB4* from *Pinus taeda*, are orthologs of *MYB46/83* and perform the same function as *MYB46/83* from *Arabidopsis* in SCW biosynthesis (Figure 2). In *Populus* developing wood, *PtMYB3* and *PtMYB20* are highly expressed in vessels and fibers and can regulate the biosynthesis of lignin, cellulose and xylan (McCarthy et al., 2010). *Eucalyptus EgMYB2* is identified based on a quantitative trait locus (QTL) for lignin content. *EgMYB2* can specifically bind to the promoters of lignin biosynthetic genes, such as *CINNAMOYL-COENZYME A REDUCTASE* (*CCR*) and *CINNAMYL ALCOHOL DEHYDROGENASE* (*CAD*). Overexpression of *EgMYB2* enhanced SCW thickness in transgenic tobacco (Goicoechea et al., 2005). In loblolly pine, *Pinus taeda MYB4* (*PtMYB4*), the homolog of *Arabidopsis MYB46/83*, is expressed in lignifying xylem cells. *PtMYB4* can bind to AC elements and activate the expression of target genes (Patzlaff et al., 2003a). Collectively, these results suggest that the orthologs of *MYB46/83* function conservatively as the

second layer master regulators in SCW biosynthesis in woody plants.

THE THIRD LAYER OF REGULATORY NETWORK IN SCW FORMATION

In addition to the master switches in the second layer of SCW regulatory network, there are TFs that regulate SCW biosynthesis, whose expression are regulated by the master switches *MYB46/83*, and act as downstream TFs in the third layer of SCW regulatory network (Figure 1). Most of these TFs belong to the MYB gene family. The first identified lignin-specific TFs were *MYB58*, *MYB63*, and *MYB85* (Zhong et al., 2008; Zhou et al., 2009). Most monolignol biosynthetic genes contain AC elements in their promoter region and are direct target of *MYB58* (Zhou et al., 2009). Moreover, *MYB6*, *MYB20*, *MYB42*, *MYB43*, *MYB52*, *MYB54*, *MYB61*, *MYB103*, etc. are also developmentally associated with cells undergoing SCW thickening (Zhong et al., 2008; Romano et al., 2012). *MYB52*, *MYB54*, *MYB85* and *MYB103* are able to induce SCW biosynthetic genes. Overexpression of *MYB85* led to ectopic lignin deposition in epidermal and cortical cells; overexpression of *MYB103* increased SCW thickening in fibers; whereas dominant repression of *MYB52*, *MYB54*, *MYB85*, or *MYB103* reduced SCW thickening in fiber cells (Zhong et al., 2008). In contrast, *MYB61* plays multiple regulatory roles in plant development, including lignification, dark-photomorphogenesis (Newman et al., 2004), stomatal aperture (Liang et al., 2005) and seed coat mucilage deposition (Penfield et al., 2001). Analysis of loss-of-function mutant of *MYB61*, *atmyb61*, showed that *MYB61* can activate the expression of *CAFFEYOYL-COA 3-O-METHYLTRANSFERASE* (*CCoAOMT*) and *PECTIN METHYLESTERASE* (*PME*) and affect xylem formation and xylem cell structure (Romano et al., 2012).

While most of these TFs activate the expression of their targets and positively regulate SCW biosynthesis, several members in MYB family play negative roles in SCW biosynthesis. *Arabidopsis MYB4* is induced by UV-B. Overexpression of *MYB4* can repress the transcription of *4CL*, *C4H* and *CAD* in tobacco (Jin et al., 2000). *MYB7* and *MYB32* share high sequence similarity with *MYB4*, act as repressors, and are strongly activated by *MYB46* (Ko et al., 2009). *MYB32* negatively regulates lignin pathway through repressing other targets, such as *COMT* (Preston et al., 2004). In addition, there is a feedback regulation between *MYB32* and SWNs. The transcription of *MYB32* is repressed in the *nst1 nst3* double mutant (Mitsuda et al., 2007). A later study based on *in vitro* trans-activation assays and electrophoretic mobility shift assay (EMSA) further confirmed that *MYB32* is directly regulated by *SND1* (Wang et al., 2011). Furthermore, *SND1* is negatively regulated by *MYB32* (Wang et al., 2011), implying that both positive and negative feedforward loop exist in SCW regulatory network.

Populus PttMYB21a, a homolog of *MYB52* (Figure 2), can negatively regulate the expression of *CCoAOMT* and the acid soluble lignin content (Karpinska et al., 2004). In grapevine, *VvMYB5a* can regulate both anthocyanin/proanthocyanidin

biosynthesis and lignin biosynthesis (Deluc et al., 2006). In *Eucalyptus gunnii*, EgMYB1 binds to the promoter of *CCR* and *CAD* to repress the monolignol biosynthesis (Legay et al., 2007). In *Pinus taeda*, PtMYB1, closely related to *Arabidopsis* MYB42, MYB43 and MYB20, is most abundantly expressed in differentiating xylem and functions as a transcriptional activator through binding the AC elements (Patzlaff et al., 2003b). Loquat (*Eriobotrya japonica*) EjMYB1 (ortholog of MYB58 and MYB63) functions as transcriptional activator and can activate both *Arabidopsis* and loquat lignin biosynthetic genes. EjMYB2 (ortholog of MYB4) functions as a repressor and can counter the induction by EjMYB1 (Xu et al., 2014). The large abundance of TFs in the third layer provide multiple opportunities to connect to the master switches in the first and second layers and the structural genes in SCW biosynthesis, and to fine tuning the pathways.

REGULATORS ASSOCIATED WITH THE THIRD LAYER OF TRANSCRIPTION FACTORS

Several genes in other TF families cooperate with MYBs or act independently to regulate the SCW biosynthesis (Figure 1). KNOTTED ARABIDOPSIS THALIANA7 (KNAT7) is a Knotted-like homeobox (KNOX) protein, is a target of MYB46 (Ko et al., 2009) and SND1 (Zhong et al., 2008), and can also be regulated by MYB61 (Romano et al., 2012). Dominant repression of *KNAT7* reduced SCW thickening in vessels and fibers (Zhong et al., 2008). In *Nicotiana*, virus-induced silence and RNAi of *NbKNAT7* inhibited the thickening of fiber cell walls and repressed the expression of lignin, cellulose and xylan biosynthetic genes (Pandey et al., 2016). *KNAT7* was known as a transcriptional repressor that negatively regulates SCW biosynthesis, and it can physically interact with MYB75, OFP4 and BLH6 (Li et al., 2011; Li E. Y. et al., 2012; Bhargava et al., 2013; Liu et al., 2014). *Arabidopsis* MYB75 positively regulates anthocyanin biosynthesis, but it functions as a repressor in SCW biosynthesis. A loss-of-function mutant *myb75-1* enhanced the expression of lignin, cellulose and xylan biosynthetic genes and increased SCW thickness in xylary and interfascicular fibers (Bhargava et al., 2010). In *Arabidopsis*, the *KNAT7*-MYB75 complex is involved in modulating SCW formation in both inflorescence stem and seed coat (Bhargava et al., 2013). OFP4 is an Ovate Family Protein transcriptional co-regulator and can interact with *KNAT7* and enhance the repression activity of *KNAT7* in SCW biosynthesis (Li et al., 2011). BLH6 is a BELL1-LIKE HOMEODOMAIN protein and functions as a transcriptional repressor. It specifically interacts with *KNAT7* to enhance its repression activity. BLH6 and *KNAT7* can repress the expression of *REV* through directly binding to its promoter (Liu et al., 2014).

In *Populus*, *KNAT7* functions as a repressor in a negative feedback loop in SCW formation (Li E. Y. et al., 2012). However, a recent study indicates that *KNAT7* positively regulates xylan biosynthesis through directly activating the expression of *IRX9* (He et al., 2018). Another member in KNOX

family, *BERVIPEDICELLUS* (BP)/*KNAT1*, also plays a role in the lignin pathway. BP binds to the promoters of genes in the lignin pathway (*COMT*, *CCoAOMT*, etc.) and overexpressing BP significantly decreases the SCW lignification (Mele et al., 2003). In addition, the tandem CCCH zinc finger (TZF) TF, *C3H14*, is able to activate SCW biosynthetic genes and is directly regulated by MYB46 and SND1 (Ko et al., 2009). Its orthologs in *Populus deltoides*, *PdC3H17* and *PdC3H18*, also positively regulate SCW formation in both *Populus* and *Arabidopsis*, and are direct targets of PdMYB3 and PdMYB21 (Chai et al., 2014a). These regulators associated with the third layer of transcription factors provide opportunities for fine tuning SCW biosynthesis at the very downstream level.

ETHYLENE RELATED TFs IN SCW BIOSYNTHESIS

Recently, a class of ethylene signaling-related TFs have attracted the attention of researchers due to their function in wood development (Figure 1). Ethylene is the smallest phytohormone with the simple structure C_2H_4 , and is involved in various plant developmental processes including leaf development, senescence, fruit ripening, germination, stress responses, etc. (Dubois et al., 2018). Notably, ethylene is also involved in multiple process during wood formation, including cambial growth, xylem cell morphogenesis, and vessels/fibers/rays ontogenesis (Little and Savidge, 1987). In angiosperm trees, ethylene, as an important signaling molecule, is involved in the remodeling of wood formation upon tension wood induction. Exogenous application of ethylene or its precursor 1-aminocyclopropane-1-carboxylic acid (ACC) enhances xylem growth in hybrid aspen (*Populus tremula* × *P. tremuloides*) (Love et al., 2009). In addition, gene expression and enzyme activity of ACC oxidase are up-regulated on the tension wood surface (Andersson-Gunneras et al., 2003).

The ethylene perception and signal transduction cascades depend on ethylene-induced Ethylene Response Factors (ERFs). In *Arabidopsis*, *ERF1*, *ERF018* and *ERF109* are involved in the vascular cell division (Etchells et al., 2012), suggesting that ERFs-mediated ethylene signaling is important for vascular development. Vahala et al. (2013) performed a genome-wide screen for ERFs in hybrid aspen stem. Among 170 ERFs in *Populus*, 50 ERFs were induced greater than five-fold by ethylene. During tension wood formation, 17 and 8 ERFs were induced greater than two-fold and ten-fold, respectively. Subsequently, the function of these ERFs was further confirmed in transgenic *Populus* (Vahala et al., 2013). Overexpression of *ERF18*, *ERF21*, *ERF30*, *ERF85* and *ERF139* in wood-forming tissues modified the wood chemotype in hybrid aspen. Overexpression of *ERF139* repressed longitudinal and lateral growth with altered wood development, overexpression of *ERF18*, *ERF34*, and *ERF105* enhanced diameter growth, whereas overexpression of *ERF71* and *ERF85* suppressed diameter growth.

Despite this work the role of ERFs-mediated ethylene signaling in the SCW regulatory network remains elusive (Figure 1). Liu et al. (2017) reported that *Populus* ERF type TF, *PsnSHN2*, is predominantly expressed in xylem tissues, and that

it positively regulates cellulose and hemicellulose biosynthesis but negatively regulates lignin biosynthesis. Recently, Seyfferth et al. (2018) constructed an ethylene-related gene expression network during SCW formation, *ETHYLENE INSENSITIVE 3D* (*EIN3D*) and 11 *ERFs* were identified as hub genes. Interestingly, a *VNI2* homolog is highly associated with *EIN3D*, suggesting *EIN3D* may act upstream or together with *VNI2* during SCW formation. How to precisely position these unresolved TFs into the SCW transcriptional regulatory network deserves further investigation.

POST-TRANSCRIPTIONAL REGULATION OF TFs INVOLVED IN SCW FORMATION

The activity of transcriptional regulators and the gene expression are also affected by post-transcriptional regulation. Alternative splicing is an important model of post-transcriptional regulation. It plays important roles for enhancing proteomic diversity in diverse cellular processes (Chen and Manley, 2009). In plants, more than 60% of intron-containing genes undergo alternative splicing (Syed et al., 2012). However, the knowledge of alternative splicing in wood formation is limited. By analyzing the xylem transcriptome in 20 *P. trichocarpa* genotypes, Bao et al. (2013) found that about 36% of the genes expressed in xylem undergo alternative splicing, especially those cell wall biosynthetic genes including glycosyl transferases and C₂H₂ TFs.

Interestingly, most key TFs in the first layer of SCW regulatory network undergo alternative splicing. In *Populus*, a “stem-differentiating xylem”-specific variant of *SND1*, *PtrSND1-A2^{IR}*, was identified as a dominant-negative regulator of *SND1*-mediated pathway (Li Q. Z. et al., 2012). The retained intron 2 in *PtrSND1-A2^{IR}* cDNA introduces a premature stop codon resulting in a truncated protein lacking the activation domain. Hence *PtrSND1-A2^{IR}* loses DNA binding and transactivation abilities, and it represses the transcription of *PtrSND1* members and *PtrMYB021* via its retained dimerization capability. This is the first report on the auto-repression of a TF family by its own splice variant in plants. Subsequently, Zhao et al. (2014) compared the function of the two isoforms, *PtrSND1-A2* (also named *PtrWND1B-s*) and *PtrSND1-A2^{IR}* (also named *PtrWND1B-l*), during wood formation. Overexpression of *PtrWND1B-s* or *PtrWND1B-l* oppositely regulate fiber SCW thickening in *Populus*. This alternative splicing type was also detected in *SND1* ortholog in *Eucalyptus grandis* (*Eucgr.E01053*), but not in *Arabidopsis*, implying that the alternative splicing regulation of *SND1* may be different between woody species and herbaceous plants (Li Q. Z. et al., 2012; Zhao et al., 2014). Recently, Lin et al. (2017) reported that another key TF in the first layer of SCW regulatory network, *VND6*, also undergoes alternative splicing during wood formation. Its splice variant retained intron 2, *PtrVND6-C1^{IR}*, which suppresses the protein function of all *PtrVND6* and *PtrSND1* family members, including *PtrSND1-A2*. In addition, *PtrVND6-C1* can also be suppressed by *PtrSND1-A2^{IR}*. *PtrVND6-C1^{IR}* and *PtrSND1-A2^{IR}* function together for reciprocal cross-regulation of *VND* and *SND* members to maintain homeostasis for xylem differentiation and plant development. Whether other key TFs in SCW regulatory

network also undergo alternative splicing is still an open question. This intron-retained splice variant-introduced reciprocal cross-regulation provides an additional insight for studying the regulation mechanism of SCW formation and appears to be woody species-specific.

In addition to alternative splicing, the TFs and structural genes in SCW regulatory network are regulated by non-coding RNAs (ncRNAs). In the past few decades, ncRNAs have been shown to play key regulatory roles in various biological processes of development and stress response (Mallory and Vaucheret, 2006; Wierzbicki, 2012). Plant ncRNAs can be classified into various types according to their molecular structures, including microRNA (miRNA), small interfering RNA (siRNA), long ncRNA (lncRNA), circular RNA (circRNA), etc. (Sunkar et al., 2007; Kim and Sung, 2012). Here, we focus on the role of miRNA and lncRNA in SCW formation, in particular on their regulation of SCW-related TFs. Lu et al. (2005) identified 21 miRNA families from the developing xylem of *P. trichocarpa* stems. Among them, 11 miRNA families have conserved sequences in *Arabidopsis* but exhibit species-specific developmental expression patterns, while 10 *Populus*-specific miRNA families might be involved in tree-specific processes. Several members in miRNA families have been reported to play important roles in SCW formation. miRNA165/166 are known to target HD-Zip III TFs, and control xylem differentiation through modulating the *PHB* gradients in the stele to maintain *PHB* at a low dosage in protoxylem and a high dosage in metaxylem differentiation (Carlsbecker et al., 2010; Miyashima et al., 2011). In hybrid aspen (*Populus tremula* × *P. alba*), *Pta-miRNA166* targets *PtaHBI1*, a homolog of *REV*, to regulate secondary growth (Ko et al., 2006). In a gain-of-function *Arabidopsis MIR166a* mutant, the transcript level of *HB15* was decreased and xylem and interfascicular region were expanded in vascular tissue (Kim et al., 2005). In *Populus*, synthetic miRNA knock-down of *POPCORONA* (*PCN*), an ortholog of *HB15*, disturbed the lignification of pith cells, whereas overexpression of a miRNA-resistant *PCN* delayed the lignification of xylem and phloem fibers (Du et al., 2011). Laccases (*LAC*) belong to the blue copper oxidase family and polymerize monolignols into lignin. Among the 49 *LAC* genes in the *Populus* genome, 29 were predicted as the targets of *ptr-miRNA397a*. Overexpression of *Ptr-MIRNA397a* reduced lignin content without changing monolignol biosynthesis in *Populus* (Lu et al., 2013). Recently, another miRNA, *miRNA319*, was also shown to be able to target *TCP4* and decrease the SCW formation in *Arabidopsis* stem. *TCP4* TF can directly activate the expression of *VND7* via binding to its promoter (Sun et al., 2017). lncRNAs are also involved in wood formation. Chen et al. (2015) performed a genome-wide analysis and compared the expression profiles of lncRNA in the xylem of normal wood, opposite wood and tension wood in *Populus tomentosa*. A total of 16 genes in cellulose or lignin biosynthetic pathways were targeted by lncRNAs. Combining whole-genome resequencing with growth and wood-property traits of 435 *P. tomentosa* individuals, Zhou et al. (2017) further identified 8 lncRNAs and 15 potential target genes in the phenylpropanoid pathway. These diversified post-transcriptional regulatory mechanisms offer new perspectives to the

SCW regulatory network through modifying gene expression or protein diversity of the key TFs.

CONCLUSION

In this review, we provide a summary of current knowledge of the transcriptional regulation of SCW biosynthesis in woody species and contrast to what is known in other plant species, particularly in the model plant *Arabidopsis*. Woody species and the herbaceous model plant *Arabidopsis* share conserved master switches in the SCW transcriptional regulatory network, especially in the first and second layers of the network. However, the large abundance of TFs in the third layer and diversified post-transcriptional regulatory mechanisms make the SCW regulatory network more complex in woody plants. For example, the alternative splicing events of SND and VND genes appeared to be woody species-specific. This poses more challenges for fully revealing the SCW regulatory mechanism in woody species. Recent advances in high-throughput sequencing provide great potentials for improving the genome annotation and identifying alternative splicing events and lncRNAs during SCW formation. In addition, expression quantitative trait loci (eQTL) analysis provides an effective and efficient way to identify novel regulators, especially in tree species with

long life cycle. Recently, Zhang et al. (2018) identified a *Populus* hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase PtHCT2 controlling caffeoylquinic acid biosynthesis and its upstream regulators through eQTL analysis, which provides a new strategy to identify novel transcriptional regulators in woody plants. Considering the ecological and economic values of woody species, it is important to understand the woody species-specific transcriptional regulation of SCW formation, and this is a fruitful area for further research.

AUTHOR CONTRIBUTIONS

JZ collected and synthesized data from literature and wrote the manuscript. MX, GT, WM and J-GC revised the manuscript.

FUNDING

This research was supported by the Center for Bioenergy Innovation (CBI). CBI is supported by the Office of Biological and Environmental Research (BER) in the U.S. Department of Energy Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract Number DE-AC05-00OR22725.

REFERENCES

- Almeida, T., Menendez, E., Capote, T., Ribeiro, T., Santos, C., and Goncalves, S. (2013). Molecular characterization of *Quercus suber* MYB1, a transcription factor up-regulated in cork tissues. *J. Plant Physiol.* 170, 172–178. doi: 10.1016/j.jplph.2012.08.023
- Andersson-Gunneras, S., Hellgren, J. M., Bjorklund, S., Regan, S., Moritz, T., and Sundberg, B. (2003). Asymmetric expression of a poplar ACC oxidase controls ethylene production during gravitational induction of tension wood. *Plant J.* 34, 339–349. doi: 10.1046/j.1365-3113X.2003.01727.x
- Bao, H., Li, E., Mansfield, S. D., Cronk, Q. C., El-Kassaby, Y. A., and Douglas, C. J. (2013). The developing xylem transcriptome and genome-wide analysis of alternative splicing in *Populus trichocarpa* (black cottonwood) populations. *BMC Genomics* 14:359. doi: 10.1186/1471-2164-14-359
- Bedon, F., Grima-Pettenati, J., and Mackay, J. (2007). Conifer R2R3-MYB transcription factors: sequence analyses and gene expression in wood-forming tissues of white spruce (*Picea glauca*). *BMC Plant Biol.* 7:17. doi: 10.1186/1471-2229-7-17
- Bhargava, A., Ahad, A., Wang, S. C., Mansfield, S. D., Haughn, G. W., Douglas, C. J., et al. (2013). The interacting MYB75 and KNAT7 transcription factors modulate secondary cell wall deposition both in stems and seed coat in *Arabidopsis*. *Planta* 237, 1199–1211. doi: 10.1007/s00425-012-1821-9
- Bhargava, A., Mansfield, S. D., Hall, H. C., Douglas, C. J., and Ellis, B. E. (2010). MYB75 Functions in regulation of secondary cell wall formation in the *Arabidopsis* inflorescence stem. *Plant Physiol.* 154, 1428–1438. doi: 10.1104/pp.110.162735
- Bomal, C., Bedon, F., Caron, S., Mansfield, S. D., Levasseur, C., Cooke, J. E. K., et al. (2008). Involvement of *Pinus taeda* MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis. *J. Exp. Bot.* 59, 3925–3939. doi: 10.1093/jxb/ern234
- Carlsbecker, A., Lee, J. -Y., Roberts, C. J., Dettmer, J., Lehesranta, S., Zhou, J., et al. (2010). Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature* 465, 316–321. doi: 10.1038/nature08977
- Chai, G., Qi, G., Cao, Y., Wang, Z., Yu, L., Tang, X., et al. (2014a). Poplar PdC3H17 and PdC3H18 are direct targets of PdMYB3 and PdMYB21, and positively regulate secondary wall formation in *Arabidopsis* and poplar. *New Phytol.* 203, 520–534. doi: 10.1111/nph.12825
- Chai, G., Wang, Z., Tang, X., Yu, L., Qi, G., Wang, D., et al. (2014b). R2R3-MYB gene pairs in *Populus*: evolution and contribution to secondary wall formation and flowering time. *J. Exp. Bot.* 65, 4255–4269. doi: 10.1093/jxb/eru196
- 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
- Chen, J., Quan, M., and Zhang, D. (2015). Genome-wide identification of novel long non-coding RNAs in *Populus tomentosa* tension wood, opposite wood and normal wood xylem by RNA-seq. *Planta* 241, 125–143. doi: 10.1007/s00425-014-2168-1
- Chen, M., and Manley, J. L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* 10, 741–754. doi: 10.1038/nrm2777
- Courtois-Moreau, C. L., Pesquet, E., Sjödin, A., Muñoz, L., Bollhöner, B., Kaneda, M., et al. (2009). A unique program for cell death in xylem fibers of *Populus* stem. *Plant J.* 58, 260–274. doi: 10.1111/j.1365-3113X.2008.03777.x
- Cseke, L. J., Zheng, J., and Podila, G. K. (2003). Characterization of PTM5 in aspen trees: a MADS-box gene expressed during woody vascular development. *Gene* 318, 55–67. doi: 10.1016/S0378-1119(03)00765-0
- del Pozo, J. C., Boniotti, M. B., and Gutierrez, C. (2002). *Arabidopsis* E2F functions in cell division and is degraded by the ubiquitin-SCFAtSKP2 pathway in response to light. *Plant Cell* 14, 3057–3071. doi: 10.1105/tpc.006791
- del Pozo, J. C., Diaz-Trivino, S., Cisneros, N., and Gutierrez, C. (2006). The balance between cell division and endoreplication depends on E2FC-DPB, transcription factors regulated by the ubiquitin-SCFAtSKP2 pathway in *Arabidopsis*. *Plant Cell* 18, 2224–2235. doi: 10.1105/tpc.105.039651
- Deluc, L., Barrieu, F., Marchive, C., Lauvergeat, V., Decendit, A., Richard, T., et al. (2006). Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. *Plant Physiol.* 140, 499–511. doi: 10.1104/pp.105.067231
- Demura, T., Tashiro, G., Horiguchi, G., Kishimoto, N., Kubo, M., Matsuoka, N., et al. (2002). Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem

- cells. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15794–15799. doi: 10.1073/pnas.232590499
- Du, J., Mansfield, S. D., and Groover, A. T. (2009). The *Populus* homeobox gene ARBORKNOX2 regulates cell differentiation during secondary growth. *Plant J.* 60, 1000–1014. doi: 10.1111/j.1365-313X.2009.04017.x
- Du, J., Miura, E., Robischon, M., Martinez, C., and Groover, A. (2011). The *Populus* class III HD ZIP transcription factor POPCORONA affects cell differentiation during secondary growth of woody stems. *PLoS One* 6:e17458. doi: 10.1371/journal.pone.0017458
- Du, Q., Avci, U., Li, S., Gallego-Giraldo, L., Pattathil, S., Qi, L., et al. (2015). Activation of miR165b represses *AtHB15* expression and induces pith secondary wall development in *Arabidopsis*. *Plant J.* 83, 388–400. doi: 10.1111/tpj.12897
- Dubois, M., Van Den Broeck, L., and Inze, D. (2018). The pivotal role of ethylene in plant growth. *Trends Plant Sci.* 23, 311–323. doi: 10.1016/j.tplants.2018.01.003
- Etchells, J. P., Provost, C. M., and Turner, S. R. (2012). Plant vascular cell division is maintained by an interaction between PXY and ethylene signalling. *PLoS Genet.* 8:e1002997. doi: 10.1371/journal.pgen.1002997
- Floyd, S. K., Zalewski, C. S., and Bowman, J. L. (2006). Evolution of class III homeodomain-leucine zipper genes in streptophytes. *Genetics* 173, 373–388. doi: 10.1534/genetics.105.054239
- Fornale, S., Shi, X. H., Chai, C. L., 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
- Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., et al. (2005). EgMYB2, a new transcriptional activator from *Eucalyptus xylem*, regulates secondary cell wall formation and lignin biosynthesis. *Plant J.* 43, 553–567. doi: 10.1111/j.1365-313X.2005.02480.x
- Groover, A. T., Mansfield, S. D., Difazio, S. P., Dupper, G., Fontana, J. R., Millar, R., et al. (2006). The *Populus* homeobox gene ARBORKNOX1 reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium. *Plant Mol. Biol.* 61, 917–932. doi: 10.1007/s11103-006-0059-y
- Grotewold, E., Drummond, B. J., Bowen, B., and Peterson, T. (1994). The MYB-homologous P-gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell* 76, 543–553. doi: 10.1016/0092-8674(94)90117-1
- Guillaumie, S., Mzid, R., Mechid, V., Leon, C., Hichri, I., Destrac-Irvine, A., et al. (2010). The grapevine transcription factor WRKY2 influences the lignin pathway and xylem development in tobacco. *Plant Mol. Biol.* 72, 215–234. doi: 10.1007/s11103-009-9563-1
- He, J. B., Zhao, X. H., Du, P. Z., Zeng, W., Beahan, C. T., Wang, Y. Q., et al. (2018). KNAT7 positively regulates xylan biosynthesis by directly activating *IRX9* expression in *Arabidopsis*. *J. Integr. Plant Biol.* 60, 514–528. doi: 10.1111/jipb.12638
- Heckmann, S., Lermontova, I., Berckmans, B., De Veylder, L., Baumlein, H., and Schubert, I. (2011). The E2F transcription factor family regulates CENH3 expression in *Arabidopsis thaliana*. *Plant J.* 68, 646–656. doi: 10.1111/j.1365-313X.2011.04715.x
- Houston, K., Tucker, M. R., Chowdhury, J., Shirley, N., and Little, A. (2016). The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. *Front. Plant Sci.* 7:984. doi: 10.3389/fpls.2016.00984
- Jensen, J. K., Kim, H., Cocuron, J. C., Orler, R., Ralph, J., and Wilkerson, C. G. (2011). The DUF579 domain containing proteins IRX15 and IRX15-L affect xylan synthesis in *Arabidopsis*. *Plant J.* 66, 387–400. doi: 10.1111/j.1365-313X.2010.04475.x
- 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
- Karpinska, B., Karlsson, M., Srivastava, M., Stenberg, A., Schrader, J., Sterky, F., et al. (2004). MYB transcription factors are differentially expressed and regulated during secondary vascular tissue development in hybrid aspen. *Plant Mol. Biol.* 56, 255–270. doi: 10.1007/s11103-004-3354-5
- Kawabe, H., Ohtani, M., Kurata, T., Sakamoto, T., and Demura, T. (2018). Protein S-nitrosylation regulates xylem vessel cell differentiation in *Arabidopsis*. *Plant Cell Physiol.* 59, 17–29. doi: 10.1093/pcp/pcx151
- Kim, E. -D., and Sung, S. (2012). Long noncoding RNA: unveiling hidden layer of gene regulatory networks. *Trends Plant Sci.* 17, 16–21. doi: 10.1016/j.tplants.2011.10.008
- Kim, J., Jung, J. H., Reyes, J. L., Kim, Y. S., Kim, S. Y., Chung, K. S., et al. (2005). microRNA-directed cleavage of ATHB15 mRNA regulates vascular development in *Arabidopsis* inflorescence stems. *Plant J.* 42, 84–94. doi: 10.1111/j.1365-313X.2005.02354.x
- Kim, W. C., Kim, J. Y., Ko, J. H., Kang, H., and Han, K. H. (2014a). Identification of direct targets of transcription factor MYB46 provides insights into the transcriptional regulation of secondary wall biosynthesis. *Plant Mol. Biol.* 85, 589–599. doi: 10.1007/s11103-014-0205-x
- Kim, W. C., Ko, J. H., and Han, K. H. (2012). Identification of a cis-acting regulatory motif recognized by MYB46, a master transcriptional regulator of secondary wall biosynthesis. *Plant Mol. Biol.* 78, 489–501. doi: 10.1007/s11103-012-9880-7
- Kim, W. C., Ko, J. H., Kim, J. Y., Kim, J., Bae, H. J., and Han, K. H. (2013). MYB46 directly regulates the gene expression of secondary wall-associated cellulose synthases in *Arabidopsis*. *Plant J.* 73, 26–36. doi: 10.1111/j.1365-313X.2012.05124.x
- Kim, W. C., Reca, I. B., Kim, Y., Park, S., Thomashow, M., Keegstra, K., et al. (2014b). Transcription factors that directly regulate the expression of CSLA9 encoding mannan synthase in *Arabidopsis thaliana*. *Plant Mol. Biol.* 84, 577–587. doi: 10.1007/s11103-013-0154-9
- Ko, J. H., Jeon, H. W., Kim, W. C., Kim, J. Y., and Han, K. H. (2014). The MYB46/MYB83-mediated transcriptional regulatory programme is a gatekeeper of secondary wall biosynthesis. *Ann. Bot.* 114, 1099–1107. doi: 10.1093/aob/mcu126
- 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, 649–665. doi: 10.1111/j.1365-313X.2009.03989.x
- Ko, J. H., Prassinios, C., and Han, K. H. (2006). Developmental and seasonal expression of *PtaHB1*, a *Populus* gene encoding a class III HD-Zip protein, is closely associated with secondary growth and inversely correlated with the level of microRNA (miR166). *New Phytol.* 169, 469–478. doi: 10.1111/j.1469-8137.2005.01623.x
- Kosugi, S., and Ohashi, Y. (2002). Interaction of the *Arabidopsis* E2F and DP proteins confers their concomitant nuclear translocation and transactivation. *Plant Physiol.* 128, 833–843. doi: 10.1104/pp.010642
- Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19, 1855–1860. doi: 10.1101/gad.1331305
- Legay, S., Lacombe, E., Goicoechea, M., Briere, C., Seguin, A., Mackay, J., et al. (2007). Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Sci.* 173, 542–549. doi: 10.1016/j.plantsci.2007.08.007
- Legay, S., Sivadon, P., Blervacq, A. S., Pavy, N., Baghdady, A., Tremblay, L., et al. (2010). EgMYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in *Arabidopsis* and poplar. *New Phytol.* 188, 774–786. doi: 10.1111/j.1469-8137.2010.03432.x
- Li, C. F., Wang, X. Q., Ran, L. Y., Tian, Q. Y., Fan, D., and Luo, K. M. (2015). PtoMYB92 is a transcriptional activator of the lignin biosynthetic pathway during secondary cell wall formation in *Populus tomentosa*. *Plant Cell Physiol.* 56, 2436–2446. doi: 10.1093/pcp/pcv157
- Li, E. Y., Bhargava, A., Qiang, W. Y., Friedmann, M. C., Forneris, N., Savidge, R. A., et al. (2012). The class II KNOX gene KNAT7 negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in *Populus*. *New Phytol.* 194, 102–115. doi: 10.1111/j.1469-8137.2011.04016.x
- Li, E. Y., Wang, S. C., Liu, Y. Y., Chen, J. G., and Douglas, C. J. (2011). OVATE FAMILY PROTEIN4 (OPF4) interaction with KNAT7 regulates secondary cell wall formation in *Arabidopsis thaliana*. *Plant J.* 67, 328–341. doi: 10.1111/j.1365-313X.2011.04595.x
- Li, Q. Z., Lin, Y. C., Sun, Y. H., Song, J., Chen, H., Zhang, X. H., et al. (2012). Splice variant of the SND1 transcription factor is a dominant negative of SND1 members and their regulation in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14699–14704. doi: 10.1073/pnas.1212977109
- Liang, Y. K., Dubos, C., Dodd, I. C., Holroyd, G. H., Hetherington, A. M., and Campbell, M. M. (2005). AtMYB61, an R2R3-MYB transcription factor

- controlling stomatal aperture in *Arabidopsis thaliana*. *Curr. Biol.* 15, 1201–1206. doi: 10.1016/j.cub.2005.06.041
- Lin, Y. -C., Li, W., Sun, Y. -H., Kumari, S., Wei, H., Li, Q., et al. (2013). SND1 transcription factor–directed quantitative functional hierarchical genetic regulatory network in wood formation in *Populus trichocarpa*. *Plant Cell* 25, 4324–4341. doi: 10.1105/tpc.113.117697
- Lin, Y. J., Chen, H., Li, Q., Li, W., Wang, J. P., Shi, R., et al. (2017). Reciprocal cross-regulation of VND and SND multigene TF families for wood formation in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9722–E9729. doi: 10.1073/pnas.1714422114
- Little, C., and Savidge, R. (1987). *The Role Of Plant Growth Regulators In Forest Tree Cambial Growth*. Berlin: Springer.
- Liu, Y., Wei, M., Hou, C., Lu, T., Liu, L., Wei, H., et al. (2017). Functional characterization of *Populus PsnSHN2* in coordinated regulation of secondary wall components in tobacco. *Sci. Rep.* 7:42. doi: 10.1038/s41598-017-00093-z
- Liu, Y. Y., You, S. J., Taylor-Teeples, M., Li, W. L., Schuetz, M., Brady, S. M., et al. (2014). Bel1-like homeodomain6 and knotted *Arabidopsis thaliana*7 interact and regulate secondary cell wall formation via repression of *revoluta*. *Plant Cell* 26, 4843–4861. doi: 10.1105/tpc.114.128322
- Love, J., Bjorklund, S., Vahala, J., Hertzberg, M., Kangasjarvi, J., and Sundberg, B. (2009). Ethylene is an endogenous stimulator of cell division in the cambial meristem of *Populus*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5984–5989. doi: 10.1073/pnas.0811660106
- Lu, S., Sun, Y. -H., Shi, R., Clark, C., Li, L., and Chiang, V. L. (2005). Novel and mechanical stress–responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. *Plant Cell* 17, 2186–2203. doi: 10.1105/tpc.105.033456
- Lu, S. F., Li, Q. Z., Wei, H. R., Chang, M. J., Tunlaya-Anukit, S., Kim, H., et al. (2013). Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10848–10853. doi: 10.1073/pnas.1308936110
- Mallory, A. C., and Vaucheret, H. (2006). Functions of microRNAs and related small RNAs in plants. *Nat. Genet.* 38, S31–S36. doi: 10.1038/ng1791
- McCarthy, R. L., Zhong, R. Q., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., et al. (2010). The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. *Plant Cell Physiol.* 51, 1084–1090. doi: 10.1093/pcp/pcq064
- McCarthy, R. L., Zhong, R. Q., 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
- Mele, G., Ori, N., Sato, Y., and Hake, S. (2003). The knotted1-like homeobox gene BREVIPEDICELLUS regulates cell differentiation by modulating metabolic pathways. *Genes Dev.* 17, 2088–2093. doi: 10.1101/gad.1120003
- 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
- Mitsuda, N., and Ohme-Takagi, M. (2008). NAC transcription factors NST1 and NST3 regulate pod shattering in a partially redundant manner by promoting secondary wall formation after the establishment of tissue identity. *Plant J.* 56, 768–778. doi: 10.1111/j.1365-313X.2008.03633.x
- Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2005). The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* 17, 2993–3006. doi: 10.1105/tpc.105.036004
- Miyashima, S., Koi, S., Hashimoto, T., and Nakajima, K. (2011). Non-cell-autonomous microRNA165 acts in a dose-dependent manner to regulate multiple differentiation status in the *Arabidopsis* root. *Development* 138, 2303–2313. doi: 10.1242/dev.060491
- 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
- Nuoendagula, Tsuji, Y., Takata, N., Sakamoto, S., Nakagawa-Izumi, A., Taniguchi, T., et al. (2017). Change in lignin structure, but not in lignin content, in transgenic poplar overexpressing the rice master regulator of secondary cell wall biosynthesis. *Physiol. Plant.* 163, 170–182. doi: 10.1111/ppl.12684
- Ohashi-Ito, K., Iwamoto, K., and Fukuda, H. (2018). LOB DOMAIN-CONTAINING PROTEIN 15 positively regulates expression of VND7, a master regulator of tracheary elements. *Plant Cell Physiol.* 59, 989–996. doi: 10.1093/pcp/pcy036
- Ohashi-Ito, K., Oda, Y., and Fukuda, H. (2010). *Arabidopsis* vascular-related nac-domain6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. *Plant Cell* 22, 3461–3473. doi: 10.1105/tpc.110.075036
- Ohtani, M., Nishikubo, N., Xu, B., Yamaguchi, M., Mitsuda, N., Goue, N., et al. (2011). A NAC domain protein family contributing to the regulation of wood formation in poplar. *Plant J.* 67, 499–512. doi: 10.1111/j.1365-313X.2011.04614.x
- Olsen, A. N., Ernst, H. A., Leggio, L. L., and Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. *Trends Plant Sci.* 10, 79–87. doi: 10.1016/j.tplants.2004.12.010
- Omer, S., Kumar, S., and Khan, B. M. (2013). Over-expression of a subgroup 4 R2R3 type MYB transcription factor gene from *Leucaena leucocephala* reduces lignin content in transgenic tobacco. *Plant Cell Rep.* 32, 161–171. doi: 10.1007/s00299-012-1350-9
- Pandey, S. K., Nookaraju, A., Fujino, T., Pattathil, S., and Joshi, C. P. (2016). Virus-induced gene silencing (VIGS)-mediated functional characterization of two genes involved in lignocellulosic secondary cell wall formation. *Plant Cell Rep.* 35, 2353–2367. doi: 10.1007/s00299-016-2039-2
- Pascual, M. B., Llebres, M. T., Craven-Bartle, B., Canas, R. A., Canovas, F. M., and Avila, C. (2018). PpNAC1, a main regulator of phenylalanine biosynthesis and utilization in maritime pine. *Plant Biotechnol. J.* 16, 1094–1104. doi: 10.1111/pbi.12854
- Patzlaff, A., Mcinnis, S., Courtenay, A., Surman, C., Newman, L. J., Smith, C., et al. (2003a). Characterisation of a pine MYB that regulates lignification. *Plant J.* 36, 743–754. doi: 10.1046/j.1365-313X.2003.01916.x
- Patzlaff, A., Newman, L. J., Dubos, C., Whetten, R., Smith, C., Mcinnis, S., et al. (2003b). Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Mol. Biol.* 53, 597–608. doi: 10.1023/B:PLAN.0000019066.07933.d6
- Penfield, S., Meissner, R. C., Shoue, D. A., Carpita, N. C., and Bevan, M. W. (2001). MYB61 is required for mucilage deposition and extrusion in the *Arabidopsis* seed coat. *Plant Cell* 13, 2777–2791. doi: 10.1105/tpc.010265
- 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
- Pyo, H., Demura, T., and Fukuda, H. (2007). TERE; a novel cis-element responsible for a coordinated expression of genes related to programmed cell death and secondary wall formation during differentiation of tracheary elements. *Plant J.* 51, 955–965. doi: 10.1111/j.1365-313X.2007.03180.x
- Ragauskas, A. J., Beckham, G. T., Biddy, M. J., Chandra, R., Chen, F., Davis, M. F., et al. (2014). Lignin valorization: improving lignin processing in the biorefinery. *Science* 344:1246843. doi: 10.1126/science.1246843
- 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
- Richmond, T. (2000). Higher plant cellulose synthases. *Genome Biol.* 1, reviews3001.1–reviews3001.6. doi: 10.1186/gb-2000-1-4-reviews3001
- Robischon, M., Du, J., Miura, E., and Groover, A. (2011). The *Populus* class III HD ZIP, popREVOLUTA, influences cambium initiation and patterning of woody stems. *Plant Physiol.* 155, 1214–1225. doi: 10.1104/pp.110.167007
- 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
- Sakamoto, S., Takata, N., Oshima, Y., Yoshida, K., Taniguchi, T., and Mitsuda, N. (2016). Wood reinforcement of poplar by rice NAC transcription factor. *Sci. Rep.* 6:19925. doi: 10.1038/srep19925
- Seyferth, C., Wessels, B., Jokipii-Lukkari, S., Sundberg, B., Delhomme, N., Felten, J., et al. (2018). Ethylene-related gene expression networks in wood formation. *Front. Plant Sci.* 9:272. doi: 10.3389/fpls.2018.00272
- Sonoda, T., Koita, H., Nakamoto-Ohta, S., Kondo, K., Suezaki, T., Kato, T., et al. (2009). Increasing fiber length and growth in transgenic tobacco

- plants overexpressing a gene encoding the *Eucalyptus camaldulensis* HD-Zip class II transcription factor. *Plant Biotechnol.* 26, 115–120. doi: 10.5511/plantbiotechnology.26.115
- Soyano, T., Thitamadee, S., Machida, Y., and Chua, N. H. (2008). *ASYMMETRIC LEAVES2-LIKE19/LATERAL ORGAN BOUNDARIES DOMAIN30* and *ASL20/LBD18* regulate tracheary element differentiation in *Arabidopsis*. *Plant Cell* 20, 3359–3373. doi: 10.1105/tpc.108.061796
- Sun, X., Wang, C., Xiang, N., Li, X., Yang, S., Du, J., et al. (2017). Activation of secondary cell wall biosynthesis by miR319-targeted TCP4 transcription factor. *Plant Biotechnol. J.* 15, 1284–1294. doi: 10.1111/pbi.12715
- Sundell, D., Street, N. R., Kumar, M., Mellerowicz, E. J., Kucukoglu, M., Johnsson, C., et al. (2017). AspWood: high-spatial-resolution transcriptome profiles reveal uncharacterized modularity of wood formation in *Populus tremula*. *Plant Cell* 29, 1585–1604. doi: 10.1105/tpc.17.00153
- Sunkar, R., Chinnusamy, V., Zhu, J., and Zhu, J.-K. (2007). Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci.* 12, 301–309. doi: 10.1016/j.tplants.2007.05.001
- Syed, N. H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J. W. (2012). Alternative splicing in plants—coming of age. *Trends Plant Sci.* 17, 616–623. doi: 10.1016/j.tplants.2012.06.001
- Tang, X., Zhuang, Y., Qi, G., Wang, D., Liu, H., Wang, K., et al. (2015). Poplar PtoMYB221 is involved in the direct and indirect regulation of secondary wall biosynthesis during wood formation. *Sci. Rep.* 5:12240. doi: 10.1038/srep12240
- 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
- Tian, Q., Wang, X., Li, C., Lu, W., Yang, L., Jiang, Y., et al. (2013). Functional characterization of the poplar R2R3-MYB transcription factor PtoMYB216 involved in the regulation of lignin biosynthesis during wood formation. *PLoS One* 8:e76369. doi: 10.1371/journal.pone.0076369
- Tuskan, G. (1998). Short-rotation woody crop supply systems in the United States: what do we know and what do we need to know? *Biomass Bioenergy* 14, 307–315. doi: 10.1016/S0961-9534(97)10065-4
- Vahala, J., Felten, J., Love, J., Gorzsás, A., Gerber, L., Lamminmäki, A., et al. (2013). A genome-wide screen for ethylene-induced Ethylene Response Factors (ERFs) in hybrid aspen stem identifies ERF genes that modify stem growth and wood properties. *New Phytol.* 200, 511–522. doi: 10.1111/nph.12386
- Wang, H. H., Tang, R. J., Liu, H., Chen, H. Y., Liu, J. Y., Jiang, X. N., et al. (2013). Chimeric repressor of PtsND2 severely affects wood formation in transgenic *Populus*. *Tree Physiol.* 33, 878–886. doi: 10.1093/treephys/tpt058
- Wang, H. Z., Avci, U., Nakashima, J., Hahn, M. G., Chen, F., and Dixon, R. A. (2010). Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22338–22343. doi: 10.1073/pnas.1016436107
- Wang, H. Z., Zhao, Q., Chen, F., Wang, M. Y., and Dixon, R. A. (2011). NAC domain function and transcriptional control of a secondary cell wall master switch. *Plant J.* 68, 1104–1114. doi: 10.1111/j.1365-313X.2011.04764.x
- Wang, S. C., Li, E. Y., Porth, I., Chen, J. G., Mansfield, S. D., and Douglas, C. J. (2014). Regulation of secondary cell wall biosynthesis by poplar R2R3 MYB transcription factor PtrMYB152 in *Arabidopsis*. *Sci. Rep.* 4:5054. doi: 10.1038/srep05054
- Wierzbicki, A. T. (2012). The role of long non-coding RNA in transcriptional gene silencing. *Curr. Opin. Plant Biol.* 15, 517–522. doi: 10.1016/j.pbi.2012.08.008
- Xiao, W. H., Yang, Y., and Yu, J. J. (2018). ZmNST3 and ZmNST4 are master switches for secondary wall deposition in maize (*Zea mays* L.). *Plant Sci.* 266, 83–94. doi: 10.1016/j.plantsci.2017.03.012
- Xu, C., Fu, X., Liu, R., Guo, L., Ran, L., Li, C., et al. (2017). PtoMYB170 positively regulates lignin deposition during wood formation in poplar and confers drought tolerance in transgenic *Arabidopsis*. *Tree Physiol.* 37, 1713–1726. doi: 10.1093/treephys/tpx093
- Xu, Q., Yin, X. R., Zeng, J. K., Ge, H., Song, M., Xu, C. J., et al. (2014). Activator- and repressor-type MYB transcription factors are involved in chilling injury induced flesh lignification in loquat via their interactions with the phenylpropanoid pathway. *J. Exp. Bot.* 65, 4349–4359. doi: 10.1093/jxb/eru208
- Yamaguchi, M., Kubo, M., Fukuda, H., and Demura, T. (2008). Vascular-related NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in *Arabidopsis* roots and shoots. *Plant J.* 55, 652–664. doi: 10.1111/j.1365-313X.2008.03533.x
- Yamaguchi, M., Mitsuda, N., Ohtani, M., Ohme-Takagi, M., Kato, K., and Demura, T. (2011). VASCULAR-RELATED NAC-DOMAIN 7 directly regulates the expression of a broad range of genes for xylem vessel formation. *Plant J.* 66, 579–590. doi: 10.1111/j.1365-313X.2011.04514.x
- 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
- Yang, L., Hou, Y., Zhao, X., Lu, W., Li, Y., Yang, F., et al. (2015). Identification and characterization of a wood-associated NAC domain transcription factor PtoVNS11 from *Populus tomentosa* Carr. *Trees* 29, 1091–1101. doi: 10.1007/s00468-015-1188-1
- Yang, L., Zhao, X., Ran, L. Y., Li, C. F., Fan, D., and Luo, K. M. (2017). PtoMYB156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. *Sci. Rep.* 7:41209. doi: 10.1038/srep41209
- Yang, L., Zhao, X., Yang, F., Fan, D., Jiang, Y., and Luo, K. (2016). PtrWRKY19, a novel WRKY transcription factor, contributes to the regulation of pith secondary wall formation in *Populus trichocarpa*. *Sci. Rep.* 6:18643. doi: 10.1038/srep18643
- Yordanov, Y. S., Regan, S., and Busov, V. (2010). Members of the LATERAL ORGAN BOUNDARIES DOMAIN transcription factor family are involved in the regulation of secondary growth in *Populus*. *Plant Cell* 22, 3662–3677. doi: 10.1105/tpc.110.078634
- Zhang, J., Nieminen, K., Serra, J. A., and Helariutta, Y. (2014). The formation of wood and its control. *Curr. Opin. Plant Biol.* 17, 56–63. doi: 10.1016/j.pbi.2013.11.003
- Zhang, J., Serra, J. A. A., and Helariutta, Y. (2015). Wood development: Growth through knowledge. *Nat. Plants* 1:15060. doi: 10.1038/nplants.2015.60
- Zhang, J., Yang, Y., Zheng, K., Xie, M., Feng, K., Jawdy, S. S., et al. (2018). Genome-wide association studies and expression-based quantitative trait loci analyses reveal roles of HCT2 in caffeoylquinic acid biosynthesis and its regulation by defense-responsive transcription factors in *Populus*. *New Phytol.* 220, 502–516. doi: 10.1111/nph.15297
- Zhao, C. S., Avci, U., Grant, E. H., Haigler, C. H., and Beers, E. P. (2008). XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulose synthesis and programmed cell death in xylem. *Plant J.* 53, 425–436. doi: 10.1111/j.1365-313X.2007.03350.x
- Zhao, C. S., Lassez, T., Bako, L., Kong, D. Y., Zhao, B. Y., Chanda, B., et al. (2017). XYLEM NAC DOMAIN1, an angiosperm NAC transcription factor, inhibits xylem differentiation through conserved motifs that interact with RETINOBLASTOMA-RELATED. *New Phytol.* 216, 76–89. doi: 10.1111/nph.14704
- Zhao, Q., Nakashima, J., Chen, F., Yin, Y. B., Fu, C. X., Yun, J. F., et al. (2013). LACCASE is necessary and nonredundant with PEROXIDASE for lignin polymerization during vascular development in *Arabidopsis*. *Plant Cell* 25, 3976–3987. doi: 10.1105/tpc.113.117770
- Zhao, Q. A., Gallego-Giraldo, L., Wang, H. Z., Zeng, Y. N., Ding, S. Y., Chen, F., et al. (2010). An NAC transcription factor orchestrates multiple features of cell wall development in *Medicago truncatula*. *Plant J.* 63, 100–114. doi: 10.1111/j.1365-313X.2010.04223.x
- Zhao, Y. J., Sun, J. Y., Xu, P., Zhang, R., and Li, L. G. (2014). Intron-mediated alternative splicing of WOOD-ASSOCIATED NAC TRANSCRIPTION FACTOR1B regulates cell wall thickening during fiber development in *Populus* species. *Plant Physiol.* 164, 765–776. doi: 10.1104/pp.113.231134
- Zhong, R., McCarthy, R. L., Haghighat, M., and Ye, Z. H. (2013). The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PLoS One* 8:e69219. doi: 10.1371/journal.pone.0069219
- 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. Q., 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. Q., Lee, C. H., and Ye, Z. H. (2010a). Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends Plant Sci.* 15, 625–632. doi: 10.1016/j.tplants.2010.08.007
- Zhong, R. Q., Lee, C. H., and Ye, Z. H. (2010b). Functional characterization of poplar wood-associated NAC domain transcription factors. *Plant Physiol.* 152, 1044–1055. doi: 10.1104/pp.109.148270
- Zhong, R. Q., Lee, C. H., and Ye, Z. H. (2010c). Global analysis of direct targets of secondary wall NAC master switches in *Arabidopsis*. *Mol. Plant* 3, 1087–1103. doi: 10.1093/mp/ssq062
- Zhong, R. Q., Lee, C. H., Zhou, J. L., 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. Q., and Ye, Z. H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.* 53, 368–380. doi: 10.1093/pcp/pcr185
- Zhou, D. L., Du, Q. Z., Chen, J. H., Wang, Q. S., and Zhang, D. Q. (2017). Identification and allelic dissection uncover roles of lncRNAs in secondary growth of *Populus tomentosa*. *DNA Res.* 24, 473–486. doi: 10.1093/dnares/dsx018
- Zhou, J. L., Lee, C. H., Zhong, R. Q., 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. L., Zhong, R. Q., 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
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2018 Zhang, Xie, Tuskan, Muchero and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Regulation of Lignin Biosynthesis and Its Role in Growth-Defense Tradeoffs

Meng Xie^{1,2,3}, Jin Zhang^{1,2}, Timothy J. Tschaplinski^{1,2}, Gerald A. Tuskan^{1,2}, Jin-Gui Chen^{1,2*} and Wellington Muchero^{1,2*}

¹ Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ² Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ³ Department of Plant Sciences, University of Tennessee, Knoxville, Knoxville, TN, United States

OPEN ACCESS

Edited by:

Laura Elizabeth Bartley,
The University of Oklahoma,
United States

Reviewed by:

Huanzhong Wang,
University of Connecticut,
United States
Javier Agusti,
Instituto de Biología Molecular y
Celular de Plantas (IBMCP), Spain

*Correspondence:

Jin-Gui Chen
chenj@ornl.gov
Wellington Muchero
mucherow@ornl.gov

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 13 June 2018

Accepted: 07 September 2018

Published: 28 September 2018

Citation:

Xie M, Zhang J, Tschaplinski TJ,
Tuskan GA, Chen J-G and
Muchero W (2018) Regulation
of Lignin Biosynthesis and Its Role
in Growth-Defense Tradeoffs.
Front. Plant Sci. 9:1427.
doi: 10.3389/fpls.2018.01427

Plant growth-defense tradeoffs are fundamental for optimizing plant performance and fitness in a changing biotic/abiotic environment. This process is thought to involve readjusting resource allocation to different pathways. It has been frequently observed that among secondary cell wall components, alteration in lignin biosynthesis results in changes in both growth and defense. How this process is regulated, leading to growth or defense, remains largely elusive. In this article, we review the canonical lignin biosynthesis pathway, the recently discovered tyrosine shortcut pathway, and the biosynthesis of unconventional C-lignin. We summarize the current model of the hierarchical transcriptional regulation of lignin biosynthesis. Moreover, the interface between recently identified transcription factors and the hierarchical model are also discussed. We propose the existence of a transcriptional co-regulation mechanism coordinating energy allowance among growth, defense and lignin biosynthesis.

Keywords: phenylpropanoid, lignin, transcription factor, growth-defense tradeoffs, secondary cell wall, transcriptional co-regulation

INTRODUCTION

Lignin is a heterogeneous polymer of monolignols and is polymerized at the surface of the cell walls. The three essential monolignols of plant lignin are p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units (Boerjan et al., 2003). Lignin is important for terrestrial plants by providing structural support for the upward growth of plants and enabling the long-distance water transportation, which are essential for the evolutionary adaptation of plants from the aquatic to terrestrial environment. Lignin also provides physical and chemical protection for plants against pathogen invasion.

The biosynthesis process of lignin has attracted much attention as lignin is the major contributor to the recalcitrance of biomass feedstocks (Studer et al., 2011), dramatically increasing the cost of biomass deconstruction and biofuels production. To increase the digestibility of biomass, genetic engineering of lignin has been an active research area in the past decade (Vanholme et al., 2008) and efforts have been guided toward reducing the cost of biofuel production. On the other hand, the lignin residues after biomass saccharification can be used to produce biodegradable plastic and lignin-derived value-added solvents and chemicals (Doherty et al., 2011; Ragauskas, 2016).

Given the extensive biological and industrial importance of lignin, the understanding of lignin biosynthesis in plants is beneficial for both agricultural and industrial purposes. In this review,

we summarize the current understanding of the lignin biosynthesis process and its transcriptional regulation. We then offer a working hypothesis on the transcriptional coordination of energy flux among plant growth, defense, and lignin biosynthesis.

CANONICAL LIGNIN BIOSYNTHESIS PATHWAY

In plants, there are two major steps to produce lignin: monolignol biosynthesis and monolignol polymerization via free radical coupling. Enzymes catalyzing monolignol biosynthesis have been well defined in the model plant *Arabidopsis*. Genetic modulation of these enzymes has been shown to dramatically alter the accumulation and/or composition of lignin.

In *Arabidopsis*, monolignols are synthesized from phenylalanine via the phenylpropanoid pathway. Therefore, most key phenylpropanoid biosynthetic enzymes are also critical for lignin biosynthesis. A recent study of 221 independent transgenic *Populus* lines has demonstrated the importance of phenylpropanoid biosynthetic enzymes for lignin biosynthesis in *Populus* (Wang et al., 2018). The phenylpropanoid pathway is essential in plants, providing precursors for numerous secondary metabolites, including monolignols, flavonoids, and coumarins (Fraser and Chapple, 2011). In the phenylpropanoid pathway, three enzymes: phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), and 4-coumarate: CoA ligase (4CL), catalyze the first three steps in sequence to provide precursors for all of the downstream metabolites (Fraser and Chapple, 2011). In *Arabidopsis* and *Populus*, genetic inhibition of *PAL*, *C4H*, and *4CL* genes has been shown to significantly decrease lignin content (Rohde et al., 2004; Chen et al., 2006; Vanholme et al., 2008; Wang et al., 2018). In addition to these three enzymes, other phenylpropanoid biosynthetic enzymes, such as quinate/shikimate *p*-hydroxycinnamoyltransferase (HCT), *p*-coumaroylshikimate 3'-hydroxylase (C3'H), caffeoyl shikimate esterase (CSE), caffeic acid *O*-methyltransferase (COMT), and caffeoyl-CoA *O*-methyltransferase (CCoAOMT), which work downstream of 4CL, are also indispensable for normal lignin biosynthesis (Figure 1). In *Arabidopsis*, the down-regulation of *HCT* and *C3'H* has been shown to induce the enrichment of H units in lignin polymers (Shadle et al., 2007; Pu et al., 2009). Similarly, the down-regulation of *HCT* gene families dramatically increased the accumulation of H units in *Populus* (Wang et al., 2018). In *Arabidopsis*, *Populus*, and *Medicago truncatula*, CSE mutants were shown to deposit less lignin (Vanholme et al., 2013; Ha et al., 2016; Saleme et al., 2017). COMT has been found to be critical for the synthesis of S units (Goujon et al., 2003). The *Arabidopsis* null allele of *CCoAOMT* (*ccomt1*) also exhibits reduced lignin content, as well as reduced G units (Do et al., 2007). It is notable that these early steps of monolignol biosynthesis may have variations in monocots. A systematic study of lignin biosynthesis in switchgrass has demonstrated that the conversion of *p*-coumaroyl CoA to caffeoyl CoA in switchgrass may have an alternative route, which involves the formation of quinate esters catalyzed by HCT-like enzymes (Shen et al., 2013). In addition, the direct involvement of CCoAOMT

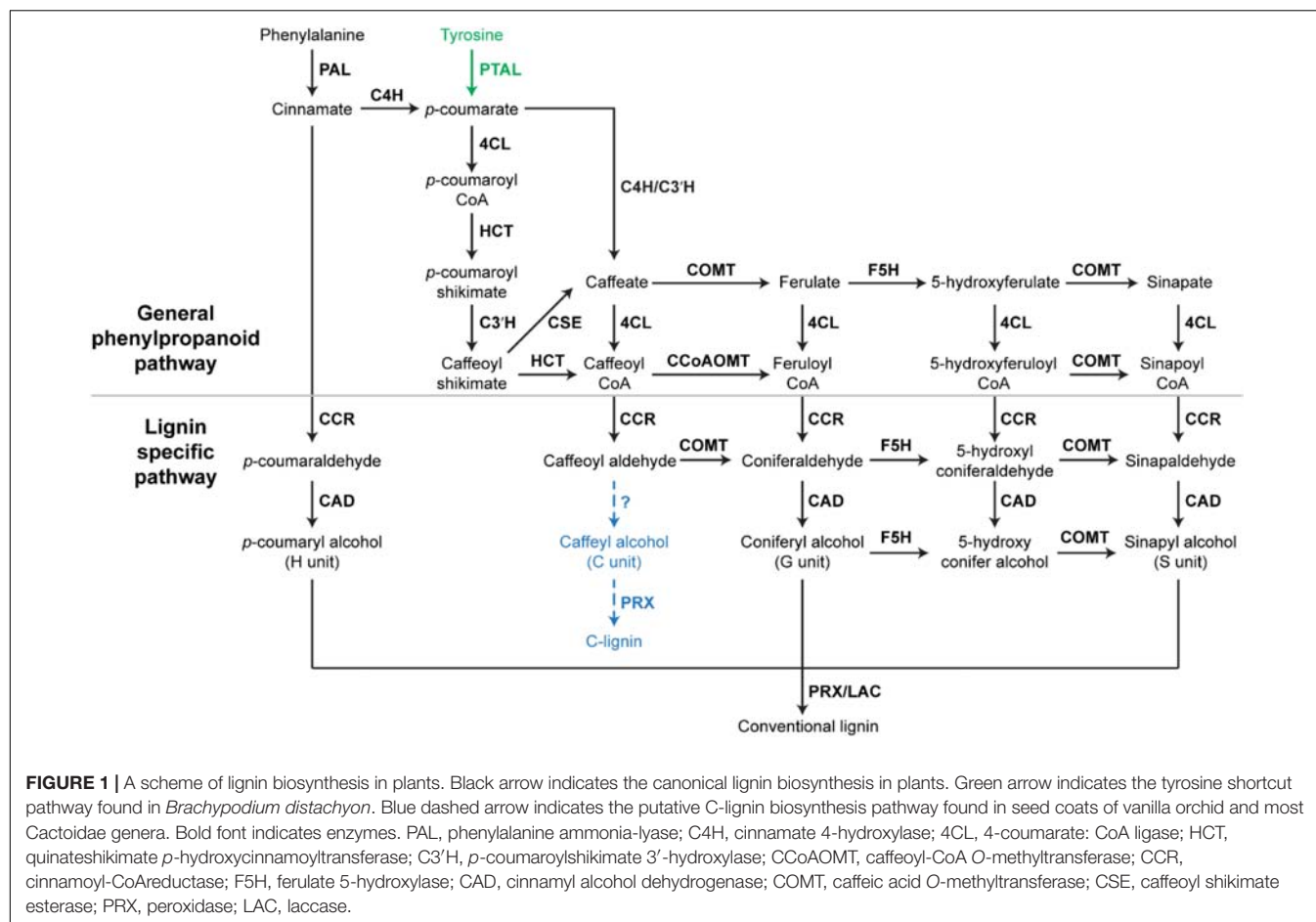
in lignin biosynthesis was questioned because the knockdown of *CCoAOMT* genes (*CCoAOMT1* and *CCoAOMT2*) did not change lignin content in switchgrass (Shen et al., 2013).

In addition to enzymes involved in the general phenylpropanoid pathway, enzymes specific for lignin biosynthesis have been identified and characterized, including cinnamoyl-CoA reductase (CCR), ferulate 5-hydroxylase (F5H), cinnamyl alcohol dehydrogenase (CAD) (Barros et al., 2015). These lignin biosynthesis-specific enzymes act downstream of phenylpropanoid biosynthetic enzymes and catalyze the biosynthesis of specific monolignols (Figure 1). The down-regulation of *F5H* can result in significant increase of G units in lignin (Chen et al., 2006). *Arabidopsis* mutants of *CAD* exhibit reduced lignin content, but increased accumulation of G and S units (Sibout et al., 2005). Besides CCR, F5H, and CAD, several lignin biosynthesis-specific steps are catalyzed by COMT, including the conversions of caffeoyl aldehyde to coniferaldehyde, 5-hydroxyl coniferaldehyde to sinapaldehyde, and 5-hydroxy conifer alcohol to sinapyl alcohol (Figure 1).

After biosynthesis, monolignols are polymerized to form lignin. It has been suggested that peroxidases and laccases are key enzymes catalyzing monolignol polymerization though experimental evidence is incomplete. Genetic studies of genes encoding peroxidase and laccase in *Arabidopsis* have illustrated the close relationship between these enzymes and lignin accumulation in secondary cell walls. Shigeto et al. (2015) found that double mutants of *Arabidopsis* peroxidases (*atprx2/atprx25*, *atprx2/atprx71*, and *atprx25/atprx71*) have 11–25% reduction in lignin content (Shigeto et al., 2015). Additionally, another peroxidase PRX17 has been found to be critical for lignin accumulation in a recent genetic study in *Arabidopsis* (Cosio et al., 2017). Similarly, laccases have been shown to affect lignin accumulation. A 20–40% reduction in lignin content has been observed in knockout mutants of two laccase genes (*LAC4* and *LAC17*) in *Arabidopsis* (Berthet et al., 2011). Moreover, the triple mutant of *LAC4*, *LAC11*, and *LAC17* completely lost lignin deposition in roots (Zhao et al., 2013).

TYROSINE SHORTCUT PATHWAY

A recent study of the model grass *Brachypodium distachyon* defined a monolignol biosynthesis process with fewer steps (Barros et al., 2016). In this process, monolignols are produced from tyrosine, which is directly converted into *p*-coumarate by a grass bifunctional phenylalanine and tyrosine ammonia-lyase (PTAL) (Barros et al., 2016). Consequently, this tyrosine shortcut of monolignol biosynthesis in grasses does not contain the steps catalyzed by C4H (Figure 1). With fewer steps, the tyrosine shortcut pathway is energetically more efficient than the canonical lignin biosynthesis pathway. The tyrosine shortcut pathway is capable of guiding carbon and electrons into the biosynthesis of lignin via skipping the production of cinnamate, which is the essential precursor of benzenoid volatiles and salicylic acid. Therefore, the discovery of the tyrosine shortcut pathway provides an alternative approach to optimize the energy



investment of lignin production. Although PTAL activity has been detected in some non-grass species, such as tobacco callus and castor bean endosperm (Gregor, 1976; Beaudoin-Eagan and Thorpe, 1985), the involvement of tyrosine shortcut pathway in lignin biosynthesis of non-grass species remains unstudied.

C-LIGNIN PATHWAY

In addition to the H, G, and S units of lignin, a natural lignin (called C-lignin) solely containing an unusual Catechyl (C) unit was found in seed coats of vanilla orchid and most Cactoidae genera (Chen et al., 2012, 2013). In C-lignin, caffeoyl alcohols are linearly connected by benzodioxane bonds via radical coupling reactions putatively catalyzed by a peroxidase (Figure 1; Chen et al., 2013). Although the detailed caffeoyl alcohol biosynthesis and polymerization processes remain unclear, the study in dicot plants demonstrated that the biosynthesis of C-lignin and conventional lignin may be controlled differently (Tobimatsu et al., 2013). C-lignin and conventional lignin were found to be synthesized in a spatially and/or temporally separated manner in seed coats of several dicot plants (Tobimatsu et al., 2013). Without side chains, the linear lignin has less cross-linking with cellulose and hemicellulose, and is capable of enhancing the digestibility of biomass. Further understanding

of the mechanism of C-lignin biosynthesis will provide an alternative bioengineering approach to generate better biomass for biofuel production. In addition, the linear C-lignin may be an ideal natural material to replace fossil-fuel based materials for the production of carbon fibers. Compared with fossil-fuel-based feedstocks, the renewable and degradable features of C-lignin make it more environmentally attractive.

TRANSCRIPTIONAL REGULATION OF LIGNIN BIOSYNTHESIS

A key finding of lignin biosynthesis studies is that AC elements widely exist in the promoters of major phenylpropanoid and lignin biosynthetic genes (Rogers and Campbell, 2004). AC elements are DNA motifs rich in adenosine and cytosine. Such significant enrichment of AC elements suggests that the phenylpropanoid and lignin biosynthesis may be under the control of specific types of transcription factors. In the past two decades, key transcription factors regulating the carbon flux into phenylpropanoid and lignin biosynthesis pathways have been identified and a hierarchical transcriptional network connecting these transcription factors has been established (Nakano et al., 2015; Figure 2).

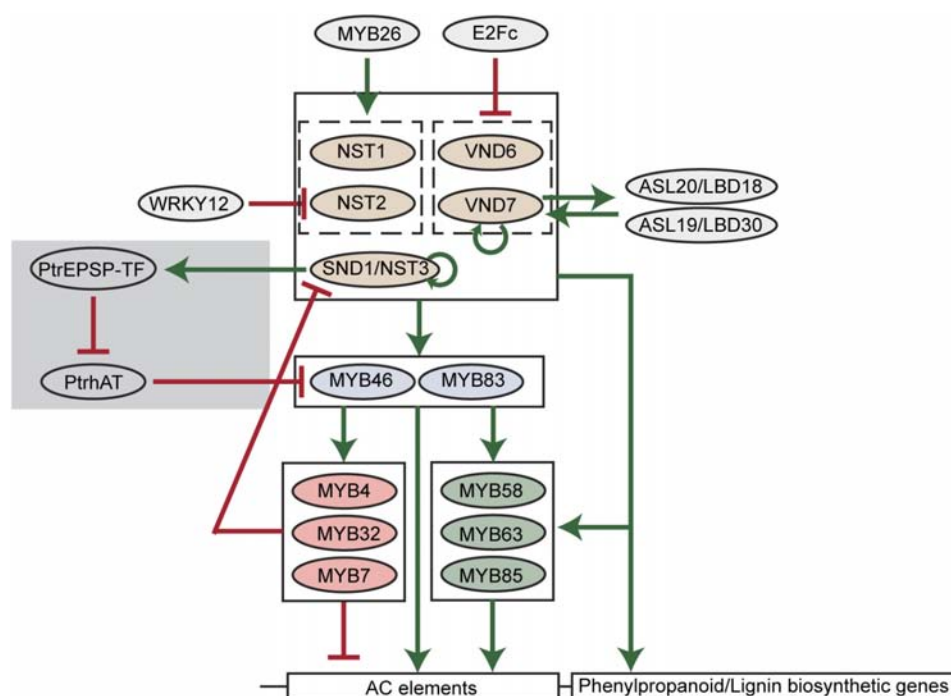


FIGURE 2 | A scheme of transcriptional regulation of lignin biosynthetic genes in plants. Green arrows indicate transcriptional activation. Red blunt arrows indicate transcriptional repression. The gray shade indicates the woody plant-specific regulatory loop.

Among transcription factors regulating phenylpropanoid and lignin biosynthesis, MYB46 and its close homolog MYB83 are well-studied in various plant species. In *Arabidopsis*, MYB46 and MYB83 have been found to activate the expression of *PAL1*, *C4H*, *4CL1*, *C3'H1*, *HCT*, *CCoAOMT*, *CCR1*, *F5H1*, *CAD6* genes via binding to AC elements in the promoters of these genes (Zhong and Ye, 2012; Kim et al., 2014). The pine and *Eucalyptus* MYB46 homologs (PtMYB4 and EgMYB2) were also found to be functional in the regulation of lignin biosynthesis (Patzlaff et al., 2003; Goicoechea et al., 2005). Moreover, four MYB46 homologs in *Populus* (PtrMYB002, PtrMYB003, PtrMYB020, and PtrMYB021) were capable of triggering ectopic lignin deposition in *Arabidopsis* and *Populus* plants (Wilkins et al., 2009; McCarthy et al., 2010; Zhong et al., 2013). In addition to lignin, the biosynthesis of other secondary cell wall components, cellulose and xylan, can also be activated by MYB46 (Zhong and Ye, 2012). Downstream of MYB46/MYB83, multiple MYB transcription factors have been identified as specific regulators of lignin biosynthesis in *Arabidopsis*, including MYB58, MYB63, MYB85, MYB4, MYB32, and MYB7. Among them, MYB58 and MYB63 are close homologs and specifically activate lignin biosynthesis via targeting AC elements (Zhou et al., 2009). By inducing *4CL* expression, MYB85 is also capable of activating monolignol biosynthesis (Zhong et al., 2008). In contrast, MYB4, MYB32, and MYB7 were found to negatively regulate the expression of lignin biosynthetic genes (Jin et al., 2000; Preston et al., 2004; Wang and Dixon, 2012).

Upstream of MYB46/MYB83, the NAC transcription factor SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN

1/NAC SECONDARY WALL THICKENING PROMOTING FACTOR 3 (SND1/NST3) and its close homologs NST1, NST2, VASCULAR-RELATED NAC DOMAIN6 (VND6), and VND7 regulate lignin biosynthesis in *Arabidopsis* (Kubo et al., 2005; Mitsuda et al., 2005; Zhong et al., 2006). Similar to MYB46/MYB83, these NAC transcription factors also regulate the biosynthesis of cellulose and xylan. In *Arabidopsis*, the regulation of secondary cell wall biosynthesis by these NAC transcription factors is cell-type specific. For example, VND6 and VND7 are specifically expressed in xylem vessel cells and regulate secondary cell wall thickening. Knockdown of VND6 or VND7 using a dominant chimeric repressor specifically inhibits the formation of metaxylem and protoxylem vessels (Kubo et al., 2005). In contrast, NST transcription factors were found to preferentially regulate the secondary cell wall biosynthesis of fiber cells [SND1/NST3 and NST1, (Zhong et al., 2006; Mitsuda et al., 2007)], silique cells [SND1/NST3 and NST1, (Mitsuda and Ohme-Takagi, 2008)], and anther endothecium [NST1 and NST2 (Mitsuda et al., 2005)]. However, the cell-type specificity seems to be absent in the woody plant *Populus* and monocots, suggesting distinct regulatory mechanisms may exist. In *Populus*, both VND and NST homologs were found to be expressed in vessels and fibers, though only homologs of VND transcription factors were found to be expressed in primary xylem vessels which have no secondary growth (Ohtani et al., 2011). In monocots, such as rice and maize, homologs of VND and NST transcription factors are named SECONDARY WALL-ASSOCIATED NAC (SWN) due to their regulatory roles in the secondary cell wall biosynthesis of vessel and fiber cells (Zhong et al., 2011).

Because of the essential regulatory roles during secondary cell wall formation, the hierarchical network comprised of NAC and MYB transcription factors is thought to be the major transcriptional regulatory mechanism of lignin and secondary cell wall biosynthesis (Nakano et al., 2015). In this network, feed-forward loops are prevalent. For example, NAC transcription factors directly activate MYB46/MYB83. Meanwhile, NAC transcription factors and MYB46/MYB83 directly activate the downstream MYB58, MYB63, MYB85, and lignin biosynthetic genes (Figure 2). Similarly, MYB46/MYB83 directly activates MYB58, MYB63, and MYB85, and they together directly activate lignin biosynthetic genes (Figure 2). These feed-forward loops are probably to ensure the robust regulation of lignin biosynthesis. Besides feed-forward loops, the NAC-MYB network also contains feed-back regulations to fine-tune lignin biosynthesis (Figure 2). For example, SND1 and VND7 were found to directly target and activate the expression of themselves in *Arabidopsis*, representing positive feed-back regulation (Wang et al., 2011; Endo et al., 2015). A negative feed-back regulatory loop, in which SND1 is targeted and down-regulated by downstream MYB transcription factors including MYB4, MYB7, and MYB32, was also defined in *Arabidopsis* (Wang et al., 2011).

Recent discoveries of new transcription factors in *Arabidopsis*, and other plant species, suggest that the regulatory network of lignin biosynthesis in plants extends beyond the established NAC-MYB network (Rao and Dixon, 2018). Among these new transcription factors, several regulate members of the NAC-MYB network. A transcriptional repressor E2Fc has been reported to directly target VND6 and VND7 in a large-scale yeast one hybrid screening (Taylor-Teeple et al., 2015). The knockdown of E2Fc using RNAi induced ectopic lignin deposition in *Arabidopsis* roots (Taylor-Teeple et al., 2015), further demonstrating that E2Fc is a negative regulator of lignin biosynthesis. Moreover, *Arabidopsis* WRKY12 was found to repress lignin biosynthesis via direct repression of the NST2 gene (Wang et al., 2010). Knockdown of WRKY12 in *Medicago* was found to enhance lignin deposition in cell walls (Wang et al., 2010). The positive regulators of SND1 and its close homologs have also been identified in *Arabidopsis*. *ASYMMETRIC LEAVES2-LIKE20/LATERAL ORGAN BOUNDARIES DOMAIN18* (ASL20/LBD18) and *ASL19/LBD30* genes are activated by VND6 and VND7 (Soyano et al., 2008). Overexpression of ASL20/LBD18 and ASL19/LBD30, in turn, enhances the expression of VND7, suggesting a positive feedback loop amplifying VND7 expression (Soyano et al., 2008). In addition, overexpression of MYB26 was found to increase lignin deposition and the expression of NST1 and NST2 (Yang et al., 2007). The discovery of these positive and negative regulators demonstrates that the NAC-MYB network is regulated to spatially and/or temporally switch on or off lignin biosynthesis.

Using genome-wide association studies (GWAS), a novel transcription factor was identified that regulates the expression of *PtrMYB021* (homolog of MYB46) in *Populus* (Xie et al., 2018). This transcription factor (PtrEPSP-TF) is a *Populus* homolog of 5-enolpyruvylshikimate 3-phosphate (EPSP) synthase, an

enzyme catalyzing the shikimate pathway. Aside from the well-established enzyme function of EPSP synthase, PtrEPSP-TF can act as a transcriptional repressor. Genetic and molecular studies further revealed that PtrEPSP-TF activates *PtrMYB021* expression and lignin biosynthesis by inhibiting the expression of a transcriptional repressor of *PtrMYB021*, which is named PtrhAT (Xie et al., 2018). Different from the ancestral EPSP synthase, the PtrEPSP-TF protein contains an additional helix-turn-helix (HTH) motif at its N-terminus, which is indispensable for the nuclear accumulation and transcriptional function of PtrEPSP-TF (Xie et al., 2018). However, the HTH motif is almost entirely missing in EPSP synthases of non-vascular plants, algae, and monocots (Xie et al., 2018). As opposed to herbaceous plants (e.g., *Arabidopsis*), woody perennial plants (e.g., *Populus*) have extensive secondary cell wall thickening over multiple growing seasons. The discovery of the additional regulatory loop of MYB46 in *Populus* supports the existence of woody plant-specific regulatory mechanisms in lignin biosynthesis. Moreover, the discovery of an activator (PtrEPSP-TF) and repressor (PtrhAT) of MYB46 provides alternative approaches to fine tune lignin biosynthesis.

LIGNIN BIOSYNTHESIS HAS COMPLEX CROSSTALK WITH GROWTH AND DEFENSE

Crosstalk among biological processes is widespread. Genetic studies in *Arabidopsis* have demonstrated crosstalk among lignin biosynthesis, growth, and defense. Meanwhile, the complexity of the crosstalk is demonstrated by these studies. To date, the relationship of lignin content with growth and defense remains unpredictable due to the limited understanding of the underlying mechanisms.

Lignin is thought to be indispensable for plant growth. The disruption of lignin biosynthesis by knocking down lignin biosynthetic genes, such as *C4H* and *CCR1*, was found to co-occur with the suppression of growth rate (Vanholme et al., 2012). However, recent studies suggest that lignin-growth crosstalk is more complex. A study of one *C3'H* mutant (*ref8-1*) illustrated that the growth suppression of *ref8-1* might be due to the flavonoid hyperaccumulation, rather than the impaired lignin biosynthesis (Besseau et al., 2007; Bonawitz et al., 2014). On the other hand, increased lignin accumulation is also harmful for plant growth. The ectopic lignin deposition induced by the overexpression of NAC and MYB transcription factors, such as SND1, MYB46/83, and MYB58/63, was observed together with inhibited growth in *Arabidopsis* (Zhong et al., 2006, 2007; Zhou et al., 2009).

The crosstalk between lignin biosynthesis and defense is even more complicated. Lignin is a well-known defense polymer, which has antimicrobial activity, forms a physical barrier to block pathogen invasion, and prevents the ingress or diffusion of toxins from pathogens (Sattler and Funnell-Harris, 2013). It has been widely observed that lignin deposition and lignin biosynthetic genes are induced during the attack of various pathogens. However, the genetic suppression of

multiple lignin biosynthesis-related genes was shown to enhance pathogen resistance. For example, *Arabidopsis* MYB46 mutants (with impaired lignin biosynthesis) displayed enhanced disease resistance, which was thought to be caused by cell wall integrity damage-triggered high immunity level (Ramírez et al., 2011). The repression of *HCT* in *Arabidopsis* and *Medicago* was found to trigger the accumulation of defense hormone salicylic acid (SA) and the expression of *PATHOGENESIS-RELATED* (*PR*) genes (Gallego-Giraldo et al., 2011a,b).

Studies on the *Arabidopsis* *HCT* mutant further illustrated the crosstalk among lignin biosynthesis, growth, and defense. *HCT* mutant has reduced lignin content, stunt growth, and enhanced defense (Gallego-Giraldo et al., 2011a). The reduced lignin content was found to trigger the accumulation of SA and the expression of *PR* genes, which subsequently enhance the pathogen resistance (Gallego-Giraldo et al., 2011a). Impaired lignin biosynthesis is generally thought to cause growth deficiency. However, in *HCT* mutant, the stunt growth was found to be unrelated with the reduced lignin content, but caused by the increased SA level (Gallego-Giraldo et al., 2011a).

THE TRANSCRIPTIONAL CO-REGULATION OF LIGNIN BIOSYNTHESIS, GROWTH, AND DEFENSE

Transcriptional co-regulation has been identified as a mechanism balancing growth and defense. In a recent study, Campos et al. (2016) found that transcriptional networks downstream of jasmonic acid (JA) and phytochrome B signaling form a hub to coordinate growth and defense. The down-regulation of both JA and phytochrome signaling (*jazQ phyB* double mutant) has been shown to uncouple growth and defense, resulting in plants displaying more growth and more defense (Campos et al., 2016).

Besides the co-regulation of growth and defense, evidence has accumulated that indicates the possible existence of transcriptional co-regulatory mechanisms coordinating growth and lignin biosynthesis, as well as defense and lignin biosynthesis. For example, the semi-dominant mutant (*ref4-3*) of one subunit of the transcriptional co-regulatory complex Mediator (*MED5b*) was found to exhibit the dwarfism phenotype, as well as repressed phenylpropanoid and lignin production (Bonawitz et al., 2012). Disruption of *MED5b* is capable of rescuing both the growth and lignin biosynthesis deficiencies of the *C3'H* mutant (Bonawitz et al., 2014), further demonstrating the involvement of transcriptional mechanisms in the co-regulation of growth and lignin biosynthesis. On the other hand, several transcription factors have been found to affect lignin biosynthesis, as well as defense responses. Overexpression of one *Medicago* WRKY transcription factor gene *WRKY W109669* in tobacco was shown to enhance the accumulation of lignin and the transcription of *PATHOGENESIS-RELATED 2* (*PR2*) gene (Naoumkina et al., 2008), which is a defense gene induced in response to virus infection (Ward et al., 1991). A recent genetic study of a *Gossypium barbadense* ethylene response-related

factor gene (*GbERF1-like*) also demonstrated the transcriptional co-regulation of defense and lignin biosynthesis. Overexpression of *GbERF1-like* in *Arabidopsis* significantly enhanced the transcription of phenylpropanoid/lignin biosynthetic genes (e.g., *PAL3*, *C4H*, *C3'H*, *HCT*, *CCoAOMT1*, *CCR1*), as well as pathogenesis-related genes including *PR3*, *PR4*, and *PLANT DEFENSIN1.2* (*PDF1.2*) (Guo et al., 2016). Moreover, the *Arabidopsis* *myb46* mutants were found to be resistant to the necrotrophic pathogen *Botrytis cinerea* and to increase *PR3* and *PDF1.2* expression after *Botrytis cinerea* infection (Ramírez et al., 2011).

More importantly, genetic studies in *Arabidopsis* have demonstrated the possibility of the transcriptional co-regulation of growth, defense, and lignin biosynthesis. In mutants of *C4H* and *CCR1*, which display significantly reduced lignin accumulation, many defense-responsive genes and many growth-related genes (i.e., auxin response genes) are down-regulated (Vanholme et al., 2012). A MADS-box transcription factor AGAMOUS-LIKE15 (*AGL15*) was found to regulate the expression of miRNA156 (Serivichyaswat et al., 2015), which negatively regulates inflorescence development and positively regulates tolerance to recurring environmental stress (Wang et al., 2009; Stief et al., 2014). A recent study further identified the regulatory role of *AGL15* in lignin biosynthesis. In a study by Cosio et al. (2017), *AGL15* was found to directly repress *PRX17* expression to regulate lignin formation.

Collectively, genetic studies have illustrated the existence of the transcriptional co-regulation mechanism coordinating growth, defense, and lignin biosynthesis. It is generally recognized that changes in lignin biosynthesis affect the structure and integration of cell walls, which in turn affects growth and defense. However, little is known about the genetic factors co-regulating growth, defense, and lignin biosynthesis or factors bridging the transcriptional co-regulations of growth-lignin biosynthesis and defense-lignin biosynthesis. This is a much-needed research area for future studies.

A PROPOSED STRATEGY TO MITIGATE GROWTH-DEFENSE TRADEOFFS BY MANIPULATING LIGNIN BIOSYNTHESIS

Growing in a dynamic environment, plants have evolved sophisticated strategies to determine the energy allowance between growth (to compete for light) and defense (to fight against pathogens). The competition for energy suggests that growth and defense of plants may have a negative relationship, which means the activation of defense processes negatively affect the growth and reproduction of plants. Such tradeoffs between growth and defense have been observed in many studies (Agrawal et al., 2012; Züst et al., 2015). Although the tradeoff between growth and defense is fundamental for plant survival in a changing environment, it is detrimental for plant yield.

Lignin biosynthesis is traditionally thought as one indivisible part of plant growth and defense, to provide structural support, transport water, and act as physical barrier. However, discoveries of transcriptional mechanisms underlying the crosstalk of lignin

biosynthesis, growth, and defense suggest the possibility to unlock lignin biosynthesis from growth and defense. On the other hand, lignin accounts for approximately 30% of the organic carbon in the biosphere (Cesarino et al., 2012). This large amount of energy invested in lignin and its precursors has the potential to compensate the costly expenditure of defense, which consequently would mitigate the tradeoff between growth and defense. Collectively, the genetic modification of the co-regulatory mechanism represents a potential strategy to overcome growth-defense tradeoffs.

PROSPECTIVE

To date, the lignin biosynthesis process and its regulatory mechanism in the model plant *Arabidopsis* are well established. Although the lignin biosynthesis process has been well-studied and monolignol biosynthetic enzymes have been systematically analyzed in woody plants, such as *Populus*, the regulation of lignin biosynthesis remains poorly understood. The regulatory mechanism in a woody perennial plant is much more complex than that in *Arabidopsis*, due to the complex genome and long life-cycle. With the application of genome-wide approaches, such as GWAS, expression quantitative trait loci (eQTL), and expression quantitative trait nucleotide (eQTN) mapping, in plant studies, it is plausible to identify causal genes affecting complex traits including lignin biosynthesis. Together with efficient post-GWAS characterization and validation, species-specific mechanisms of lignin biosynthesis can be defined.

The concept of a transcriptional hub to coordinate carbon and energy flux into growth, defense, and lignin biosynthesis is just emerging based on studies in *Arabidopsis*. A comprehensive understanding of this transcriptional hub will have significant impacts on the field of bioengineering. Current studies of NAC-MYB network are restricted in secondary cell wall biosynthesis, although RNA-seq analyses have illustrated that NAC and MYB transcription factors can induce broader gene expression changes. By identifying genome-wide targets of secondary cell wall-related NAC and MYB transcription factors using ChIP-seq, molecular mechanisms linking lignin biosynthesis, growth, and defense are expected to be discovered. In the woody plant *Populus*, although data is relatively limited, our

recent study demonstrated the transcriptional co-regulation of lignin biosynthesis and defense. A defense-responsive WRKY transcription factor was found to regulate the expression of a *Populus HCT* gene (*PtHCT2*) (Zhang et al., 2018). With the identification and characterization of transcription factors with multiple functions in the regulation of growth, defense, and lignin biosynthesis, a transcriptional network needs to be established to unveil how growth, defense, and lignin biosynthesis are co-regulated. With this knowledge, the resource flux in plants can be fine-tuned depending on human demand, which will greatly reduce the cost of agricultural and forestry biofuels production.

AUTHOR CONTRIBUTIONS

MX drafted the manuscript. JZ, TT, GT, J-GC, and WM revised the manuscript.

FUNDING

This work was supported by the Center for Bioenergy Innovation and the Plant-Microbe Interfaces Scientific Focus Area by the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

ACKNOWLEDGMENTS

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

REFERENCES

- Agrawal, A. A., Hastings, A. P., Johnson, M. T., Maron, J. L., and Salminen, J.-P. (2012). Insect herbivores drive real-time ecological and evolutionary change in plant populations. *Science* 338, 113–116. doi: 10.1126/science.1225977
- Barros, J., Serk, H., Granlund, I., and Pesquet, E. (2015). The cell biology of lignification in higher plants. *Ann. Bot.* 115, 1053–1074. doi: 10.1093/aob/mcv046
- Barros, J., Serrani-Yarce, J. C., Chen, F., Baxter, D., Venables, B. J., and Dixon, R. A. (2016). Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nat. Plants* 2:16050. doi: 10.1038/nplants.2016.50
- Beaudoin-Eagan, L. D., and Thorpe, T. A. (1985). Tyrosine and phenylalanine ammonia lyase activities during shoot initiation in tobacco callus cultures. *Plant Physiol.* 78, 438–441.
- Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cezard, L., Le Bris, P., et al. (2011). Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell* 23, 1124–1137. doi: 10.1105/tpc.110.082792
- Besseau, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B., and Legrand, M. (2007). Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. *Plant Cell* 19, 148–162. doi: 10.1105/tpc.106.044495
- Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi: 10.1146/annurev.arplant.54.031902.134938
- Bonawitz, N. D., Kim, J. I., Tobimatsu, Y., Ciesielski, P. N., Anderson, N. A., Ximenes, E., et al. (2014). Disruption of mediator rescues the stunted growth of a lignin-deficient *Arabidopsis* mutant. *Nature* 509, 376–380. doi: 10.1038/nature13084
- Bonawitz, N. D., Soltan, W. L., Blatchley, M. R., Powers, B. L., Hurlock, A. K., Seals, L. A., et al. (2012). REF4 and RFR1, subunits of the transcriptional coregulatory complex mediator, are required for phenylpropanoid homeostasis in *Arabidopsis*. *J. Biol. Chem.* 287, 5434–5445. doi: 10.1074/jbc.M111.312298

- Campos, M. L., Yoshida, Y., Major, I. T., De Oliveira Ferreira, D., Weraduwa, S. M., Froehlich, J. E., et al. (2016). Rewiring of jasmonate and phytochrome B signalling uncouples plant growth-defense tradeoffs. *Nat. Commun.* 7:12570. doi: 10.1038/ncomms12570
- Cesarino, I., Araújo, P., Domingues Júnior, A. P., and Mazzafera, P. (2012). An overview of lignin metabolism and its effect on biomass recalcitrance. *Braz. J. Bot.* 35, 303–311. doi: 10.1016/j.biotechadv.2012.03.003
- Chen, F., Srinivasa Reddy, M. S., Temple, S., Jackson, L., Shadle, G., and Dixon, R. A. (2006). Multi-site genetic modulation of monolignol biosynthesis suggests new routes for formation of syringyl lignin and wall-bound ferulic acid in alfalfa (*Medicago sativa* L.). *Plant J.* 48, 113–124. doi: 10.1111/j.1365-313X.2006.02857.x
- Chen, F., Tobimatsu, Y., Havkin-Frenkel, D., Dixon, R. A., and Ralph, J. (2012). A polymer of caffeoyl alcohol in plant seeds. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1772–1777. doi: 10.1073/pnas.1120992109
- Chen, F., Tobimatsu, Y., Jackson, L., Nakashima, J., Ralph, J., and Dixon, R. A. (2013). Novel seed coat lignins in the cactaceae: structure, distribution and implications for the evolution of lignin diversity. *Plant J.* 73, 201–211. doi: 10.1111/tpj.12012
- Cosio, C., Ranocha, P., Francoz, E., Burlat, V., Zheng, Y., Perry, S. E., et al. (2017). The class III peroxidase PRX17 is a direct target of the MADS-box transcription factor AGAMOUS-LIKE15 (AGL15) and participates in lignified tissue formation. *New Phytol.* 213, 250–263. doi: 10.1111/nph.14127
- Do, C. T., Pollet, B., Thevenin, J., Sibout, R., Denoue, D., Barriere, Y., et al. (2007). Both caffeoyl Coenzyme A 3-O-methyltransferase 1 and caffeic acid O-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in Arabidopsis. *Planta* 226, 1117–1129. doi: 10.1007/s00425-007-0558-3
- Doherty, W. O. S., Mousavioun, P., and Fellows, C. M. (2011). Value-adding to cellulosic ethanol: lignin polymers. *Ind. Crops Prod.* 33, 259–276. doi: 10.1016/j.indcrop.2010.10.022
- Endo, H., Yamaguchi, M., Tamura, T., Nakano, Y., Nishikubo, N., Yoneda, A., et al. (2015). Multiple classes of transcription factors regulate the expression of VASCULAR-RELATED NAC-DOMAIN7, a master switch of xylem vessel differentiation. *Plant Cell Physiol.* 56, 242–254. doi: 10.1093/pcp/pcu134
- Fraser, C. M., and Chapple, C. (2011). The phenylpropanoid pathway in Arabidopsis. *Arabidopsis Book* 9:e0152. doi: 10.1199/tab.0152
- Gallego-Giraldo, L., Escamilla-Trevino, L., Jackson, L. A., and Dixon, R. A. (2011a). Salicylic acid mediates the reduced growth of lignin down-regulated plants. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20814–20819. doi: 10.1073/pnas.1117873108
- Gallego-Giraldo, L., Jikumaru, Y., Kamiya, Y., Tang, Y., and Dixon, R. A. (2011b). Selective lignin downregulation leads to constitutive defense response expression in alfalfa (*Medicago sativa* L.). *New Phytol.* 190, 627–639. doi: 10.1111/j.1469-8137.2010.03621.x
- Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., et al. (2005). EgMYB2, a new transcriptional activator from eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J.* 43, 553–567. doi: 10.1111/j.1365-313X.2005.02480.x
- Goujon, T., Sibout, R., Pollet, B., Maba, B., Nussaume, L., Bechtold, N., et al. (2003). A new *Arabidopsis thaliana* mutant deficient in the expression of O-methyltransferase impacts lignins and sinapoyl esters. *Plant Mol. Biol.* 51, 973–989.
- Gregor, H. D. (1976). Studies on tyrosine ammonia-lyase activity in castor bean endosperm. *Zeitschrift Fur Pflanzenphysiologie* 77, 372–375.
- Guo, W., Jin, L., Miao, Y., He, X., Hu, Q., Guo, K., et al. (2016). An ethylene response-related factor, GBERF1-like, from *Gossypium barbadense* improves resistance to *Verticillium dahliae* via activating lignin synthesis. *Plant Mol. Biol.* 91, 305–318. doi: 10.1007/s11103-016-0467-6
- Ha, C. M., Escamilla-Trevino, L., Yancey, J. C., Kim, H., Ralph, J., Chen, F., et al. (2016). An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *Plant J.* 86, 363–375. doi: 10.1111/tpj.13177
- 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
- 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, 589–599. doi: 10.1007/s11103-014-0205-x
- Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19, 1855–1860. doi: 10.1101/gad.131305
- McCarthy, R. L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., et al. (2010). The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. *Plant Cell Physiol.* 51, 1084–1090. doi: 10.1093/pcp/pcq064
- 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
- Mitsuda, N., and Ohme-Takagi, M. (2008). NAC transcription factors NST1 and NST3 regulate pod shattering in a partially redundant manner by promoting secondary wall formation after the establishment of tissue identity. *Plant J.* 56, 768–778. doi: 10.1111/j.1365-313X.2008.03633.x
- Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2005). The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* 17, 2993–3006. doi: 10.1105/tpc.105.036004
- Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N. A., and Ohtani, M. (2015). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* 6:288. doi: 10.3389/fpls.2015.00288
- Naoumkina, M. A., He, X., and Dixon, R. A. (2008). Elicitor-induced transcription factors for metabolic reprogramming of secondary metabolism in *Medicago truncatula*. *BMC Plant Biol.* 8:132. doi: 10.1186/1471-2229-8-132
- Ohtani, M., Nishikubo, N., Xu, B., Yamaguchi, M., Mitsuda, N., Goue, N., et al. (2011). A NAC domain protein family contributing to the regulation of wood formation in poplar. *Plant J.* 67, 499–512. doi: 10.1111/j.1365-313X.2011.04614.x
- Patzlaff, A., Mcinnis, S., Courtenay, A., Surman, C., Newman, L. J., Smith, C., et al. (2003). Characterisation of a pine MYB that regulates lignification. *Plant J.* 36, 743–754. doi: 10.1046/j.1365-313X.2003.01916.x
- 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
- Pu, Y. Q., Chen, F., Ziebell, A., Davison, B. H., and Ragauskas, A. J. (2009). NMR characterization of C3H and HCT down-regulated alfalfa lignin. *Bioenergy Res.* 2, 198–208. doi: 10.1007/s12155-009-9056-8
- Ragauskas, A. J. (2016). Challenging/interesting lignin times. *Biofuels Bioprod. Biorefining Biofr* 10, 489–491. doi: 10.1002/bbb.1714
- Ramirez, V., Garcia-Andrade, J., and Vera, P. (2011). Enhanced disease resistance to *Botrytis cinerea* in myb46 Arabidopsis plants is associated to an early down-regulation of CesA genes. *Plant Signal. Behav.* 6, 911–913.
- 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
- Rogers, L. A., and Campbell, M. M. (2004). The genetic control of lignin deposition during plant growth and development. *New Phytol.* 164, 17–30. doi: 10.1111/j.1469-8137.2004.01143.x
- Rohde, A., Morreel, K., Ralph, J., Goeminne, G., Hostyn, V., De Rycke, R., et al. (2004). Molecular phenotyping of the pal1 and pal2 mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. *Plant Cell* 16, 2749–2771. doi: 10.1105/tpc.104.023705
- Saleme, M. L. S., Cesarino, I., Vargas, L., Kim, H., Vanholme, R., Goeminne, G., et al. (2017). Silencing CAFFEYOYL SHIKIMATE ESTERASE affects lignification and improves saccharification in poplar. *Plant Physiol.* 175, 1040–1057. doi: 10.1104/pp.17.00920
- Sattler, S. E., and Funnell-Harris, D. L. (2013). Modifying lignin to improve bioenergy feedstocks: strengthening the barrier against pathogens? *Front. Plant Sci.* 4:70. doi: 10.3389/fpls.2013.00070
- Serivichyaswat, P., Ryu, H. S., Kim, W., Kim, S., Chung, K. S., Kim, J. J., et al. (2015). Expression of the floral repressor miRNA156 is positively regulated by the AGAMOUS-like proteins AGL15 and AGL18. *Mol. Cells* 38, 259–266. doi: 10.14348/molcells.2015.2311

- Shadle, G., Chen, F., Srinivasa Reddy, M. S., Jackson, L., Nakashima, J., and Dixon, R. A. (2007). Down-regulation of hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry* 68, 1521–1529. doi: 10.1016/j.phytochem.2007.03.022
- Shen, H., Mazarei, M., Hisano, H., Escamilla-Trevino, L., Fu, C., Pu, Y., et al. (2013). A genomics approach to deciphering lignin biosynthesis in switchgrass. *Plant Cell* 25, 4342–4361. doi: 10.1105/tpc.113.118828
- Shigeto, J., Itoh, Y., Hirao, S., Ohira, K., Fujita, K., and Tsutsumi, Y. (2015). Simultaneously disrupting AtPrx2, AtPrx25 and AtPrx71 alters lignin content and structure in Arabidopsis stem. *J. Integr. Plant Biol.* 57, 349–356. doi: 10.1111/jipb.12334
- Sibout, R., Eudes, A., Mouille, G., Pollet, B., Lapiere, C., Jouanin, L., et al. (2005). CINNAMYL ALCOHOL DEHYDROGENASE-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. *Plant Cell* 17, 2059–2076. doi: 10.1105/tpc.105.030767
- Soyano, T., Thitamadee, S., Machida, Y., and Chua, N. H. (2008). ASYMMETRIC LEAVES2-LIKE19/LATERAL ORGAN BOUNDARIES DOMAIN30 and ASL20/LBD18 regulate tracheary element differentiation in Arabidopsis. *Plant Cell* 20, 3359–3373. doi: 10.1105/tpc.108.061796
- Stief, A., Altmann, S., Hoffmann, K., Pant, B. D., Scheible, W. R., and Baurle, I. (2014). Arabidopsis miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell* 26, 1792–1807. doi: 10.1105/tpc.114.123851
- Studer, M. H., Demartini, J. D., Davis, M. F., Sykes, R. W., Davison, B., Keller, M., et al. (2011). Lignin content in natural populus variants affects sugar release. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6300–6305. doi: 10.1073/pnas.1009252108
- Taylor-Teeple, 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
- Tobimatsu, Y., Chen, F., Nakashima, J., Escamilla-Trevino, L. L., Jackson, L., Dixon, R. A., et al. (2013). Coexistence but independent biosynthesis of catechyl and guaiacyl/syringyl lignin polymers in seed coats. *Plant Cell* 25, 2587–2600. doi: 10.1105/tpc.113.113142
- Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., et al. (2013). Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis. *Science* 341, 1103–1106. doi: 10.1126/science.1241602
- Vanholme, R., Morreel, K., Ralph, J., and Boerjan, W. (2008). Lignin engineering. *Curr. Opin. Plant Biol.* 11, 278–285. doi: 10.1016/j.pbi.2008.03.005
- Vanholme, R., Storme, V., Vanholme, B., Sundin, L., Christensen, J. H., Goeminne, G., et al. (2012). A systems biology view of responses to lignin biosynthesis perturbations in Arabidopsis. *Plant Cell* 24, 3506–3529. doi: 10.1105/tpc.112.102574
- Wang, H., Avci, U., Nakashima, J., Hahn, M. G., Chen, F., and Dixon, R. A. (2010). Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22338–22343. doi: 10.1073/pnas.1016436107
- 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, H. Z., Zhao, Q., Chen, F., Wang, M. Y., and Dixon, R. A. (2011). NAC domain function and transcriptional control of a secondary cell wall master switch. *Plant J.* 68, 1104–1114. doi: 10.1111/j.1365-3113X.2011.04764.x
- Wang, J. P., Matthews, M. L., Williams, C. M., Shi, R., Yang, C., Tunlaya-Anukit, S., et al. (2018). Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat. Commun.* 9:1579. doi: 10.1038/s41467-018-03863-z
- Wang, J. W., Czech, B., and Weigel, D. (2009). miR156-regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. *Cell* 138, 738–749. doi: 10.1016/j.cell.2009.06.014
- Ward, E. R., Payne, G. B., Moyer, M. B., Williams, S. C., Dincher, S. S., Sharkey, K. C., et al. (1991). Differential regulation of beta-1,3-glucanase messenger RNAs in response to pathogen infection. *Plant Physiol.* 96, 390–397.
- Wilkins, O., Nahal, H., Foong, J., Provart, N. J., and Campbell, M. M. (2009). Expansion and diversification of the populus R2R3-MYB family of transcription factors. *Plant Physiol.* 149, 981–993. doi: 10.1104/pp.108.13.2795
- Xie, M., Muchero, W., Bryan, A. C., Yee, K., Guo, H. B., Zhang, J., et al. (2018). A 5-enolpyruvylshikimate 3-phosphate synthase functions as a transcriptional repressor in populus. *Plant Cell* 30, 1645–1660. doi: 10.1105/tpc.18.00168
- Yang, C., Xu, Z., Song, J., Conner, K., Vizcay Barrena, G., and Wilson, Z. A. (2007). Arabidopsis MYB26/MALE STERILE35 regulates secondary thickening in the endothecium and is essential for anther dehiscence. *Plant Cell* 19, 534–548. doi: 10.1105/tpc.106.046391
- Zhang, J., Yang, Y., Zheng, K., Xie, M., Feng, K., Jawdy, S. S., et al. (2018). Genome-wide association studies and expression-based quantitative trait loci analyses reveal roles of HCT 2 in caffeoylquinic acid biosynthesis and its regulation by defense-responsive transcription factors in populus. *New Phytol.* doi: 10.1111/nph.15297 [Epub ahead of print].
- Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J., et al. (2013). Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in Arabidopsis. *Plant Cell* 25, 3976–3987. doi: 10.1105/tpc.113.117770
- Zhong, R., Demura, T., and Ye, Z. H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. *Plant Cell* 18, 3158–3170. doi: 10.1105/tpc.106.047399
- 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., 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., McCarthy, R. L., Haghighat, M., and Ye, Z. H. (2013). The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PLoS One* 8:e69219. doi: 10.1371/journal.pone.0069219
- 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.
- Zhong, R., and Ye, Z. H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.* 53, 368–380. doi: 10.1093/pcp/pcr185
- 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
- Züst, T., Rasman, S., and Agrawal, A. A. (2015). Growth–defense tradeoffs for two major anti-herbivore traits of the common milkweed *Asclepias syriaca*. *Oikos* 124, 1404–1415.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Xie, Zhang, Tschaplinski, Tuskan, Chen and Muchero. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Balancing Strength and Flexibility: How the Synthesis, Organization, and Modification of Guard Cell Walls Govern Stomatal Development and Dynamics

Yue Rui^{1,2†}, Yintong Chen^{1,3}, Baris Kandemir⁴, Hojae Yi⁵, James Z. Wang⁴,
Virendra M. Puri⁵ and Charles T. Anderson^{1,2,3*}

OPEN ACCESS

Edited by:

Marisa Otegui,
University of Wisconsin-Madison,
United States

Reviewed by:

June M. Kwak,
Daegu Gyeongbuk Institute of Science
and Technology (DGIST), South Korea
Caspar Christian Cedric Chater,
University of Sheffield,
United Kingdom

*Correspondence:

Charles T. Anderson
cta3@psu.edu

† Present Address:

Yue Rui,
Department of Biology, Stanford
University, Stanford, CA, United States

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 30 May 2018

Accepted: 26 July 2018

Published: 20 August 2018

Citation:

Rui Y, Chen Y, Kandemir B, Yi H,
Wang JZ, Puri VM and Anderson CT
(2018) Balancing Strength and
Flexibility: How the Synthesis,
Organization, and Modification of
Guard Cell Walls Govern Stomatal
Development and Dynamics.
Front. Plant Sci. 9:1202.
doi: 10.3389/fpls.2018.01202

¹ Department of Biology, The Pennsylvania State University, University Park, PA, United States, ² Intercollege Graduate Degree Program in Plant Biology, The Pennsylvania State University, University Park, PA, United States, ³ Intercollege Graduate Degree Program in Molecular Cellular and Integrative Biosciences, The Pennsylvania State University, University Park, PA, United States, ⁴ College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA, United States, ⁵ Department of Agricultural and Biological Engineering, The Pennsylvania State University, University Park, PA, United States

Guard cells are pairs of epidermal cells that control gas diffusion by regulating the opening and closure of stomatal pores. Guard cells, like other types of plant cells, are surrounded by a three-dimensional, extracellular network of polysaccharide-based wall polymers. In contrast to the walls of diffusely growing cells, guard cell walls have been hypothesized to be uniquely strong and elastic to meet the functional requirements of withstanding high turgor and allowing for reversible stomatal movements. Although the walls of guard cells were long underexplored as compared to extensive studies of stomatal development and guard cell signaling, recent research has provided new genetic, cytological, and physiological data demonstrating that guard cell walls function centrally in stomatal development and dynamics. In this review, we highlight and discuss the latest evidence for how wall polysaccharides are synthesized, deposited, reorganized, modified, and degraded in guard cells, and how these processes influence stomatal form and function. We also raise open questions and provide a perspective on experimental approaches that could be used in the future to shed light on the composition and architecture of guard cell walls.

Keywords: guard cells, plant cell wall, stomatal development, stomatal function, pectin, hemicellulose, cellulose

INTRODUCTION

One of the most crucial adaptations for plants to colonize land is the innovation of stomata over 400 million years ago (Edwards et al., 1992; Berry et al., 2010). With an earlier appearance than vascular tissues and roots (Peterson et al., 2010; Chen et al., 2017), stomata are thought to have evolved once (Raven, 2002) and exist in almost all terrestrial plants except liverworts, although some liverwort species have a 16-cell barrel-shaped structure called the air pore complex that might

serve a function similar to stomata (Jones and Dolan, 2017). The myriad programs of epidermal growth and development adopted by different species result in a diversity of stomatal ontogeny (Rudall et al., 2013). For example, in some moss species such as *Physcomitrella patens* and *Funaria hygrometrica*, guard mother cells undergo incomplete cytokinesis, which results in a single guard cell encasing a stomatal pore (Sack and Paolillo, 1985; Chater et al., 2016). In vascular plants, stomatal guard cells exist in pairs: grass species typically have two dumbbell-shaped guard cells flanked by specialized subsidiary cells, and their stomata exist in a developmental gradient along the proximodistal leaf axis, which is convergently analogous to progressive stomatal development in hornwort sporophytes (Renzaglia et al., 2017). In contrast, guard cells in most eudicots are kidney-shaped without surrounding subsidiary cells, and the stomata in eudicot leaves vary in age and are oriented randomly within the same leaf region (Rudall et al., 2013; **Figure 1**). Despite a debatable function in gas exchange vs. sporophyte dehiscence in moss species (Merced, 2015; Chater et al., 2016), stomatal complexes in land plants are canonically thought to serve as epidermal valves that open and close repeatedly and reversibly to respond to various stimuli in a changing terrestrial environment. For more thorough and detailed overviews on the cell differentiation and division events during stomatal development and the signal transduction networks that underlie stomatal movements, we recommend other nicely written reviews (Fan et al., 2004; Bergmann and Sack, 2007; Casson and Hetherington, 2010; Kim et al., 2010; Pillitteri and Torii, 2012; Hepworth et al., 2018). In this update, we will focus on the walls that surround guard cells and discuss their functions and dynamics during pore formation and stomatal movements.

THE PRIMARY WALL

Growing plant cells are encased by a three-dimensional cell wall called the primary wall, wherein cellulose is embedded in a matrix containing hemicelluloses, pectins, and structural proteins (Somerville et al., 2004). The biosynthesis, modification, degradation, and reorganization of wall polymers and their interactions make the primary wall quite complex and dynamic (Voiniciuc et al., 2018). The composition of the primary wall is diverse across plant species. For instance, there are two major types of primary walls in flowering plants, based on the relative amounts and types of matrix polymers. Most eudicots such as *Arabidopsis thaliana* (*Arabidopsis*), and non-commelinoid monocots, possess a Type I cell wall, with xyloglucan being the predominant hemicellulose and pectins composing 20–35% dry weight of the wall; in contrast, Type II cell walls are typical in commelinoid monocots such as grasses, and contain xylans and mixed-linkage glucans as the major hemicelluloses and much less pectin than Type I cell walls (Jones et al., 2005; Vogel, 2008).

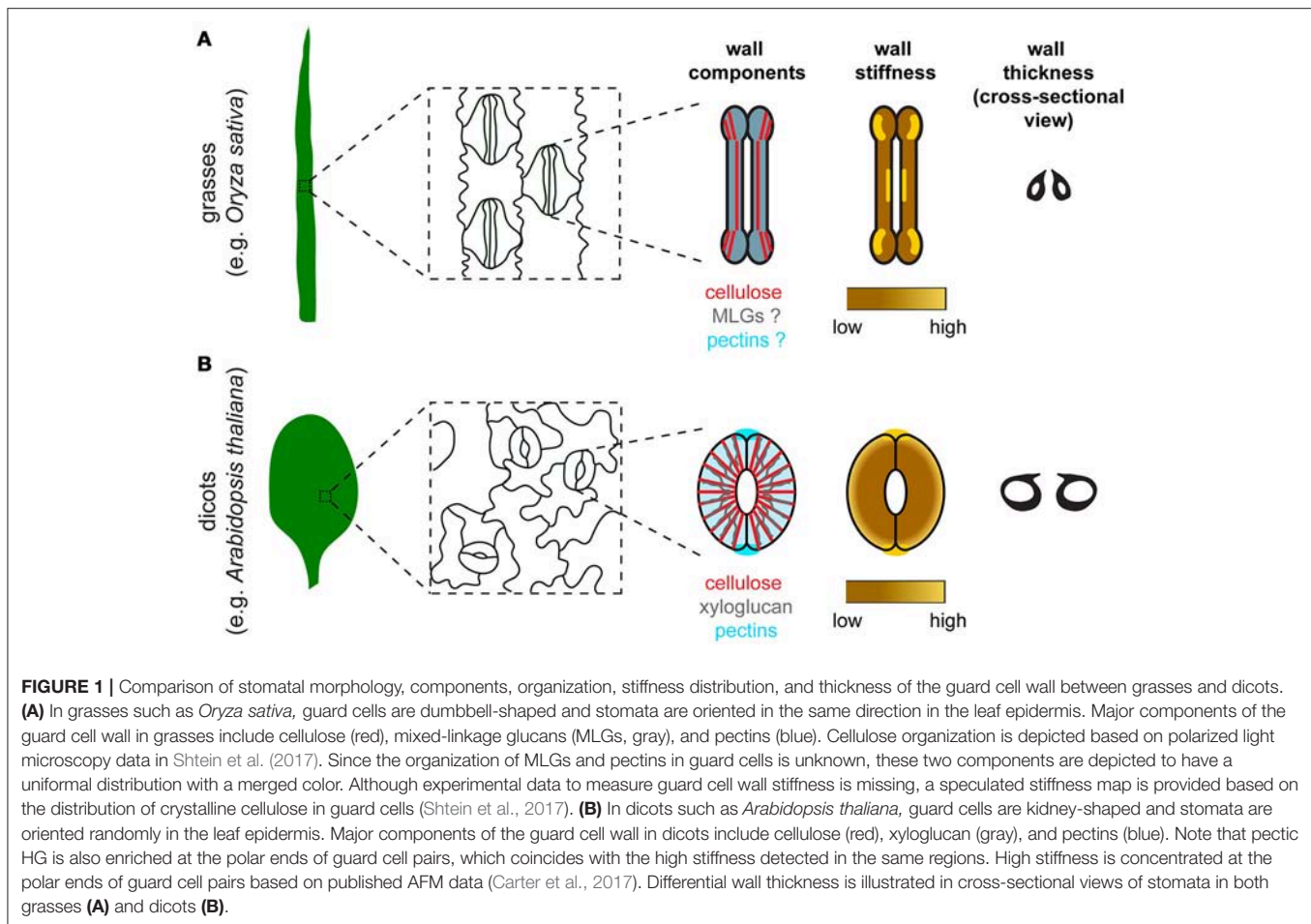
For a given plant cell, wall composition undergoes spatiotemporal changes during cell development and differentiation, with older polymers such as middle lamellar pectins being deposited earlier and thus being farther from the

plasma membrane, and nascent materials being laid down later and thus being closer to the cell surface (Keegstra, 2010). Cell growth in the short term, such as over a few minutes, can involve large-scale reorientations of wall components (Anderson et al., 2010).

Cellulose is synthesized at the cell surface by plasma membrane-localized cellulose synthase complexes (CSCs) (Paredez et al., 2006). CSCs move along linear trajectories that co-align with cortical microtubules (MTs), but the presence of MTs is not a prerequisite for CSC motility (Paredez et al., 2006). Cellulose is the most ordered wall polymer and is often oriented transversely to the growth axis of a cell, providing tensile strength to the wall (Green, 1962). Hemicelluloses (e.g., xyloglucan) and pectins are synthesized in the Golgi and secreted to the apoplast (Wolf et al., 2009; Pauly and Keegstra, 2016). Xyloglucan can intertwine with cellulose, forming junctions that serve as mechanical hotspots for wall loosening (Park and Cosgrove, 2012a,b). Xyloglucan in extended conformations can also bind to the hydrophobic faces of cellulose (Zheng et al., 2018).

Pectins are structurally complex polymers composed of the following domains: homogalacturonan (HG), rhamnogalacturonan-I (RG-I), rhamnogalacturonan-II (RG-II), xylogalacturonan, and apiogalacturonan (Mohnen, 2008). HG is the simplest and most abundant pectin domain. HG is synthesized and methyl-esterified in the Golgi by galacturonosyltransferases (GAUTs) and pectin methyltransferases (PMTs), respectively (Mohnen, 2008; Wolf et al., 2009). Highly methyl-esterified HG is exocytosed to the wall where it is then de-methyl-esterified by pectin methylesterases (PMEs) (Wolf et al., 2009). The methyl-esterification status of HG is also affected by endogenous pectin methylesterase inhibitors (PMEIs), which antagonize the activity of PMEs (Jolie et al., 2010). Different de-methyl-esterification patterns can lead to opposing effects on wall mechanics: blockwise de-methyl-esterification usually facilitates HG crosslinking via Ca^{2+} , thus contributing to wall stiffening, whereas random de-methyl-esterification makes HG susceptible to degradation by polygalacturonases (PGs) or pectate lyases (PLs), resulting in wall loosening (Hocq et al., 2017; **Figure 2**). In model species such as *Arabidopsis*, genes encoding these pectin-modifying and -degrading enzymes all exist in large families (McCarthy et al., 2014), few of which have been functionally and/or biochemically characterized.

We have gained our knowledge of the primary wall predominantly from studies in tissue types that undergo irreversible expansion, such as roots (Anderson et al., 2010, 2012), etiolated hypocotyls (Paredez et al., 2006; Desprez et al., 2007), and shoots (Peaucelle et al., 2011; Braybrook and Peaucelle, 2013), but less so in guard cells that undergo reversible shape changes, despite the longstanding hypothesis that guard cell walls must possess unique material properties to allow for cycles of stomatal movements (Wu and Sharpe, 1979). Below, we will present an update on recent research and suggest future directions to advance our understanding of the molecular details of how guard cell walls are built to allow for fast and reversible stomatal movements.



COMPOSITION AND SYNTHESIS OF THE GUARD CELL WALL

The polysaccharide components of guard cell walls have been identified mostly by imaging experiments, including polarized light microscopy for cellulose, and immunolabeling coupled with fluorescence microscopy or transmission electron microscopy in thin sections for hemicelluloses and pectins (Palevitz and Hepler, 1976; Majewska-Sawka et al., 2002; Jones et al., 2003; Merced and Renzaglia, 2014; Amsbury et al., 2016; Giannoutsou et al., 2016; Shtein et al., 2017; **Figure 1**). These techniques, although they are not quantitative, can reveal differences in wall composition between guard cells and neighboring cells. For example, LM15 antibody-labeled xyloglucan is more enriched in guard cells than in neighboring epidermal cells in *Arabidopsis* (Amsbury et al., 2016), and LM6-labeled 1,5- α -L-arabinan is present in guard cells but not subsidiary cells in *Zea mays* (Giannoutsou et al., 2016). Polysaccharide components of the guard cell wall that are conserved across species include cellulose, HG, and RG-I. Pectic arabinan, in particular, has been demonstrated to maintain the flexibility of guard cell walls in various species, since exogenous treatment with arabinanase in epidermal strips prevents stomatal

opening or closure in species such as *Commelina communis* and *Vicia faba* (Jones et al., 2003, 2005).

Cell wall structural proteins have also been found to be present in the guard cell wall by functional characterizations or immunolabeling approaches. In *Arabidopsis*, *FUSED OUTER CUTICULAR LEDGE1* (*FOCL1*) encodes a putative cell wall glycoprotein that is required for the formation of the stomatal outer cuticular ledge (Hunt et al., 2017). Plants lacking *FOCL1* have larger stomata and are impaired in controlling stomatal aperture and transpiration rate (Hunt et al., 2017), suggesting that wall structural proteins and cuticular ledges might affect stomatal dynamics. In *Z. mays*, arabinogalactan proteins have been detected in the walls encasing guard cells, but not in subsidiary cells (Giannoutsou et al., 2016).

Despite the visualization of representative wall components, a global, quantitative analysis of guard cell wall composition is still missing. This is largely due to technical difficulties in isolating and enriching enough guard cell wall materials for quantitative compositional assays. Knowing the relative amount of each wall component will aid the comparison of wall constitution between guard cells and other cell types, between dicots and monocots, and between wild type and mutant plants.

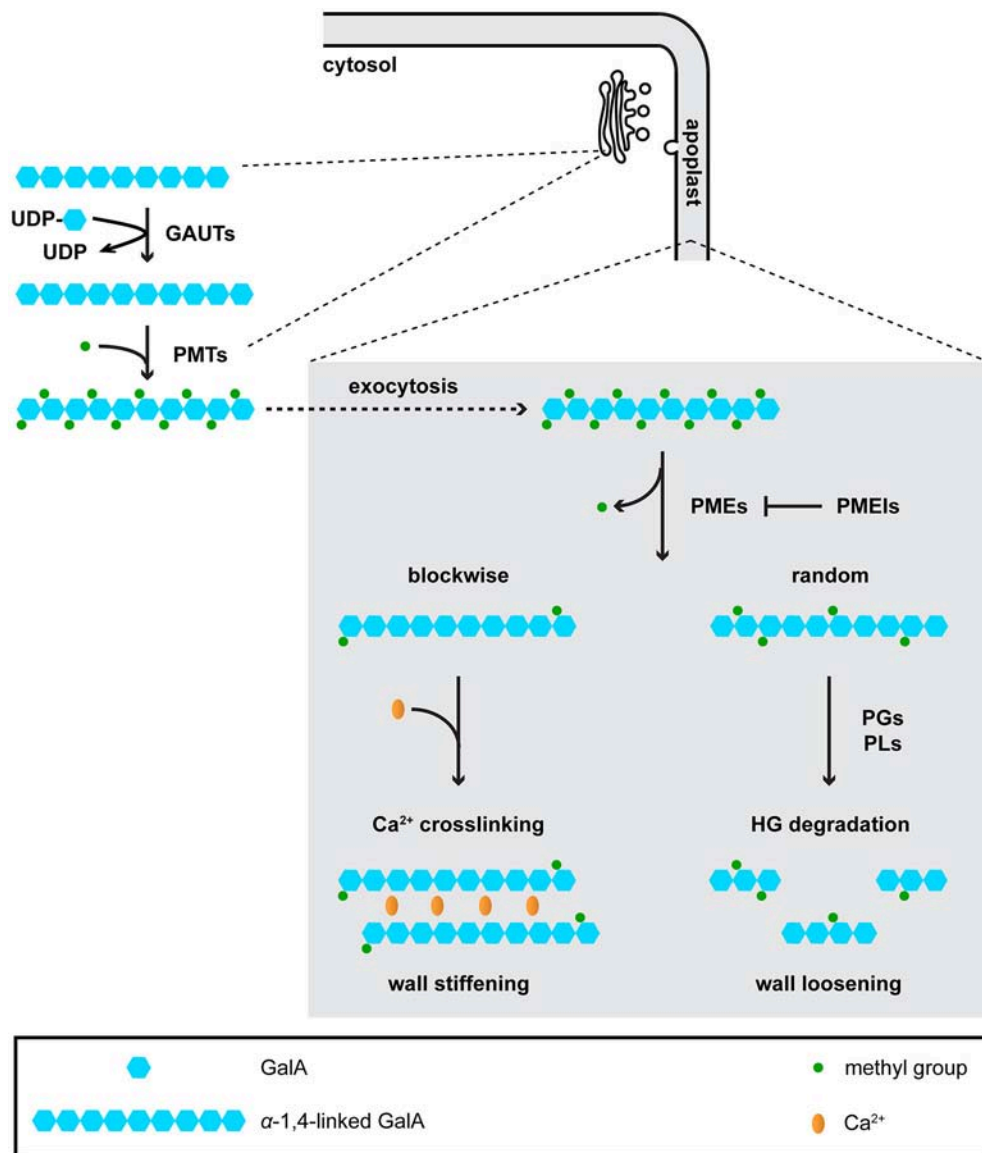


FIGURE 2 | Homogalacturonan (HG) is synthesized in the Golgi, and is de-methyl-esterified and degraded in the apoplast. In the Golgi, galacturonosyltransferases (GAUTs) transfer galacturonic acid (GalA) residues onto existing α -1,4-linked GalA chains. Pectin methyltransferases (PMTs) add methyl groups onto GalA residues. Although it is currently unknown whether PMTs function after GAUTs or PMTs and GAUTs act as a protein complex, the first scenario is shown in the figure. Highly methyl-esterified HG is then exocytosed to the apoplast, where it is de-methyl-esterified by pectin methyl-esterases (PMEs). De-methyl-esterified HG can be crosslinked by Ca^{2+} , or subject to degradation by polygalacturonases (PGs) and pectate lyases (PLs).

Guard cell walls are synthesized and deposited in the apoplast during stomatal development (**Movie S1**). As a result, their thickness gradually and differentially increases at different regions as stomata mature, with outer and inner periclinal walls eventually being thicker than ventral and dorsal walls (Zhao and Sack, 1999; Merced and Renzaglia, 2014; **Figure 1**). However, the molecular details of how these differentially thickened walls are synthesized and how their synthesis is spatially controlled are not clearly understood, raising several questions, e.g., which glycosyltransferases are expressed during stomatal development?; how are their activities spatiotemporally regulated?; and is there

a bias of their subcellular localization at different regions within a guard cell? Transcriptomic datasets of stomatal lineage cells (Hachez et al., 2011; Adrian et al., 2015) are open resources for the search of genes encoding such glycosyltransferases, but proteomic analyses will be required to globally predict their activity levels at each stage of stomatal development.

Cellulose is actively synthesized in young guard cells, and is likely to contribute to the build-up of wall strength to withstand the high turgor pressure inside a guard cell. In *Arabidopsis*, although genes encoding primary wall-associated cellulose synthases (CESAs) are not highly expressed in stomatal

lineage cells (Adrian et al., 2015), fluorescent protein (FP)-tagged CESAs in guard cells in young tissues are actively moving along linear trajectories, which mirror the distribution pattern of cortical MTs that radiate out from the stomatal pore (Rui and Anderson, 2016). Upon a short-term dark treatment, the co-localization between FP-CESAs and MTs is reduced in guard cells in young tissues, suggesting that some CSCs might dissociate from MT “rails” during stomatal closure (Rui and Anderson, 2016). Future studies of other glycosyltransferase families such as the cellulose synthase-like C (CSLC) family and the GAUT family, which are required for the synthesis of xyloglucan and pectins, respectively (Cocuron et al., 2007; Mohnen, 2008), will shed light on how matrix polysaccharides are produced during stomatal development.

ORGANIZATION OF THE GUARD CELL WALL

Some components of the guard cell wall are conserved across terrestrial plants, but their distribution patterns can be distinct in different species. For example, in ferns, dicots, and some monocots where stomata are kidney-shaped, cellulose exhibits an overall radial arrangement (Palevitz and Hepler, 1976; Fujita and Wasteneys, 2014; Rui and Anderson, 2016; Shtein et al., 2017), whereas in grasses where stomata are dumbbell-shaped, radially oriented cellulose is evident only in the polar regions (Shtein et al., 2017; **Figure 1**). Compared to cellulose, there are fewer studies on the spatial organization of matrix polysaccharides in guard cells, although stretches of de-methyl-esterified HG have been reported to be diffusely distributed in the periclinal wall, but enriched at the polar ends of guard cell pairs in *Arabidopsis* (Carter et al., 2017; Rui et al., 2017; **Figure 1**).

Visualization of the organization of the guard cell wall is challenging, partly due to the particularly thick cuticles that prevent penetration of many probes (Voiniciuc et al., 2018) and the projection or two-dimensional information gained by some imaging approaches such as polarized microscopy, field emission scanning electron microscopy (FESEM), and immunolabeling in thin sections. To address these issues and to learn how individual components of guard cell walls are distributed in 3D, one direction is to develop and apply a library of small fluorescent dyes (Anderson and Carroll, 2014) that can penetrate the cuticle and bind to specific wall components in guard cells. Dyes that are compatible with super-resolution imaging techniques such as structured illumination microscopy (SIM) (Gustafsson, 2005) and stochastic optical reconstruction microscopy (STORM) (Huang et al., 2008) would facilitate more finely detailed investigations of cell wall organization in intact, hydrated guard cells. Results from these dye-based imaging experiments should be interpreted with the caution that the dye might alter the function of the wall component to which it binds.

Given that stomata open and close on a time scale of minutes and that synthesizing and/or degrading substantial amounts of wall components during every cycle of stomatal movement would be metabolically expensive (Zhang et al., 2011), one might wonder which dynamic process(es) occur in the guard cell wall

to allow for rapid changes in stomatal shape. One hypothesis we have raised is that stomatal movement is accompanied by the dynamic reorganization of wall components in guard cells (**Figure 3**; **Movie S1**). One piece of data that supports this hypothesis is that cellulose microfibrils in intact guard cells of *Arabidopsis* exhibit a relatively even distribution when stomata are open, but become more bundled and evidently fibrillar when stomata are closed (**Figure 3**; **Movie S1**) (Rui and Anderson, 2016). Such a change in cellulose organization is aberrant in the *CELLULOSE SYNTHASE3* (*CESA3*) mutant, *cesa3^{je5}*, that is deficient in cellulose (Desprez et al., 2007) and in a double mutant that lacks the expression of *XYLOGLUCAN XYLOSYLTRANSFERASE1* (*XXT1*) and *XXT2*, *xxt1 xxt2*, which is deficient in xyloglucan (Cavaler et al., 2008). In addition, stomatal apertures during stomatal movements are larger in *cesa3^{je5}* mutants but smaller in *xxt1 xxt2* mutants compared to wild type controls (Rui and Anderson, 2016). These observations suggest that the construction of a wall that facilitates cellulose reorganization and proper control of stomatal aperture depends on sufficient levels of cellulose and xyloglucan (**Movie S1**) (Rui and Anderson, 2016). In addition to cellulose reorganization, we also proposed that during stomatal movements, pectins might undergo remodeling from being un-crosslinked in the open state to crosslinked in the closed state (**Figure 3**; **Movie S1**) (Rui et al., 2017), a process that should be distinguished from the metabolic turnover of pectins (i.e., their synthesis, deposition in the apoplast, and degradation). To further test the above hypotheses, high-resolution imaging of multiple wall constituents in living guard cells during the movements of individual stomata will be needed to reveal and quantify any spatiotemporal changes in nanoscale wall organization.

POST-SYNTHESIS MODIFICATION OF THE GUARD CELL WALL

Comparative transcriptomic analyses between wild type *Arabidopsis* and mutants lacking or overexpressing a transcription factor that governs cell differentiation during stomatal development such as *STOMATAL CARPENTER1* (*SCAP1*) (Negi et al., 2013) or between different developmental stages in the stomatal lineage (Adrian et al., 2015) make possible the identification of genes encoding cell wall-modifying proteins that are likely to function in stomata. For example, *PECTIN METHYLESTERASE6* (*PME6*) is downregulated in a loss-of-function mutant of *SCAP1*, a transcription factor that is essential for maintaining the proper shape of guard cells and controlling stomatal conductance (Negi et al., 2013). A transposon insertional mutant of *PME6* exhibits a narrower range of stomatal conductance in response to changes in CO₂ level or light intensity, as demonstrated by two independent studies (Negi et al., 2013; Amsbury et al., 2016). Although *PME6* has not yet been biochemically confirmed to act as a PME, methyl-esterified pectin epitopes are more abundant in *pme6* guard cell walls than in wild type controls, suggesting a link between cell wall composition and stomatal function in guard cells (Amsbury et al., 2016). Alternatively, cell wall-modifying

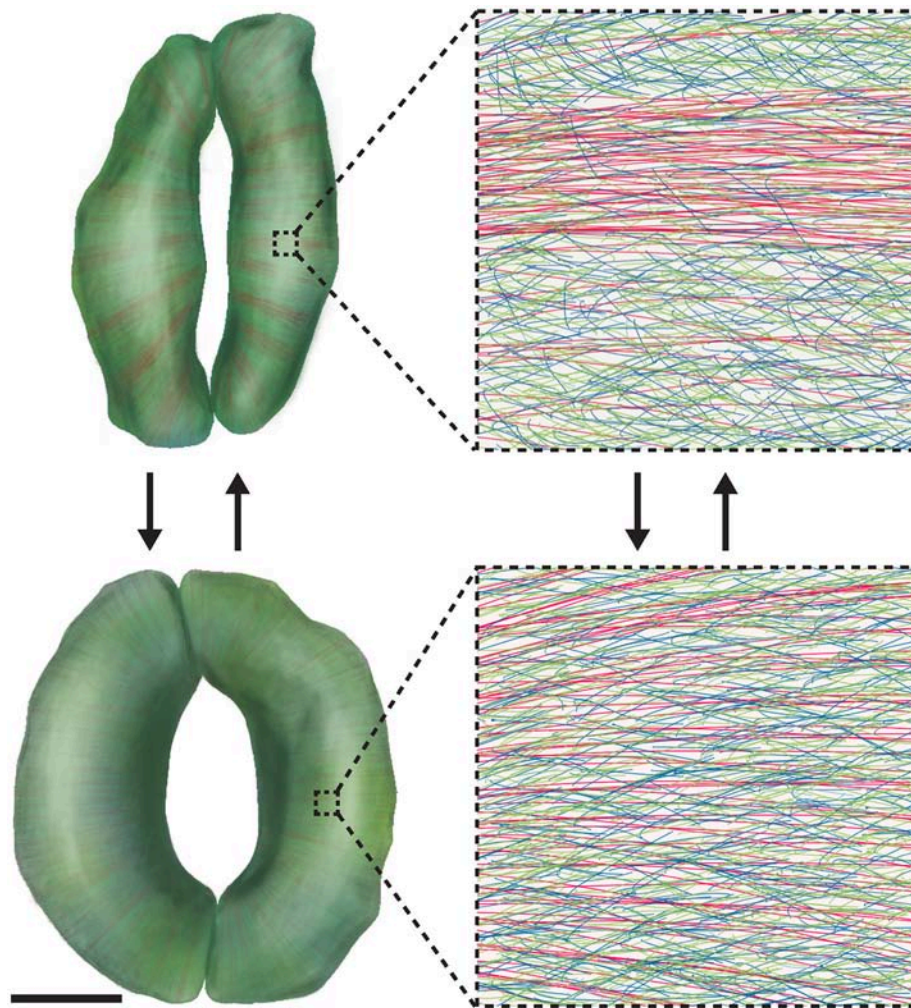


FIGURE 3 | Wall Reorganization in Guard Cells Might Underlie Stomatal Dynamics. Mature guard cells might use wall materials that have been deposited to meet their functional requirement for repeated elastic deformation. 3D renderings of closed (**Top**) and open (**Bottom**) stomatal complexes are shown on the left, with cellulose shown in red, xyloglucan shown in green, and pectins shown in blue. Note that 3D renderings were made via computerized segmentation of 3D microscopic images of actual guard cells. Scale bar for images on the left is 5 μm ; insets on the right show the detailed molecular organization of guard cell walls in $1 \times 1 \mu\text{m}$ regions. During stomatal opening, cellulose becomes diffusely distributed, and pectic homogalacturonan networks are uncoupled. During stomatal closure, cellulose becomes bundled and pectic homogalacturonan is crosslinked by Ca^{2+} . See also **Movie S1**.

genes that are more highly expressed in guard cells than in guard mother cells are candidates for functional characterizations in stomatal maturation (Adrian et al., 2015).

Genes encoding wall-modifying proteins can also be identified by screening using assays to test stomatal physiology. In a thermotolerance screen using a collection of *Arabidopsis* PME mutants, a *pme34* mutant was found to be less tolerant to heat stress than wild type controls (Huang et al., 2017). *PME34* is expressed in guard cells and encodes a biochemically active PME as indicated by changes in total PME activity *in planta* in *pme34* mutants and *PME34* overexpression lines (Huang et al., 2017). These data suggest that disrupting pectin modification by *PME34* might impair water evaporation through stomata during heat stress. Because de-methyl-esterified HG produced by PMEs can have contrasting mechanical effects on the wall depending on

how the de-methyl-esterification occurs (Hocq et al., 2017), it is currently unknown how these PMEs alter the biomechanics of the guard cell wall. Because PMEs enzymatically remove methyl groups from the pectin backbone in the apoplast whereas putative PMTs are localized in the Golgi (Mouille et al., 2007; Kim et al., 2015; Xu et al., 2017; **Figure 2**), it is unlikely that a de-methylesterification/methylesterification cycle occurs on pectins in the apoplast during fast and reversible stomatal movements in mature guard cells. Instead, the de-methyl-esterification patterns of HG generated by PME activity and their mechanical effects on the guard cell wall are more likely established during guard cell morphogenesis (**Movie S1**). Such effects persist in the walls of mature guard cells to allow for repetitive stomatal movements.

Other cell wall-modifying proteins act non-enzymatically, and include expansins, which can cause pH-dependent wall

loosening and extension (Cosgrove, 2000). *EXPANSIN1* (*EXPA1*) is expressed in Arabidopsis guard cells, and its overexpression accelerates light-induced stomatal opening by reducing the volumetric elastic modulus of the guard cell, which likely reflects the effect of EXPA1 on the mechanical properties of the cell wall (Zhang et al., 2011). Other expansin gene candidates that might function in the guard cell wall include Arabidopsis *EXPANSIN4*, *EXPANSIN5* (Zhang et al., 2011), and *EXPANSIN9* (Negi et al., 2013). These experimental data open up the possibility that expansin-mediated wall loosening might be another dynamic process that acts independently or synergistically with the reorganization of wall components during stomatal movement. However, it remains to be tested what process would occur during stomatal closure to revert the loosening effect of expansin, an effect we would call “wall tightening” for which there is currently little experimental evidence.

DEGRADATION OF THE GUARD CELL WALL IN RELATION TO STOMATAL DEVELOPMENT AND MOVEMENT

Plant cell walls can be degraded by exogenous or endogenous glycoside hydrolases (GHs). Exogenous treatment with arabinanase in epidermal strips prevents stomatal opening or closure in species such as *C. communis* and *V. faba* (Jones et al., 2003, 2005). However, stomatal function remains normal in Arabidopsis mutants that lack an endogenous arabinan biosynthetic gene, *ARABINAN DEFICIENT1* (*ARAD1*), and have a 25% reduction in arabinan content in leaves (Harholt et al., 2006). Given that *ARAD1* is in a subgroup of glycosyltransferase family 47 (GT47) that has seven other members (Li et al., 2004; Harholt et al., 2006), it would be interesting to see whether mutants with more severe deficiencies in arabinan biosynthesis or plants that overexpress endogenous arabinanase-encoding genes have any defects in stomatal function.

The final step of stomatal development requires partial separation of the wall between sister guard cells (Bergmann and Sack, 2007), which based on analogous cell separation events in other plant tissues (Liljegren, 2012) likely involves pectin degradation in the middle lamella. However, virtually no data have been reported to support this hypothesis, although pectic strands have been shown to be present in newly formed stomata (Carr et al., 1980). Recently, our group characterized the function of *POLYGALACTURONASE INVOLVED IN EXPANSION3* (*PGX3*) in Arabidopsis stomata. In cotyledons, GFP-tagged *PGX3* is enriched at sites of stomatal pore initiation and *PGX3* expression is associated with pore dimensions, suggesting that pectin degradation by *PGX3* contributes to the controlled cell separation between sister guard cells during stomatal pore formation (Rui et al., 2017). It remains to be tested whether additional mechanisms other than pectin degradation exist to separate sister guard cells at pore initiation sites, and how pectin degradation is spatially restricted to facilitate pore formation while retaining strong connections between the ends of sister guard cells at their poles (Carter et al., 2017).

Overexpression of an apple *POLYGALACTURONASE* (*PG*) leads to malfunctioning stomata, possibly due to the presence of smaller pectins and/or holes at one or both ends of stomata in transgenic leaves (Atkinson et al., 2002), but it is unclear which phenotype is the cause of defects in stomatal function in transgenic plants. In adult true leaves of Arabidopsis, *PGX3* regulates stomatal dynamics by fine-tuning the abundance of de-methyl-esterified HG and pectin molecular size, providing a molecular explanation for how pectins maintain the flexibility of guard cell walls during stomatal movement (**Movie S1**) (Rui et al., 2017). In addition to PGs, it would also be worthwhile to extend functional characterizations to genes encoding other wall-degrading enzymes, such as pectate lyases and glucanases.

APPROACHES TO STUDYING THE MECHANICS OF GUARD CELL WALLS

In addition to conventional methods such as fluorescence/electron microscopy and functional characterization of genes to investigate the guard cell wall, there are many approaches that have been applied in the cell wall field, but have not been fully exploited to investigate the biomechanical properties of the guard cell wall in particular. Atomic force microscopy (AFM) has been used to visualize the pattern and movement of cellulose microfibrils on the nanoscale in onion epidermis (Zhang et al., 2013, 2016, 2017). Unfortunately, cellulose microfibrils in the guard cell wall cannot be directly probed by AFM due to the presence of cuticles in aerial tissues. However, stiffness distribution on the cellular scale in guard cells and neighboring pavement cells can still be revealed by AFM (Sampathkumar et al., 2014; Carter et al., 2017). Recently, Carter et al. reported that guard cells are stiffer at polar regions than along their outer periphery (**Figure 1**) and that exogenous PG treatment weakens polar stiffness, indicating that de-methyl-esterified pectins likely contribute to the polar stiffening of guard cells, which might help to fix stomatal poles during stomatal opening (Carter et al., 2017; Woolfenden et al., 2018).

There has been growing interest in assessing responses to mechanical stress with cellular resolution in live tissues such as Arabidopsis cotyledons (Bringmann and Bergmann, 2017; Robinson et al., 2017). Bringmann and Bergmann applied stretch forces to whole cotyledons using elastic strips and observed that the distribution of polarity markers for stomatal stem cells follows the direction of tissue-wide tensile stress (Bringmann and Bergmann, 2017). This finding should be interpreted with the caveats that polarity markers are not imaged simultaneously when stretch is applied and that the amount of force applied to the cotyledon is undefined. Recently, an Automated Confocal Micro-Extensometer (ACME) that allows real-time quantification of strain response to known stresses in 3D has been developed (Robinson et al., 2017). A future avenue to understanding the stress-strain relationship of the guard cell wall is to apply ACME in leaves or epidermal peels, with the caution that mechanical stress is applied to whole tissue

rather than to individual stomata. In addition, the setup of an ACME for this purpose will require expertise and perhaps some customization.

Stomatal opening and closure are driven by turgor changes in guard cells, but direct measurement of guard cell turgor is technically challenging mainly due to the small size of guard cells and their vacuoles (Franks et al., 1995). To our knowledge, pressure values of guard cells have been determined by pressure probe in only a few species (e.g., *Tradescantia virginiana* and *V. faba*) (Franks et al., 1995, 1998, 2001), all of which have much larger stomata than *Arabidopsis*. Another factor limiting the wide use of the pressure probe is that the number of reported successful measurements of guard cell turgor in a given species is very small. Therefore, the development of new methodologies to either directly measure or indirectly calculate guard cell turgor will be required, which could be inspired by turgor pressure measurements in other cell types. For example, a combination of micro-indentation and osmotic treatments has recently been used to estimate turgor pressure in tobacco BY-2 cells (Weber et al., 2015). In single cells, measuring changes in cell volume or dimensions under external osmotic stress could help determine mechanical parameters of the cell, such as cell stiffness as measured in the A7 stem cell line (Guo et al., 2017), and Young's modulus of the cell wall and turgor pressure as measured in fission yeast (Atilgan et al., 2015).

An emerging tool for investigating stomatal mechanics is 3D finite element modeling, although the earliest finite element models of stomata were reported in the 1970s (Cooke et al., 1976). Thus far, all published stomatal models focus on kidney-shaped guard cells during stomatal opening (Cooke et al., 1976, 2008; Rui et al., 2016; Carter et al., 2017; Marom et al., 2017; Shtein et al., 2017; Woolfenden et al., 2017), leaving the modeling of dumbbell-shaped guard cells and stomatal closure process an uncharted area. However, results from those models lead to discrepant conclusions about which guard cell features are crucial for stomatal opening. Cooke et al. took into account neighboring cells and found that stomatal opening is a consequence of the elliptical geometry of guard cells and changes in their cross-sectional shape, whereas differential thickness of the guard cell wall and radially arranged cellulose microfibrils are not essential (Cooke et al., 1976, 2008).

More recently, Woolfenden et al. constructed stomatal models based on 2D geometric parameters of guard cells and stomatal pores in *V. faba*, and argued that circumferential reinforcement by radially oriented cellulose microfibrils is required for stomatal opening (Woolfenden et al., 2017). However, this conclusion might need to be further substantiated due to the following reasons: (1) the authors modeled guard cells with idealized geometries; (2) their experimental data in *Arabidopsis* were not consistent with the modeling results for *V. faba*. Using the same framework of stomatal models, Carter et al. argued that differential wall thickness plays a minimal role in stomatal opening, which is consistent with what Cooke et al. found, and suggested that de-methyl-esterified HG-based stiffening at guard cell ends might function as “pins” to help to fix stomatal complex length during stomatal opening (Carter et al., 2017). Testing

the effect of highly localized perturbation of HG de-methyl-esterification at polar regions will be useful to further test this hypothesis.

Good finite element models depend on high-quality inputs including geometry, boundary conditions, and material models (Bidhendi and Geitmann, 2017). The aforementioned experimental approaches to measure mechanical properties of the guard cell wall would be useful to improve the quality of finite element models. A step further is to adopt a multiscale, multiphysics modeling strategy and to incorporate interactions between wall polymers at the molecular scale and interactions between guard cells and neighboring cells at the cellular scale into the model.

PERSPECTIVE

The past few years of research have brought increased attention to the structural and functional complexity of guard cell walls. We predict that the dialog between experimental data generated from genetic, cytological, and biomechanical approaches, and mechanical modeling of stomata, will shed light on the following outstanding questions: (1) Is cell wall degradation an essential molecular mechanism underlying stomatal pore formation? (2) How is the enrichment of pectin-degrading enzymes and pectins at stomatal pore initiation sites determined and established during stomatal development? (3) What is the composition and architecture of the guard cell wall? (4) Is there any difference in wall composition between guard cells and neighboring cells? (5) Is there any difference in wall composition and/or architecture between kidney-shaped guard cells and dumbbell-shaped guard cells? (6) What are the turgor pressure values in guard cells at different developmental stages, at different opening/closed states, and in different plant species? (7) During leaf senescence, mesophyll cells undergo a reduction in chloroplast number and degradation of cortical MTs (Zeiger and Schwartz, 1982; Keech et al., 2010), whereas guard cells retain chloroplasts and MT network (Zeiger and Schwartz, 1982; Ozuna et al., 1985; Hurng et al., 1988; Willmer et al., 1988; Thomas et al., 1991). Although stomatal conductance diminishes (Willmer et al., 1988) and cuticular occlusions increase as leaves and plants age (England and Attiwill, 2005), do guard cell walls likewise “wear out” and fail, and can guard cells self-repair and continue to function?

Future studies on how the amazingly strong and elastic cell walls of stomatal guard cells are constructed and how their dynamics are spatiotemporally regulated will help elucidate structure-function relationships during reversible cell expansion in plants, knowledge of which can be translated into applications such as creating elastic biomimetic materials (Li and Wang, 2016) and generating crop species with improved gas exchange efficiency.

AUTHOR CONTRIBUTIONS

YR and CA generated **Movie S1** and wrote the manuscript with input from YC, BK, HY, JW, and VP. YC, BK, HY, and JW

generated **Figure 3**. BK and JW developed the computer image segmentation algorithm for guard cell images that has made possible the 3D renderings shown in the illustrations in **Figure 3**.

ACKNOWLEDGMENTS

We thank members of the Anderson lab, Dr. Sarah M. Assmann, and Dr. Gabriele Monshausen for helpful discussions.

REFERENCES

- Adrian, J., Chang, J., Ballenger, C. E., Bargmann, B. O., Alassimone, J., Davies, K. A., et al. (2015). Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. *Dev. Cell* 33, 107–118. doi: 10.1016/j.devcel.2015.01.025
- Amsbury, S., Hunt, L., Elhaddad, N., Baillie, A., Lundgren, M., Verhertbruggen, Y., et al. (2016). Stomatal function requires Pectin De-methyl-esterification of the guard cell wall. *Curr. Biol.* 26, 1–8. doi: 10.1016/j.cub.2016.08.021
- Anderson, C. T., and Carroll, A. (2014). Identification and use of fluorescent dyes for plant cell wall imaging using high-throughput screening. *Methods Mol. Biol.* 1056, 103–109. doi: 10.1007/978-1-62703-592-7_10
- Anderson, C. T., Carroll, A., Akhmetova, L., and Somerville, C. (2010). Real-time imaging of cellulose reorientation during cell wall expansion in Arabidopsis roots. *Plant Physiol.* 152, 787–796. doi: 10.1104/pp.109.150128
- Anderson, C. T., Wallace, I. S., and Somerville, C. R. (2012). Metabolic click-labeling with a fucose analog reveals pectin delivery, architecture, and dynamics in Arabidopsis cell walls. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1329–1334. doi: 10.1073/pnas.1120429109
- Atilgan, E., Magidson, V., Khodjakov, A., and Chang, F. (2015). Morphogenesis of the fission yeast cell through cell wall expansion. *Curr. Biol.* 25, 2150–2157. doi: 10.1016/j.cub.2015.06.059
- Atkinson, R. G., Schröder, R., Hallett, I. C., Cohen, D., and MacRae, E. A. (2002). Overexpression of polygalacturonase in transgenic apple trees leads to a range of novel phenotypes involving changes in cell adhesion. *Plant Physiol.* 129, 122–133. doi: 10.1104/pp.010986
- Bergmann, D. C., and Sack, F. D. (2007). Stomatal development. *Annu. Rev. Plant Biol.* 58, 163–181. doi: 10.1146/annurev.arplant.58.032806.104023
- Berry, J. A., Beerling, D. J., and Franks, P. J. (2010). Stomata: key players in the earth system, past and present. *Curr. Opin. Plant Biol.* 13, 233–240. doi: 10.1016/j.pbi.2010.04.013
- Bidhendi, A. J., and Geitmann, A. (2017). Finite element modeling of shape changes in plant cells. *Plant Physiol.* 176, 41–56. doi: 10.1104/pp.17.01684
- Braybrook, S. A., and Peaucelle, A. (2013). Mechano-chemical aspects of organ formation in Arabidopsis thaliana: the relationship between auxin and pectin. *PLoS ONE* 8:e57813. doi: 10.1371/journal.pone.0057813
- Bringmann, M., and Bergmann, D. C. (2017). Tissue-wide mechanical forces influence the polarity of stomatal stem cells in Arabidopsis. *Curr. Biol.* 27, 877–883. doi: 10.1016/j.cub.2017.01.059
- Carr, D. J., Carr, S. G. M., and Jahnke, R. (1980). Intercellular strands associated with stomata: stomatal pectic strands. *Protoplasma* 102, 177–182. doi: 10.1007/BF01276958
- Carter, R., Woolfenden, H., Baillie, A., Amsbury, S., Carroll, S., Healicon, E., et al. (2017). Stomatal opening involves polar, not radial, stiffening of guard cells. *Curr. Biol.* 27, 2974–2983.e2. doi: 10.1016/j.cub.2017.08.006
- Casson, S. A., and Hetherington, A. M. (2010). Environmental regulation of stomatal development. *Curr. Opin. Plant Biol.* 13, 90–95. doi: 10.1016/j.pbi.2009.08.005
- Cavaler, D. M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., et al. (2008). Disrupting two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell* 20, 1519–1537. doi: 10.1105/tpc.108.059873
- Chater, C. C., Caine, R. S., Tomek, M., Wallace, S., Kamisugi, Y., Cumming, A. C., et al. (2016). Origin and function of stomata in the moss *Physcomitrella patens*. *Nature Plants* 2:16179. doi: 10.1038/nplants.2016.179
- Chen, Z. H., Chen, G., Dai, F., Wang, Y., Hills, A., Ruan, Y. L., et al. (2017). Molecular evolution of grass stomata. *Trends Plant Sci.* 22, 124–139. doi: 10.1016/j.tplants.2016.09.005
- Cocuron, J. C., Lerouxel, O., Drakakaki, G., Alonso, A. P., Liepman, A. H., Keegstra, K., et al. (2007). A gene from the cellulose synthase-like C family encodes a beta-1,4 glucan synthase. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8550–8555. doi: 10.1073/pnas.0703133104
- Cooke, J. R., De Baerdemaeker, J. G., Rand, R. H., and Mang, H. A. (1976). A finite element shell analysis of guard cell deformation. *Trans. Am. Soc. Agric. Eng.* 19, 1107–1121. doi: 10.13031/2013.36186
- Cooke, J. R., Rand, R. H., Mang, H. A., De Baerdemaeker, J. G., and Lee, J. Y. (2008). “Shell analysis of elliptical guard cells in higher plants: a review,” in *International Conference on Computation of Shell and Spatial Structures IASS-IACM 2008: “Spanning Nano to Mega.”* (Ithaca, NY).
- Cosgrove, D. J. (2000). Loosening of plant cell walls by expansins. *Nature* 407, 321–326. doi: 10.1038/35030000
- Desprez, T., Juraniec, M., Crowell, E. F., Jouy, H., Pochylova, Z., Parcy, F., et al. (2007). Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15572–15577. doi: 10.1073/pnas.0706569104
- Edwards, D., Davies, K. L., and Axe, L. (1992). A vascular conducting strand in the early land plant Cooksonia. *Nature* 357, 683–685. doi: 10.1038/357683a0
- England, J. R., and Attiwill, P. M. (2005). Changes in leaf morphology and anatomy with tree age and height in the broadleaved evergreen species, *Eucalyptus regnans* F. Muell. *Trees* 20, 79–90. doi: 10.1007/s00468-005-0015-5
- Fan, L. M., Zhao, Z., and Assmann, S. M. (2004). Guard cells: a dynamic signaling model. *Curr. Opin. Plant Biol.* 7, 537–546. doi: 10.1016/j.pbi.2004.07.009
- Franks, P. J., Buckley, T. N., Shope, J. C., and Mott, K. A. (2001). Guard cell volume and pressure measured concurrently by confocal microscopy and the cell pressure probe. *Plant Physiol.* 125, 1577–1584. doi: 10.1104/pp.125.4.1577
- Franks, P. J., Cowan, I. R., and Farquhar, G. D. (1998). A study of stomatal mechanics using the cell pressure probe. *Plant Cell Environ.* 21, 94–100. doi: 10.1046/j.1365-3040.1998.00248.x
- Franks, P. J., Cowan, I. R., Tyerman, S. D., Cleary, A. L., Lloyd, J., and Farquhar, G. D. (1995). Guard cell pressure/aperture characteristics measured with the pressure probe. *Plant Cell Environ.* 18, 795–800. doi: 10.1111/j.1365-3040.1995.tb00583.x
- Fujita, M., and Wasteneys, G. O. (2014). A survey of cellulose microfibril patterns in dividing, expanding, and differentiating cells of Arabidopsis thaliana. *Protoplasma* 251, 687–698. doi: 10.1007/s00709-013-0571-2
- Giannoutsou, E., Apostolakis, P., and Galatis, B. (2016). Spatio-temporal diversification of the cell wall matrix materials in the developing stomatal complexes of Zea mays. *Planta* 244, 1125–1143. doi: 10.1007/s00425-016-2574-7
- Green, P. B. (1962). Mechanism for plant cellular morphogenesis. *Science* 138, 1404–1405.
- Guo, M., Pegoraro, A. F., Mao, A., Zhou, E. H., Arany, P. R., Han, Y., et al. (2017). Cell volume change through water efflux impacts cell stiffness and stem cell fate. *Proc. Natl. Acad. Sci. U.S.A.* 114, E8618–E8627. doi: 10.1073/pnas.1705179114
- Gustafsson, M. G. (2005). Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13081–13086. doi: 10.1073/pnas.0406877102
- Hachez, C., Ohashi-Ito, K., Dong, J., and Bergmann, D. C. (2011). Differentiation of Arabidopsis guard cells: analysis of the networks incorporating the basic

- helix-loop-helix transcription factor, FAMA. *Plant Physiol.* 155, 1458–1472. doi: 10.1104/pp.110.167718
- Harholt, J., Jensen, J. K., Sorensen, S. O., Orfila, C., Pauly, M., and Scheller, H. V. (2006). ARABINAN DEFICIENT 1 is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in *Arabidopsis*. *Plant Physiol.* 140, 49–58. doi: 10.1104/pp.105.072744
- Hepworth, C., Caine, R. S., Harrison, E. L., Sloan, J., and Gray, J. E. (2018). Stomatal development: focusing on the grasses. *Curr. Opin. Plant Biol.* 41, 1–7. doi: 10.1016/j.pbi.2017.07.009
- Hocq, L., Pelloux, J., and Lefebvre, V. (2017). Connecting homogalacturonan-type pectin remodeling to acid growth. *Trends Plant Sci.* 22, 20–29. doi: 10.1016/j.tplants.2016.10.009
- Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* 319, 810–813. doi: 10.1126/science.1153529
- Huang, Y. C., Wu, H. C., Wang, Y. D., Liu, C. H., Lin, C. C., Luo, D. L., et al. (2017). PECTIN METHYLESTERASE34 contributes to heat tolerance through its role in promoting stomatal movement. *Plant Physiol.* 174, 748–763. doi: 10.1104/pp.17.00335
- Hunt, L., Amsbury, S., Baillie, A., Movahedi, M., Mitchell, A., Afsharinafar, M., et al. (2017). Formation of the stomatal outer cuticular ledge requires a guard cell wall proline-rich protein. *Plant Physiol.* 174, 689–699. doi: 10.1104/pp.16.01715
- Hung, W. P., Lin, T. L., Ren, S. S., Chen, J. C., Chen, Y. R., and Kao, C. H. (1988). Senescence of rice leaves XVIII. Changes of stomatal aperture during senescence. *Plant Cell Physiol* 29, 27–31.
- Jolie, R. P., Duvetter, T., Van Loeys, A. M., and Hendrickx, M. E. (2010). Pectin methylesterase and its proteinaceous inhibitor: a review. *Carbohydr. Res.* 345, 2583–2595. doi: 10.1016/j.carres.2010.10.002
- Jones, L., Milne, J. L., Ashford, D., McCann, M. C., and McQueen-Mason, S. J. (2005). A conserved functional role of pectic polymers in stomatal guard cells from a range of plant species. *Planta* 221, 255–264. doi: 10.1007/s00425-004-1432-1
- Jones, L., Milne, J. L., Ashford, D., and McQueen-Mason, S. J. (2003). Cell wall arabinan is essential for guard cell function. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11783–11788. doi: 10.1073/pnas.1832434100
- Jones, V. A., and Dolan, L. (2017). MpWIP regulates air pore complex development in the liverwort *Marchantia polymorpha*. *Development* 144, 1472–1476. doi: 10.1242/dev.144287
- Keech, O., Pesquet, E., Gutierrez, L., Ahad, A., Bellini, C., Smith, S. M., et al. (2010). Leaf senescence is accompanied by an early disruption of the microtubule network in *Arabidopsis*. *Plant Physiol.* 154, 1710–1720. doi: 10.1104/pp.110.163402
- Keegstra, K. (2010). Plant cell walls. *Plant Physiol.* 154, 483–486. doi: 10.1104/pp.110.161240
- Kim, S. J., Held, M. A., Zemelis, S., Wilkerson, C., and Brandizzi, F. (2015). CGR2 and CGR3 have critical overlapping roles in pectin methylesterification and plant growth in *Arabidopsis thaliana*. *Plant J.* 82, 208–220. doi: 10.1111/tjp.12802
- Kim, T. H., Böhmer, M., Hu, H., Nishimura, N., and Schroeder, J. I. (2010). Guard cell signal transduction network: advances in understanding abscisic acid, CO₂, and Ca²⁺ signaling. *Annu. Rev. Plant Biol.* 61, 561–591. doi: 10.1146/annurev-arplant-042809-112226
- Li, S., and Wang, K. W. (2016). Plant-inspired adaptive structures and materials for morphing and actuation: a review. *Bioinspir. Biomim.* 12:011001. doi: 10.1088/1748-3190/12/1/011001
- Li, X., Cordero, I., Caplan, J., Molhoj, M., and Reiter, W. D. (2004). Molecular analysis of 10 coding regions from *Arabidopsis* that are homologous to the MUR3 xyloglucan galactosyltransferase. *Plant Physiol.* 134, 940–950. doi: 10.1104/pp.103.036285
- Liljegren, S. J. (2012). Organ abscission: exit strategies require signals and moving traffic. *Curr. Opin. Plant Biol.* 15, 670–676. doi: 10.1016/j.pbi.2012.09.012
- Majewska-Sawka, A., Münster, A., and Rodríguez-García, M. I. (2002). Guard cell wall: immunocytochemical detection of polysaccharide components. *J. Exp. Bot.* 53, 1067–1079. doi: 10.1093/jexbot/53.371.1067
- Marom, Z., Shtein, I., and Bar-On, B. (2017). Stomatal Opening: The role of cell-wall mechanical anisotropy and its analytical relations to the bio-composite characteristics. *Front. Plant Sci.* 8:2061. doi: 10.3389/fpls.2017.02061
- McCarthy, T. W., Der, J. P., Honaas, L. A., dePamphilis, C. W., and Anderson, C. T. (2014). Phylogenetic analysis of pectin-related gene families in *Physcomitrella patens* and nine other plant species yields evolutionary insights into cell walls. *BMC Plant Biol.* 14:79. doi: 10.1186/1471-2229-14-79
- Merced, A. (2015). Novel insights on the structure and composition of pseudostomata of *Sphagnum*. *Am. J. Bot.* 102, 329–335. doi: 10.3732/ajb.1400564
- Merced, A., and Renzaglia, K. (2014). Developmental changes in guard cell wall structure and pectin composition in the moss *Funaria*: implications for function and evolution of stomata. *Ann. Bot.* 114, 1001–1010. doi: 10.1093/aob/mcu165
- Mohnen, D. (2008). Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* 11, 266–277. doi: 10.1016/j.pbi.2008.03.006
- Mouille, G., Ralet, M. C., Cavelier, C., Eland, C., Effroy, D., Hematy, K., et al. (2007). Homogalacturonan synthesis in *Arabidopsis thaliana* requires a Golgi-localized protein with a putative methyltransferase domain. *Plant J.* 50, 605–614. doi: 10.1111/j.1365-3113X.2007.03086.x
- Negi, J., Moriwaki, K., Konishi, M., Yokoyama, R., Nakano, T., Kusumi, K., et al. (2013). A Dof transcription factor, SCAP1, is essential for the development of functional stomata in *Arabidopsis*. *Curr. Biol.* 23, 479–484. doi: 10.1016/j.cub.2013.02.001
- Ozuna, R., Yera, R., Ortega, K., and Tallman, G. (1985). Analysis of guard cell viability and action in senescing leaves of *Nicotiana glauca* (Graham), tree tobacco. *Plant Physiol.* 79, 7–10. doi: 10.1104/pp.79.1.7
- Palevitz, B. A., and Hepler, P. K. (1976). Cellulose microfibril orientation and cell shaping in developing guard cells of *Allium*: the role of microtubules and ion accumulation. *Planta* 132, 71–93. doi: 10.1007/BF00390333
- Paredes, A. R., Somerville, C. R., and Ehrhardt, D. W. (2006). Visualization of cellulose synthase demonstrates functional association with microtubules. *Science* 312, 1491–1495. doi: 10.1126/science.1126551
- Park, Y. B., and Cosgrove, D. J. (2012a). Changes in cell wall biomechanical properties in the xyloglucan-deficient xxt1/xt2 mutant of *Arabidopsis*. *Plant Physiol.* 158, 465–475. doi: 10.1104/pp.111.189779
- Park, Y. B., and Cosgrove, D. J. (2012b). A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant Physiol.* 158, 1933–1943. doi: 10.1104/pp.111.192880
- Pauly, M., and Keegstra, K. (2016). Biosynthesis of the plant cell wall matrix polysaccharide xyloglucan. *Annu. Rev. Plant Biol.* 67, 235–259. doi: 10.1146/annurev-arplant-043015-112222
- Peaucelle, A., Braybrook, S. A., Le Guillou, L., Bron, E., Kuhlemeier, C., and Höfte, H. (2011). Pectin-induced changes in cell wall mechanics underlie organ initiation in *Arabidopsis*. *Curr. Biol.* 21, 1720–1726. doi: 10.1016/j.cub.2011.08.057
- Peterson, K. M., Rychel, A. L., and Torii, K. U. (2010). Out of the mouths of plants: the molecular basis of the evolution and diversity of stomatal development. *Plant Cell* 22, 296–306. doi: 10.1105/tpc.109.072777
- Pillitteri, L. J., and Torii, K. U. (2012). Mechanisms of stomatal development. *Annu. Rev. Plant Biol.* 63, 591–614. doi: 10.1146/annurev-arplant-042811-105451
- Raven, J. A. (2002). Selection pressures on stomatal evolution. *New Phytol.* 153, 371–386. doi: 10.1046/j.0028-646X.2001.00334.x
- Renzaglia, K. S., Villarreal, J. C., Piatkowski, B. T., Lucas, J. R., and Merced, A. (2017). Hornwort stomata: architecture and fate shared with 400-million-year-old fossil plants without leaves. *Plant Physiol.* 174, 788–797. doi: 10.1104/pp.17.00156
- Robinson, S., Huflejt, M., Barbier de Reuille, P., Braybrook, S. A., Schorderet, M., Reinhardt, D., et al. (2017). An automated confocal micro-extensometer enables *in vivo* quantification of mechanical properties with cellular resolution. *Plant Cell* 29, 2959–2973. doi: 10.1105/tpc.17.00753
- Rudall, P. J., Hilton, J., and Bateman, R. M. (2013). Several developmental and morphogenetic factors govern the evolution of stomatal patterning in land plants. *New Phytol.* 200, 598–614. doi: 10.1111/nph.12406
- Rui, Y., and Anderson, C. T. (2016). Functional analysis of cellulose and xyloglucan in the walls of stomatal guard cells of *Arabidopsis*. *Plant Physiol.* 170, 1398–1419. doi: 10.1104/pp.15.01066
- Rui, Y., Xiao, C., Yi, H., Kandemir, B., Wang, J. Z., Puri, V. M., et al. (2017). POLYGALACTURONASE INVOLVED IN EXPANSION3 functions in

- seedling development, rosette growth, and stomatal dynamics in *Arabidopsis thaliana*. *Plant Cell* 29, 2413–2432. doi: 10.1105/tpc.17.00568
- Rui, Y., Yi, H., Kandemir, B., Wang, J. Z., Puri, V. M., and Anderson, C. T. (2016). Integrating cell biology, image analysis, and computational mechanical modeling to analyze the contributions of cellulose and xyloglucan to stomatal function. *Plant Signal. Behav.* 11:e1183086. doi: 10.1080/15592324.2016.1183086
- Sack, F. D., and Paolillo, D. J. (1985). Incomplete cytokinesis in *Funaria* stomata. *Am. J. Bot.* 72, 1325–1333. doi: 10.1002/j.1537-2197.1985.tb08389.x
- Sampathkumar, A., Krupinski, P., Wightman, R., Milani, P., Berquand, A., Boudaoud, A., et al. (2014). Subcellular and supracellular mechanical stress prescribes cytoskeleton behavior in *Arabidopsis* cotyledon pavement cells. *Elife* 3:e01967. doi: 10.7554/eLife.01967
- Shtein, I., Shelef, Y., Marom, Z., Zelinger, E., Schwartz, A., Popper, Z. A., et al. (2017). Stomatal cell wall composition: distinctive structural patterns associated with different phylogenetic groups. *Ann. Bot.* 119, 1021–1033. doi: 10.1093/aob/mcw275
- Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., et al. (2004). Toward a systems approach to understanding plant cell walls. *Science* 306, 2206–2211. doi: 10.1126/science.1102765
- Thomas, C., Davis, S. D., and Tallman, G. (1991). Responses of stomata of senescing and nonsenescing leaves of *Nicotiana glauca* to changes in intercellular concentrations of leaf carbon dioxide. *Plant Cell Environ.* 14, 971–978. doi: 10.1111/j.1365-3040.1991.tb00967.x
- Vogel, J. (2008). Unique aspects of the grass cell wall. *Curr. Opin. Plant Biol.* 11, 301–307. doi: 10.1016/j.pbi.2008.03.002
- Voiniciuc, C., Pauly, M., and Usadel, B. (2018). Monitoring polysaccharide dynamics in the plant cell wall. *Plant Physiol.* 176, 2590–2600. doi: 10.1104/pp.17.01776
- Weber, A., Braybrook, S., Huflejt, M., Mosca, G., Routier-Kierzkowska, A. L., and Smith, R. S. (2015). Measuring the mechanical properties of plant cells by combining micro-indentation with osmotic treatments. *J. Exp. Bot.* 66, 3229–3241. doi: 10.1093/jxb/erv135
- Willmer, C. M., Wilson, A. B., and Jones, H. G. (1988). Changing responses of stomata to abscisic acid and CO₂ as leaves and plants age. *J. Exp. Bot.* 39, 401–410. doi: 10.1093/jxb/39.4.401
- Wolf, S., Mouille, G., and Pelloux, J. (2009). Homogalacturonan methyl-esterification and plant development. *Mol. Plant* 2, 851–860. doi: 10.1093/mp/ssp066
- Woelfenden, H. C., Baillie, A. L., Gray, J. E., Hobbs, J. K., Morris, R. J., and Fleming, A. J. (2018). Models and mechanisms of stomatal mechanics. *Trends Plant Sci.* doi: 10.1016/j.tplants.2018.06.003. [Epub ahead of print].
- Woelfenden, H. C., Bourdais, G., Kopischke, M., Miedes, E., Molina, A., Robatzek, S., et al. (2017). A computational approach for inferring the cell wall properties that govern guard cell dynamics. *Plant J.* 92, 5–18. doi: 10.1111/tpj.13640
- Wu, H., and Sharpe, P. J. H. (1979). Stomatal mechanics II: material properties of guard cell walls. *Plant Cell Environ.* 2, 235–244. doi: 10.1111/j.1365-3040.1979.tb00075.x
- Xu, Y., Sechet, J., Wu, Y., Fu, Y., Zhu, L., Li, J., et al. (2017). Rice sucrose partitioning mediated by a putative pectin methyltransferase and homogalacturonan methylesterification. *Plant Physiol.* 174, 1595–1608. doi: 10.1104/pp.16.01555
- Zeiger, E., and Schwartz, A. (1982). Longevity of guard cell chloroplasts in falling leaves: implication for stomatal function and cellular aging. *Science* 218, 680–682.
- Zhang, T., Mahgoudy-Louyeh, S., Tittmann, B., and Cosgrove, D. J. (2013). Visualization of the nanoscale pattern of recently-deposited cellulose microfibrils and matrix materials in never-dried primary walls of the onion epidermis. *Cellulose* 21, 853–862. doi: 10.1007/s10570-013-9996-1
- Zhang, T., Vavylonis, D., Durachko, D. M., and Cosgrove, D. J. (2017). Nanoscale movements of cellulose microfibrils in primary cell walls. *Nat Plants* 3:17056. doi: 10.1038/nplants.2017.56
- Zhang, T., Zheng, Y., and Cosgrove, D. J. (2016). Spatial organization of cellulose microfibrils and matrix polysaccharides in primary plant cell walls as imaged by multichannel atomic force microscopy. *Plant J.* 85, 179–192. doi: 10.1111/tpj.13102
- Zhang, X. Q., Wei, P. C., Xiong, Y. M., Yang, Y., Chen, J., and Wang, X. C. (2011). Overexpression of the *Arabidopsis* alpha-expansin gene *ATEXPA1* accelerates stomatal opening by decreasing the volumetric elastic modulus. *Plant Cell Rep.* 30, 27–36. doi: 10.1007/s00299-010-0937-2
- Zhao, L., and Sack, F. D. (1999). Ultrastructure of stomatal development in *Arabidopsis* (Brassicaceae) leaves. *Am. J. Bot.* 86, 929–939. doi: 10.2307/2656609
- Zheng, Y., Wang, X., Chen, Y., Wagner, E., and Cosgrove, D. J. (2018). Xyloglucan in the primary cell wall: assessment by FESEM, selective enzyme digestions and nanogold affinity tags. *Plant J.* 93, 211–226. doi: 10.1111/tpj.13778

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Rui, Chen, Kandemir, Yi, Wang, Puri and Anderson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



New Insights Into Wall Polysaccharide O-Acetylation

Markus Pauly and Vicente Ramírez*

Institute for Plant Cell Biology and Biotechnology – Cluster of Excellence on Plant Sciences, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

OPEN ACCESS

Edited by:

Georgia Drakakaki,
University of California, Davis,
United States

Reviewed by:

Yumiko Sakuragi,
University of Copenhagen, Denmark
Ian S. Wallace,
University of Nevada, Reno,
United States
Olga A. Zabolina,
Iowa State University, United States

*Correspondence:

Vicente Ramírez
ramirezg@hhu.de

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 30 May 2018

Accepted: 27 July 2018

Published: 21 August 2018

Citation:

Pauly M and Ramírez V (2018) New
Insights Into Wall Polysaccharide
O-Acetylation.
Front. Plant Sci. 9:1210.
doi: 10.3389/fpls.2018.01210

The extracellular matrix of plants, algae, bacteria, fungi, and some archaea consist of a semipermeable composite containing polysaccharides. Many of these polysaccharides are O-acetylated imparting important physiochemical properties to the polymers. The position and degree of O-acetylation is genetically determined and varies between organisms, cell types, and developmental stages. Despite the importance of wall polysaccharide O-acetylation, only recently progress has been made to elucidate the molecular mechanism of O-acetylation. In plants, three protein families are involved in the transfer of the acetyl substituents to the various polysaccharides. In other organisms, this mechanism seems to be conserved, although the number of required components varies. In this review, we provide an update on the latest advances on plant polysaccharide O-acetylation and related information from other wall polysaccharide O-acetylating organisms such as bacteria and fungi. The biotechnological impact of understanding wall polysaccharide O-acetylation ranges from the design of novel drugs against human pathogenic bacteria to the development of improved lignocellulosic feedstocks for biofuel production.

Keywords: O-acetylation, cell wall, polysaccharides, biosynthesis, mechanism

PLANT CELL WALL POLYMERS ARE O-ACETYLATED

The biomass of plants contains considerable amounts of esterified acetate. For example, poplar wood contains 5% of its weight as acetate (Johnson et al., 2017), while corn stover contains 4.5% (w/w; Chundawat et al., 2010). Upon processing of the plant biomass the acetate is often released (Selig et al., 2009) not only acidifying the resulting material, but also presenting a potent inhibitor for further downstream microbial fermentation (Helle et al., 2003) such as for the production for biofuels.

The predominant portion of the bound acetate found in plant biomass is present in the cell wall material in the form of O-linked acetate on many wall polysaccharides (references below), and to a minor extent on the polyphenol lignin (Ralph, 1996; Del Río et al., 2007). While cellulose, callose, mixed-linkage glucans, and structural glycoproteins are not O-acetylated, the dominant matrix polysaccharides including the various pectic polysaccharides and hemicelluloses such as xylan, xyloglucan, and mannans can be O-acetylated (Gille and Pauly, 2012). The position and degree of acetylation depends on the wall polymer and can differ not only between plant species, but also cell types, and/or the developmental stage of the plant (Del Río et al., 2007; Obel et al., 2009; Pauly and Keegstra, 2010; Gille and Pauly, 2012; Lourenço et al., 2016). Both the polymer

glycan-backbone and/or the side-chain sugar moieties can be O-acetylated (Kiefer et al., 1989; Ishii, 1991, 1997; Pauly, 1999; Teleman et al., 2000; Lundqvist et al., 2002; Kabel et al., 2003; O'Neill et al., 2004; Gibeaut et al., 2005; Hoffman et al., 2005; Jia et al., 2005; Hsieh and Harris, 2009; Sengkhamparn et al., 2009). For example, the hemicellulose xyloglucan (XyG) predominantly found in dicot species contains O-acetyl moieties on the galactosyl side-chain residues, while in Solanaceous plants and grasses, the glucan-backbone of xyloglucan is O-acetylated. In addition, several wall polymers contain sugar-residues that can be mono-/or di-O-acetylated (reviewed in Gille and Pauly, 2012).

POLYSACCHARIDE O-ACETYLATION MECHANISM

Several lines of evidence suggest that O-acetylation of wall polysaccharides takes place as part of the polysaccharide biosynthesis in the Golgi lumen. First, acetylated xyloglucan can be isolated from microsomal preparations suggesting that O-acetylation takes place before the wall polysaccharides are secreted into the apoplast (Obel et al., 2009). Second, pectic polysaccharides can be O-acetylated *in vitro* in isolated plant microsomes (Pauly and Scheller, 2000). Third, all proteins involved in this modification (see below) are predicted to be located in the Golgi membrane with the putative catalytic domains facing the Golgi lumen (Gille et al., 2011b; Lee et al., 2011; Manabe et al., 2011; Yuan et al., 2013; Schultink et al., 2015; Gao et al., 2017). However, it should be noted that the degree and pattern of polysaccharide O-acetylation is also determined by apoplastic plant O-acetyltransferases, presumably post-deposition in the wall (Gou et al., 2012; Orfila et al., 2012; de Souza et al., 2014; Zhang et al., 2017).

The identification of plant mutants affected in the O-acetylation of wall polysaccharides has been instrumental in our understanding of the molecular mechanism of polysaccharide O-acetylation. Based on these findings, so far three different protein families are involved in polysaccharide O-acetylation (**Figure 1**). One of these protein families is the TRICHOME-BIREFRINGENCE-LIKE (TBL) protein family comprising 46 members in the model species *Arabidopsis thaliana*. Members of the TBL family have been shown to participate in the O-acetylation of specific wall polymers. Loss-of-function of the *Arabidopsis* *ALTERED XYLOGLUCAN 4* (*AXY4/TBL27*) gene results in a complete lack of O-acetyl substituents on the hemicellulose XyG without affecting the acetylation status of the other wall polymers (Gille et al., 2011b). Its paralogous gene – *AXY4-like* (*AXY4L/TBL22*) – appears to have the same function but specifically in seeds, indicating that *AXY4* and *AXY4L* are XyG-specific acetyltransferases, although the biochemical activity of both proteins remains to be experimentally demonstrated.

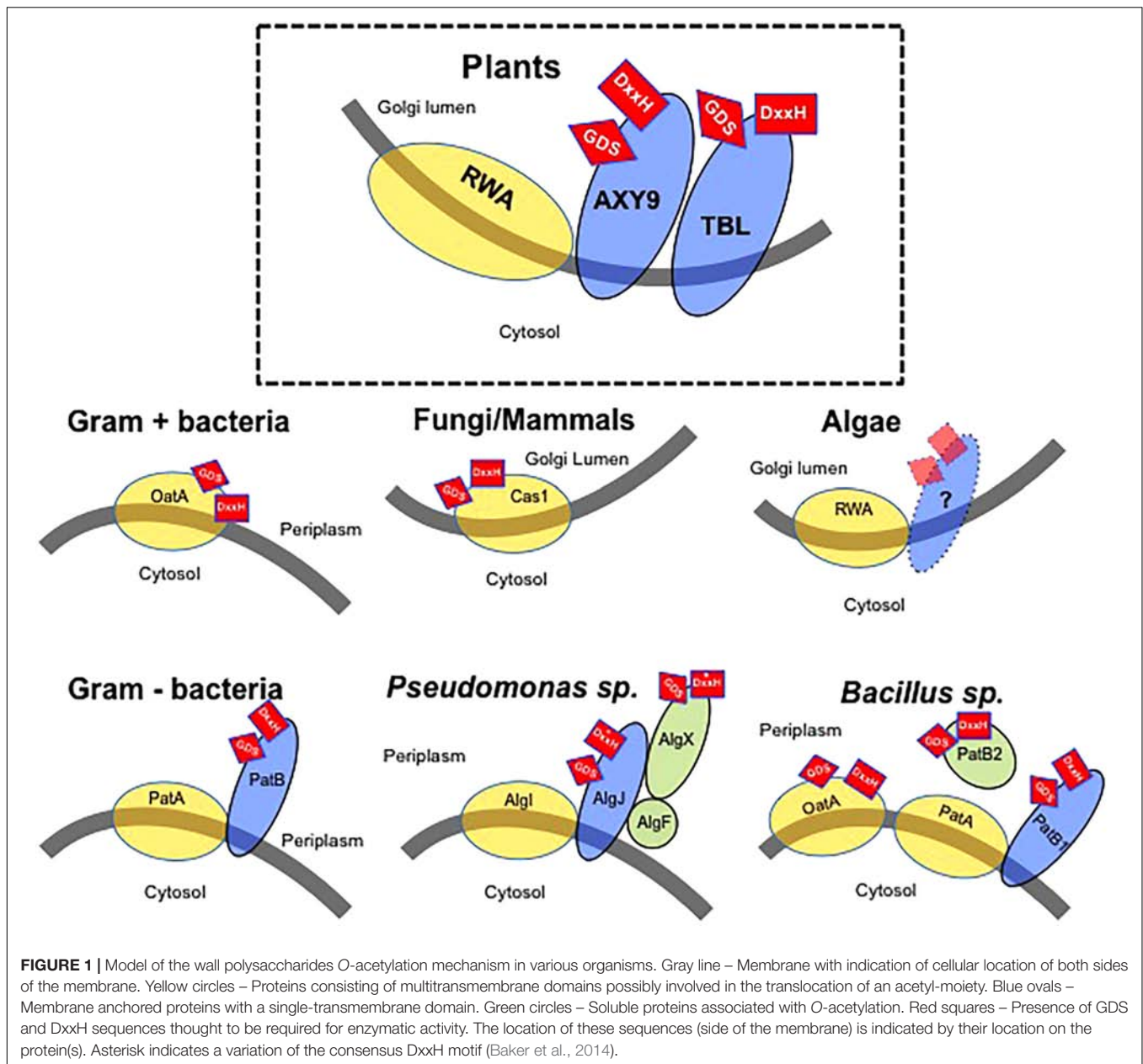
Another well studied example is the *Arabidopsis* *tbl29/eskimo1* mutant that was shown to reduce xylan O-acetylation by 46% in the stem (Xiong et al., 2013). The corresponding TBL29/ESKIMO1 protein was found to catalyze the transfer of O-acetyl groups to β -(1→4) xylooligosaccharides *in vitro* thus

confirming its role as a xylan O-acetyltransferase (Urbanowicz et al., 2014). Recently, the xylan O-acetyltransferase activities of other TBL proteins and their regiospecificity of xylose 2-O- and/or 3-O-acetylation has been demonstrated in *Arabidopsis*, rice, and poplar (Zhong et al., 2017, 2018a,b). In summary, in *Arabidopsis*, 9 TBLs lead to xylan 2-O-, 3-O-monoacetylation or 2,3-di-O-acetylation (Zhong et al., 2017). In rice, 66 TBL genes have been identified (Gao et al., 2017). Among these, 14 TBL proteins show xylan 2-O- and 3-O-acetyltransferase activity (OsXOAT1-14). OsXOAT1-7 are able to complement the defects in xylan O-acetylation of the *Arabidopsis* *esk1/tbl29* mutant (Zhong et al., 2018a). In poplar 64 TBLs were identified, 12 of those proteins were shown to O-acetylate xylan when heterologously expressed (Zhong et al., 2018b). Other members of the TBL family are thought to be involved in pectin O-acetylation such as AtPMR5/AtTBL44, AtTBR, and AtTBL3 (Vogel et al., 2004; Bischoff et al., 2010a) or mannan O-acetylation in the case of AtTBL25/AtTBL26 (Gille et al., 2011a). However, in all of these cases enzymatic activity and specificity remains to be demonstrated.

TRICHOME-BIREFRINGENCE-LIKE proteins contain three characteristic protein signatures (**Figure 1**) (Bischoff et al., 2010b). A N-terminus transmembrane domain and two plant-specific domains, DUF231 and TBL. The DUF231 is a domain of unknown function containing a conserved DxxH motif while the TBL motif is characterized by the presence of an esterase GDS motif. The Ser residue from the GDS motif and the Asp and His residues of the DxxH motif are essential for the function of TBL29/ESK1 as mutations of these residues result in a loss of enzyme activity (Zhong et al., 2017).

A second family of proteins involved in polysaccharide O-acetylation is represented by *ALTERED XYLOGLUCAN 9* (*AXY9*; **Figure 1**). *Arabidopsis* mutants affected in *AXY9* expression show a strong reduction in total wall O-acetylation in stems and leaf tissues (Schultink et al., 2015). Interestingly, unlike the large, diversified TBL gene family, *AXY9* seems to be present in the genome of land plants only as a single copy. Contrary to the polysaccharide substrate specificity of TBL proteins, *AXY9* seems to be non-specific in polysaccharide O-acetylation as the corresponding *axy9* mutant plants display reductions in O-acetylation of multiple hemicelluloses such as xyloglucan or xylan but not pectin. Due to these unique features, *AXY9* has been suggested to be involved in the generation of an intermediate acetyl donor substrate used later by TBL proteins (Schultink et al., 2015). The *AXY9* protein contains a N-terminus transmembrane domain and a C-terminus facing the Golgi lumen containing GDS and DxxH motifs (**Figure 1**) suggesting that it could also be an O-acetyltransferase although if this protein harbors any enzyme activity has yet to be determined.

REDUCED WALL O-ACETYLATION (RWA) is the third group of proteins involved in plant polysaccharide O-acetylation (Manabe et al., 2011) (**Figure 1**). The *Arabidopsis* genome contains four RWA proteins required for O-acetylation of both pectic and non-pectic polysaccharides including xyloglucan, xylan, and mannan. Quadruple *rwa* mutant plants exhibit a 63% reduction in total wall O-acetylation (Manabe et al., 2013). Similarly, down-regulation of the four RWA genes found in



hybrid aspen (*Populus tremula x tremuloides*) results in reduced wood xylan and xyloglucan O-acetylation, suggesting that RWA function is conserved among plant species (Pawar et al., 2017). RWA proteins are characterized by the presence of 10 predicted transmembrane domains (Manabe et al., 2011). In contrast, AXY9 and TBL proteins contain a single transmembrane domain anchoring the protein to the Golgi membrane while the C-terminus of these proteins is oriented toward the Golgi lumen containing the putative catalytic motifs. Despite the lack of amino acid similarity, all these enzymes are predicted to have a short N-terminal cytoplasmic region that has been proposed to act as a signal for retention in the Golgi in the case of other proteins such as glycosyltransferases (Banfield, 2011), although no experimental evidence has been obtained so far for

AXY9 or TBL proteins. Also, microsomal preparations isolated from potato cells incubated with radio-labeled acetyl-CoA are able to incorporate and transfer radioactive acetate to proteins and cell wall polysaccharides suggesting that acetyl-CoA is a donor-substrate for the O-acetylation of wall polysaccharides (Pauly and Scheller, 2000). As acetyl-CoA cannot diffuse through membranes and the Golgi is not able to produce it (Oliver et al., 2009), it has been proposed that RWA is responsible for the translocation of acetyl-groups across the membrane in order to supply the substrate to the other two families of O-acetyltransferases (AXY9 and the various TBLs). Although no experimental evidence has been reported yet, the existence of intermediary acetyl donor(s) is a likely option (Lee et al., 2011; Manabe et al., 2011, 2013; Schultink et al., 2015). In any case, the

cytosolic pool of acetyl-CoA is likely the source used by plants for the O-acetylation of polysaccharides as it is for alkaloids, anthocyanins, isoprenoids, or phenolics (Fatland et al., 2005; Oliver et al., 2009).

SIMILARITIES WITH OTHER POLYSACCHARIDE O-ACETYLATING ORGANISMS

All Gram-positive and most Gram-negative bacteria O-acetylate extracellular polysaccharides such as their cell wall peptidoglycan (PG) polymer. This heteropolymer is the main component of the bacterial wall, and consists of alternating N-acetylglucosaminyl-(β -1,4)-N-acetylmuramoyl residues cross-linked with stem peptides. PG O-acetylation can occur in 20–70% of the MurNAc residues, depending on the species and growth conditions and provides protection against lytic enzymes such as lysozyme (Moynihan and Clarke, 2011). In the last few years, a great effort has been made to identify and characterize the proteins involved in the O-acetylation of PG and other secondary cell wall polysaccharides due to the importance of this modification for the virulence of human pathogens such as *Neisseria gonorrhoeae*, *Bacillus anthracis*, or *Streptococcus pneumoniae* (Moynihan and Clarke, 2010; Moynihan et al., 2014; Sychantha et al., 2017, 2018). One can find surprising similarities of those systems with the polysaccharide O-acetylation mechanisms in plants indicating common ancestry.

In Gram-positive bacteria, OatA proteins consist of a N-terminal RWA-like multitransmembrane domain fused to a globular extracytoplasmic C-terminal domain containing a SGNH/GDSL esterase motif with similarity to plant TBL proteins (Figure 1). Hence, Gram-positive bacteria seem to be simultaneously translocating the acetyl groups from a cytoplasmic source and O-acetylating the N-acetylmuramoyl residues in the extracellular PG polysaccharide using a single bimodular protein. Several OatA homologs have been identified and characterized but only recently the crystal structure of the C-terminal domain of OatA has been resolved and point mutations in the DxxH and GDS motifs demonstrated that these amino acids are essential for catalyzing O-acetylation of PG in *Streptococcus pneumoniae* and *Staphylococcus aureus* (Sychantha et al., 2017). A similar protein combination consisting of a globular O-acetyltransferase domain combined with multiple transmembrane domains is also observed in fungi and mammals. The fungal CnCas1p protein is responsible for the O-acetylation of capsular glucuronoxylomannans in *Cryptococcus neoformans* (Janbon et al., 2001) (Figure 1). Although its activity has not been determined experimentally, the human HsCasD1 protein, showing high similarity and structure to CnCas1p, has been demonstrated to be essential and sufficient for O-acetylation of sialic acids, a family of nine-carbon monosaccharides typically found capping the glycan chains attached to cell surface glycoproteins and glycolipids in mammals including humans (Arming et al., 2011; Baumann et al., 2015). Similarly to bacterial OatA, activity assays showed that a N-terminus globular domain of HsCasD1 containing the SGNH/GDSL motif catalyzes the

9-O-acetylation of sialic acids *in vitro* (Baumann et al., 2015). These results suggest an ancient functional fusion between the multitransmembrane and globular domains in a single protein as a common mechanism to O-acetylate extracellular polysaccharides in Gram-positive bacteria, fungi, and mammals (Janbon et al., 2001; Anantharaman and Aravind, 2010; Baumann et al., 2015).

In Gram-negative bacteria multitransmembrane proteins have also been involved in O-acetylation of extracellular polysaccharides, such as Noll that O-acetylates lipo-chitin oligosaccharides in *Rhizobium* species, or GumG and GumF involved in the acetylation of the mannose residues of xanthan gum produced by *Xanthomonas oryzae* (Pacios Bras et al., 2000; Kim et al., 2009). However, the O-acetylation machinery of Gram-negative bacteria consists of multiple proteins as has been observed in plants (Figure 1). A multitransmembrane protein might translocate the acetyl moieties from a cytoplasmic source into the periplasm, where one or more plasma membrane-anchored proteins containing a SGNH/GDSL motif facing the periplasm might transfer the acetyl-moiety to the polysaccharide (Figure 1). This two-component mechanism involves the coordinated expression of multiple components arranged in operons. A model was originally proposed based on the O-acetylation of alginate, a linear exopolysaccharide consisting of 1-4-linked L-mannuronyl and D-glucuronyl residues present in *Pseudomonas aeruginosa* (Figure 1) (Clarke et al., 2000). In this bacterial species, the multi transmembrane domain protein AlgI has been suggested to play a similar role as OatA or RWA, exporting the acetyl groups from the cytoplasm. The available acetate would then be used by AlgJ and AlgF proteins, both containing a SGNH/GDSL motif. Although AlgJ and AlgF are both required for alginate O-acetylation, their precise functions have not been experimentally demonstrated and it has been proposed that they would not transfer acetyl groups directly to alginate. Instead, they would form a complex that could be acting as an intermediary step providing acetyl groups to AlgX, a protein shown to be able to O-acetylate the mannuronyl alginate residues *in vitro* (Baker et al., 2014). According to this model, the intermediate proteins AlgJ and AlgF might be analogous to AXY9 in plants, whereas AlgX would be catalyzing the final step in the O-acetylation of alginate, playing a similar role as the TBL protein family in plants. A similar mechanism has been postulated for other Gram-negative bacteria, including *N. gonorrhoeae* or *Campylobacter jejuni* (Figure 1) (Weadge et al., 2005; Moynihan and Clarke, 2010; Ha et al., 2016). In these Gram-negative bacteria, several homologs of AlgI (i.e., PatA proteins) are supposed to translocate the acetyl groups through the plasma membrane, whereas PatB proteins catalyze the transfer to the C6 hydroxyl groups of the PG muramoyl residues.

Despite the presence of proteins containing multiple transmembrane domains in both one- and multiple-component polysaccharide O-acetylating systems, proteins such as OatA, RWA2 or AlgI share very limited sequence homology with PatA. For example, SaOatA and NgPatA share only 15.1% sequence identity and 23.6% similarity. A similar situation occurs when comparing the O-acetyltransferase domain of plant TBL or

Gram-positive bacterial OatA proteins with the Gram-negative AlgX or PatB proteins. For example, there is only 15.4% identity and 18.3% similarity between the globular domain of SaOatA and HgPatB (Sychantha et al., 2017). This low degree of sequence similarity suggests different evolutive origins.

Interestingly, some *Bacillus* species seem to have two independent machineries to O-acetylate extracellular polysaccharides (Figure 1). On the one hand, a bimodal OatA homolog has been characterized exhibiting a mechanism as described above, involving the simultaneous translocation of acetyl groups and PG O-acetylation (Laaberki et al., 2011), whereas another system consisting of PatA1 and PatA2 multitransmembrane proteins and the PatB1 periplasmic O-acetyltransferase is responsible of the O-acetylation of secondary cell wall polysaccharides (Sychantha et al., 2017). Additionally, a second periplasmic protein with demonstrated acetyltransferase activity -PatB2- has also been involved in O-acetylation of additional cell wall components although the exact donor/acceptor substrate remains to be discovered (Sychantha et al., 2017). Hence, these organisms seem to have developed two different, independent systems for the translocation of acetyl-groups to then specifically O-acetylate the various wall polysaccharides utilizing members of two or more O-acetyl transferase families.

EVOLUTION OF PLANT POLYSACCHARIDE O-ACETYLATION MACHINERY

Gram-positive bacteria, fungi, and mammals developed a one component machinery to O-acetylate extracellular polymers. These systems use a single protein combining a multiple transmembrane domain translocating acetyl groups from the cytoplasm fused to a globular domain, containing a SGNH/GDSL-like catalytic motif. In plants, the protein domains and thus functionalities evolved into separate proteins (RWA, AXY9, and TBL protein families, respectively). As plants contain multiple wall polymers an expansion and increased diversification of the TBL protein family might have become necessary. Interestingly, although plant RWA proteins belong to the same sugar acyltransferase superfamily containing 10 transmembrane domains as bacterial OatAs, CnCas1p, and HsCasD1, they lack the globular O-acetyltransferase domain, indicating that plants need the additional involvement of other components such as members of the TBL family and/or AXY9 in order to O-acetylate their wall polysaccharides. Accordingly, the globular domain of OatA or CnCas1p proteins contains the GDS and DxxH motifs similar to plant TBLs and AXY9 hinting their analogous functions. A similar separate, multiple component mechanism was also developed by Gram-negative bacteria in order to O-acetylate extracellular polysaccharides, albeit likely arising through convergent evolution. The development of a multiple component system in these bacteria could reflect again a more complex wall with a variety of extracellular O-acetylated polysaccharides. In these bacteria, an increased diversification

of the O-acetyltransferases is also observed (e.g., PatB1/PatB2 in *B. anthracis* of AlgF, AlgJ and AlgX in *P. aeruginosa*).

All three families of proteins involved in O-acetylation of plant wall polysaccharides can be found in vascular plants but also in pteridophytes and bryophytes, including hornworts, mosses, and liverworts (Figure 2). A sequence comparison of nine representative embryophytic species showed that AXY9, TBL29, and RWA2 proteins seems to be highly conserved in dicots (*Arabidopsis thaliana* and *Populus trichocarpa*), monocots (*Oryza sativa*) and gymnosperms (*Pinus radiata*) sharing identities higher than 50% and similarities around 75% with the *Arabidopsis* representatives. Primitive plants such as *Equisetum hyemale*, liverworts (*Marchantia polymorpha*), hornworts (*Phaeoceros carolinianus*) and mosses (*Physcomitrella patens*) also contain highly conserved sequences sharing identity and similarity values around 40% and 60%, respectively.

Land plants evolved from Charophyte green algae after their separation from Chlorophyte green algae (Lewis and McCourt, 2004; Becker and Marin, 2009). Although during the transition from an aquatic to terrestrial environment cell walls in both algae and plants have evolved independently (Niklas, 2004), it is still likely that some of the wall components have a common ancestry (Popper and Tuohy, 2010). Accordingly, the biosynthetic machinery of some of the polysaccharides present in a typical plant wall (e.g., xylan) can be traced back to the Charophyte green algae (Jensen et al., 2018). When probing algal genomes with the *Arabidopsis* RWA2 sequence, homologs can be found in dozens of green algae species including members of both the Chlorophyta (e.g., *Volvox aureus* or *Nephroselmis pyriformis*) and Charophyta (e.g., *Klebsormidium subtile* or *Coleochaete scutata*) divisions. However, algae seem not to encode proteins with sequence similarity to AXY9 or TBL29. Since algal RWA orthologs do not contain a GxxH and/or GSD domain required for polysaccharide O-acetylation algae might harbor additional, hitherto unidentified proteins that would be necessary for O-acetylation to occur. These results indicate that RWA proteins emerged earlier than AXY9 and the TBLs and suggest that green algae may also use a polysaccharide O-acetylation system based on RWA. The walls of several Chlorophyta and Charophyta species have been reported to contain plant-type wall polysaccharides such as xylan, mannans or XyG (Painter, 1983; Lahaye et al., 1994; Lahaye and Robic, 2007; Popper et al., 2011). Unfortunately, information about the O-acetylation status of these organisms is missing probably due to the alkali-based methods used during wall isolation.

BIOLOGICAL SIGNIFICANCE OF POLYSACCHARIDE O-ACETYLATION

O-acetylation of polysaccharides, including the various hemicelluloses and the pectic polysaccharides homogalacturonan and rhamnogalacturonan I, influences the polymer's physiochemical properties. Addition of O-acetyl-moieties contribute to the gelling properties and viscosity of the isolated polymers an important issue for food applications (Rombouts and Thibault, 1986; Huang et al., 2002). O-acetyl

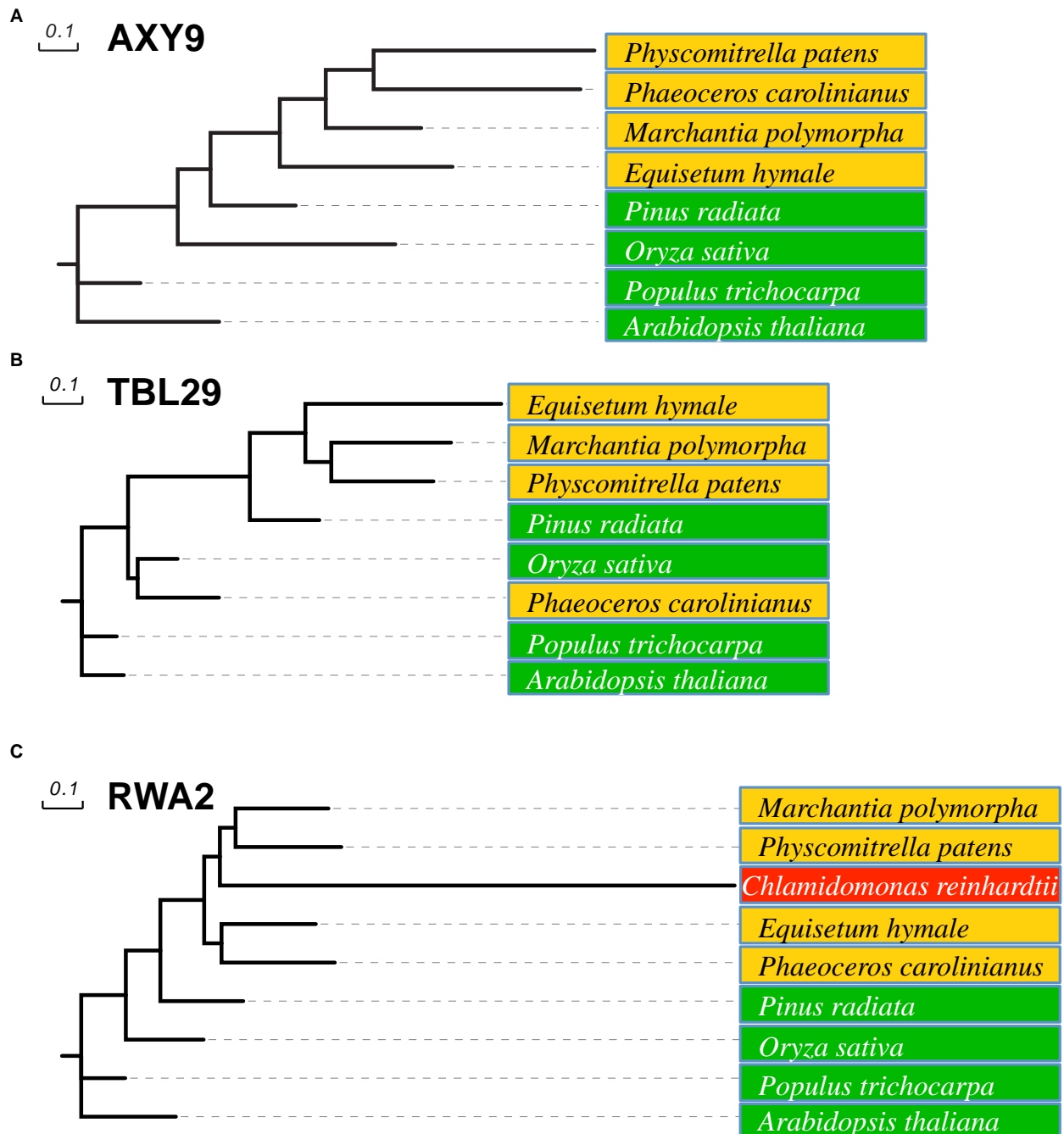


FIGURE 2 | Phylogenetic tree of AXY9, TBL, and RWA proteins. Likelihood tree of AXY9 **(A)**, TBL29 **(B)**, and RWA2 **(C)** protein homologs constructed from sequence alignment of selected species. Green: embryophytes (*Arabidopsis thaliana* and *Populus trichocarpa*, dicots; *Oryza sativa*, monocot; and *Pinus radiata*, gymnosperm). Orange: Bryophyta (*Marchantia polymorpha*, liverwort; *Phaeoceros carolinianus*, hornwort; and *Physcomitrella patens*, moss) and Pteridophyta (*Equisetum hymale*, horsetail). Red: Algae (*Chlamydomonas reinhardtii*, green algae). *Arabidopsis thaliana* AXY9, TBL29, and RWA2 protein sequences (UniProtKB references Q9M9N9-1, Q9LY46-1, and Q0WW17-4, respectively) were used in Basic Alignment Search tool protein (BLASTp) against the 1,000 Plants Initiative databases (Matasci et al., 2014; <https://db.cngb.org/blast/blastp/>) with default parameters and the best hits for every specie were selected for phylogenetic analysis using the Phylogeny.fr web tool with default settings (Dereeper et al., 2008). This tool uses MUSCLE to align the sequences and the Gblocks program to eliminate poorly aligned positions and divergent regions. Phylogenetic trees were then constructed using PhyML using default parameters (Approximate Likelihood-Ratio Test) and the Evolview tool (<http://www.evolgenius.info>) was used to edit the graphical representation.

substituents increase polysaccharide hydrophobicity and lead to conformational changes that influences interactions with other polymers, either supporting binding or due to steric hindrance abolish interaction (Busse-Wicher et al., 2014). As a result, de-O-acetylation through, e.g., alkali-treatments leads often to a decrease in solubility in aqueous environments and precipitation of polymers (Gibeaut et al., 2005; Busse-Wicher et al., 2016). Moreover, enzymatic attack of the polymer by glycosyl hydrolases is restricted due to steric hindrance in the vicinity of the target glycosidic bond (reviewed by Biely et al., 2016). As an application example in the wood industry, biomass chemical treatments include chaotropic alkali and acetic anhydride treatments in order to de-acetylate and re-acetylate the lignocellulosic polysaccharides, respectively, to modify the wood properties. De-acetylation improves properties for pulping, saccharification and fermentation due to the properties mentioned above. On the other hand, chemical acetylation of wood increases mechanical strength, durability and resistance to fungi, bacteria, and termites, as acetylation of xylan and mannan increases the stiffness and allows interactions with hydrophobic substances such as lignin (reviewed in Pawar et al., 2013). However, in non-lignified tissues, de-acetylation of primary wall polysaccharides (e.g., pectic polysaccharides) has been associated with increased cell wall stiffness probably due to a close spatial association between pectin and cellulose microfibrils (Gou et al., 2012; Orfila et al., 2012).

In *planta* the biological significance of a particular polysaccharide O-acetylation pattern is diverse and in many cases not clear. For example, a complete lack of XyG sidechain O-acetylation has no apparent impact on plant growth and development as the wild-type (WT)-like phenotypes of *Arabidopsis axy4* and *axy4L* knockout mutants demonstrate. Reinforcing this notion, a natural *Ty-0* *Arabidopsis* accession displays an almost complete lack of XyG O-acetylation without detrimental plant morphological and developmental side-effects when grown in its native environment in the Highlands of Scotland (Gille et al., 2011b). However, O-acetylation seems to affect the aluminum binding capacity of XyG as demonstrated by an increased aluminum content in the hemicellulose fraction in *axy4* mutant roots compared to the WT when growing in the presence of this metal (Zhu et al., 2014). Yet, one cannot rule out the possibility that XyG O-acetylation may play a role in other environmental adaptation processes including specific stresses and/or growing conditions yet to be identified. In contrast to dicots such as *Arabidopsis* or poplar, in the grasses and plant members of the Solanaceae (such as tomato, tobacco, etc.) the glucan-backbone of XyG is partially O-acetylated (Gibeaut et al., 2005; Jia et al., 2005). This is caused by XyG O-acetyltransferases such as the *Brachypodium BdXyBAT1* that O-acetylate the non-xylosylated glucosyl backbone residues (Jia et al., 2005; Liu et al., 2016). When *BdXyBAT1* is expressed in *Arabidopsis*, the backbone of XyG becomes O-acetylated reducing the degree of xylosylation of XyG indicating that O-acetylation impacts negatively the addition of other substitutions (Liu et al., 2016). A reduction of the size of glycosyl side-chains of XyG lacking, e.g., the fucosyl and galactosyl residues leads to retarded plant growth (Pauly et al., 2013; Schultink et al., 2013).

It is thought that this dwarfism is caused by a distorted matrix polysaccharide secretion system caused by the poor solubility of the less substituted XyG (Jensen et al., 2012; Kong et al., 2015). However, the addition of backbone O-acetyl substituents to this lowly substituted XyG in the *Arabidopsis* mutant results in a reversion of the retarded growth (Liu et al., 2016). These results indicate that O-acetylation of the XyG glucan-backbone is functionally equivalent to glycosyl-sidechains and might represent an energetically favorable strategy by replacing C5 and C6 carbon sugars with C2 acetates (Gibeaut et al., 2005; Jia et al., 2005; Gille and Pauly, 2012; Liu et al., 2016).

Mutants affected in xylan O-acetylation display multiple pleiotropic phenotypes including dwarfism, altered plant architecture and constitutive stress-related responses associated with a vascular collapse. Xylan is a major component of the walls present in the water conducting xylem. Hypoacetylation of xylan seems to affect the physical strength of the xylem walls, as they are not able to resist the negative water pressure generated during water transport. As a consequence, mutants affected in members of the *AXY9*, *RWA* or particular TBLs that impact xylan O-acetylation all display alterations in plant growth and development. In *Arabidopsis*, the *axy9-2* mutant shows an 80% reduction in xylan O-acetylation and a strong growth arrest (Schultink et al., 2015), whereas the quadruple *rwa* mutant shows a 42% reduction in xylan O-acetylation with a reduction in secondary wall thickening and collapsed xylem morphology (Lee et al., 2011). Regarding the TBL family, only *tbl29/esk1* single mutant alleles, with a 40% reduction in xylan O-acetylation, show a clear irregular xylem phenotype. Several other *tbl* single mutants with only minor reductions in xylan O-acetylation show only additive effects in the corresponding double, triple or multiple mutant combinations in vascular development and plant growth in several plant species (Yuan et al., 2016a,b,c; Gao et al., 2017).

In addition to the xylem collapse and growth arrest, xylan hypoacetylation has also been associated with other developmental and stress-related phenotypes. *tbl29/esk1* mutant alleles also show stress-related pleiotropic phenotypes such as increased tolerance to drought, salt or freezing, likely an indirect consequence of the collapsed xylem (Xin and Browse, 1998; Xin et al., 2007; Bouchabke-Coussa et al., 2008; Lefebvre et al., 2011; Ramirez et al., 2018). Intriguingly, several lines of evidence seem to indicate that low xylan acetylation may not be directly responsible for these observed phenotypes. For example, expression of fungal acetyl-esterases in *Arabidopsis* and *Brachypodium* causes post-synthetic de-acetylation of xylan but it does not impact plant development or xylem morphology (Pogorelko et al., 2013b). Most recently, the identification of two *tbl29* suppressors, where the xylem collapse and growth arrest are recovered but the wall/xylan acetate remains reduced, strongly supports these observations. *KAKTUS* (*KAK*) loss of function increases stem diameter and activates the development of larger tracheary elements. As a consequence, *kak* mutations are able to recover almost completely from *tbl29/esk1*-associated dwarfism without affecting wall acetate content. Although *KAK* has been described previously as an endoreduplication repressor affecting trichome morphology, the mechanism how it regulates

vascular development is not known (Downes et al., 2003; El Refy et al., 2003; Bensussan et al., 2015). Altered biosynthesis and/or perception of some plant hormones (e.g., abscisic acid; ABA) have been suggested to play a role in the pleiotropic phenotype of *tbl29/esk1*. *tbl29/esk1* alleles show increased ABA levels and enhanced expression of several ABA-dependent genes, but genetic evidence discarded that this hormonal pathway is directly responsible for the phenotypes of *tbl29/esk1* plants. Double mutants blocking ABA biosynthesis or perception in a *tbl29/esk1* background are not able to recover the developmental defects shown by the *tbl29/esk1* single mutant. Moreover, increased ABA perception and *tbl29/esk1* down-regulation seems to have additive effects on drought tolerance, suggesting that they affect independent pathways (Lefebvre et al., 2011). These findings indicate that altered ABA levels are more a consequence of the pleiotropic phenotype of the *tbl29/esk1* mutant rather than the cause. In contrast, it has been shown that blocking strigolactone (SL) synthesis in *tbl29/esk1* plants (i.e., *tbl29 max4* double mutants) is able to completely suppress both developmental defects and increased freezing tolerance without affecting the reduced acetate content (Ramírez et al., 2018). In addition, exogenous applications of a synthetic SL to *tbl29 max4* plants result in dwarfism and collapsed xylem, further confirming that these phenotypes are SL-dependent. This suggests that an altered SL pathway could be directly involved in leading to the pleiotropic phenotypes associated to the *tbl29/esk1* mutants. As SLs are hormones involved in the regulation of multiple plant processes including stem elongation, secondary growth, leaf expansion and adaptation to abiotic stress (reviewed in Waters et al., 2017), this opens the possibility that xylan hypoacetylation could be perceived by an unknown mechanism triggering the activation a SL-dependent response regulating xylem development (Ramírez et al., 2018).

In addition to acetate, xylan can also be substituted with (methyl-)glucuronic acid (methyl-GlcA) residues by xylan glucuronosyltransferases termed GUX. Together, these decorations have been found to be important for xylan-cellulose binding (Mortimer et al., 2010; Rennie et al., 2012; Bromley et al., 2013; Busse-Wicher et al., 2014). Actually, vascular plants seem to generate a specific xylan decoration pattern as acetate and GlcA are found spaced on even-numbered residues in the xylan backbone (Busse-Wicher et al., 2016). Recently, it has been shown that in *Arabidopsis*, TBL29/ESK1-dependent xylan O-acetylation is required for the generation of the even-patterned GlcA substitutions (Grantham et al., 2017). In a *tbl29/esk1* mutant, where xylan acetylation is reduced, GUX1 is unable to maintain the GlcA decoration pattern suggesting that a correct O-acetylation pattern is required for the addition of GlcA residues. As a consequence of this uneven substitution, xylan might not be able to acquire the typical twofold screw ribbon conformation impeding its docking onto the hydrophilic face of a cellulose microfibril to form semicrystalline xylanocellulose fibrils (Grantham et al., 2017). Intriguingly, expression of GUX1 in vascular tissue under the control of a tissue specific promoter is able to rescue the *tbl29/esk1* mutant growth defects indicating that xylan functionality is restored (Xiong et al., 2015). GUX1 is able to glucuronosylate additional available positions on the

xylan backbone due to the absence of O-acetyl-groups in *tbl29*. Glucuronosylation of xylan can thus be considered functionally equivalent to O-acetylation *in vivo*. This agrees with the notion that the addition of O-acetyl substituents (C2 units) to wall polysaccharides instead of sugars (C5-C6) could have evolved as a more energetically favorable strategy as described above for XyG. Other TBL proteins than TBL29/ESK1 participate in the regiospecific O-acetylation of xylan (Zhong et al., 2017), suggesting the existence of a precisely regulated mechanism to create a tissue-specific O-acetylation pattern in xylan in order to adequately interact with cellulose and likely other cell wall components.

O-acetylation of pectic polysaccharides has also been associated with plant signaling processes. *rwa2* mutant alleles show increased resistance to the necrotrophic fungal pathogen *Botrytis cinerea* accompanied by leaf surface defects including trichome collapse, enhanced leave permeability and altered cuticle formation (Manabe et al., 2011; Nafisi et al., 2015). As these defects have not been observed in other mutants affected in hemicellulosic polysaccharide O-acetylation, it has been speculated that these phenotypes may be caused by pectin hypoacetylation although direct evidence is still lacking (Nafisi et al., 2015). Other reports have associated reduced pectin O-acetylation with increased disease resistance. For example, plants overexpressing a fungal rhamnogalacturonan acetyltransferase constitutively activate defense responses and show increased resistance to pathogens (Pogorelko et al., 2013b). Since a similar response has been observed after application of oligogalacturonide fragments (OGs) and more efficiently by partially acetylated OGs, it has been proposed that pectin O-acetylation might be involved in a cell wall integrity maintenance system (Randoux et al., 2010; Pogorelko et al., 2013a).

Recently, pectin O-acetylation has been also proposed to regulate other important developmental processes such as photomorphogenesis (Sinclair et al., 2017). Both the *tbr* mutant, affected in a putative pectin acetyltransferase, and the *rwa2* mutant, show a photomorphogenic response when grown in the dark. This phenotype can be restored by adding small homogalacturonan fragments and thus pectin O-acetylation might regulate a dark signal involved in a complex network of light-dependent seedling development (Sinclair et al., 2017).

OPEN QUESTIONS

There are still open questions regarding the mechanism of polysaccharide O-acetylation not only in plants but also in bacteria, fungi, and mammals. First, the identity of the acetyl-donor and the detailed mechanism of translocation through the Golgi membrane are not known in any of these organisms. Although it is likely that the cytosolic acetyl-CoA pool is tapped for this purpose, acetyl-CoA itself is likely not transferred. Studying this process is challenging as the likely responsible protein contains multiple transmembrane domains. Second, the exact mechanism of the transfer of an acetyl group from a donor to the hydroxyl group of an acceptor sugar remains unknown,

although mechanistic insights became recently available from bacterial OatA proteins (Sythantha and Clarke, 2018). Recent reports in bacteria have suggested a direct acylation of the OatA protein following a ping-pong bi-bi mechanism of action where the acetyl group is covalently attached to the catalytic Ser residue of the enzyme before being transferred to the substrate. A similar acetyl-enzyme intermediate has also been proposed in the Gram-negative PatA/PatB mode of action, where PatB O-acetyltransferases could form a complex with the acetyl-bound PatA membrane proteins precluding free water from accessing the active site, preventing the hydrolysis of the translocated acetyl group ensuring an efficient acetate transfer (Moynihan and Clarke, 2010). The existence of likely intermediary steps as suggested in other Gram-negative bacteria systems (e.g., AlgI/AlgF/AlgJ/AlgX) could implicate the formation of a multiprotein complex for the O-acetylation of extracellular polysaccharides. A RWA/AXY9/TBL complex formation could also be conserved in plant systems, although mechanistic details are still missing. Comparison of the various polysaccharide O-acetylation systems raises the question of the evolution of the various O-acetylation mechanism – in essence why multiple proteins are required for this process in some species while in other species apparently a single protein suffices. Third, the transferases responsible for the O-acetylation of some wall polymers (e.g., mannans, pectins, or lignin) remain

to be discovered, albeit it is likely that members of the TBL family are involved. The identification and characterization of such proteins is not only needed to understand the wall O-acetylation mechanism of particular wall polysaccharides, but also to gain insights into the function of the O-acetyl substituent on this polymer. Fourth, the recent advent of identifying the genes responsible for polysaccharide O-acetylation and their genetic manipulation *in vivo* lead to the discovery of intriguing function of this substituent. However, at this stage the phenotypic results are rather descriptive and additional research in the future is required to ascertain causal relationships as well as mechanistic insights into polymer interactions, cellular sensing and responses.

AUTHOR CONTRIBUTIONS

MP and VR designed and wrote the manuscript.

FUNDING

This research was supported by CEPLAS (Cluster of Excellence on Plant Sciences – Deutsche Forschungsgemeinschaft EXC1028) and Marie Curie PIOF-GA-2013-623553 to VR.

REFERENCES

- Anantharaman, V., and Aravind, L. (2010). Novel eukaryotic enzymes modifying cell-surface biopolymers. *Biol. Direct* 5:1. doi: 10.1186/1745-6150-5-1
- Arming, S., Wipfler, D., Mayr, J., Merling, A., Vilas, U., Schauer, R., et al. (2011). The human Cas1 protein: a sialic acid-specific O-acetyltransferase? *Glycobiology* 21, 553–564. doi: 10.1093/glycob/cwq153
- Baker, P., Ricer, T., Moynihan, P. J., Kitova, E. N., Walvoort, M. T., Little, D. J., et al. (2014). P. aeruginosa SGNH hydrolase-like proteins AlgJ and AlgX have similar topology but separate and distinct roles in alginate acetylation. *PLoS Pathog.* 10:e1004334. doi: 10.1371/journal.ppat.1004334
- Banfield, D. K. (2011). Mechanisms of protein retention in the Golgi. *Cold Spring Harb. Perspect. Biol.* 3:a005264. doi: 10.1101/cshperspect.a005264
- Baumann, A. M., Bakkers, M. J., Buettner, F. F., Hartmann, M., Grove, M., Langereis, M. A., et al. (2015). 9-O-Acetylation of sialic acids is catalysed by CASD1 via a covalent acetyl-enzyme intermediate. *Nat. Commun.* 6:7673. doi: 10.1038/ncomms8673
- Becker, B., and Marin, B. (2009). Streptophyte algae and the origin of embryophytes. *Ann. Bot.* 103, 999–1004. doi: 10.1093/aob/mcp044
- Bensussan, M., Lefebvre, V., Ducamp, A., Trouverie, J., Gineau, E., Fortabat, M. N., et al. (2015). Suppression of dwarf and irregular xylem phenotypes generates low-acetylated biomass lines in Arabidopsis. *Plant Physiol.* 168, 452–463. doi: 10.1104/pp.15.00122
- Biely, P., Singh, S., and Puchart, V. (2016). Towards enzymatic breakdown of complex plant xylan structures: state of the art. *Biotechnol. Adv.* 34, 1260–1274. doi: 10.1016/j.biotechadv.2016.09.001
- Bischoff, V., Nita, S., Neumetzler, L., Schindelasch, D., Urbain, A., Eshed, R., et al. (2010a). TRICHOME BIREFRINGENCE and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in Arabidopsis. *Plant Physiol.* 153, 590–602. doi: 10.1104/pp.110.153320
- Bischoff, V., Selbig, J., and Scheible, W. R. (2010b). Involvement of TBL/DUF231 proteins into cell wall biology. *Plant Signal. Behav.* 5, 1057–1059.
- Bouchabke-Coussa, O., Quashie, M. L., Seoane-Redondo, J., Fortabat, M. N., Gery, C., Yu, A., et al. (2008). ESKIMO1 is a key gene involved in water economy as well as cold acclimation and salt tolerance. *BMC Plant Biol.* 8:125. doi: 10.1186/1471-2229-8-125
- Bromley, J. R., Busse-Wicher, M., Tryfona, T., Mortimer, J. C., Zhang, Z., Brown, D. M., et al. (2013). GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *Plant J.* 74, 423–434. doi: 10.1111/tjp.12135
- Busse-Wicher, M., Gomes, T. C., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N. J., et al. (2014). The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of *Arabidopsis thaliana*. *Plant J.* 79, 492–506. doi: 10.1111/tjp.12575
- Busse-Wicher, M., Li, A., Silveira, R. L., Pereira, C. S., Tryfona, T., Gomes, T. C. F., et al. (2016). Evolution of xylan substitution patterns in gymnosperms and angiosperms: implications for xylan interaction with cellulose. *Plant Physiol.* 171, 2418–2431. doi: 10.1104/pp.16.00539
- Chundawat, S. P. S., Vismeh, R., Sharma, L. N., Humpala, J. F., da Costa Sousa, L., Chambliss, C. K., et al. (2010). Multifaceted characterization of cell wall decomposition products formed during ammonia fiber expansion (AFEX™) and dilute acid based pretreatments. *Bioresour. Technol.* 101, 8429–8438. doi: 10.1016/j.biortech.2010.06.027
- Clarke, A. J., Strating, H., and Blackburn, N. T. (2000). “Pathways for the O-acetylation of bacterial cell wall polymers,” in *Glycomicrobiology*, ed. R. J. Doyle (New York, NY: Plenum Publishing Co. Ltd), 187–212.
- de Souza, A., Hull, P. A., Gille, S., and Pauly, M. (2014). Identification and functional characterization of the distinct plant pectin esterases PAE8 and PAE9 and their deletion mutants. *Planta* 240, 1123–1138. doi: 10.1007/s00425-014-2139-6
- Del Río, J. C., Marques, G., Rencoret, J., Martínez, A. T., and Gutiérrez, A. (2007). Occurrence of naturally acetylated lignin units. *J. Agric. Food Chem.* 55, 5461–5468. doi: 10.1021/jf0705264
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, W465–W469. doi: 10.1093/nar/gkn180
- Downes, B. P., Stupar, R. M., Gingerich, D. J., and Vierstra, R. D. (2003). The HECT ubiquitin-protein ligase (UPL) family in *Arabidopsis*: UPL3 has a specific role in trichome development. *Plant J.* 35, 729–742. doi: 10.1046/j.1365-3113.2003.01844.x
- El Refy, A., Perazza, D., Zekraoui, L., Valay, J. G., Bechtold, N., Brown, S., et al. (2003). The *Arabidopsis* KAKTUS gene encodes a HECT protein and controls

- the number of endoreduplication cycles. *Mol. Genet. Genomics* 270, 403–414. doi: 10.1007/s00438-003-0932-1
- Fatland, B. L., Nikolau, B. J., and Wurtele, E. S. (2005). Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP-citrate lyase in *Arabidopsis*. *Plant Cell* 17, 182–203. doi: 10.1105/tpc.104.026211
- Gao, Y., He, C., Zhang, D., Liu, X., Xu, Z., Tian, Y., et al. (2017). Two trichome birefringence-like proteins mediate xylan acetylation, which is essential for leaf blight resistance in rice. *Plant Physiol.* 173, 470–481. doi: 10.1104/pp.16.01618
- Gibeaut, D. M., Pauly, M., Bacic, A., and Fincher, G. B. (2005). Changes in cell wall polysaccharides in developing barley (*Hordeum vulgare*) coleoptiles. *Planta* 221, 729–738. doi: 10.1007/s00425-005-1481-0
- Gille, S., Cheng, K., Skinner, M. E., Liepman, A. H., Wilkerson, C. G., and Pauly, M. (2011a). Deep sequencing of voodoo lily (*Amorphophallus konjac*): an approach to identify relevant genes involved in the synthesis of the hemicellulose glucomannan. *Planta* 234, 515–526. doi: 10.1007/s00425-011-1422-z
- Gille, S., de Souza, A., Xiong, G., Benz, M., Cheng, K., Schultink, A., et al. (2011b). O-acetylation of *Arabidopsis* hemicellulose xyloglucan requires AX4 or AX4L, proteins with a TBL and DUF231 domain. *Plant Cell* 23, 4041–4053. doi: 10.1105/tpc.111.091728
- Gille, S., and Pauly, M. (2012). O-acetylation of plant cell wall polysaccharides. *Front. Plant Sci.* 3:12. doi: 10.3389/fpls.2012.00012
- Gou, J. Y., Miller, L. M., Hou, G., Yu, X. H., Chen, X. Y., and Liu, C. J. (2012). Acetyltransferase-mediated deacetylation of pectin impairs cell elongation, pollen germination, and plant reproduction. *Plant Cell* 24, 50–65. doi: 10.1105/tpc.111.092411
- Grantham, N. J., Wurman-Rodrich, J., Terrett, O. M., Lyczakowski, J. J., Stott, K., Iuga, D., et al. (2017). An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat. Plants* 3, 859–865. doi: 10.1038/s41477-017-0030-8
- Ha, R., Frirdich, E., Sychantha, D., Biboy, J., Taveirne, M. E., Johnson, J. G., et al. (2016). Accumulation of peptidoglycan O-acetylation leads to altered cell wall biochemistry and negatively impacts pathogenesis factors of *Campylobacter jejuni*. *J. Biol. Chem.* 291, 22686–22702. doi: 10.1074/jbc.M116.746404
- Helle, S., Cameron, D., Lam, J., and White, B. (2003). Effect of inhibitory compounds found in biomass hydrolysates on growth and xylose fermentation by a genetically engineered strain of *S. cerevisiae*. *Enzyme Microb. Technol.* 33, 786–792. doi: 10.1016/S0141-0229(03)00214-X
- Hoffman, M., Jia, Z., Pena, M., Cash, M., Harper, A., Blackburn, A., et al. (2005). Structural analysis of xyloglucans in the primary cell walls of plants in the subclass *Asteridae*. *Carbohydr. Res.* 340, 1826–1840. doi: 10.1016/j.carres.2005.04.016
- Hsieh, Y. S. Y., and Harris, P. J. (2009). Xyloglucans of monocotyledons have diverse structures. *Mol. Plant* 2, 943–965. doi: 10.1093/mp/ssp061
- Huang, L., Takahashi, R., Kobayashi, S., Kawase, T., and Nishinari, K. (2002). Gelation behavior of native and acetylated konjac glucomannan. *Biomacromolecules* 3, 1296–1303. doi: 10.1021/bm0255995
- Ishii, T. (1991). Acetylation at O-2 of arabinofuranose residues in feruloylated arabinoxylan from bamboo shoot cell-walls. *Phytochemistry* 30, 2317–2320. doi: 10.1016/0031-9422(91)83639-3
- Ishii, T. (1997). O-acetylated oligosaccharides from pectins of potato tuber cell walls. *Plant Physiol.* 113, 1265–1272. doi: 10.1104/pp.113.4.1265
- Janbon, G., Himmelreich, U., Moyrand, F., Improvisi, L., and Dromer, F. (2001). Cas1p is a membrane protein necessary for the O-acetylation of the *Cryptococcus neoformans* capsular polysaccharide. *Mol. Microbiol.* 42, 453–467. doi: 10.1046/j.1365-2958.2001.02651.x
- Jensen, J. K., Busse-Wicher, M., Poulsen, C. P., Fangel, J. U., Smith, P. J., Yang, J. Y., et al. (2018). Identification of an algal xylan synthase indicates that there is functional orthology between algal and plant cell wall biosynthesis. *New Phytol.* 218, 1049–1060. doi: 10.1111/nph.15050
- Jensen, J. K., Schultink, A., Keegstra, K., Wilkerson, C. G., and Pauly, M. (2012). RNA-Seq analysis of developing nasturtium seeds (*Tropaeolum majus*): identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis. *Mol. Plant* 5, 984–992. doi: 10.1093/mp/sss032
- Jia, Z., Cash, M., Darvill, A., and York, W. (2005). NMR characterization of endogenously O-acetylated oligosaccharides isolated from tomato (*Lycopersicon esculentum*) xyloglucan. *Carbohydr. Res.* 340, 1818–1825. doi: 10.1016/j.carres.2005.04.015
- Johnson, A. M., Kim, H., Ralph, J., and Mansfield, S. D. (2017). Natural acetylation impacts carbohydrate recovery during deconstruction of *Populus trichocarpa* wood. *Biotechnol. Biofuels* 10:48. doi: 10.1186/s13068-017-0734-z
- Kabel, M. A., de Waard, P., Schols, H. A., and Voragen, A. G. J. (2003). Location of O-acetyl substituents in xylo-oligosaccharides obtained from hydrothermally treated eucalyptus wood. *Carbohydr. Res.* 338, 69–77. doi: 10.1016/S0008-6215(02)00351-8
- Kiefer, L., York, W., Darvill, A., and Albersheim, P. (1989). Structure of plant-cell walls. 27. Xyloglucan isolated from suspension-cultured sycamore cell-walls is o-acetylated. *Phytochemistry* 28, 2105–2107. doi: 10.1016/S0031-9422(00)97928-7
- Kim, S. Y., Kim, J. G., Lee, B. M., and Cho, J. Y. (2009). Mutational analysis of the *gum* gene cluster required for xanthan biosynthesis in *Xanthomonas oryzae* pv *oryzae*. *Biotechnol. Lett.* 31, 265–270. doi: 10.1007/s10529-008-9858-3
- Kong, Y., Peña, M. J., Renna, L., Avci, U., Pattathil, S., Tuomivaara, S. T., et al. (2015). Galactose-depleted xyloglucan is dysfunctional and leads to dwarfism in *Arabidopsis*. *Plant Physiol.* 167, 1296–1306. doi: 10.1104/pp.114.255943
- Laaberki, M. H., Pfeffer, J., Clarke, A. J., and Dworkin, J. (2011). O-acetylation of peptidoglycan is required for proper cell separation and s-layer anchoring in *Bacillus anthracis*. *J. Biol. Chem.* 286, 5278–5288. doi: 10.1074/jbc.M110.183236
- Lahaye, M., Jegou, D., and Buléon, A. (1994). Chemical characteristics of insoluble glucans from the cell wall of the marine green alga *Ulva lactuca* (L.) Thuret. *Carbohydr. Res.* 262, 115–125. doi: 10.1016/0008-6215(94)84008-3
- Lahaye, M., and Robic, A. (2007). Structure and functional properties of ulvan, a polysaccharide from green seaweeds. *Biomacromolecules* 8, 1765–1774. doi: 10.1021/bm061185q
- Lee, C., Teng, Q., Zhong, R., and Ye, Z.-H. (2011). The four *Arabidopsis* reduced wall acetylation genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. *Plant Cell Physiol.* 52, 1289–1301. doi: 10.1093/pcp/pcr075
- Lefebvre, V., Fortabat, M. N., Ducamp, A., North, H. M., Maia-Grondard, A., Trouverie, J., et al. (2011). *ESKIMO1* disruption in *Arabidopsis* alters vascular tissue and impairs water transport. *PLoS One* 6:e16645. doi: 10.1371/journal.pone.0016645
- Lewis, L. A., and McCourt, R. M. (2004). Green algae and the origin of land plants. *Am. J. Bot.* 91, 1535–1556. doi: 10.3732/ajb.91.10.1535
- Liu, L., Hsia, M. M., Dama, M., Vogel, J., and Pauly, M. (2016). A xyloglucan backbone 6-O-acetyltransferase from *Brachypodium distachyon* modulates xyloglucan xylosylation. *Mol. Plant* 9, 615–617. doi: 10.1016/j.molp.2015.11.004
- Lourenço, A., Rencoret, J., Chemetova, C., Gominho, J., Gutierrez, A., del Rio, J. C., et al. (2016). Lignin composition and structure differs between xylem, phloem and pith in *Quercus suber*. *Front. Plant Sci.* 7:1612. doi: 10.3389/fpls.2016.01612
- Lundqvist, J., Teleman, A., Junell, L., Zacchi, G., Dahlman, O., Tjerneld, F., et al. (2002). Isolation and characterization of galactoglucomannan from spruce (*Picea abies*). *Carbohydr. Polym.* 48, 29–39. doi: 10.1021/jf303741h
- Manabe, Y., Nafisi, M., Verhertbruggen, Y., Orfila, C., Gille, S., Rautengarten, C., et al. (2011). Loss-of-function mutation of *REDUCED WALL ACETYLATION2* in *Arabidopsis* leads to reduced cell wall acetylation and increased resistance to *Botrytis cinerea*. *Plant Physiol.* 155, 1068–1078. doi: 10.1104/pp.110.168989
- Manabe, Y., Verhertbruggen, Y., Gille, S., Harholt, J., Chong, S. L., Pawar, P. M., et al. (2013). Reduced wall acetylation proteins play vital and distinct roles in cell wall O-acetylation in *Arabidopsis*. *Plant Physiol.* 163, 1107–1117. doi: 10.1104/pp.113.225193
- Matasci, N., Hung, L. H., Yan, Z., Carpenter, E. J., Wickett, N. J., Mirabab, S., et al. (2014). Data access for the 1,000 Plants (1KP) project. *Gigascience* 3, 10–17. doi: 10.1186/2047-217X-3-17
- Mortimer, J. C., Miles, G. P., Brown, D. M., Zhang, Z., Segura, M. P., Weimar, T., et al. (2010). Absence of branches from xylan in *Arabidopsis gux* mutants reveals potential for simplification of lignocellulosic biomass. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17409–17414. doi: 10.1073/pnas.1005456107
- Moynihn, P. J., and Clarke, A. J. (2010). O-acetylation of peptidoglycan in Gram-negative bacteria: identification and characterization of peptidoglycan O-acetyltransferase in *Neisseria gonorrhoeae*. *J. Biol. Chem.* 285, 13264–13273. doi: 10.1074/jbc.M110.107086
- Moynihn, P. J., and Clarke, A. J. (2011). O-Acetylated peptidoglycan: controlling the activity of bacterial autolysins and lytic enzymes of innate immune systems. *Int. J. Biochem. Cell Biol.* 43, 1655–1659. doi: 10.1016/j.biocel.2011.08.007

- Moynihan, P. J., Sychantha, D., and Clarke, A. J. (2014). Chemical biology of peptidoglycan acetylation and deacetylation. *Bioorg. Chem.* 54, 44–50. doi: 10.1016/j.bioorg.2014.03.010
- Nafisi, M., Stranne, M., Fimognari, L., Atwell, S., Martens, H. J., and Pedas, P. R. (2015). Acetylation of cell wall is required for structural integrity of the leaf surface and exerts a global impact on plant stress responses. *Front. Plant Sci.* 6:550. doi: 10.3389/fpls.2015.00550
- Niklas, K. J. (2004). The cell walls that bind the tree of life. *Bioscience* 54, 831–841. doi: 10.1641/0006-3568(2004)054[0831:TCWTBT]2.0.CO;2
- Obel, N., Erben, V., Schwarz, T., Kuhnelt, S., Fodor, A., and Pauly, M. (2009). Microanalysis of plant cell wall polysaccharides. *Mol. Plant* 2, 922–932. doi: 10.1093/mp/ssp046
- Oliver, D. J., Nikolau, B. J., and Wurtele, E. S. (2009). Acetyl CoA - life at the metabolic nexus. *Plant Sci.* 176, 597–601. doi: 10.1016/j.plantsci.2009.02.005
- O'Neill, M. A., Ishii, T., Albersheim, P., and Darvill, A. G. (2004). Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annu. Rev. Plant Biol.* 55, 109–139. doi: 10.1146/annurev.arplant.55.031903.141750
- Orfila, C., Dal Degan, F., Jorgensen, B., Scheller, H. V., Ray, P. M., and Ulvskov, P. (2012). Expression of mung bean pectin acetyl esterase in potato tubers: effect on acetylation of cell wall polymers and tuber mechanical properties. *Planta* 236, 185–196. doi: 10.1007/s00425-012-1596-z
- Pacios Bras, C., Jordai, M. A., Wijffes, A. H. M., Harteveld, M., Stuurman, N., Thomas-Oates, J. E., et al. (2000). A *Lotus japonicus* nodulation system based on heterologous expression of the fucosyl transferase NodZ and the acetyl transferase Noll in *Rhizobium leguminosarum*. *Mol. Plant Microbe Interact.* 13, 475–479. doi: 10.1094/MPMI.2000.13.4.475
- Painter, T. J. (1983). "Algal polysaccharides," in *The Polysaccharides*, Vol. 2, ed. G. O. Aspinall (New York, NY: Academic Press), 195–285.
- Pauly, M. (1999). *Development of Analytical Tools to Study Plant Cell Wall Xyloglucan*. Ph.D. thesis, Aachen, Shaker Verlag.
- Pauly, M., Gille, S., Liu, L. F., Mansoori, N., de Souza, A., Schultink, A., et al. (2013). Hemicellulose biosynthesis. *Planta* 238, 627–642. doi: 10.1007/s00425-013-1921-1
- Pauly, M., and Keegstra, K. (2010). Plant cell wall polymers as precursors for biofuels. *Curr. Opin. Plant Biol.* 13, 304–311. doi: 10.1016/j.pbi.2009.12.009
- Pauly, M., and Scheller, H. (2000). O-acetylation of plant cell wall polysaccharides: identification and partial characterization of a rhamnogalacturonan O-acetyltransferase from potato suspension-cultured cells. *Planta* 210, 659–667. doi: 10.1007/s004250050057
- Pawar, P. M., Koutaniemi, S., Tenkanen, M., and Mellerowicz, E. J. (2013). Acetylation of woody lignocellulose: significance and regulation. *Front. Plant Sci.* 4:118. doi: 10.3389/fpls.2013.00118
- Pawar, P. M., Ratke, C., Balasubramanian, V. K., Chong, S. L., Gandla, M. L., Adriasola, M., et al. (2017). Downregulation of RWA genes in hybrid aspen affects xylan acetylation and wood saccharification. *New Phytol.* 214, 1491–1505. doi: 10.1111/nph.14489
- Pogorelko, G., Lionetti, V., Bellincampi, D., and Zabolina, O. (2013a). Cell wall integrity: targeted post-synthetic modifications to reveal its role in plant growth and defense against pathogens. *Plant Signal. Behav.* 8:e25435. doi: 10.4161/psb.25435
- Pogorelko, G., Lionetti, V., Fursova, O., Sundaram, R. M., Qi, M., Whitham, S. A., et al. (2013b). *Arabidopsis* and *Brachypodium distachyon* transgenic plants expressing *Aspergillus nidulans* acetylsterases have decreased degree of polysaccharide acetylation and increased resistance to pathogens. *Plant Physiol.* 162, 9–23. doi: 10.1104/pp.113.214460
- Popper, Z. A., Gurvan, M., Hervé, C., Domozochy, 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.* 62, 567–590. doi: 10.1146/annurev-arplant-042110-103809
- Popper, Z. A., and Tuohy, M. G. (2010). Beyond the green: understanding the evolutionary puzzle of plant and algal cell walls. *Plant Physiol.* 153, 373–383. doi: 10.1104/pp.110.158055
- Ralph, J. (1996). An unusual lignin from kenaf. *J. Nat. Prod.* 59, 341–342. doi: 10.1021/np960143s
- Ramírez, V., Xiong, G., Mashiguchi, K., Yamaguchi, S., and Pauly, M. (2018). Growth- and stress-related defects associated with wall hypoacetylation are strigolactone-dependent. *Plant Direct* 2:e00062. doi: 10.1002/pld3.62
- Randoux, B., Renard-Merlier, D., Mulard, G., Rossard, S., Duyme, F., Sanssené, J., et al. (2010). Distinct defenses induced in wheat against powdery mildew by acetylated and nonacetylated oligogalacturonides. *Phytopathology* 100, 1352–1363. doi: 10.1094/PHYTO-03-10-0086
- Rennie, E. A., Hansen, S. F., Baidoo, E. E., Hadi, M. Z., Keasling, J. D., and Scheller, H. V. (2012). Three members of the Arabidopsis glycosyltransferase family 8 are xylan glucuronosyltransferases. *Plant Physiol.* 159, 1408–1417. doi: 10.1104/pp.112.200964
- Rombouts, M., and Thibault, J. F. (1986). Enzymatic and chemical degradation and the fine-structure of pectins from sugar-beet pulp. *Carbohydr. Res.* 154, 189–203. doi: 10.1016/S0008-6215(00)90032-6
- Schultink, A., Cheng, K., Park, Y. B., Cosgrove, D. J., and Pauly, M. (2013). The identification of two arabinosyltransferases from tomato reveals functional equivalency of xyloglucan side chain substituents. *Plant Physiol.* 163, 86–94. doi: 10.1104/pp.113.221788
- Schultink, A., Naylor, D., Dama, M., and Pauly, M. (2015). The role of the plant-specific ALTERED XYLOGLUCAN9 protein in Arabidopsis cell wall polysaccharide O-acetylation. *Plant Physiol.* 167, 1271–1283. doi: 10.1104/pp.114.256479
- Selig, M., Adney, W., Himmel, M., and Decker, S. R. (2009). The impact of cell wall acetylation on corn stover hydrolysis by cellulolytic and xylanolytic enzymes. *Cellulose* 16, 711–722. doi: 10.1007/s10570-009-9322-0
- Sengkhampan, N., Bakx, E. J., Verhoef, R., Schols, H. A., Sajjaanantakul, T., and Voragen, A. G. J. (2009). Okra pectin contains an unusual substitution of its rhamnosyl residues with acetyl and alpha-linked galactosyl groups. *Carbohydr. Res.* 344, 1842–1851. doi: 10.1016/j.carres.2008.11.022
- Sinclair, S. A., Larue, C., Bonk, L., Khan, A., Castillo-Michel, H., Stein, R. J., et al. (2017). Etiolated seedling development requires repression of photomorphogenesis by a small cell-wall-derived dark signal. *Curr. Biol.* 27, 3403–3418. doi: 10.1016/j.cub.2017.09.063
- Sychantha, D., and Clarke, A. J. (2018). Peptidoglycan modification by the catalytic domain of *Streptococcus pneumoniae* OatA follows a ping-pong Bi-Bi mechanism of action. *Biochemistry* 24, 2394–2401. doi: 10.1021/acs.biochem.8b00301
- Sychantha, D., Jones, C. S., Little, D. J., Moynihan, P. J., Robinson, H., Galley, N. F., et al. (2017). In vitro characterization of the antivirulence target of Gram-positive pathogens, peptidoglycan O-acetyltransferase A (OatA). *PLoS Pathog.* 13:e1006667. doi: 10.1371/journal.ppat.1006667
- Sychantha, D., Little, D. J., Chapman, R. N., Boons, G. J., Robinson, H., Howell, P. L., et al. (2018). PatB1 is an O-acetyltransferase that decorates secondary cell wall polysaccharides. *Nat. Chem. Biol.* 14, 79–85. doi: 10.1038/nchembio.2509
- Teleman, A., Lundqvist, J., Tjerneld, F., Stålbrand, H., and Dahlman, O. (2000). Characterization of acetylated 4-O-methylglucuronoxylan isolated from aspen employing ¹H and ¹³C NMR spectroscopy. *Carbohydr. Res.* 329, 807–815. doi: 10.1016/S0008-6215(00)00249-4
- Urbanowicz, B. R., Pena, M. J., Moniz, H. A., Moremen, K. W., and York, W. S. (2014). Two Arabidopsis proteins synthesize acetylated xylan *in vitro*. *Plant J.* 80, 197–206. doi: 10.1111/tpj.12643
- Vogel, J., Raab, T., Somerville, C., and Somerville, S. (2004). Mutations in PMR5 result in powdery mildew resistance and altered cell wall composition. *Plant J.* 40, 968–978. doi: 10.1111/j.1365-313X.2004.02264.x
- Waters, M. T., Gutjahr, C., Bennett, T., and Nelson, D. C. (2017). Strigolactone signaling and evolution. *Annu. Rev. Plant Biol.* 68, 291–322. doi: 10.1146/annurev-arplant-042916-040925
- Weadge, J. T., Pfeffer, J. M., and Clarke, A. J. (2005). Identification of a new family of enzymes with potential O-acetylpeptidoglycan esterase activity in both Gram-positive and Gram-negative bacteria. *BMC Microbiol.* 5:49. doi: 10.1186/1471-2180-5-49
- Xin, Z., and Browse, J. (1998). *Eskimo1* mutants of *Arabidopsis* are constitutively freezing-tolerant. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7799–7804. doi: 10.1073/pnas.95.13.7799
- Xin, Z., Mandaokar, A., Chen, J., Last, R. L., and Browse, J. (2007). Arabidopsis *ESK1* encodes a novel regulator of freezing tolerance. *Plant J.* 49, 786–799. doi: 10.1111/j.1365-313X.2006.02994.x
- Xiong, G., Cheng, K., and Pauly, M. (2013). Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the Arabidopsis mutant tbl29. *Mol. Plant* 6, 1373–1375. doi: 10.1093/mp/sst014

- Xiong, G. Y., Dama, M., and Pauly, M. (2015). Glucuronic acid moieties on xylan are functionally equivalent to O-acetyl-substituents. *Mol. Plant* 8, 1119–1121. doi: 10.1016/j.molp.2015.02.013
- Yuan, Y., Teng, Q., Zhong, R., Haghighat, M., Richardson, E. A., and Ye, Z. H. (2016a). Mutations of Arabidopsis *TBL32* and *TBL33* affect xylan acetylation and secondary wall deposition. *PLoS One* 11:e0146460. doi: 10.1371/journal.pone.0146460
- Yuan, Y., Teng, Q., Zhong, R., and Ye, Z. H. (2016b). Roles of Arabidopsis *TBL34* and *TBL35* in xylan acetylation and plant growth. *Plant Sci.* 243, 120–130. doi: 10.1016/j.plantsci.2015.12.007
- Yuan, Y., Teng, Q., Zhong, R., and Ye, Z. H. (2016c). *TBL3* and *TBL31*, two Arabidopsis DUF231 domain proteins, are required for 3-O- monoacetylation of xylan. *Plant Cell Physiol.* 57, 35–45. doi: 10.1093/pcp/pcv172
- Yuan, Y., Teng, Q., Zhong, R., and Ye, Z. H. (2013). The Arabidopsis DUF231 domain-containing protein *ESK1* mediates 2-O- and 3-O-acetylation of xylosyl residues in xylan. *Plant Cell Physiol.* 54, 1186–1199. doi: 10.1093/pcp/pct070
- Zhang, B., Zhang, L., Li, F., Zhang, D., Liu, X., Wang, H., et al. (2017). Control of secondary cell wall patterning involves xylan deacetylation by a GDSL esterase. *Nat. Plants* 3:17017. doi: 10.1038/nplants.2017.17
- Zhong, R., Cui, D., Dasher, R. L., and Ye, Z. H. (2018a). Biochemical characterization of rice xylan O-acetyltransferases. *Planta* 247, 1489–1498. doi: 10.1007/s00425-018-2882-1
- Zhong, R., Cui, D., and Ye, Z. H. (2018b). A group of *Populus trichocarpa* DUF231 proteins exhibit differential O-acetyltransferase activities toward xylan. *PLoS One* 13:e0194532. doi: 10.1371/journal.pone.0194532
- Zhong, R., Cui, D., and Ye, Z.-H. (2017). Regiospecific acetylation of xylan is mediated by a group of DUF231-containing O-acetyltransferases. *Plant Cell Physiol.* 58, 2126–2138. doi: 10.1093/pcp/pcx147
- Zhu, X. F., Sun, Y., Zhang, B. C., Mansoori, N., Wan, J. X., Liu, Y., et al. (2014). *TRICHOME BIREFRINGENCE-LIKE27* affects aluminum sensitivity by modulating the O-acetylation of xyloglucan and aluminum-binding capacity in Arabidopsis. *Plant Physiol.* 166, 181–189. doi: 10.1104/pp.114.243808

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pauly and Ramírez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Functional Analysis of Cellulose Synthase *CesA4* and *CesA6* Genes in Switchgrass (*Panicum virgatum*) by Overexpression and RNAi-Mediated Gene Silencing

Mitra Mazarej^{1,2}, Holly L. Baxter^{1,2}, Mi Li^{2,3}, Ajaya K. Biswal^{2,4}, Keonhee Kim⁵, Xianzhi Meng^{2,6}, Yunqiao Pu^{2,3}, Wegi A. Wuddineh^{1,2}, Ji-Yi Zhang^{2,7}, Geoffrey B. Turner^{2,8}, Robert W. Sykes^{2,8}, Mark F. Davis^{2,8}, Michael K. Udvardi^{2,7}, Zeng-Yu Wang^{2,7}, Debra Mohnen^{2,4}, Arthur J. Ragauskas^{2,3,6}, Nicole Labbé⁵ and C. Neal Stewart Jr.^{1,2*}

OPEN ACCESS

Edited by:

Charles T. Anderson,
The Pennsylvania State University,
United States

Reviewed by:

Jenny C. Mortimer,
Lawrence Berkeley National
Laboratory (LBNL), United States
Markus Pauly,
Heinrich-Heine-Universität Düsseldorf,
Germany

*Correspondence:

C. Neal Stewart Jr.
nealstewart@utk.edu

Specialty section:

This article was submitted to
Plant Physiology,
a section of the journal
Frontiers in Plant Science

Received: 30 April 2018

Accepted: 10 July 2018

Published: 03 August 2018

Citation:

Mazarej M, Baxter HL, Li M, Biswal AK, Kim K, Meng X, Pu Y, Wuddineh WA, Zhang J-Y, Turner GB, Sykes RW, Davis MF, Udvardi MK, Wang Z-Y, Mohnen D, Ragauskas AJ, Labbé N and Stewart CN Jr. (2018) Functional Analysis of Cellulose Synthase *CesA4* and *CesA6* Genes in Switchgrass (*Panicum virgatum*) by Overexpression and RNAi-Mediated Gene Silencing. *Front. Plant Sci.* 9:1114. doi: 10.3389/fpls.2018.01114

¹ Department of Plant Sciences, University of Tennessee, Knoxville, TN, United States, ² BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ³ Biosciences Division, Joint Institute for Biological Science, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ⁴ Complex Carbohydrate Research Center, University of Georgia, Athens, GA, United States, ⁵ Center for Renewable Carbon, University of Tennessee, Knoxville, TN, United States, ⁶ Department of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN, United States, ⁷ Noble Research Institute, Ardmore, OK, United States, ⁸ National Renewable Energy Laboratory, Golden, CO, United States

Switchgrass (*Panicum virgatum* L.) is a leading lignocellulosic bioenergy feedstock. Cellulose is a major component of the plant cell walls and the primary substrate for saccharification. Accessibility of cellulose to enzymatic breakdown into fermentable sugars is limited by the presence of lignin in the plant cell wall. In this study, putatively novel switchgrass secondary cell wall cellulose synthase *PvCesA4* and primary cell wall *PvCesA6* genes were identified and their functional role in cellulose synthesis and cell wall composition was examined by overexpression and knockdown of the individual genes in switchgrass. The endogenous expression of *PvCesA4* and *PvCesA6* genes varied among including roots, leaves, stem, and reproductive tissues. Increasing or decreasing *PvCesA4* and *PvCesA6* expression to extreme levels in the transgenic lines resulted in decreased biomass production. *PvCesA6*-overexpressing lines had reduced lignin content and syringyl/guaiacyl lignin monomer ratio accompanied by increased sugar release efficiency, suggesting an impact of *PvCesA6* expression levels on lignin biosynthesis. Cellulose content and cellulose crystallinity were decreased, while xylan content was increased in *PvCesA4* and *PvCesA6* overexpression or knockdown lines. The increase in xylan content suggests that the amount of non-cellulosic cell wall polysaccharide was modified in these plants. Taken together, the results show that the manipulation of the cellulose synthase genes alters the cell wall composition and availability of cellulose as a bioprocessing substrate.

Keywords: cellulose synthase, switchgrass, overexpression, RNAi-gene silencing, *PvCesA4*, *PvCesA6*, lignocellulosic, biofuel

INTRODUCTION

Plant cell walls consist largely of polysaccharides (cellulose, hemicellulose, pectin) and the polyphenolic compound lignin (Somerville et al., 2004). Cellulose is the most abundant constituent of primary and secondary cell walls in plants and plays a central role in plant mechanical strength and morphogenesis (Cosgrove, 2005; Liu et al., 2016). Cellulose is made up of chains containing repeated glucose residues, which together form strong microfibril structures (Somerville, 2006). In higher plants, cellulose is synthesized by a large cellulose synthase (CesA) complex located on the plasma membrane (Schneider et al., 2016). Since the first plant CesA gene was identified from cotton (Pear et al., 1996), the CesA superfamily has been characterized in many plant species, including *Arabidopsis* (Taylor et al., 2003; Desprez et al., 2007; Persson et al., 2007), rice (Tanaka et al., 2003; Wang et al., 2010), maize (Appenzeller et al., 2004), cotton (Li A. et al., 2013), barley (Burton et al., 2004), and poplar (Joshi et al., 2004; Djerbi et al., 2005). While cellulose biosynthesis is not fully understood, work with *Arabidopsis* mutants has elucidated the key genes encoding the catalytic subunits of CesA, with some involved in making primary cell walls (AtCesA1, AtCesA3, AtCesA6) and others in making secondary cell walls (AtCesA4, AtCesA7, AtCesA8) (Endler and Persson, 2011). At least three CesAs each are expressed in cells during either primary or secondary wall formation, and mutations in any one of them disrupt cellulose synthesis, indicating the non-redundant function of members of the different subclass members (Somerville, 2006).

Switchgrass (*Panicum virgatum* L.) is a promising lignocellulosic bioenergy feedstock owing to its wide adaptation, high genetic variability, and its ability to reliably produce easily-harvested aboveground biomass each year. The resistance of plant cell walls to deconstruction, defined as biomass recalcitrance, hinders the accessibility of cellulose to enzymatic breakdown into fermentable sugars for biofuel production (Himmel and Bayer, 2009). Biomass recalcitrance is mainly determined by the cell wall composition and its complex structure. Lignin is the primary contributor to biomass recalcitrance (Chen and Dixon, 2007). Genetic engineering of plant cell walls has been shown to reduce biomass recalcitrance (Nelson et al., 2017; Biswal et al., 2018). Since cellulose is a major structural component of the cell wall and the primary substrate for saccharification, manipulating the genes involved in cellulose synthesis could alter the cell wall composition and availability of cellulose as a bioprocessing substrate (Mizrachi et al., 2012; Kalluri et al., 2014; Bali et al., 2016). The practical interest of such an investigation for switchgrass is supported by studies demonstrating that a functional relationship exists between CesA structures, cellulose crystallinity and saccharification efficiency in *Arabidopsis* and rice (Harris et al., 2012; Li F. et al., 2017).

In the present study, novel switchgrass cellulose synthase *PvCesA4* and *PvCesA6* genes were identified and their functional role in cellulose synthesis was examined by overexpression and knockdown of the individual genes in switchgrass. These transgenic plants were analyzed for (i) *PvCesA* expression, (ii) growth morphology and biomass yield, (iii) cell wall

composition and properties, and (iv) sugar release efficiency. To our knowledge, this is the first functional description of switchgrass CesA genes.

MATERIALS AND METHODS

Gene Identification

Using the CesA amino acid sequences of *Arabidopsis* (*Arabidopsis thaliana*), rice (*Oryza sativa*), and maize (*Zea mays*) as heterologous probes, TBLASTN was used to identify the homologous gene sequences from switchgrass EST databases (Zhang J.Y. et al., 2013) and the draft genome (*Panicum virgatum* v1.1 DOE-JGI) at Phytozome. A phylogenetic tree of CesAs protein family members of *A. thaliana* TAIR10, *O. sativa* v7.0, *Populus trichocarpa* v3.0, *Setaria italica* v2.2, and *P. virgatum* v1.1 from Phytozome 12.0¹ was constructed by the neighbor-joining method using MEGA6 (Tamura et al., 2013). The CesA phylogenetic tree was divided into 6-clades [Clade A (CESA1), Clade B (CESA3), Clade C (CESA6), Clade D (CESA7), Clade E (CESA8), and Clade F (CESA4)] based on Kumar et al. (2009) and Kumar and Turner (2015). Pavir.Ib00804 from Clade F and Pavir.Ba01088 from Clade C were named *PvCesA4* and *PvCesA6*, respectively.

Vector Construction and Transgenic Plant Production

Overexpression cassettes were constructed by isolating the gene open reading frame (ORF) from switchgrass cDNAs of the ST1 clonal genotype of “Alamo” switchgrass using individual gene-specific primers flanking the ORF of each gene (Supplementary Figures S1, S2 and Supplementary Table S1) and subsequently cloning each into pCR8 entry vector for sequence confirmation. For RNAi cassettes, target sequences of 223 bp (*PvCesA4*) and 331 bp (*PvCesA6*) were used (Supplementary Figures S1, S2) and cloned into pCR8 vector for sequence confirmation. Sequence-confirmed fragments were then sub-cloned into pANIC-10A overexpression vector or into pANIC-8A RNAi-vector (Supplementary Figure S3) by GATEWAY recombination (Mann et al., 2012) to place each target gene under the control of the maize ubiquitin 1 (*ZmUbi1*) promoter. Embryogenic callus derived from “Alamo” switchgrass NFCX01 genotype was transformed with the expression vector construct using *Agrobacterium*-mediated transformation (Xi et al., 2009).

Plants and Growth Conditions

Transgenic and non-transgenic control plants were grown under the same conditions (16 h day/8 h night light at 26°C) in growth chambers. For growth analysis, each transgenic and non-transgenic line was propagated from a single tiller to yield three clonal replicates each (Hardin et al., 2013). The growth parameters were measured at R1 growth stage (Moore et al., 1991). Plant height was determined by measuring the five tallest tillers from each replicate. Stem diameters were measured for

¹<https://phytozome.jgi.doe.gov/>

each of these tillers with a digital caliper. For plant width, the diameter of the plant crown mid-section was measured. Tiller numbers were tallied for each plant. The fresh biomass was measured from the aboveground plant biomass cut at a similar stage of growth while the dry biomass was measured from fresh biomass dried at 42°C for 96 h.

RNA Extraction and qRT-PCR

Total RNA was extracted from root, leaf sheath, leaf blade, stem, and panicle samples at the R1 growth stage or from the shoot tips of transgenic lines at the E4 growth stage using Tri-Reagent (Sigma-Aldrich, St. Louis, MO, United States) following the manufacturer's instructions. One microgram of the purified RNA was treated with DNase-1 (Qiagen, Valencia, CA, United States) to remove any potential genomic DNA contaminants. The DNase-treated RNA was used for first-strand cDNA synthesis using High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA, United States). qRT-PCR experiments were performed with Power SYBR Green PCR Master Mix (Applied Biosystems) in an optical 96-well plate using a Quant Studio 6 Flex system (Applied Biosystems). Analysis of the relative expression was carried out by the change in Ct method. The standard curve method was used for relative transcript quantification normalized by switchgrass ubiquitin 1 (*PvUbi1*) as a reference gene (Shen et al., 2009). Primers used for transcript analysis are listed in **Supplementary Table S1**.

Cell Wall Characterization

Tillers were collected at the R1 growth stage, dried at 42°C for 96 h, and ground to 0.5 mm (40 mesh) particle size. Cell wall chemical composition was determined following the National Renewable Energy Laboratory (NREL) protocols. Briefly, approximately 3 g samples were sequentially extracted with water and ethanol using an automated extraction system (ASE 350, Dionex Corp., Sunnyvale, CA, United States) following the NREL protocol "Determination of extractives in biomass (NREL/TP 510-42619)." The extracted samples were then dried at 40°C for 3 days until constant weight. Cellulose, hemicellulose, lignin, acetyl content, and structural ash were then measured after a two-step acid hydrolysis following the NREL protocol "Determination of structural carbohydrates and lignin in biomass (NREL/TP 510-42618)." High pressure liquid chromatography (HPLC) was employed to quantify the structural monomeric sugars after the two-step of acid hydrolysis. Acid insoluble lignin was measured gravimetrically and acid soluble lignin was measured using a Genesys 10S UV-Vis Spectrophotometer (Thermo Scientific, Dubuque, IA, United States). The HPLC system for carbohydrates measurement was equipped with an Aminex HPX-87P column (300 nm × 7.8 nm, 9 μm particle sizes) (Bio-Rad, Hercules, CA, United States) attached to a micro-guard Carbo-P guard column (Bio-Rad), and a refractive index detector (Perkin Elmer, Waltham, MA, United States). The HPLC's RI detector temperature was 50°C and the oven temperature was set at 85°C. The injection volume was 30 μl with 0.25 ml/min of flow rate. Mannose peak was not detected in this study, and the total hemicellulose

content was the sum of xylose, galactose, and arabinose. The acetyl content was also measured utilizing HPLC equipped with Aminex HPX-87H column (300 nm × 7.8 nm, 9 μm particle size) (Bio-Rad). The RI detector temperature was 50°C and the oven temperature was set at 45°C. The mobile phase was 0.05 M sulfuric acid. The injection volume was 50 μl with 0.6 ml/min of flow rate. The total (unextracted biomass) and structural ash (extractives-free biomass) content was gravimetrically determined by combusting 0.5 g of biomass in a furnace (Fisher Scientific Isotemp Programmable Muffle Furnace 750, Dubuque, IA, United States) at 575°C for 24 h following the NREL protocol "Determination of ash in biomass (NREL/TP 510-42622)." Lignin composition was determined by pyrolysis molecular beam mass spectrometry (py-MBMS) using the NREL high-throughput method on extractive- and starch-free samples (Sykes et al., 2009; Decker et al., 2012). Sugar release by enzymatic hydrolysis was determined by NREL high-throughput method on extractive- and starch-free samples described by Selig et al. (2010). Briefly, cell wall residues were prepared by removing soluble extractives and starch. Samples were loaded into a 96-well plate. A hot water pretreatment was conducted in a steam chamber at 180°C for 17.5 min. After pretreatment, enzymatic hydrolysis was performed in the well plate on the pretreated slurry by incubation with Ctec2 enzyme cocktail (70 mg protein/g biomass) at 40°C for 72 h. Glucose and xylose release were determined by colorimetric assays, and total sugar release is the sum of glucose + xylose released (Studer et al., 2009).

Cellulose Characterization

Cellulose properties were determined using tillers collected at the R1 growth stage and dried at 42°C for 96 h, followed by milling to 1 mm (20 mesh) particle size. Cellulose isolation and the measurements of cellulose crystallinity and its degree of polymerization were conducted as previously described (Li M. et al., 2017). Briefly, the extractives of switchgrass were removed by extraction using ethanol:toluene mixture (1:2, v:v) for 24 h. The extractives-free switchgrass samples were delignified by peracetic acid (32% solution in acetic acid) and air-dried to obtain switchgrass holocellulose. One portion of the holocellulose was subjected to hydrochloric acid in boiling water bath to remove hemicellulose. The residue cellulose was washed with deionized water, filtered, and used to measure cellulose crystallinity using cross polarization magic angle spinning (CP/MAS) on a Bruker Avance-400 spectrometer. The cellulose crystallinity index (CrI) was determined from the areas of the crystalline and amorphous C₄ signals of cellulose. Another portion of the holocellulose was extracted with sodium hydroxide. The obtained cellulose residue, namely α-cellulose, was used to measure the weight-average molecular weight (M_w) and number-average molecular weight (M_n) of cellulose using gel permeation chromatographic analysis after tricarbanilation. The weight-average (DP_w) and number-average (DP_n) degree of polymerization of cellulose were calculated by dividing the M_w and M_n by 519 g/mol, the molecular mass of repeating unit of derivatized cellulose. Both the cellulose crystallinity and its degree of polymerization were reported as the average value of three biological replicates.

Statistics

Statistical analyses were performed in SAS version 9.4 (SAS Institute Inc., Cary, NC, United States). One-way ANOVA with Fisher's least significant difference method was used to compare means among transgenic lines and the control. Differences were considered significant when *P*-values were less than or equal to 0.05. For pairwise comparisons, each transgenic line was compared with the control using the PROC TTEST procedure in SAS. Differences were considered significant when *P*-values were less than or equal to 0.05.

RESULTS

Identification of PvCesA Homologs

The CesA orthologous amino acid sequences from Arabidopsis, rice, and maize were used to identify the switchgrass *PvCesA* sequences and CesA protein family members of *A. thaliana*, *P. trichocarpa*, *O. sativa*, *S. italica*, and *P. virgatum* from Phytozome were used to elucidate amino acid sequences profiles (Figure 1). Based on this, Pavir.Ib00804 from Clade F and Pavir.Ba01088 from Clade C were identified and named *PvCesA4* and *PvCesA6* for *P. virgatum*, respectively (Figure 1). At the time that this work was started, Pavir.Ib00804 and Pavir.Ba01088 were indicated as the genes for switchgrass *CesA4* and *CesA6*, respectively. After completion and release of the switchgrass genome, four other related *CesAs* were located in the Clade C (*CESA6*) (Figure 1).

Expression Patterns of *PvCesA4* and *PvCesA6* in Non-transgenic Switchgrass

Quantitative reverse transcription-polymerase chain reaction (qRT-PCR) results revealed detectable levels of expression for both *PvCesA4* and *PvCesA6* genes in roots, leaf sheaths, leaf blades, stems, and inflorescences at the R1 growth stage (Figures 2A,B). The expression of *PvCesA4* was highest in stem and inflorescence (Figure 2A), whereas the expression of *PvCesA6* was highest in leaves and inflorescence (Figure 2B). Furthermore, the expression levels of the other four related *PvCesA6* (Clade C, Figure 1) were also highest in leaves (Supplementary Figure S4).

Generation of Transgenic Plants Overexpressing *PvCesA4* and *PvCesA6*

Five independent transgenic lines overexpressing either *PvCesA4* or *PvCesA6* driven by the *ZmUbi1* promoter were produced (Figures 3A,C). Genomic PCR using primers specific to the transgene and the hygromycin resistance gene confirmed that the plants were stably transgenic (data not shown). *PvCesA4* was overexpressed between 19- and 41-fold in the transgenic lines compared to non-transgenic control by qRT-PCR analysis (Figure 3B) and *PvCesA6* was overexpressed between 3- and 30-fold (Figure 3D) in the transgenic lines. The level of expression of the

endogenous *PvCesA4* and *PvCesA6* was not affected and was similar to the expression levels in non-transgenic control (Figures 3B,D).

Expression Levels of Other Major Secondary and Primary Wall *CesAs* in Transgenic Plants Overexpressing *PvCesA4* and *PvCesA6*

Transcript abundance of other major secondary wall *PvCesA7* (Clade D) and *PvCesA8* (Clade E), and primary wall *PvCesA1* (Clade A) and *PvCesA3* (Clade B) (Figure 1) was determined in *PvCesA4* and *PvCesA6* overexpressing lines. The expression levels of these genes were generally unaffected as compared to those found in non-transgenic control (Supplementary Figure S5).

Generation of *PvCesA4*-RNAi and *PvCesA6*-RNAi Transgenic Plants

Five independent RNAi-transgenic lines for *PvCesA4* and *PvCesA6* driven by the *ZmUbi1* promoter were produced. Three transgenic lines for each gene that had normal growth rates and reached the R1 growth stage were selected for further characterization (Figures 4A,C). Genomic PCR using primers specific to the transgene and the hygromycin resistance gene confirmed that the plants were stably transgenic (data not shown). qRT-PCR analysis showed that the *PvCesA4* transcript level was decreased by 15–36% (Figure 4B) and the *PvCesA6* transcript level was decreased by 11–64% (Figure 4D) in the RNAi-transgenic lines.

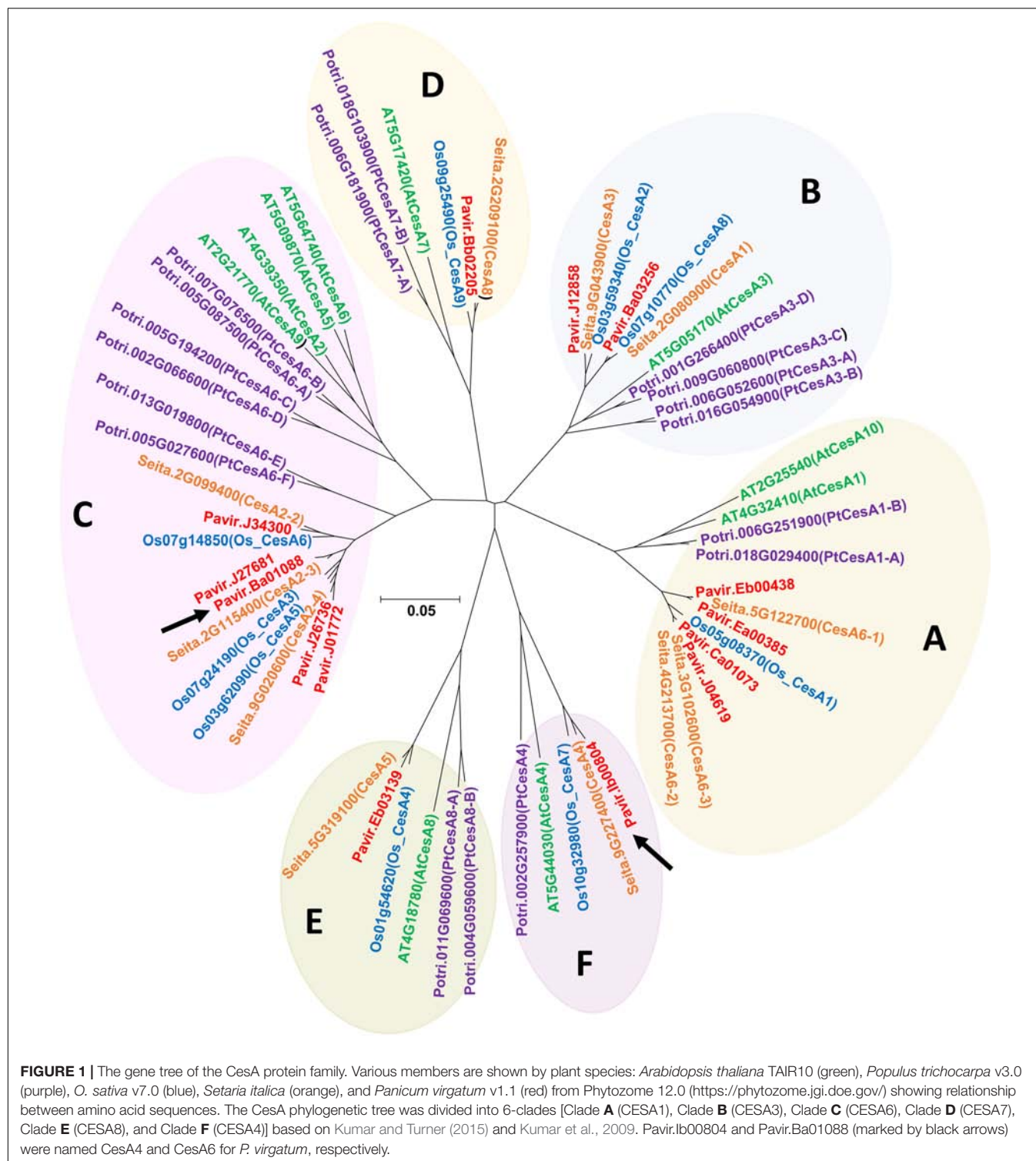
Phenotypic Characterization of Transgenic Plants

PvCesA4-Overexpressing Lines

There was a statistically significant decrease in tiller height for three transgenic lines (7, 8, and 9) and in plant width for three transgenic lines (1, 8, and 9) compared with non-transgenic controls. There were no significant differences in stem diameter between the transgenic lines and the non-transgenic controls. Tiller number was significantly increased for two transgenic lines (7 and 10) and decreased for one transgenic line (9). All transgenic lines had equivalent dry biomass relative to the non-transgenic control lines with the exception of line-8 (decreased biomass by 33%) and line-9 (decreased biomass by 58%) (Table 1A), congruent to the highest transcript level of the transgene (Figure 3B).

PvCesA6-Overexpressing Lines

Two transgenic lines (6 and 13) had consistent significant decrease in tiller height, plant width, and stem diameter compared with non-transgenic controls. These two transgenic lines (6 and 13) also had significantly decreased dry biomass weight to 59% (line-6) and 81% (line-13) relative to non-transgenic controls (Table 1B), corresponding to the highest transcript level of the transgene (Figure 3D). There were more tillers in transgenic line-2 resulting in a significant increase



in dry biomass (19% more) relative to non-transgenic controls (Table 1B).

PvCesA4-RNAi Lines

There were no significant differences in tiller height and stem diameter between the transgenics and non-transgenic

controls. There was a significant decrease in plant width for two transgenic lines (12 and 20) and an increase in tiller number for one transgenic line (15) compared with non-transgenic controls. All transgenic lines had equivalent dry biomass weight relative to the non-transgenic controls (Table 2A).

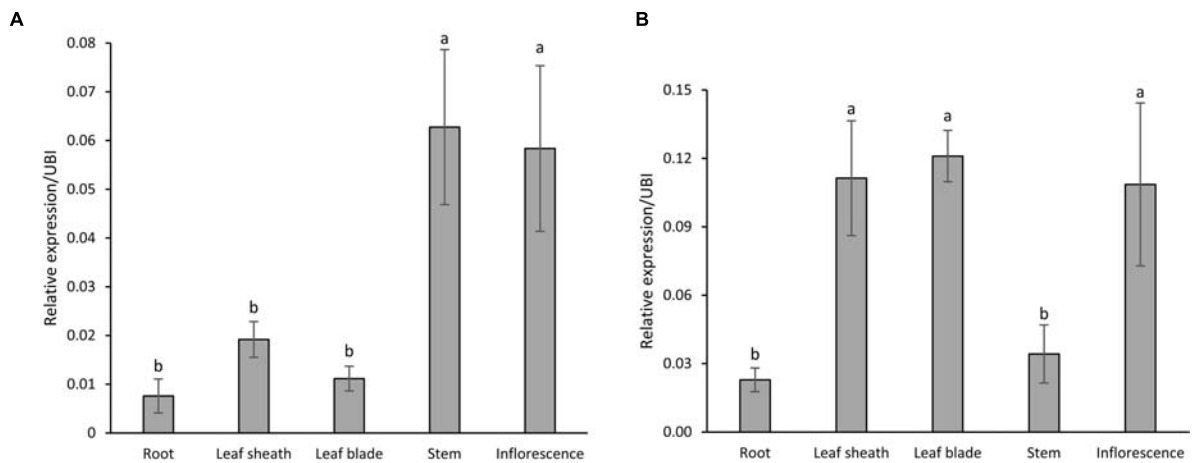


FIGURE 2 | Expression patterns of *PvCesA4* (A) and *PvCesA6* (B) in different plant tissues as determined by qRT-PCR. Plant samples for RNA extraction used in the qRT-PCR experiments were collected at R1 (reproductive stage 1) developmental stage. The relative levels of transcripts were normalized to the switchgrass ubiquitin 1 gene expression (UBI). Bars represent mean values of three biological replicates \pm standard error. Bars represented by different letters are significantly different at $P \leq 0.05$ as tested by LSD method with SAS software (SAS Institute Inc.).

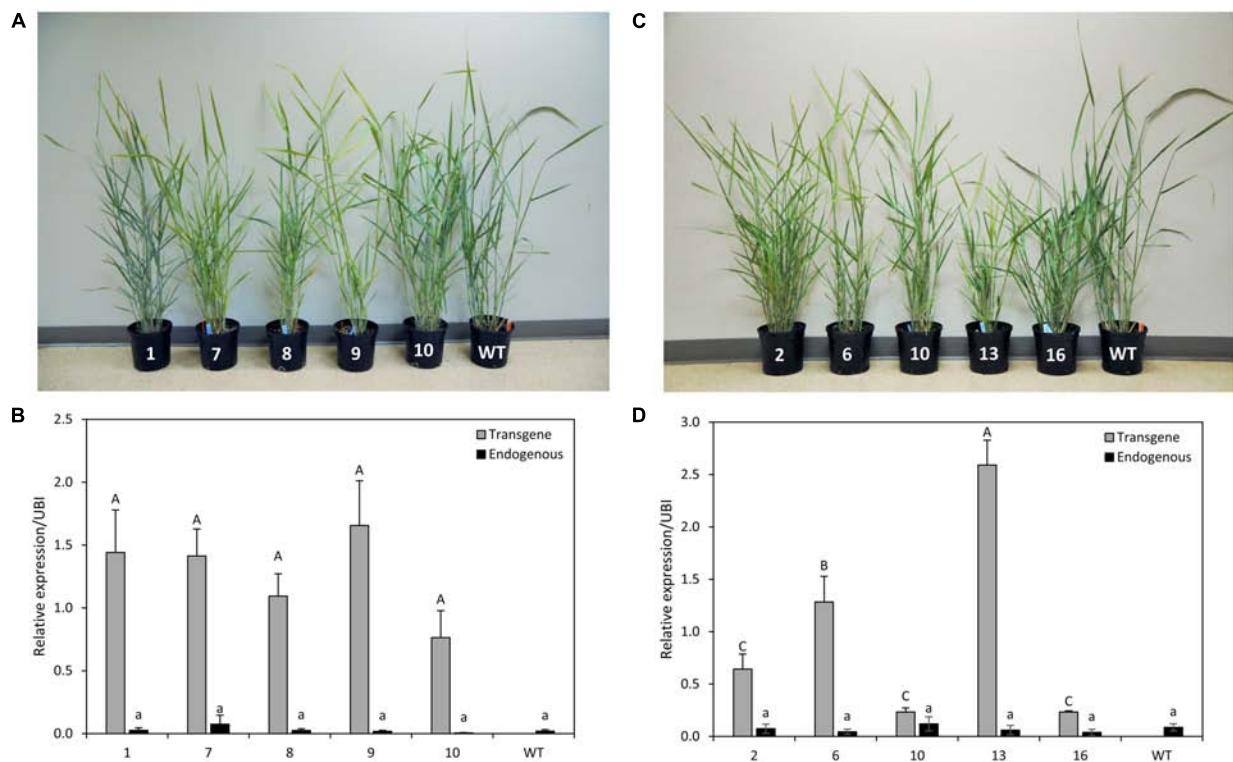


FIGURE 3 | Representative *PvCesA4*-overexpressing (A) and *PvCesA6*-overexpressing (C) transgenic lines. Relative transcript levels of *PvCesA4* (B) and *PvCesA6* (D) in overexpressing transgenic lines as determined by qRT-PCR. WT: non-transgenic control. The relative levels of transcripts were normalized to the switchgrass ubiquitin 1 gene expression (UBI). Bars represent mean values of three biological replicates \pm standard error. Bars represented by different letters are significantly different at $P \leq 0.05$ as tested by LSD method with SAS software (SAS Institute Inc.).

PvCesA6-RNAi Lines

Two transgenic lines (2 and 12) had shorter and fewer tillers (line 12 only) with smaller plant width compared with non-transgenic controls. These same two transgenic lines (2 and 12) also had

significantly less dry biomass (68% for line 2, 77% for line 12) relative to non-transgenic controls (Table 2B) and were congruent with the highest levels of transcript reduction of *PvCesA6* (Figure 4D).

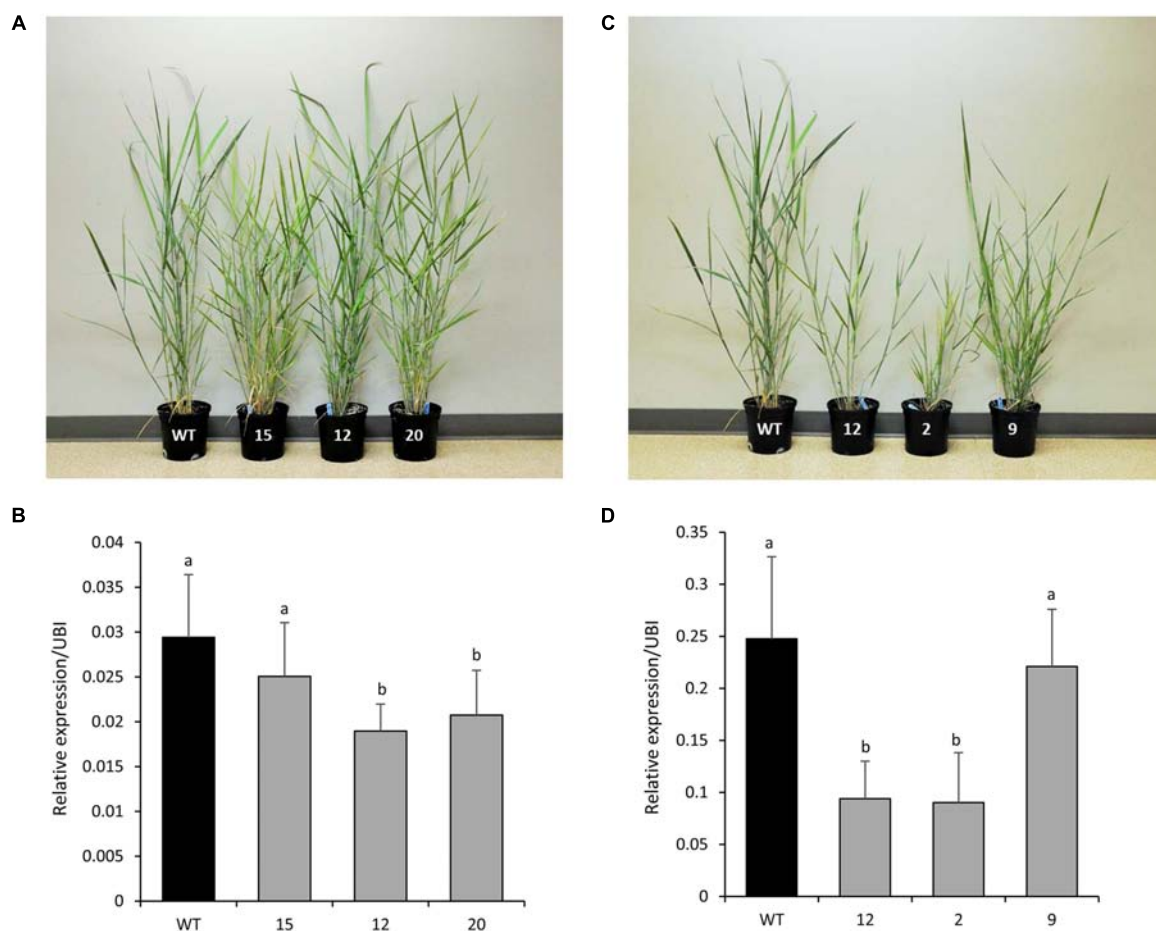


FIGURE 4 | Representative *PvCesA4*-RNAi (**A**) and *PvCesA6*-RNAi (**C**) transgenic lines. Relative transcript levels of *PvCesA4* (**B**) and *PvCesA6* (**D**) in RNAi-transgenic lines as determined by qRT-PCR. WT: non-transgenic control. The relative levels of transcripts were normalized to the switchgrass ubiquitin 1 gene expression (UBI). Bars represent mean values of three biological replicates \pm standard error. Bars represented by different letters are significantly different at $P \leq 0.05$ as tested by LSD method with SAS software (SAS Institute Inc.).

Cell Wall Chemical Composition of Transgenic Plants

PvCesA4-Overexpressing Lines

The lignin content was unchanged in transgenic lines with the exception of a significant decrease (4%) in transgenic line 7 and an increase (8%) in transgenic line 10 compared with the non-transgenic controls. There were no significant differences between the S/G ratios of the transgenic and non-transgenic control lines. Cellulose content was decreased (6–33%) and xylan content was increased (3–12%) in the transgenic lines compared with the non-transgenic controls. Galactan and arabinan contents were decreased in all transgenic lines by 17–25% and 38–48%, respectively (Table 3A). Complete chemical composition data is presented in Supplementary Table S2A.

PvCesA6-Overexpressing Lines

Lignin content was decreased significantly in all transgenic lines by 4–7% compared with the non-transgenic controls. There was a significant decrease in S/G ratio by 14% in all

transgenic lines. Transgenic lines had decreased cellulose content (8–13%) and increased xylan content (2–4%) compared with the non-transgenic controls. There were no significant differences in galactan content between the transgenic and non-transgenic control lines, whereas arabinan content was increased in transgenic line-2 (8%) and line-10 (5%) (Table 3B). Complete chemical composition data is presented in Supplementary Table S2B.

PvCesA4-RNAi Lines

Lignin content was unchanged in the transgenic lines compared with the non-transgenic controls. There were no significant differences between the S/G ratio of the transgenic and non-transgenic control lines. Cellulose content was decreased (up to 9%) and xylan content was increased (up to 4%) in transgenic lines compared with the non-transgenic controls. There was no change in galactan content, whereas arabinan content was decreased up to 30% in transgenic lines compared with non-transgenic controls (Table 4A). Complete chemical composition data is presented in Supplementary Table S3A.

TABLE 1 | Morphology and biomass yields of transgenic lines overexpressing *PvCesA4* or *PvCesA6* and non-transgenic (WT) controls.

A. <i>PvCesA4</i>-overexpression						
Lines	1	7	8	9	10	WT
Height (cm)	128.7 ± 5.9	116.0 ± 2.0*	119.9 ± 1.9*	109.4 ± 1.7*	127.0 ± 1.7	130.7 ± 3.0
Width (cm)	44.9 ± 1.7**	50.0 ± 3.4	41.5 ± 0.8**	33.9 ± 5.2*	49.1 ± 3.4	58.4 ± 1.5
Stem diameter (mm)	5.6 ± 0.1	5.7 ± 0.3	6.2 ± 0.1	6.1 ± 0.7	5.9 ± 0.1	5.9 ± 0.2
Tiller number	18.0 ± 2.1	24.7 ± 2.9*	11.0 ± 1.5	5.0 ± 2.1*	23.0 ± 3.5*	13.7 ± 0.9
Fresh weight (g)	158.5 ± 8.2	186.3 ± 21.7	92.4 ± 9.4*	65.5 ± 26.9*	169.3 ± 28.3	140.9 ± 11.2
Dry weight (g)	49.5 ± 4.3	50.2 ± 7.2	26.4 ± 1.5*	16.5 ± 7.0*	46.1 ± 7.5	39.6 ± 4.2
B. <i>PvCesA6</i>-overexpression						
Lines	2	6	10	13	16	WT
Height (cm)	110.9 ± 0.8	95.4 ± 15.0*	119.7 ± 4.1	76.5 ± 6.2**	103.6 ± 3.7**	130.7 ± 3.0
Width (cm)	59.3 ± 0.8	38.9 ± 4.2*	46.6 ± 5.9	43.2 ± 1.5**	59.3 ± 1.7	58.4 ± 1.5
Stem diameter (mm)	5.3 ± 0.2	5.0 ± 0.4*	6.0 ± 0.1	4.6 ± 0.0**	5.8 ± 0.2	5.9 ± 0.2
Tiller number	32.7 ± 1.3**	14.0 ± 5.1	14.0 ± 0.6	11.3 ± 2.6	18.3 ± 3.8	13.7 ± 0.9
Fresh weight (g)	179.1 ± 0.2*	65.7 ± 28.0*	107.3 ± 0.9	32.2 ± 6.0**	110.7 ± 27.9	140.9 ± 11.2
Dry weight (g)	47.1 ± 1.3*	16.1 ± 7.5*	25.8 ± 1.3	7.6 ± 1.8**	29.2 ± 7.3	39.6 ± 4.2

Values represent the mean of three biological replicates ± standard error. Bold values with asterisks are significantly different from controls at $P \leq 0.05^*$ and $P \leq 0.01^{**}$ as calculated using *t*-tests for pairwise comparison with SAS software (SAS Institute Inc.).

TABLE 2 | Morphology and biomass yields of *PvCesA4*-RNAi or *PvCesA6*-RNAi transgenic lines and non-transgenic (WT) controls.

A. <i>PvCesA4</i>-RNAi				
Lines	12	15	20	WT
Height (cm)	114.0 ± 15.4	117.7 ± 3.8	120.5 ± 2.2	125.9 ± 7.4
Width (cm)	35.6 ± 5.3*	48.3 ± 2.5	33.9 ± 6.1*	53.3 ± 2.6
Stem diameter (mm)	5.7 ± 0.1	5.6 ± 0.1	5.6 ± 0.1	5.8 ± 0.3
Tiller number	15.0 ± 2.3	28.0 ± 1.7**	18.3 ± 2.0	13.3 ± 0.3
Fresh weight (g)	97.8 ± 27.2	155.4 ± 7.5	128.2 ± 25.5	122.1 ± 15.9
Dry weight (g)	21.0 ± 6.4	40.6 ± 2.3	26.2 ± 8.6	33.9 ± 2.2
B. <i>PvCesA6</i>-RNAi				
Lines	2	9	12	WT
Height (cm)	83.3 ± 10.0*	114.3 ± 5.8	82.7 ± 11.4*	125.9 ± 7.4
Width (cm)	35.6 ± 5.3*	45.7 ± 2.5	31.3 ± 4.5*	53.3 ± 2.6
Stem diameter (mm)	5.2 ± 0.0	4.8 ± 0.2	5.2 ± 0.3	5.8 ± 0.3
Tiller number	10.0 ± 1.2	24.7 ± 2.7*	7.7 ± 1.2*	13.3 ± 0.3
Fresh weight (g)	50.7 ± 21.6*	103.5 ± 22.7	35.9 ± 6.4**	122.1 ± 15.9
Dry weight (g)	10.7 ± 4.8*	30.8 ± 12.9	7.9 ± 1.3**	33.9 ± 2.2

Values represent the mean of three biological replicates ± standard error. Bold values with asterisks are significantly different from controls at $P \leq 0.05^*$ and $P \leq 0.01^{**}$ as calculated using *t*-tests for pairwise comparison with SAS software (SAS Institute Inc.).

PvCesA6-RNAi Lines

Lignin content was unchanged in the transgenic lines compared with the non-transgenic controls. There were no significant differences between the S/G ratio of the transgenic and non-transgenic control lines. Cellulose content was decreased (up to 8%) and xylan content was increased (up to 4%) in transgenic lines compared with non-transgenic controls. There was no change in galactan content, whereas arabinan content was decreased up to 35% in transgenic lines compared with non-transgenic controls (Table 4B). Complete

chemical composition data is presented in Supplementary Table S3B.

Sugar Release Efficiency of Transgenic Plants

PvCesA4-Overexpressing Lines

Glucose release was significantly decreased in transgenic line 10 (19%), whereas xylose release was increased in transgenic line 7 (10%), line 9 (9%), and line 10 (19%) compared with the non-transgenic controls. No significant differences were observed

TABLE 3 | Cell wall chemical composition of transgenic lines overexpressing *PvCesA4* or *PvCesA6* and non-transgenic (WT) controls.

A. <i>PvCesA4</i>-overexpression						
Lines	1	7	8	9	10	WT
Lignin	16.9 ± 0.1	16.5 ± 0.2*	17.1 ± 0.1	17.4 ± 0.2	18.5 ± 0.1*	17.2 ± 0.1
S/G	0.7 ± 0.01	0.7 ± 0.01	0.7 ± 0.02	0.7 ± 0.03	0.7 ± 0.01	0.7 ± 0.01
Cellulose	28.1 ± 0.2*	28.0 ± 0.4*	28.7 ± 0.2*	24.9 ± 0.8*	20.4 ± 0.2*	30.5 ± 0.1
Hemicellulose	20.8 ± 0.2*	21.0 ± 0.2*	20.7 ± 0.1*	21.5 ± 0.4*	23.4 ± 0.3*	22.9 ± 0.1
Xylan	17.7 ± 0.2	17.8 ± 0.2	17.7 ± 0.1	18.2 ± 0.4*	19.9 ± 0.2*	17.7 ± 0.1
Galactan	0.9 ± 0.0*	0.9 ± 0.0*	0.9 ± 0.0*	1.0 ± 0.0*	1.0 ± 0.0*	1.2 ± 0.0
Arabinan	2.2 ± 0.0*	2.3 ± 0.0*	2.1 ± 0.0*	2.3 ± 0.0*	2.5 ± 0.0*	4.0 ± 0.0
B. <i>PvCesA6</i>-overexpression						
Lines	2	6	10	13	16	WT
Lignin	16.1 ± 0.1*	16.0 ± 0.1*	16.0 ± 0.2*	16.3 ± 0.0*	16.5 ± 0.2*	17.2 ± 0.1
S/G	0.6 ± 0.01*	0.6 ± 0.01*	0.6 ± 0.02*	0.6 ± 0.01*	0.6 ± 0.01*	0.7 ± 0.01
Cellulose	27.2 ± 0.3*	28.2 ± 0.1*	27.8 ± 0.3*	28.1 ± 0.1*	26.6 ± 0.1*	30.5 ± 0.1
Hemicellulose	23.7 ± 0.3*	23.2 ± 0.1	23.4 ± 0.0*	23.3 ± 0.1	22.8 ± 0.1	22.9 ± 0.1
Xylan	18.2 ± 0.3*	18.0 ± 0.1	18.1 ± 0.1*	18.3 ± 0.1*	17.7 ± 0.1	17.7 ± 0.1
Galactan	1.2 ± 0.0	1.1 ± 0.0	1.1 ± 0.0	1.1 ± 0.0	1.1 ± 0.0	1.2 ± 0.0
Arabinan	4.3 ± 0.0*	4.1 ± 0.0	4.2 ± 0.1*	3.9 ± 0.0	4.0 ± 0.0	4.0 ± 0.0

Values (weight% of cell wall residue) represent the mean of three biological replicates ± standard error. Bold values with asterisks are significantly different from controls at $P \leq 0.05^*$ and $P \leq 0.01^{**}$ as calculated using t-tests for pairwise comparison with SAS software (SAS Institute Inc.).

TABLE 4 | Cell wall chemical composition of *PvCesA4*-RNAi or *PvCesA6*-RNAi transgenic lines and non-transgenic (WT) controls.

A. <i>PvCesA4</i>-RNAi				
Lines	12	15	20	WT
Lignin	16.9 ± 0.1	16.5 ± 0.1	16.6 ± 0.1	16.8 ± 0.1
S/G	0.6 ± 0.01	0.6 ± 0.03	0.6 ± 0.02	0.6 ± 0.01
Cellulose	30.1 ± 0.1*	28.7 ± 0.1*	29.4 ± 0.1*	31.4 ± 0.3
Hemicellulose	21.4 ± 0.2	21.5 ± 0.1	21.3 ± 0.2	21.4 ± 0.2
Xylan	18.8 ± 0.1*	18.5 ± 0.1*	18.0 ± 0.1	18.0 ± 0.2
Galactan	1.0 ± 0.0	1.0 ± 0.0	1.1 ± 0.0	1.1 ± 0.0
Arabinan	1.6 ± 0.1*	2.0 ± 0.1*	2.2 ± 0.0	2.3 ± 0.1
B. <i>PvCesA6</i>-RNAi				
Lines	2	9	12	WT
Lignin	16.5 ± 0.5	16.3 ± 0.1	16.2 ± 0.5	16.8 ± 0.1
S/G	0.6 ± 0.02	0.6 ± 0.02	0.6 ± 0.01	0.6 ± 0.01
Cellulose	29.6 ± 0.4*	28.9 ± 0.4*	28.9 ± 0.4*	31.4 ± 0.3
Hemicellulose	22.0 ± 0.3	21.3 ± 0.3	21.1 ± 0.5	21.4 ± 0.2
Xylan	18.7 ± 0.2*	18.3 ± 0.3*	18.5 ± 0.3*	18.0 ± 0.2
Galactan	1.2 ± 0.0	1.1 ± 0.0	1.1 ± 0.0	1.1 ± 0.0
Arabinan	2.1 ± 0.1	1.9 ± 0.1*	1.5 ± 0.2*	2.3 ± 0.1

Values (weight% of cell wall residue) represent the mean of three biological replicates ± standard error. Bold values with asterisks are significantly different from controls at $P \leq 0.05^*$ and $P \leq 0.01^{**}$ as calculated using t-tests for pairwise comparison with SAS software (SAS Institute Inc.).

in total sugar release between transgenic and non-transgenic control lines with the exception of an increase in transgenic line 7 (5%) (**Figure 5A**), which was congruent with the significant reduced lignin content (**Table 3A**).

PvCesA6-Overexpressing Lines

Glucose release was significantly increased in transgenic line 13 (7%), whereas xylose release was increased in line 10 (8%) compared with non-transgenic control. However, all transgenic lines had increased total sugar release by 4–9% compared with the non-transgenic controls (**Figure 5B**) and were congruent with the significant reduction in lignin content and S/G ratio (**Table 3B**).

PvCesA4-RNAi and *PvCesA6*-RNAi Lines

There were no significant differences in glucose, xylose, and total sugar release between the *PvCesA4*-RNAi or *PvCesA6*-RNAi transgenic lines and non-transgenic controls (**Figures 6A,B**).

Cellulose Crystallinity of Transgenic Plants

PvCesA4-Overexpressing Lines

Cellulose crystallinity was significantly decreased in transgenic line 7 (7%), line 9 (6%), and line 10 (8%) compared with the non-transgenic controls (**Figure 7A**).

PvCesA6-Overexpressing Lines

All transgenic lines showed a significant decrease in cellulose crystallinity by 5–10% compared with the non-transgenic controls (**Figure 7B**).

PvCesA4-RNAi and *PvCesA6*-RNAi Lines

All *PvCesA4*-RNAi and *PvCesA6*-RNAi transgenic lines showed significantly decreased cellulose crystallinity up to 5 and 7%, respectively, compared with non-transgenic controls (**Figures 8A,B**).

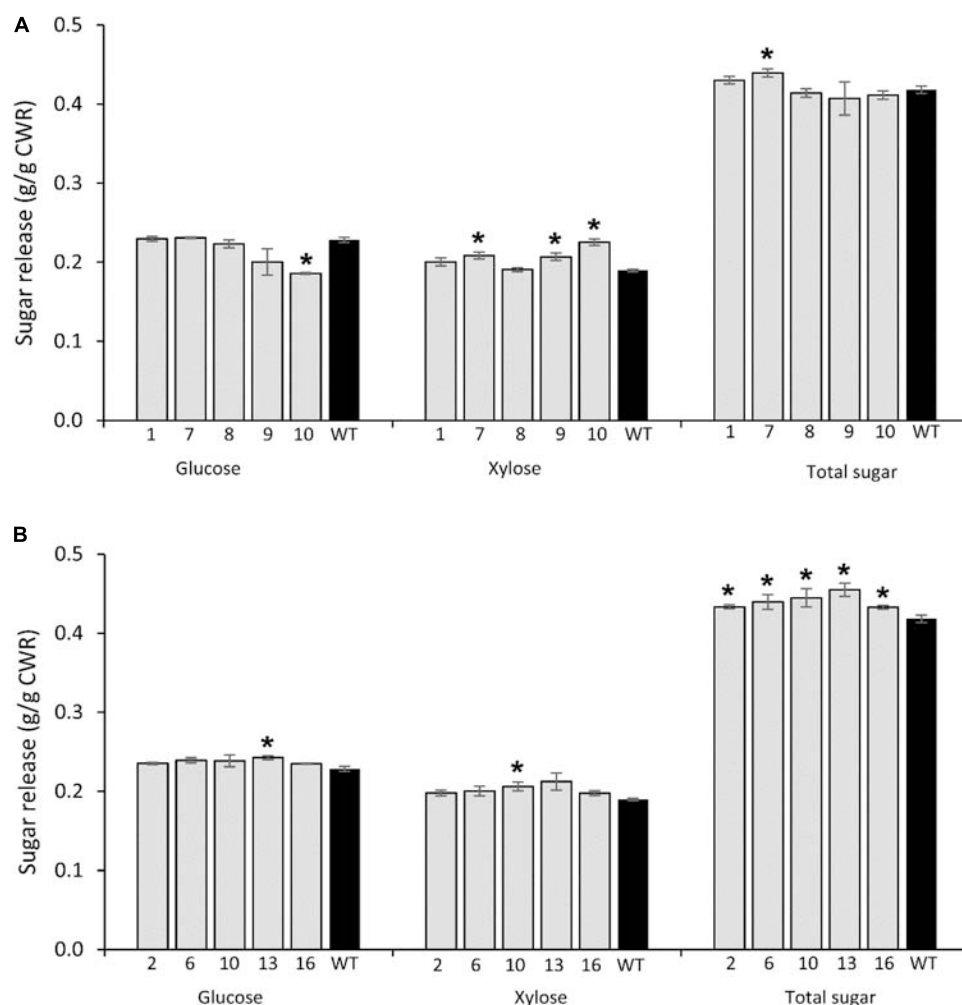


FIGURE 5 | Glucose, xylose, and total sugar release by enzymatic hydrolysis from transgenic lines overexpressing *PvCesA4* (A) and *PvCesA6* (B) and non-transgenic (WT) controls. Bars represent mean values of three biological replicates \pm standard error. Bars with asterisk are significantly different from controls at $P \leq 0.05^*$ and $P \leq 0.01^{**}$ as calculated using *t*-tests for pairwise comparison with SAS software (SAS Institute Inc.). CWR: cell wall residue.

Cellulose Characteristics of Transgenic Plants

Cellulose degree of polymerization and polydispersity were unchanged in the *PvCesA4* and *PvCesA6* overexpressing lines and in the RNAi lines compared with their respective non-transgenic controls (Supplementary Tables S4, S5).

Expression Level of Xylan Biosynthetic Genes *IRX9* and *IRX14* in Transgenic Plants Overexpressing *PvCesA4* and *PvCesA6*

Transcript abundance of *IRX9* and *IRX14* shown to be involved in xylan biosynthesis in Arabidopsis (Scheller and Ulvskov, 2010) was determined in *PvCesA4* and *PvCesA6* overexpressing lines. Expression analysis showed that *PvIRX9* transcripts were increased in the *PvCesA4* and *PvCesA6* transgenic lines, whereas *PvIRX14* transcripts were not affected as compared

to those found in non-transgenic control (Supplementary Figure S6).

DISCUSSION

Understanding the enzymes responsible for the synthesis and regulation of cellulose synthesis is a key to the engineering of biofuel crops for cellulose production and more efficient extraction of glucose from cellulose. With a view to study the CesA machinery in switchgrass, we identified switchgrass *PvCesA4* and *PvCesA6* that have orthologues with other plant species. For example, *PvCesA6* falls into the same clade as Arabidopsis AtCes2, AtCes5, AtCes6, AtCes9, and rice OsCesA3, OsCesA5, and OsCesA6. These proteins are all known to participate in primary wall cellulose synthesis (Wang et al., 2010; Endler and Persson, 2011). *PvCesA4* falls into the same clade as Arabidopsis AtCes4 and AtCes8, and rice OsCesA4 and

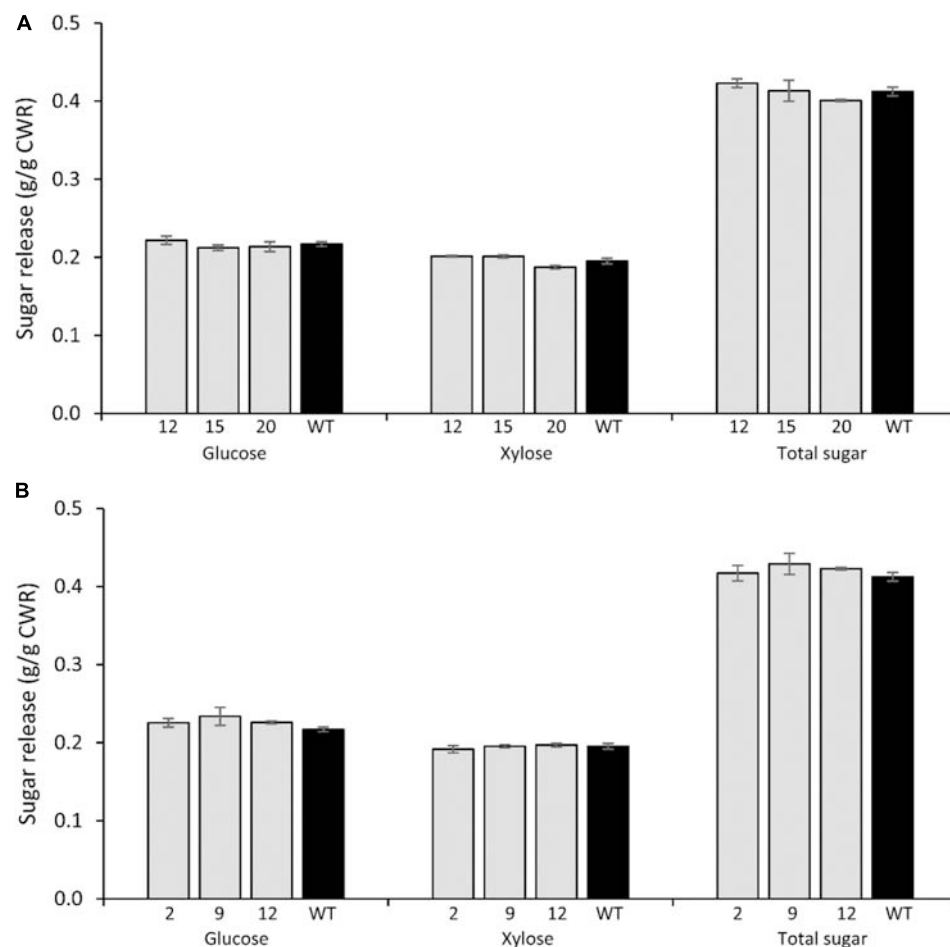


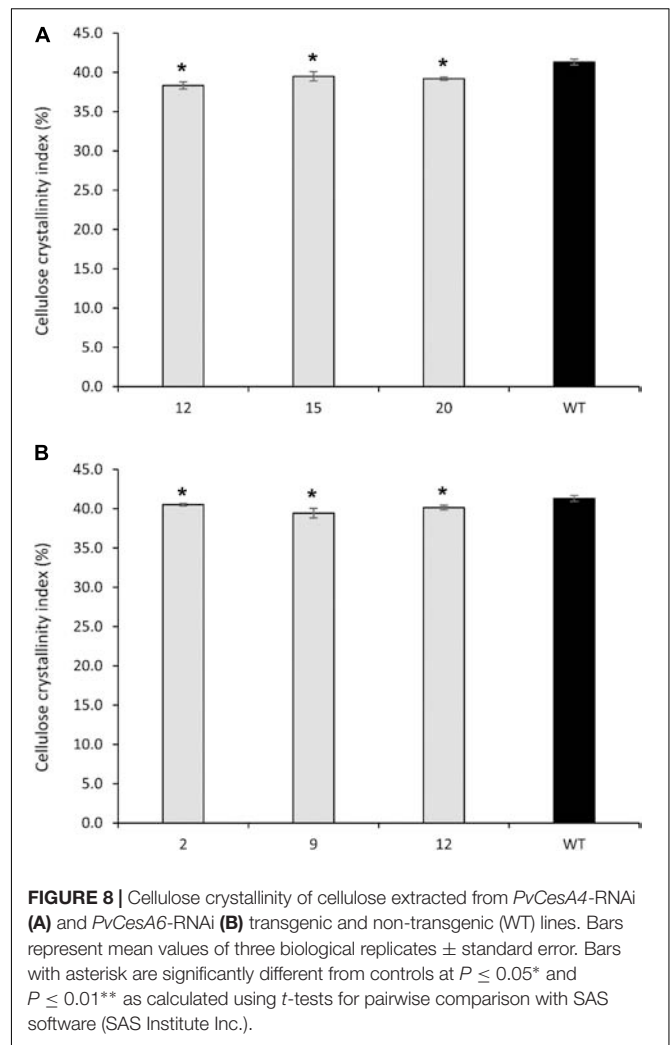
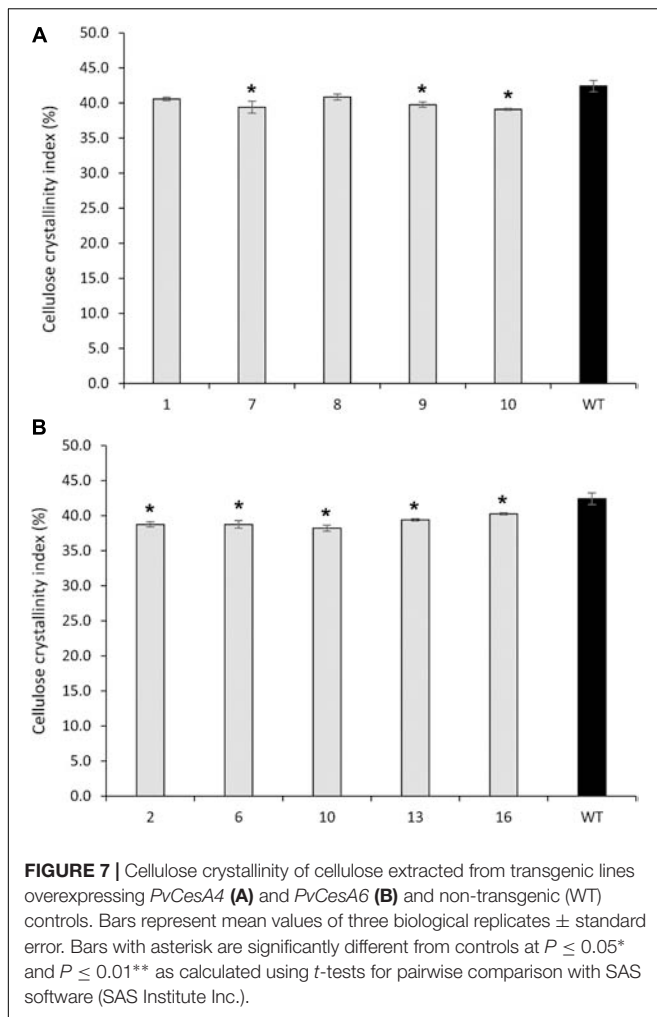
FIGURE 6 | Glucose, xylose, and total sugar release by enzymatic hydrolysis from *PvCesA4*-RNAi (A) and *PvCesA6*-RNAi (B) transgenic lines and non-transgenic (WT) controls. Bars represent mean values of three biological replicates \pm standard error. Bars with asterisk are significantly different from controls at $P \leq 0.05^*$ and $P \leq 0.01^{**}$ using *t*-tests for pairwise comparison with SAS software (SAS Institute Inc.). CWR: cell wall residue.

OsCesA7. These are known to be essential isoforms for secondary wall cellulose synthesis (Wang et al., 2010; Endler and Persson, 2011). Taken together, we conclude that *PvCesA6* plays a role in primary cell wall formation and *PvCesA4* in secondary cell wall formation in switchgrass.

Transcript expression analysis revealed that *PvCesA4* and *PvCesA6* are expressed in all the tissues tested, but that *PvCesA4* expression was highest in stems whereas *PvCesA6* expression was highest in leaves. Stem tissue is composed predominately of secondary cell walls that contain high amounts of cellulose and lignin which are valuable for biomass applications (Li F. et al., 2017). Furthermore, secondary cell walls of stems provide much of the rigidity and tensile and compression strength needed to support leaves and flowers, in contrast to the flexible primary wall of organs such as leaves (Fagard et al., 2000a). Consistently, higher secondary cell wall *CesA* gene expression in stems relative to that of primary cell wall *CesA* genes was observed in other plant species (Kotake et al., 2011; Li A. et al., 2013; Song et al., 2013; Mokshina et al., 2014; Chantreau et al., 2015; Petrik et al., 2016). A recent study involving switchgrass cell suspension

cultures also showed a higher expression of *CesA4* associated with the secondary cell wall formation (Rao et al., 2017). Our results are congruent with the function of *PvCesA4* in secondary wall formation and *PvCesA6* in primary cell wall formation in switchgrass.

There was a negative association between the level of increasing *PvCesA4* and *PvCesA6* expression and plant biomass production. For example, *PvCesA4* (lines 8 and 9) and *PvCesA6* (lines 6 and 13) overexpression lines with the greatest transcript levels (up to 41-fold increase of the transgene) had the greatest decrease (up to 81%) in biomass yield. Conversely, there was a positive association between the level of decreasing *PvCesA4* and *PvCesA6* expression and plant biomass production. For example, the two transgenic RNAi-knockdown *PvCesA6* (lines 2 and 12) with the greatest reduction (up to 64%) in expression of the gene had up to 77% decreased biomass yield. In general, the reduction in plant biomass was associated with decreased plant height and width. Yet, increasing or decreasing *PvCesA4* and *PvCesA6* expression at low to moderate levels in the transgenic lines resulted in biomass production equivalent to that of the



non-transgenic controls. Despite interest in the *CesA* family genes as potential targets for improving sugar yield in plant biomass, efforts to genetically manipulate members of the *CesA* gene family to achieve this goal have been challenging (Burton and Fincher, 2014). For example, overexpression of *CesA* genes in barley and poplar largely resulted in the silencing of both the transgenes and the endogenous genes (Joshi et al., 2011; Tan et al., 2015). Furthermore, although *CesAs* have been shown to be essential for plant growth (Persson et al., 2007), the overexpression of *CesA* genes has not generally led to improved plant growth but rather resulted in defective plant growth and reduced biomass yield in *Arabidopsis*, barley, and poplar plants (Zhong et al., 2003; Joshi et al., 2011; Tan et al., 2015). However, a recent study has demonstrated that overexpression of certain primary wall *CesA6*-like genes can improve plant growth in *Arabidopsis* (Hu H. et al., 2018). In the present study, we also observed that overexpression of *PvCesA6* (i.e., a putative primary wall *CesA6*-like gene) by moderate levels of transgene expression (line 2 with seven-fold increase) resulted in increased plant biomass largely by increasing plant tiller number, which may provide a useful trait for biomass crops such as switchgrass. Decreasing *CesA* expression in tobacco and flax by

gene knockdown via the VIGS system and in *Brachypodium* by gene knockdown using an artificial microRNA system resulted in markedly shorter plants (Burton et al., 2000; Handakumbura et al., 2013; Chantreau et al., 2015). Studies of *Arabidopsis* and rice mutants impaired in *CesA* expression also reported severe growth inhibition of plants (Fagard et al., 2000b; Ellis et al., 2002; Tanaka et al., 2003; Taylor et al., 2003; Zhong et al., 2003; Wang et al., 2006; Zhang et al., 2009; Pysh et al., 2012; Rubio-Díaz et al., 2012; Wang et al., 2012). In contrast, however, rice mutants with amino acid alterations in *CesA* showed normal plant growth (Song et al., 2013) or greater biomass production (Li F. et al., 2017). In the present study, plant growth in the RNAi-knockdown *PvCesA* lines mostly depended on the level of reduction of transcript abundance of *PvCesA* where decreased expression at high levels (up to 64%) resulted in smaller plants attributing to the up to 77% decrease in biomass production, whereas low to moderate levels of decrease (11–36%) had no effect on plant growth and biomass production. Consistent with these observations, several of the transgenic lines with more than 70% reduced *CesA* expression failed to reach the R1 growth stage (data not shown). These results

suggest that an optimized level of expression of the *CesA* genes by an inducible or appropriate promoter may be required to produce transgenic plants with the desired growth and cellulose content.

Cell wall chemical analyses showed that both the upregulation and downregulation of *PvCesA* were associated with a decrease in cellulose content in the transgenic lines. Cellulose content was also reduced in VIGS *CesA*-silenced tobacco and flax (Burton et al., 2000; Chantreau et al., 2015) and in the Arabidopsis, rice, and barley mutants (Ellis et al., 2002; Tanaka et al., 2003; Taylor et al., 2003; Zhong et al., 2003; Chen et al., 2005; MacKinnon et al., 2006; Paredes et al., 2008; Zhang et al., 2009; Burton et al., 2010; Kotake et al., 2011; Harris et al., 2012; Rubio-Díaz et al., 2012; Song et al., 2013; Li F. et al., 2017). Our results involving the reduced cellulose content in the RNAi *PvCesA*-silenced plants are in line with those reported studies, which may further suggest that *PvCesAs* are functional orthologues of *CesA* genes in the other plant species. The reduced cellulose content was also observed in the *PvCesA*-overexpressing lines. Co-expression of at least three *CesA* genes is essential for cellulose biosynthesis (Somerville, 2006), thus, the individual overexpression of *PvCesA* possibly interferes with the machinery controlling cellulose biosynthesis in switchgrass. Consistent with this, reduced cellulose content was shown in *CesA*-overexpressing transgenic Arabidopsis, barley, and poplar plants (Zhong et al., 2003; Joshi et al., 2011; Tan et al., 2015). In contrast, however, there is a recent study demonstrating that the overexpression of specific individual primary wall *CesA6*-like genes results in increased cellulose content in Arabidopsis (Hu H. et al., 2018). In regard to the hemicellulose content (i.e., non-cellulosic cell wall polysaccharides), xylan was increased while galactan and arabinan were mostly reduced or unchanged in *PvCesA* overexpressing or RNAi knockdown lines. Xylan is a major hemicellulose in cell walls of mature tissues of grasses while galactan and arabinan are more abundant in woody plants (Girio et al., 2010). Our expression analysis involving *IRX9* and *IRX14*, shown to be involved in xylan biosynthesis in Arabidopsis (Scheller and Ulvskov, 2010), showed that while *PvIRX9* expression was increased in the *PvCesA4* and *PvCes6* transgenic lines, the expression of *PvIRX14* was unaffected. It has been shown that *IRX14*, rather than *IRX9*, is required for xylan backbone synthesis in primary cell walls of Arabidopsis (Mortimer et al., 2015). Our results may suggest differential functions of these *IRXs* genes in xylan extension in switchgrass. It may also suggest that the increase in xylan content in *PvCesA4* and *PvCes6* transgenic lines is regulated at a transcriptional level (if a part).

Xylan content was also shown to positively affect the enzymatic digestibility of biomass by reducing cellulose crystallinity (Li F. et al., 2013). Thus increased xylan content is considered as another possible means to enhance lignocellulose saccharification in bioenergy crops. It is possible that the increased xylan content in *PvCesA*-overexpressing or RNAi knockdown lines was a compensation response to decreased cellulose production. In other plant species, reduced cellulose content was also associated with an increase in the non-cellulosic cell wall-related sugars (Burton et al., 2000; Kotake et al., 2011;

Song et al., 2013; Chantreau et al., 2015; Li F. et al., 2017). The present study further supports a connection between the cellular machinery controlling cellulose and hemicellulose biosynthesis in switchgrass.

Increasing or decreasing *PvCesA* expression resulted in reduced cellulose crystallinity in the transgenic lines. Cellulose crystallinity has been demonstrated as a factor that negatively impacts saccharification efficiency (Himmel et al., 2007; Harris et al., 2012; Zhang W. et al., 2013; Li F. et al., 2017). Cellulose concentration is positively associated with cellulose crystallinity, which is negatively associated with biomass saccharification in most plant species (Wang et al., 2016). Therefore, there may exist an upper limit to direct cellulose synthesis in cell walls. Where examined, the reduction of cellulose content in *CesA* mutants and *CesA*-overexpressed transgenic plants is consistently associated with reduced cellulose crystallinity (Wang et al., 2016). Increased xylan content also leads to reduced cellulose crystallinity (Li F. et al., 2013). Thus, the reduced cellulose crystallinity observed in the present study could be the result of decreased cellulose content and/or increased xylan content in the transgenic lines. Although we must note that there was not a consistent association between the decreased cellulose crystallinity and sugar release efficiency in *PvCesA* overexpressing or RNAi knockdown lines.

The accessibility of cellulose to enzymatic breakdown into fermentable sugars is also limited by the presence of lignin, and genetic modifications of lignin have been shown to enhance biomass saccharification (Chen and Dixon, 2007; Baxter et al., 2014, 2015; Bonawitz et al., 2014; Eudes et al., 2014; Wilkerson et al., 2014; Hu Z. et al., 2018). Notably, lignin content and S/G ratio were decreased only in *PvCesA6*-overexpressing lines that also had an accompanying increase in sugar release efficiency. Lignin content was shown to be unchanged in the *CesA* rice mutants (Kotake et al., 2011; Li F. et al., 2017) and increased in *CesA*-overexpressing Arabidopsis (Hu H. et al., 2018) plants. Decreased lignin content and modified composition in the *PvCesA6*-overexpressing lines reported here might suggest an impact of *PvCesA6* expression on lignin biosynthesis and reduced recalcitrance in switchgrass. A goal for the genetic engineering of plant cell walls for improved biofuel production is to develop feedstock with increased biomass saccharification properties without effecting plant growth and biomass accumulation (Abramson et al., 2010). In line with this goal, we have shown that lines with low-to-moderate *PvCesA6*-overexpression had decreased recalcitrance, increased sugar release, and that some lines also grew better than wild type.

CONCLUSION

In conclusion, we have shown that genetic manipulation of *PvCesA* can affect cellulose, hemicellulose, and lignin content and cellulose crystallinity to result in improved biomass digestibility without negatively, and in some cases even positively, affecting plant growth. These results validate *PvCesA* as cellulose synthesis genes and provide further insights into the effects of specific

over-expression and knockdown expression of *CesA* on cell wall content and plant growth in switchgrass. In addition, our results suggest a direct or indirect association between cellulose, xylan, and lignin expression. These results may suggest a possible role for cellulose-xylan-lignin polymer covalent or non-covalent physical interaction in switchgrass biomass. Therefore, further study of the *PvCesA4* and *PvCesA6* transgenic plants to identify specific wall fractions that may contain cellulose-xylan-lignin polymer cross-links or strong association is expected to yield insight into the mechanisms involved in cellulose synthesis and cell wall architecture of plant cell walls in switchgrass. Switchgrass is among several high-biomass grass species. Certainly, key cell wall enzymes in these species play a role in the fast growth of their herbaceous biomass. It would be interesting to use switchgrass *CesA* genes in complementation studies of genomically-characterized grasses such as rice and *Brachypodium*, in which native *CesA* orthologues have been knocked out. One could imagine the simultaneous directed overexpression of switchgrass *CesA* genes in rice while knocking out the rice orthologues, both singularly and in combinations. Such an approach could further elucidate functionality and perhaps serve to engineer cereal crops for higher productivity.

AUTHOR CONTRIBUTIONS

MM designed the experiments, performed the expression studies, participated in plant phenotyping and preparation of plant samples for cell wall analysis, analyzed the data, and wrote the manuscript. HB performed plant phenotyping and statistical analysis, participated in preparation of plant samples for cell wall analysis and RNA isolation. ML, XM, YP, and AR performed cellulose analyses. AB and DM performed phylogenetic tree work and contributed to the conception of the study. KK and NL performed cell wall chemical analyses. WW assisted with designing gene specific primers. J-YZ and MU performed cloning of the target genes. GT, RS, and MD performed lignin and sugar

release analyses. Z-YW produced the transgenic plants. CNS conceived of the study and its design, coordination and assisted with interpretation of results, and revisions to the manuscript. All authors contributed to text, data analysis, read, and approved the final manuscript.

FUNDING

This work was supported by funding from the BioEnergy Science Center. The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the U.S. Department of Energy. The work was also partially supported by the Agriculture and Food Research Initiative (United States Department of Agriculture) and Southeastern Partnership for Integrated Biomass and Supply Systems (The IBSS Partnership). Funding was also provided by the Ivan Racheff Endowment and a USDA Hatch grant to CNS.

ACKNOWLEDGMENTS

We thank Hayley Rideout for her assistance with maintaining plants and RNA isolation. We thank Erica Gjersing and Crissa Doeppke of NREL for their assistance with cell wall characterization and Susan Holladay for her assistance with data entry into LIMS. We also thank the two reviewers for helpful comments that greatly improved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.01114/full#supplementary-material>

REFERENCES

- Abramson, M., Shoseyov, O., and Shani, Z. (2010). Plant cell wall reconstruction toward improved lignocellulosic production and processability. *Plant Sci.* 178, 61–72. doi: 10.1016/j.plantsci.2009.11.003
- Appenzeller, L., Doblin, M., Barreiro, R., Wang, H. Y., Niu, X. M., 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
- Bali, G., Khunsupat, R., Akinoshio, H., Payyavula, R. S., Samuel, R., Tuskan, G. A., et al. (2016). Characterization of cellulose structure of *Populus* plants modified in candidate cellulose biosynthesis genes. *Biomass Bioenergy* 94, 146–154. doi: 10.1016/j.biombioe.2016.08.013
- Baxter, H. L., Mazarei, M., Labbé, N., Kline, L. M., Cheng, Q., Windham, M. T., et al. (2014). Two-year field analysis of reduced recalcitrance transgenic switchgrass. *Plant Biotechnol. J.* 12, 914–924. doi: 10.1111/pbi.12195
- 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
- Biswal, A. K., Atmodjo, M. A., Li, M., Baxter, H. L., Yoo, C. G., Pu, Y., et al. (2018). Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. *Nat. Biotechnol.* 36, 249–257. doi: 10.1038/nbt.4067
- Bonawitz, N. D., Kim, J. I., Tobimatsu, Y., Ciesielski, P. N., Anderson, N. A., Ximenes, E., et al. (2014). Disruption of mediator rescues the stunted growth of a lignin-deficient *Arabidopsis* mutant. *Nature* 509, 376–380. doi: 10.1038/nature13084
- Burton, R. A., and Fincher, G. B. (2014). Plant cell wall engineering: applications in biofuel production and improved human health. *Curr. Opin. Biotechnol.* 26, 79–84. doi: 10.1016/j.copbio.2013.10.007
- Burton, R. A., Gibeaut, D. M., Bacic, A., Findlay, K., Roberts, K., Hamilton, A., et al. (2000). Virus-induced silencing of a plant cellulose synthase gene. *Plant Cell* 12, 691–705. doi: 10.1105/tpc.12.5.691
- Burton, R. A., Ma, G., Baumann, U., Harvey, A. J., Shirley, N. J., Taylor, J., et al. (2010). A customized gene expression microarray reveals that the brittle stem phenotype fs2 of barley is attributable to a retroelement in the *HvCesA4* cellulose synthase gene. *Plant Physiol.* 153, 1716–1728. doi: 10.1104/pp.110.158329
- Burton, R. A., Shirley, N. J., King, B. J., Harvey, A. J., and Fincher, G. B. (2004). The *CesA* gene family of barley. Quantitative analysis of transcripts reveals two

- groups of co-expressed genes. *Plant Physiol.* 134, 224–236. doi: 10.1104/pp.103.032904
- Chantreau, M., Chabbert, B., Billiard, S., Hawkins, S., and Neutelings, G. (2015). Functional analyses of cellulose synthase genes in flax (*Linum usitatissimum*) by virus-induced gene silencing. *Plant Biotechnol. J.* 13, 1312–1324. doi: 10.1111/pbi.12350
- Chen, F., and Dixon, R. A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.* 25, 759–761. doi: 10.1038/nbt1316
- Chen, Z., Hong, X., Zhang, H., Wang, Y., Li, X., Zhu, J. K., et al. (2005). Disruption of the cellulose synthase gene, *AtCesA8/IRX1*, enhances drought and osmotic stress tolerance in *Arabidopsis*. *Plant J.* 43, 273–283. doi: 10.1111/j.1365-313X.2005.02452.x
- Cosgrove, D. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6, 850–861. doi: 10.1038/nrm1746
- Decker, S. R., Carlile, M., Selig, M. J., Doeppke, C., Davis, M., Sykes, R., et al. (2012). Reducing the effect of variable starch levels in biomass recalcitrance screening. *Methods Mol. Biol.* 908, 181–195. doi: 10.1007/978-1-61779-956-3_17
- Desprez, T., Juraniec, M., Crowell, E. F., Jouy, H., Pochylova, Z., Parcy, F., et al. (2007). Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15572–15577. doi: 10.1073/pnas.0706569104
- Djerbi, S., Lindskog, M., Arvestad, L., Sterky, F., and Teeril, T. T. (2005). The genome sequence of black cottonwood (*Populus trichocarpa*) reveals 18 conserved cellulose synthase (CesA) genes. *Planta* 221, 739–746. doi: 10.1007/s00425-005-1498-4
- Ellis, C., Karafyllidis, I., Wasternack, C., and Turner, J. G. (2002). The *Arabidopsis* mutant *cevl* links cell wall signaling to jasmonate and ethylene responses. *Plant Cell* 14, 1557–1566. doi: 10.1105/tpc.002022
- Endler, A., and Persson, S. (2011). Cellulose synthases and synthesis in *Arabidopsis*. *Mol. Plant* 4, 199–211. doi: 10.1093/mp/ssq079
- Eudes, A., Liang, Y., Mitra, P., and Loqué D. (2014). Lignin bioengineering. *Curr. Opin. Biotechnol.* 26, 189–198. doi: 10.1016/j.copbio.2014.01.002
- Fagard, M., Desnos, T., Desprez, T., Goubet, F., Refregier, G., Mouille, G., et al. (2000a). PROCUSTE1 encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of *Arabidopsis*. *Plant Cell* 12, 2409–2424. doi: 10.1105/tpc.12.12.2409
- Fagard, M., Höfte, H., and Vernhettes, S. (2000b). Cell wall mutants. *Plant Physiol. Biochem.* 38, 15–25. doi: 10.1016/S0981-9428(00)00171-6
- Girio, F. M., Fonseca, C., Carvalho, F., Duarte, L. C., Marques, S., and Bogel-Lukasik, R. (2010). Hemicelluloses for fuel ethanol: a review. *Bioresour. Technol.* 101, 4775–4800. doi: 10.1016/j.biortech.2010.01.088
- 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 A4 or 7 results in abnormal cell walls. *BMC Plant Biol.* 13:131. doi: 10.1186/1471-2229-13-131
- Hardin, C. F., Fu, C., Hisano, H., Xiao, X., Shen, H., Stewart, C. N. Jr., et al. (2013). Standardization of switchgrass sample collection for cell wall and biomass trait analysis. *BioEnergy Res.* 6, 755–762. doi: 10.1007/s12155-012-9292-1
- Harris, D. M., Corbin, K., Wang, T., Gutierrez, R., Bertolo, A. L., Petti, C., et al. (2012). Cellulose microfibril crystallinity is reduced by mutating C-terminal transmembrane region residues CES1A903V and CES3T942I of cellulose synthase. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4098–4103. doi: 10.1073/pnas.1200352109
- Himmel, M. E., and Bayer, E. A. (2009). Lignocellulose conversion to biofuels: current challenges, global perspectives. *Curr. Opin. Biotechnol.* 20, 316–317. doi: 10.1016/j.copbio.2009.05.005
- Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., et al. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315, 804–807. doi: 10.1126/science.1137016
- Hu, H., Zhang, R., Feng, S., Wang, Y., Wang, Y., Fan, C., et al. (2018). Three *AtCesA6*-like members enhance biomass production by distinctively promoting cell growth in *Arabidopsis*. *Plant Biotechnol. J.* 16, 976–988. doi: 10.1111/pbi.12842
- Hu, Z., Zhang, G., Muhammad, A., Abdul Samad, R., Wang, Y., Walton, J. D., et al. (2018). Genetic loci simultaneously controlling lignin monomers and biomass digestibility of rice straw. *Sci. Rep.* 8:3636. doi: 10.1038/s41598-018-21741-y
- Joshi, C. P., Bhandari, S., Ranjan, P., Kalluri, U. C., Liang, X., and Fujino, T. (2004). Genomics of cellulose biosynthesis in poplars. *New Phytol.* 164, 53–61. doi: 10.1111/j.1469-8137.2004.01155.x
- Joshi, C. P., Thammannagowda, S., Fujino, T., Gou, J. Q., Avci, U., Haigler, C. H., et al. (2011). Perturbation of wood cellulose synthesis causes pleiotropic effects in transgenic aspen. *Mol. Plant* 4, 331–345. doi: 10.1093/mp/ssq081
- Kalluri, U. C., Yin, H., Yang, X., and Davison, B. H. (2014). Systems and synthetic biology approaches to alter plant cell walls and reduce biomass recalcitrance. *Plant Biotechnol. J.* 12, 1207–1216. doi: 10.1111/pbi.12283
- Kotake, T., Aohara, T., Hirano, K., Sato, A., Kaneko, Y., Tsumuraya, Y., et al. (2011). Rice Brittle culm 6 encodes a dominant-negative form of CesA protein that perturbs cellulose synthesis in secondary cell walls. *J. Exp. Bot.* 62, 2053–2062. doi: 10.1093/jxb/erq395
- Kumar, M., and Turner, S. (2015). Plant cellulose synthesis: CESA proteins crossing kingdoms. *Phytochemistry* 112, 91–99. doi: 10.1016/j.phytochem.2014.07.009
- Kumar, M., Thammannagowda, S., Bulone, V., Chiang, V., Han, K. H., Joshi, C. P., et al. (2009). An update on the nomenclature for the cellulose synthase genes in *Populus*. *Trends Plant Sci.* 14, 248–254. doi: 10.1016/j.tplants.2009.02.004
- Li, A., Xia, T., Xu, W., Chen, T. T., Li, X. L., Fan, J., et al. (2013). An integrative analysis of four CESA isoforms specific for fiber cellulose production between *Gossypium hirsutum* and *Gossypium barbadense*. *Planta* 237, 1585–1597. doi: 10.1007/s00425-013-1868-2
- Li, F., Ren, S., Zhang, W., Xu, Z., Xie, G., Chen, Y., et al. (2013). Arabinose substitution degree in xylan positively affects lignocellulose enzymatic digestibility after various NaOH/H₂SO₄ pretreatments in *Miscanthus*. *Bioresour. Technol.* 130, 629–637. doi: 10.1016/j.biortech.2012.12.107
- Li, F., Xie, G., Huang, J., Zhang, R., Li, Y., Zhang, M., et al. (2017). OsCESA9 conserved-site mutation leads to largely enhanced plant lodging resistance and biomass enzymatic saccharification by reducing cellulose DP and crystallinity in rice. *Plant Biotechnol. J.* 15, 1093–1104. doi: 10.1111/pbi.12700
- Li, M., Pu, Y., Yoo, C. G., Gjersing, E., Decker, S. R., Doeppke, C., et al. (2017). Study of traits and recalcitrance reduction of field-grown COMT down-regulated switchgrass. *Biotechnol. Biofuels* 10, 12. doi: 10.1186/s13068-016-0695-7
- Liu, Z., Schneider, R., Kesten, C., Zhang, Y., Somssich, M., Zhang, Y., et al. (2016). Cellulose-microtubule uncoupling proteins prevent lateral displacement of microtubules during cellulose synthesis in *Arabidopsis*. *Dev. Cell* 38, 305–315. doi: 10.1016/j.devcel.2016.06.032
- MacKinnon, I. M., Sturcová, A., Sugimoto-Shirasu, K., His, I., McCann, M. C., and Jarvis, M. C. (2006). Cell-wall structure and anisotropy in procuste, a cellulose synthase mutant of *Arabidopsis thaliana*. *Planta* 224, 438–448. doi: 10.1007/s00425-005-0208-6
- Mann, D. G. J., LaFayette, P. R., Abercrombie, L. L., King, Z. R., Mazarei, M., Halter, M. C., et al. (2012). Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum* L.) and other monocot species. *Plant Biotechnol. J.* 10, 226–236. doi: 10.1111/j.1467-7652.2011.00658.x
- Mizrachi, E., Mansfield, S. D., and Myburg, A. A. (2012). Cellulose factories: advancing bioenergy production from forest trees. *New Phytol.* 194, 54–62. doi: 10.1111/j.1469-8137.2011.03971.x
- Mokshina, N., Gorshkova, T., and Deyholos, M. K. (2014). Chitinase-like (CTL) and cellulose synthase (CESA) gene expression in gelatinous-type cellulosic walls of flax (*Linum usitatissimum* L.) bast fibers. *PLoS One* 9:e97949. doi: 10.1371/journal.pone.0097949
- Moore, K. J., Moser, L. E., Vogel, K. P., Waller, S. S., Johnson, B. E., and Pedersen, J. F. (1991). Describing and quantifying growth stages of perennial forage grasses. *Agron. J.* 83, 1073–1077. doi: 10.2134/agronj1991.00021962008300060027x
- Mortimer, J. C., Faria-Blanc, N., Yu, X., Tryfona, T., Sorieul, M., Ng, Y. Z., et al. (2015). An unusual xylan in *Arabidopsis* primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14. *Plant J.* 83, 413–426. doi: 10.1111/tpj.12898
- Nelson, R. S., Stewart, C. N. Jr., Gou, J., Holladay, S., Gallego-Giraldo, L., Flanagan, A., et al. (2017). Development and use of a switchgrass (*Panicum virgatum* L.) transformation pipeline by the BioEnergy Science Center to evaluate plants for reduced cell wall recalcitrance. *Biotechnol. Biofuels* 10:309. doi: 10.1186/s13068-017-0991-x
- Paredez, A. R., Persson, S., Ehrhardt, D. W., and Somerville, C. R. (2008). Genetic evidence that cellulose synthase activity influences microtubule cortical array organization. *Plant Physiol.* 147, 1723–1734. doi: 10.1104/pp.108.120196

- Pear, J. R., Kawagoe, Y., Schreckengost, W. E., Delmer, D. P., and Stalker, D. M. (1996). Higher plants contain homologs of the bacterial *celA* genes encoding the catalytic subunit of cellulose synthase. *Proc. Natl. Acad. Sci. U.S.A.* 93, 12637–12642. doi: 10.1073/pnas.93.22.12637
- Persson, S., Paredes, A., Carroll, A., Palsdottir, H., Doblin, M., Poindexter, P., et al. (2007). Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15566–15571. doi: 10.1073/pnas.0706592104
- Petrík, D. L., Cass, C. L., Padmakshan, D., Foster, C. E., Vogel, J. P., Karlen, S. D., et al. (2016). BdCESA7, BdCESA8, and BdPMT utility promoter constructs for targeted expression to secondary cell-wall-forming cells of grasses. *Front. Plant Sci.* 7:55. doi: 10.3389/fpls.2016.00055
- Pysh, L., Alexander, N., Swatzyna, L., and Harbert, R. (2012). Four alleles of AtCESA3 form an allelic series with respect to root phenotype in *Arabidopsis*. *Physiol. Planta* 144, 369–381. doi: 10.1111/j.1399-3054.2012.01575.x
- Rao, X., Shen, H., Pattathil, S., Hahn, M. G., Gelineo-Albersheim, I., Mohnen, D., et al. (2017). Dynamic changes in transcriptome and cell wall composition underlying brassinosteroid-mediated lignification of switchgrass suspension cells. *Biotechnol. Biofuels* 10, 266. doi: 10.1186/s13068-017-0954-2
- Rubio-Díaz, S., Pérez-Pérez, J. M., González-Bayón, R., Muñoz-Viana, R., Borrega, N., Mouille, G., et al. (2012). Cell expansion-mediated organ growth is affected by mutations in three EXIGUA Genes. *PLoS One* 7:e36500. doi: 10.1371/journal.pone.0036500
- Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315
- Schneider, R., Hanak, T., Persson, S., and Voigt, C. A. (2016). Cellulose and callose synthesis and organization in focus, what's new? *Curr. Opin. Plant Biol.* 34, 9–16. doi: 10.1016/j.pbi.2016.07.007
- Selig, M. J., Tucker, M. P., Sykes, R. W., Reichel, K. L., Brunecky, R., Himmel, M. E., et al. (2010). Lignocellulose recalcitrance screening by integrated high-throughput hydrothermal pretreatment and enzymatic saccharification. *Ind. Biotechnol.* 6, 104–111. doi: 10.1089/ind.2010.0009
- Shen, H., Fu, C., Xiao, X., Ray, T., Tang, Y., Wang, Z., et al. (2009). Developmental control of lignification in stems of lowland switchgrass variety Alamo and the effects on saccharification efficiency. *BioEnergy Res.* 2, 233–245. doi: 10.1007/s12155-009-9058-6
- Somerville, C. (2006). Cellulose synthesis in higher plants. *Annu. Rev. Cell Dev. Biol.* 22, 53–78. doi: 10.1146/annurev.cellbio.22.022206.160206
- Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., et al. (2004). Toward a systems approach to understanding plant cell walls. *Science* 306, 2206–2211. doi: 10.1126/science.1102765
- Song, X. Q., Liu, L. F., Jiang, Y. J., Zhang, B. C., Gao, Y. P., Liu, X. L., et al. (2013). Disruption of secondary wall cellulose biosynthesis alters cadmium translocation and tolerance in rice plants. *Mol. Plant* 6, 768–780. doi: 10.1093/mp/sst025
- Studer, M. H., DeMartini, J. D., Brethauer, S., McKenzie, H. L., and Wyman, C. E. (2009). Engineering of a high-throughput screening system to identify cellulosic biomass, pretreatments, and enzyme formulations that enhance sugar release. *Biotechnol. Bioeng.* 105, 231–238. doi: 10.1002/bit.22527
- Sykes, R., Yung, M., Novaes, E., Kirst, M., Peter, G., and Davis, M. (2009). High-throughput screening of plant cell-wall composition using pyrolysis molecular beam mass spectroscopy. *Methods Mol. Biol.* 581, 169–183. doi: 10.1007/978-1-60761-214-8_12
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tan, H., Shirley, N. J., Singh, R. R., Henderson, M., Dhugga, K. S., Mayo, G. M., et al. (2015). Powerful regulatory systems and post-transcriptional gene silencing resist increases in cellulose content in cell walls of barley. *BMC Plant Biol.* 15:62. doi: 10.1186/s12870-015-0448-y
- Tanaka, K., Murata, K., Yamazaki, M., Onosato, K., Miyao, A., and Hirochika, H. (2003). Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. *Plant Physiol.* 133, 73–83. doi: 10.1104/pp.103.022442
- Taylor, N. G., Howells, R. M., Huttly, A. K., Vickers, K., and Turner, S. R. (2003). Interactions among three distinct Cesa proteins essential for cellulose synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1450–1455. doi: 10.1073/pnas.0337628100
- Wang, D., Yuan, S., Yin, L., Zhao, J., Guo, B., Lan, J., et al. (2012). A missense mutation in the transmembrane domain of CESA9 affects cell wall biosynthesis and plant growth in rice. *Plant Sci.* 196, 117–124. doi: 10.1016/j.plantsci.2012.08.002
- Wang, J., Howles, P. A., Cork, A. H., Birch, R. J., and Richard E. Williamson, R. E. (2006). Chimeric proteins suggest that the catalytic and/or C-terminal domains give Cesa1 and Cesa3 access to their specific sites in the cellulose synthase of primary walls. *Plant Physiol.* 142, 685–695. doi: 10.1104/pp.106.084004
- Wang, L., Guo, K., Li, Y., Tu, Y., Hu, H., Wang, B., et al. (2010). Expression profiling and integrative analysis of the CESA/CSL superfamily in rice. *BMC Plant Biol.* 10, 282. doi: 10.1186/1471-2229-10-282
- Wang, Y., Fan, C., Hu, H., Li, Y., Sun, D., Wang, Y., et al. (2016). Genetic modification of plant cell walls to enhance biomass yield and biofuel production in bioenergy crops. *Biotechnol. Adv.* 34, 997–1017. doi: 10.1016/j.biotechadv.2016.06.001
- Wilkerson, C. G., Mansfield, S. D., Lu, F., Withers, S., Park, J. Y., Karlen, S. D., et al. (2014). Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. *Science* 344, 90–93. doi: 10.1126/science.1250161
- Xi, Y., Fu, C., Ge, Y., Nandakumar, R., Hisano, H., Bouton, J., et al. (2009). Agrobacterium-mediated transformation of switchgrass and inheritance of the transgenes. *Bioenergy Res.* 2, 275–283. doi: 10.1007/s12155-009-9049-7
- Zhang, B., Deng, L., Qian, Q., Xiong, G., Zeng, D., Li, R., et al. (2009). A missense mutation in the transmembrane domain of CESA4 affects protein abundance in the plasma membrane and results in abnormal cell wall biosynthesis in rice. *Plant Mol. Biol.* 71, 509–524. doi: 10.1007/s11103-009-9536-4
- Zhang, J. Y., Lee, Y. C., Torres-Jerez, I., Wang, M., Yin, Y., Chou, W. C., et al. (2013). Development of an integrated transcript sequence database and a gene expression atlas for gene discovery and analysis in switchgrass (*Panicum virgatum* L.). *Plant J.* 74, 160–173. doi: 10.1111/tpj.12104
- Zhang, W., Yi, Z., Huang, J., Li, F., Hao, B., Li, M., et al. (2013). Three lignocellulose features that distinctively affect biomass enzymatic digestibility under NaOH and H₂SO₄ pretreatments in *Miscanthus*. *Bioresour. Technol.* 130, 30–37. doi: 10.1016/j.biortech.2012.12.029
- Zhong, R. Q., Morrison, W. H., Freshour, G. D., Hahn, M. G., and Ye, Z. H. (2003). Expression of a mutant form of cellulose synthase AtCesa7 causes dominant negative effect on cellulose biosynthesis. *Plant Physiol.* 132, 786–795. doi: 10.1104/pp.102.019331

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Mazarei, Baxter, Li, Biswal, Kim, Meng, Pu, Wuddineh, Zhang, Turner, Sykes, Davis, Udvardi, Wang, Mohnen, Ragauskas, Labbé and Stewart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Profusion of Molecular Scissors for Pectins: Classification, Expression, and Functions of Plant Polygalacturonases

Yang Yang^{1,2}, Youjian Yu^{1,3}, Ying Liang^{1,2}, Charles T. Anderson^{4,5} and Jiashu Cao^{1,2*}

¹ Laboratory of Cell and Molecular Biology, Institute of Vegetable Science, Zhejiang University, Hangzhou, China, ² Key Laboratory of Horticultural Plant Growth, Development and Quality Improvement, Ministry of Agriculture – Zhejiang Provincial Key Laboratory of Horticultural Plant Integrative Biology, Hangzhou, China, ³ Department of Horticulture, College of Agriculture and Food Science, Zhejiang A & F University, Hangzhou, China, ⁴ Department of Biology, The Pennsylvania State University, University Park, Pennsylvania, PA, United States, ⁵ Center for Lignocellulose Structure and Formation, The Pennsylvania State University, University Park, Pennsylvania, PA, United States

OPEN ACCESS

Edited by:

Dan Szymanski,
Purdue University, United States

Reviewed by:

Olga A. Zabotina,
Iowa State University, United States
Kim Johnson,
La Trobe University, Australia

*Correspondence:

Jiashu Cao
jshcao@zju.edu.cn

Specialty section:

This article was submitted to
Plant Cell Biology,
a section of the journal
Frontiers in Plant Science

Received: 27 February 2018

Accepted: 27 July 2018

Published: 14 August 2018

Citation:

Yang Y, Yu Y, Liang Y, Anderson CT
and Cao J (2018) A Profusion of Molecular Scissors for Pectins: Classification, Expression, and Functions of Plant Polygalacturonases.
Front. Plant Sci. 9:1208.
doi: 10.3389/fpls.2018.01208

In plants, the construction, differentiation, maturation, and degradation of the cell wall are essential for development. Pectins, which are major constituents of primary cell walls in eudicots, function in multiple developmental processes through their synthesis, modification, and degradation. Several pectin modifying enzymes regulate pectin degradation via different modes of action. Polygalacturonases (PGs), which function in the last step of pectin degradation, are a crucial class of pectin-modifying enzymes. Based on differences in their hydrolyzing activities, PGs can be divided into three main types: exo-PGs, endo-PGs, and rhamno-PGs. Their functions were initially investigated based on the expression patterns of PG genes and measurements of total PG activity in organs. In most plant species, PGs are encoded by a large, multigene family. However, due to the lack of genome sequencing data in early studies, the number of identified PG genes was initially limited. Little was initially known about the evolution and expression patterns of PG family members in different species. Furthermore, the functions of PGs in cell dynamics and developmental processes, as well as the regulatory pathways that govern these functions, are far from fully understood. In this review, we focus on how recent studies have begun to fill in these blanks. On the basis of identified PG family members in multiple species, we review their structural characteristics, classification, and molecular evolution in terms of plant phylogenetics. We also highlight the diverse expression patterns and biological functions of PGs during various developmental processes, as well as their mechanisms of action in cell dynamic processes. How PG functions are potentially regulated by hormones, transcription factors, environmental factors, pH and Ca^{2+} is discussed, indicating directions for future research into PG function and regulation.

Keywords: polygalacturonase, pectin modification, cell wall, classification, expression pattern, function

INTRODUCTION

The cell walls of plants are unique extracellular structures, the construction, differentiation, maturation, and degradation of which lay the foundations for tissue differentiation, organ patterning and developmental transitions. Their structural variability in different plant species not only reflects the phylogenetic diversity of plants, but is also associated with the complexity of plant development and plant resistance to biotic and abiotic stresses (Somerville et al., 2004; Cosgrove, 2005; Braybrook and Joensson, 2016). In primary cell walls, 90% of the non-water mass is composed of cellulose, hemicelluloses, and pectins (Albersheim et al., 1996). Pectins, which are acidic polysaccharides that surround cellulose and hemicelluloses in a matrix, control wall porosity, wall hydration, and intercellular adhesion (Daher and Braybrook, 2015; Anderson, 2016). They can be classified into homogalacturonan (HG), rhamnogalacturonan-I, rhamnogalacturonan-II, and xylogalacturonan domains (Sénéchal et al., 2014). As major constituents of primary cell walls, the pollen intine, pollen tube walls, and the middle lamella (Aouali et al., 2001; Dardelle et al., 2010), pectins and their synthesis and degradation influence tissue elongation, pollen development, fruit ripening, and organ abscission (Willats et al., 2001; Sénéchal et al., 2014). The metabolism of pectins in the cell wall is regulated by different classes of pectin-modifying enzymes (Willats et al., 2001; Sénéchal et al., 2014). The polygalacturonases (PGs), which comprise one of these classes, catalyze the hydrolysis of pectins and are involved in numerous developmental processes. Based on differences in hydrolyzing activity, PGs can be divided into three main types: exo-PGs, endo-PGs, and rhamno-PGs. Therefore, illuminating the functions of PGs is of great importance in understanding the dynamics of cell walls during plant growth and strengthening breeding strategies for improving the productivity of crop varieties.

Early studies succeeded in cloning PG genes from species like maize (*Zea mays*) (Allen and Lonsdale, 1993), tobacco (*Nicotiana tabacum*) (Tebbutt et al., 1994), peach (*Prunus persica*) (Lester et al., 1996), and tomato (*Solanum lycopersicum*) (Kalaitzis et al., 1997). Total PG activity was measured in various organs of these species, indicating the functions of PGs in fruit ripening, pollen development, and organ abscission, in line with the expression patterns of these genes. Hadfield and Bennett (1998) reviewed this initial progress and proposed the first classification system for PGs (Hadfield et al., 1998). However, without genome data, little was known about the composition, classification, and evolution of the PG family. Although the sequences of some PG genes were reported, only a few of these were characterized in terms of their expression patterns and functions. At the turn of the 21st century, it remained unclear how pectin hydrolysis by PGs alters cell wall structure, modulates cell growth, and regulates organ growth.

In the last 18 years, the expression patterns and putative functions of hundreds of PG genes have been delineated. Some regulatory factors for PGs at both the transcriptional and post-translational levels have also been reported. Innovative research methods and a flood of genome sequencing data have both contributed to the systematic identification and phylogenetic

analysis of PG families in numerous plant species. This review highlights the latest advances in the structural analysis and classification of PG families, the molecular evolution of PG genes in the context of plant evolution, and the expression, functions, and regulators of PGs during different developmental processes in plants. It also explores the basis for a better understanding of how cell wall dynamics influence cell and plant growth.

STRUCTURAL CHARACTERISTICS AND CLASSIFICATION OF POLYGALACTURONASES

Structural Characteristics of Polygalacturonases

Polygalacturonases belong to glycoside hydrolase family 28 and contain at least one GH28 (Pfam00295) domain (Markovič and Janeček, 2001; Kim et al., 2006). This domain is replaced by a Pectate Lyase 3 domain (Pfam12708) in *Arabidopsis* (*Arabidopsis thaliana*) QRT3 and its homologs in other species (Rhee et al., 2003; Yu et al., 2014). The encoded protein of *Arabidopsis* QRT3 exhibits PG activity, as demonstrated by heterologous expression in *Saccharomyces cerevisiae* (Rhee et al., 2003). Most PG genes (91.2% of PG genes in *Arabidopsis* and 87.9% of PG genes in Chinese cabbage (*Brassica campestris*, syn. *Brassica rapa*) are predicted to encode a signal peptide upstream of the GH28 domain. Since signal peptides typically function in guiding proteins through secretory pathways that end with exocytosis (Babu et al., 2013), their presence implies that most PGs should be located in the apoplast. This hypothesis has been supported in recent studies of PG localization (Irshad et al., 2008; Xiao et al., 2014, 2017; Rui et al., 2017). In addition to the apoplast, one PG with a signal peptide has been localized in Golgi bodies and vesicles *in vitro*, presumably as part of its secretory journey (Nakashima et al., 2004).

In plant and fungal PGs, there are four commonly conserved, functional domain motifs known as SPNTDG (motif I), GDDC (motif II), CGPGHGISIGSLG (motif III), and RIK (motif IV), although domain III shows lower conservation and is missing in rhamno-PGs (Park et al., 2008). Typically, a PG gene encodes at least one of these domains (Torki et al., 2000). Within these four motifs, NTD, DD, GHG, and RIK amino acid segments are conserved active-site residues in plant, fungal, bacterial, and insect PGs, and have been demonstrated to be essential for fungal PG activity (Bussink et al., 1991; Markovič and Janeček, 2001). Within motifs I and II, the aspartic acid (D) residues in NTD and DD are components of the catalytic site (Rexovabenkova, 1990). In motif III, the histidine (H) residue acts in catalysis (Rao et al., 1996; Markovič and Janeček, 2001). Finally, the positively charged motif IV is thought to form interactions with the carboxylate groups of pectate substrates (Bussink et al., 1991).

Crystal structures of PGs can reveal the mechanisms by which they select substrates or act in catalysis. Although crystal structures of plant PGs are not yet published, we can infer their modes of action by inspecting the reported structures of fungal and bacterial enzymes due to the high conservation of

functional domains between plant, fungal, and bacterial PGs. A typical PG displays 10 complete turns of β -structure, which is formed by four parallel sheets extending along the longitudinal axis. This structure can be used to distinguish PGs from polysaccharide lyases, which have three-sheet topology (Kluskens et al., 2005; Abbott and Boraston, 2007). Endo-PGs and exo-PGs show significant differences in crystal structures, indicating their different substrates and modes of action (**Figure 1A**). The active site of a typical endo-PG is a tunnel-like substrate-binding cleft lying between two loop regions. Hence, endo-PGs can potentially bind polysaccharides in either direction and produce oligosaccharides with varying degrees of polymerization (Jenkins et al., 1998). The adjacent loop regions help identify substrates and guide them to the active site (Pickersgill et al., 1998). For example, in *Achaetomium* sp. Xz8 endo-PG, the Asn94 residue of the T3 loop binds to substrates in the active site cleft by forming a hydrogen bond, which ensures correct positioning of substrates (Tu et al., 2015). In contrast, exo-PGs have a closed-pocket active site that only binds to the non-reducing ends of pectins due to inserted stretches of amino acids (Abbott and Boraston, 2007). Rhamno-PGs (RGs), which hydrolyze GalA-rhamnose bonds of rhamnogalacturonan-I, can be further divided into exo-RGs and endo-RGs (Mutter et al., 1998; Choi et al., 2004; Damak et al., 2015). However, tertiary structures of exo-RGs remain unreported. A predicted structure of an endo-RG has only been modeled in *Aspergillus aculeatus* (Choi et al., 2004). Compared with endo-PGs, endo-RGs are also predicted to have a tunnel-like active-site with two open ends. Differences in loop structure are likely to provide endo-RGs with more space in the active site to bind more complex substrates. The most significant difference between the structures of endo-PGs and endo-RGs is that endo-RGs have long tails of 19 and 45 residues at the N-terminus and C-terminus, respectively, whereas endo-PGs lack these tails (Choi et al., 2004). As a result of their unique structures, exo-PGs can only remove galacturonic acid residues from the non-reducing ends of HG chains; endo-PGs catalyze the random hydrolytic cleavage of α -1,4 glycosidic bonds in HG chains; and rhamno-PGs catalyze the hydrolytic cleavage of α -1,2 glycosidic bonds randomly within or from the non-reducing ends of rhamnogalacturonan-I main chains (**Figure 1A**) (Markovič and Janeček, 2001; Park et al., 2010).

Classification and Molecular Evolution of Polygalacturonases

As mentioned above, PGs can be divided into exo-PGs, endo-PGs and rhamno-PGs, based on their modes of hydrolysis. PGs of these different types emerged at different times during plant evolution. Rhamno-PGs, which are regarded as the earliest type, appear in both algae and land plants, and endo-PGs exist across land plants, whereas exo-PGs only appear in angiosperms (Park et al., 2010). The PG family in land plants had five common ancestral genes rather than a single one (McCarthy et al., 2014), which might be explained by this early divergence of rhamno-PGs and endo-PGs.

Using bioinformatics, PG genes can be grouped by their phylogenetic relationships. Two main classification systems have

been proposed (**Figure 1B**) for analyzing these relationships according to amino acid sequence. The first system was put forward by Hadfield et al. (1998), who grouped three PG genes from melon (*Cucumis melo*) and 17 homologous genes from other plants and fungi into Clades A–C. In agreement with Hadfield's system, Torki et al. (2000), Markovič and Janeček (2001), and Park et al. (2008) grouped PG genes into five clades: Clades A–E. Clades A, B, and C were found to have invariant conserved residues (Gly264 and Phe294 in Clade A, Asn104 in Clade B, and Lys176 in Clade C), revealing distinct structural characteristics in different clades. However, Clades D and E lack exclusively invariant residues (Markovič and Janeček, 2001). With the accumulation of genomic data for numerous species, PG classification entered into a new phase in which cluster analysis of all PG gene family members for a species could be performed. Park et al. (2010) divided 225 genes into six clades, which contain eight PG gene subfamilies from algae to angiosperms. Later, Liang et al. (2015) classified PG genes from five species ranging from algae to eudicots into seven clades, as supported by the classification of 557 PG genes from five grass and five eudicot species (Liang et al., 2016). Arabidopsis *QRT3* and its homologous genes in core angiosperms are grouped into the new clade (Clade G). The second classification system was proposed by Kim et al. (2006) and divided 125 PG genes from Arabidopsis and rice (*Oryza sativa*) into three classes, A–C. The subsequent studies that used this system also grouped 75 PG genes from poplar (*Populus trichocarpa*) and 100 PG genes from Chinese cabbage into three classes (Yang et al., 2013; Duan et al., 2016).

Each classification system has advantages and shortcomings for analyzing molecular evolution. The first system is more suitable for analyzing the emergence of PG genes over time and the compositional changes in PG families. PG genes in different clades emerged at different times (**Figure 2**). Based on this system, PG genes in Clade E exist from algae to angiosperms, those in Clades A and B appear in land plants, and the genes in Clades C, D, F, and G only appear in flowering plants (Park et al., 2010). Clade emergence shows a pattern that is consistent with species evolution (**Figure 2**). The dominant clade in non-vascular plants is Clade E, which diversifies into Clades B, D, and E in monocots. The proportions of Clades B and E decrease in eudicots [except in soybean (*Glycine max*)], and Clades C, D and F are instead predominant. Further analysis shows that Clade D is the principal clade in Cruciferae, Cucurbitaceae, and Solanaceae, whereas the most-represented clades in poplar and soybean are Clades C and E, respectively. The second system also reveals different appearance times for different PG classes. However, PG genes within the same class also emerged at different times in this system. For example, the two major subgroups of Class A, A1, and A2, exist in angiosperms and non-vascular plants, respectively (Duan et al., 2016), and the proportion of Class C is lower than Classes A and B in species from moss (*Physcomitrella patens*) to vascular plants. The representation of Class A outpaces Class B in moss and vascular plants, but the converse is true in lycophytes (*Selaginella moellendorffii*). The various expansion rates of PG classes were likely driven by different selection pressures on those classes, accounting for their different proportionalities across taxa (Kim et al., 2006; Yang et al., 2013; Wang et al., 2016).

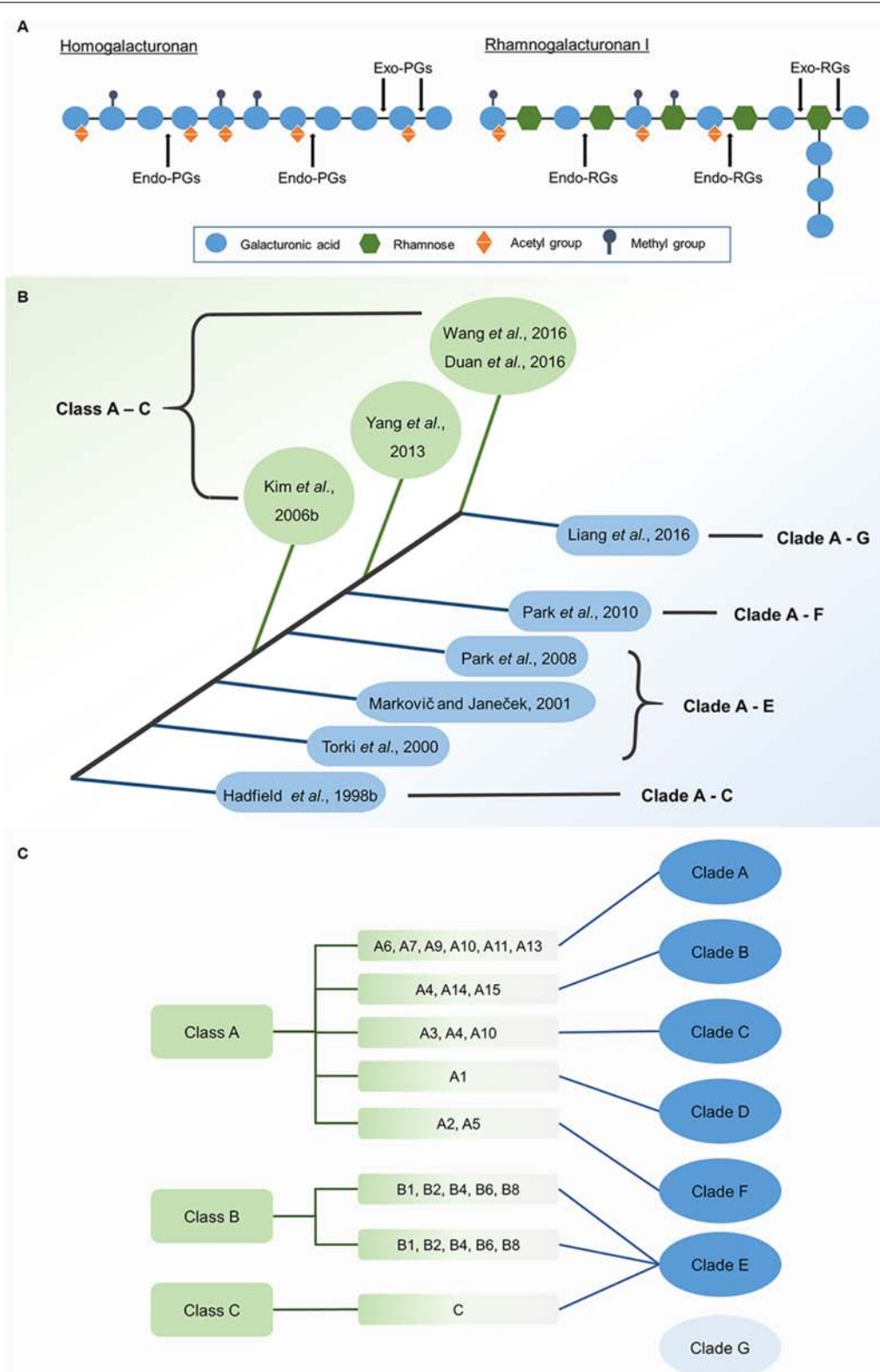
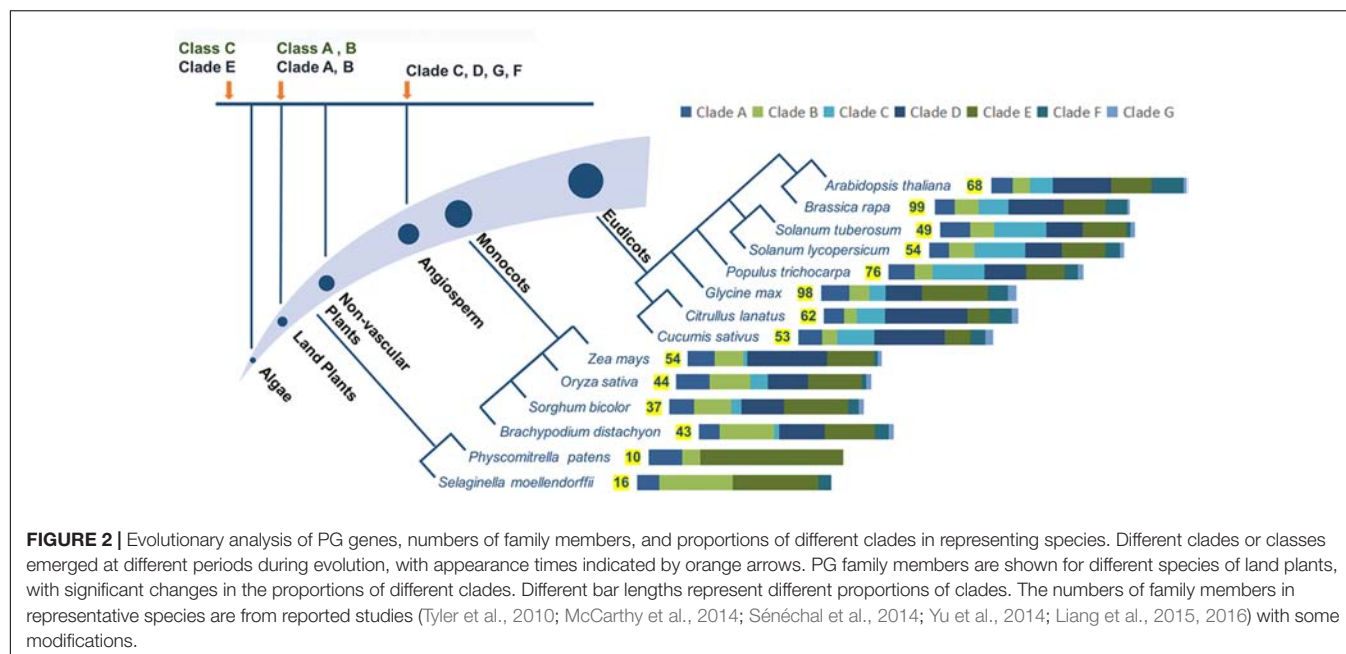


FIGURE 1 | Modes of action of PGs of different types **(A)** and classification of PG genes **(B)**. **(A)** Different types of PGs differ in the selection of substrates and the action site of pectins main chain. **(B)** Based on the system first proposed by Kim, PG genes were all grouped into three classes (with green background). Based on the system first proposed by Hadfield, PG genes were divided into three to seven clades in different studies (with blue background). **(C)** A comparison of Kim's system (with green background) and Liang's system (with blue background) in the classification of Arabidopsis PG gene family.



The first classification system has the advantage of analyzing large numbers of PG genes, which contributes to a comprehensive understanding of the phylogenetic relationships between genes across multiple species. Neighbor-joining and maximum parsimony algorithms were mainly used in this system, enabling concise and fast calculations to analyze large quantities of sequence data. However, the increased amount of data might lead to lower confidence in the existence of a new clade and reduce statistical accuracy. Based on the first system, the number of PG family members shows an increasing trend from non-vascular plants to monocots to eudicots. Gene duplication and loss are both likely to be contributors to these differences in gene number. For example, more old and recent duplications exist in Clade B in grasses, likely explaining the larger numbers of genes in this family in grasses than in eudicots (Liang et al., 2016). The difference in wall composition between grasses and eudicots, wherein pectins are less abundant in grasses than in eudicots, is another likely explanation for the different scales of PG families (McCarthy et al., 2014; Liang et al., 2016). In comparison, the second system built high-value phylogenetic trees, mainly by applying maximum likelihood algorithms. Although 212 PG genes from five species were successfully classified in the second system (Yang et al., 2013), these time-consuming calculations are inadequate for analyzing complex phylogenetic relationships among PG genes across a large number of species.

Instead, the second system proposed by Kim has the potential to reveal orthologous relationships and ancestral PGs between two species. In total, 21, 39, and 54 common ancestral PG genes were found between *Arabidopsis* and rice, poplar, and Chinese cabbage, respectively (Kim et al., 2006; Yang et al., 2013; Duan et al., 2016). The difference in ancestral gene number can be explained by the phylogenetic distance between these species. Genes clustered in the same group are orthologs. Therefore, the possible biological functions of candidate PG genes from a given

species can be inferred by referring to the function of orthologous genes from other species, if this information is available.

EXPRESSION, FUNCTIONS, AND MECHANISMS OF ACTION OF POLYGALACTURONASES

Expression Patterns of Polygalacturonase Genes

Numerous PG genes have been identified in various species. Here, we summarize PG genes with defined accession numbers and expression patterns as analyzed by qRT-PCR and/or promoter activity (Table 1 and Supplementary Table 1). According to their expression patterns, these genes are divided into five groups: ubiquitously expressed in multiple organs, specifically expressed in flowers and pollen, expressed during fruit ripening, expressed at sites of organ abscission and dehiscence, and expressed in other organs. Expression patterns among different clades/classes and PG genes with close phylogenetic relationships are discussed below.

Polygalacturonase Genes Among Different Clades Are Conserved and Diversified in Expression

Previous studies revealed that the expression patterns of PG genes are conserved in the same clade, but diversified between clades. PG genes in Clades A and B are expressed in fruits and abscission and dehiscence zones, and genes in Clade C are expressed in flowers (Hadfield et al., 1998). This pattern is supported by the expression of most identified genes. PG genes such as *PS-2*, *ADPG1*, *ADPG2*, *RDPG1*, and *SDPG* (Table 1 and Supplementary Table 1D) are expressed during organ abscission and dehiscence, and belong to Clade B with the conserved motif

TABLE 1 | Expression and functions of identified PG genes.

Gene	Accession number	Species	Expression	Gene function	PG activity	Reference(s)
<i>NIMNA</i>	At2g33160	<i>Arabidopsis thaliana</i>	Reproductive organ (flowers, pollen, and siliques)	Suspensor elongation	Not tested	Babu et al., 2013
<i>PGX1</i>	At3g26610	<i>Arabidopsis thaliana</i>	Seedlings, roots, leaves, and flowers	Flower organ patterning, hypocotyl elongation, leaf expansion	✓	Xiao et al., 2014
<i>PGX2</i>	At1g78400	<i>Arabidopsis thaliana</i>	Seedlings and in roots, stems, leaves, flowers, and siliques	Hypocotyl elongation, leaf expansion, stem lignification, mechanical stiffening, and lodging	✓	Xiao et al., 2017
<i>PGX3</i>	At1g48100	<i>Arabidopsis thaliana</i>	Hypocotyls, roots, and at lateral root initiation sites, cotyledons, rosette leaves, flowers, siliques, seed funiculi, and stomata	Root and hypocotyl elongation, rosette expansion, the shape and the stomatal size of cotyledons, the stomatal dynamic in true leaves	✓	Rui et al., 2017
<i>QRT3</i>	At4g20050	<i>Arabidopsis thaliana</i>	Tapeta in tetrad stage	Degrading pollen mother cell wall to allow microspore separation	✓	Rhee et al., 2003
<i>BcMF2</i>	ABW24665	<i>Brassica campestris</i>	Tapeta and pollen after tetrad stage	Pollen wall intine layer development	Not tested	Huang et al., 2009a
<i>LLP-PG</i>	AY834213	<i>Lilium longiflorum</i>	The late stage of pollen maturation	Degrading the walls of stylar cells to allow penetration of the pollen tube	✓	Chiang et al., 2006
<i>BcMF6</i>	ACP74159.1	<i>Brassica campestris</i>	Flower buds from stage III to stage V, open flowers and green silique	Pollen maturation and pollen tube growth	Not tested	Zhang et al., 2008
<i>GmPG031</i>	AP011817	<i>Glycine max</i>	Flowers, pollen grains, and pollen tube	Flowers and silique development	Not tested	Wang et al., 2016
<i>MdPG1</i>	L27743	<i>Malus domestica</i>	Ripening fruit from ripening early stage	Fruit ripening, reduce cellular adhesion and influence fruit texture characters; leaf morphology, stomatal structure and function, and leaf attachment	✓	Atkinson et al., 2002, 2012; Wakasa et al., 2006; Costa et al., 2010; Li et al., 2010
<i>FaPG1</i>	DQ458990	<i>Fragaria ananassa</i>	Ripening fruit of firmer cultivar	Fruit ripening	✓	Villarreal et al., 2008, 2009; Pose et al., 2013
<i>PG</i>	Ca10g18920	<i>Capsicum annuum</i>	Softening fruit	Fruit softening	Not tested	Kim S. et al., 2014
<i>PGAZAT (ADPG2)</i>	AT2G41850	<i>Arabidopsis thaliana</i>	Abscission zone of the sepals, petals, and stamens of flowers; anther during microspore separation; dehiscence zone of maturing siliques	Flower tissues abscission; anther and silique dehiscence; pollen separation	✓	Ogawa et al., 2009; González-Carranza et al., 2012
<i>ADPG1</i>	AT3G57510	<i>Arabidopsis thaliana</i>	Dehiscence zone of silique; seed abscission zones; anther before anthesis	Anther and silique dehiscence; pollen separation	✓	Ogawa et al., 2009
<i>QRT2</i>	At3g07970	<i>Arabidopsis thaliana</i>	Abscission zones of sepal, petal, and stamen in flowers	Flower organs abscission; anther dehiscence; pollen separation	✓	Ogawa et al., 2009
<i>PS-2</i>	EU111748	<i>Solanum lycopersicum</i>	Dehiscence zone of anther	Anther dehiscence	Not tested	Gorguet et al., 2009

FGAKGDG (Rodriguez-Llorente et al., 2004); PG genes such as *PcPG1*, *VvPG1*, *VvPG2* and *MAPG3*, which are expressed during fruit ripening, belong to Clade B, and *sPG* belongs to Clade A (**Supplementary Table 1C**). PG genes such as *BcMF2*, *BcMF6*, *BcMF9*, *BcMF16*, *BcMF24*, and *PGA4* (**Table 1** and **Supplementary Table 1B**), are expressed in flowers and pollen

and belong to Clade C. However, *FaPG1*, which belongs to Clade C and has a cysteine residue conserved in pollen-specific genes, is expressed in the middle-late stage of fruit ripening and is involved in flesh softening (Villarreal et al., 2008).

Genes from the same PG clade can also sometimes show distinct expression patterns. Hadfield's hypothesis that

orthologous genes are expressed similarly has proven to be partly incorrect. Expression patterns of PG genes are conserved in Clades D and E, with expression in inflorescences and ubiquitous expression in multiple organs, respectively (Yu et al., 2014; Liang et al., 2015, 2016). These results illustrate that these PG genes play crucial roles during multiple developmental processes. Members of Clades C and F, which are mainly expressed in reproductive organs in grasses and eudicots, are also sometimes expressed during root and seed development. A few eudicot PG genes of Clade C show expression in roots and root nodules. In some members of Clade F, expression can be detected in embryos and developing seeds as well. However, expression patterns are diversified in other clades. Members of Clades A, B, and G are either expressed specifically in different organs or widely across various organs. Some Clade G members, for example Arabidopsis *QRT3* and its homologs, are expressed either in one specific organ (such as seeds, stems, flowers, or siliques) or in multiple organs (Liang et al., 2016). Cucumber (*Cucumis sativus*) PG genes in Clade A are related to fruit development. However, these PG genes are expressed ubiquitously instead of showing fruit-specific expression (Yu et al., 2014).

Based on Kim's system, Classes B and C have particular expression patterns in moss, lycophytes, rice, Arabidopsis, and poplar. However, members of Class A, which contains PG genes from lycophytes to angiosperms, are expressed with variable patterns (Kim et al., 2006; Yang et al., 2013). The requirement of angiosperm PG genes to function in the flower, which is the unique organ of angiosperms, might partly explain the diverse expression patterns of Class A genes. This diversity also likely relies on lower selective pressure compared with other classes (Yang et al., 2013). Specific expression patterns for genes within Class A, the largest class, can be found in its subgroups (Kim et al., 2006; Duan et al., 2016). To explore the relationship between these two systems of PG classification in a model species, we analyzed the Arabidopsis PG family using Kim's system (Kim et al., 2006) and Liang's system (Liang et al., 2015). Interestingly, almost all PG genes within the same subgroup of Kim's system are grouped into a specific clade of Liang's system, and different classes are composed of certain clade(s) (**Figure 1C**). This relationship can explain the specific expression of Classes B and C, which only contain PG genes from the deeply conserved Clade E (**Figure 2**). The complex composition of Class A accounts for its various expression patterns. Hence, the classification system of Hadfield has some advantages in connecting expression patterns to possible functions.

It is noteworthy that the conservation of expression patterns for PG genes is relative. PG genes expressed in the same general organs can still differ in specific expression patterns when they are studied in more detail. For example, *PcPG1* and *PcPG3* are both expressed in pear (*Pyrus communis*) fruit, but are expressed at different stages of fruit ripening and function separately in affecting fruit ripening and flesh texture, respectively (Sekine et al., 2006). Promoter activity analysis of *ADPG2* and its three closest homologs reveal that three of these genes are specifically expressed in different organs, and that another one had no detectable expression, even though *ADPG2* and its three closest homologs are all grouped into Clade

B (González-Carranza et al., 2007). Expression specificity is also supported by the expression patterns of PG genes in other species. For example, 16 fruit-specific PG genes in cucumber can be divided into three subgroups, with highest expression occurring for each subgroup at 3 days after pollination, 6–9 days after pollination, or 27 days after pollination, respectively (Yu et al., 2014). In Arabidopsis, 47 abscission-related PG genes have nine significantly different expression patterns during five stages of floral organ abscission (Kim and Patterson, 2006). The diversity of expression patterns in the same organ implies that these genes play different roles in the development of a single organ, but this idea requires further research to be validated.

Duplicated Polygalacturonase Genes Show Divergent Expression

Duplicated PG genes, which are highly similar in sequence, are more likely to have different expression patterns. These gene pairs result from duplication events such as whole genome duplication (WGD) and tandem duplication (TD) during plant evolution. Two tandemly duplicated PG genes, *PpendoPGM* and *PpendoPGF*, promote the softening of peach flesh, but only the latter is involved in stone adhesion (Gu et al., 2016). *MaPG1* and *MaPG2*, of which the cDNA sequences are 98% similar, are expressed in different organs: the former is expressed in roots, stem, leaves, and flowers, whereas the latter is expressed in the late stage of fruit ripening (Mbéguié-A-Mbéguié et al., 2009). PG paralog pairs generated by TD are more likely to be expressed differently than pairs generated by WGD. For example, based on microarray and RNA-seq-based data, paralogous PG pairs with different expression patterns constitute 71.4%, 75%, and 75% of pairs resulting from TD in rice, poplar, and cucumber (Kim et al., 2006; Yang et al., 2013; Yu et al., 2014), whereas PG copies with diversified expression patterns cover 56% and 67.8% of PG copies caused by WGD in poplar and soybean, respectively (Yang et al., 2013; Wang et al., 2016). Expression diversity of duplicated genes is also demonstrated by variable expression between paralogs in Arabidopsis and Chinese cabbage, which underwent WGD twice and tree times, respectively (Liang et al., 2015). The difference in expression of duplicated PG copies highlights the diversity of their functions and exemplifies the duplicate retention mechanisms known as neofunctionalization and subfunctionalization (Innan and Kondrashov, 2010).

Functions of Polygalacturonases in Pectin Degradation, Cell Dynamics, and Plant Development

Polygalacturonases and Pectin Degradation

Homogalacturonan, which is the most abundant pectin domain, has been extensively studied in terms of its synthesis, modification, and degradation. HG is synthesized in the Golgi apparatus, and its polysaccharide chain is extended by galacturonic acid residues being added to the non-reducing end. Under the action of pectin methyltransferases (PMTs) and acetyltransferases (PATs), HG can be modified with methyl-ester or acetyl groups. Newly synthesized HG polymers

are apportioned into vesicles and transported to the plasma membrane, and finally are secreted into the apoplast for incorporation into the cell wall (Ridley et al., 2001; Sénéchal et al., 2014).

In the wall, HG can be metabolized by the action of HG modifying-enzymes (HGMEs), and participates in the regulation of plant development and responses to external stimuli. PGs, pectin methylesterases (PMEs), pectin acetylsterases (PAEs), and pectin/pectate lyase-like proteins (PLLs, including pectin lyases and pectate lyases) all are HGMEs. Their coding genes have distinct expression patterns (Sénéchal et al., 2014), and they modify pectin in different ways. PMEs and PAEs remove modifications from the HG backbone, whereas PGs and PLLs cleave or depolymerize HG. More specifically, PMEs control the degree of methylesterification of HG by removing methyl-ester groups, resulting in negatively charged galacturonic acid residues (Jolie et al., 2010). PAEs hydrolyze *O*-acetyl groups and produce linear HG (Gou et al., 2012). PGs hydrolyze the α -1,4 glycosidic bonds of demethylesterified HG, releasing oligogalacturonides (OGs) or galacturonic acid monomers (Marković and Janeček, 2001). In contrast to PGs, pectin lyases and pectate lyases can cleave the α -1,4 glycosidic bonds of methylesterified and demethylesterified HG, respectively, by β -elimination (Mayans et al., 1997; Herron et al., 2000).

Models of HG remodeling have been proposed that take distinct cell wall microenvironments into consideration (Sénéchal et al., 2014; Hocq et al., 2017). In an acidic cell wall context, HG that is randomly demethylesterified by acidic PMEs can be a substrate for PGs. Basic PMEs, which can be easily trapped by free carboxyl groups in this circumstance, would inefficiently remove methyl-ester groups from HG. In a slightly alkaline context, H^+ diffuse into the cytoplasm by producing IAAH, due to the absence of auxin (IAA^-). As a result, the cell wall can be high in the concentrations of ionic cations. Under this condition, basic PMEs play the main role in HG de-methylesterification, whereas acidic PMEs will be charged and unable to bind HG. Continuous stretches of demethylesterified HG can crosslink via Ca^{2+} to form stable “egg-box” structures, which restrict cell wall loosening (Figure 3). Other regions of HG can be hydrolyzed by PGs, releasing OGs. HG has been proposed to go through a cycle of crosslinking, modification, cell wall loosening, and cell growth (Sénéchal et al., 2014; Boyer, 2016; Hocq et al., 2017).

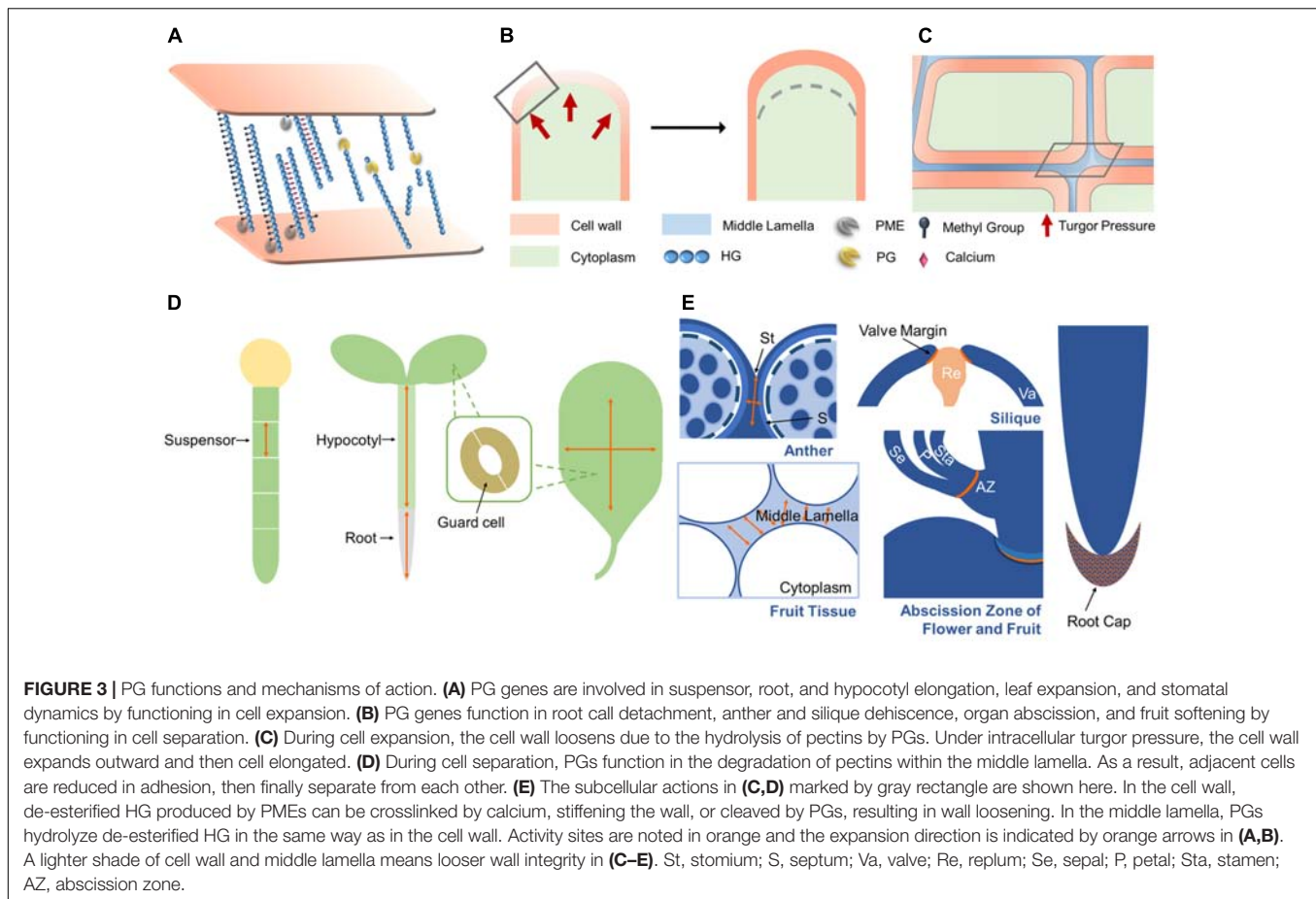
Polygalacturonase Mechanisms of Action in Cell Dynamics

Cell proliferation, expansion, and separation form the foundation of plant morphogenesis and development. During cell expansion and separation, pectin is degraded in the primary cell wall and middle lamella, decreasing cell wall stiffness and increasing wall fluidity. The identified roles of PGs in plant development reveal that they can function in both cell separation and expansion. Here, we summarize the possible mechanisms of action of PGs in cell dynamics.

Separation between cells in the abscission zone (AZ) results in organ abscission and dehiscence. After receiving the abscission signal in the AZ, the cell wall and middle lamella begin

to detach via the action of cell wall degrading enzymes (Figures 3A,C). Cells separate along the middle lamella, allowing for organ separation and dehiscence (Kim et al., 2015). After organ detachment, cells in AZ can elongate, implying that cell elongation is also involved in organ abscission (Patterson and Bleecker, 2004). AZ-expressed PG genes function in root cap detachment, anther and fruit dehiscence, leaf shedding, and fruit maturation, highlighting their central roles in cell separation (Figure 3E). Crosslinks between adjacent cells are formed under the control of Ca^{2+} , and cell adhesion is maintained by HG methylesterification (Daher and Braybrook, 2015). During cell separation, PMEs remove the methyl-ester groups from HG to provide substrates for PGs. PGs can then cleave HG backbones and reduce the HG-mediated cell adhesion that is maintained by Ca^{2+} . Finally, cells detach from each other (Figures 3A,C) (Daher and Braybrook, 2015). This model is supported by the possible role of MdPG1 during leaf shedding. In MdPG1 overexpression lines, more abundant low-esterified pectins are found in AZs with increased PG activity in leaves (Atkinson et al., 2002). Noticeably, tetrad separation is a special case of tissue separation. Instead of cell separation, the release of microspores results from the degradation of the primary pollen mother wall. PGs like QRT3, which are secreted from tapetum cells during early microspore stage, are responsible for micropore separation by degrading this cell wall (Francis et al., 2006).

Cell expansion is required by numerous developmental processes such as suspensor elongation, hypocotyl elongation, stomatal opening and closure, and leaf expansion (Figure 3D). Therefore, PG genes like *POLYGALACTURONASE INVOLVED IN EXPANSION* (PGX) genes, which affect the elongation and expansion of several organs, might have their mechanism of action explained as facilitating cell expansion (Xiao et al., 2014, 2017; Rui et al., 2017). Lower demethylesterified HG level, higher PG activity and smaller HG size in PGX3 overexpression lines provide further evidence for the functions of PGs in regulating stomata dynamics and leaf expansion (Rui et al., 2017). The aforementioned fact that cells in the AZ elongate after organ shedding (Patterson and Bleecker, 2004) suggests that abscission-related PGs might influence both cell separation and cell expansion. Cell expansion can be regulated by intracellular turgor pressure and mechanical interactions between cell wall components. The direction of cell expansion is influenced by both the orientation of cellulose microfibrils (Cosgrove, 2005; Palin and Geitmann, 2012) and the modification of pectins (Peaucelle et al., 2015). Turgor pressure, in combination with increased cell wall extensibility, is thought to be the driving force for cell expansion (Figure 3B). Pectins affect cell expansion by modulating the mechanical properties of cell walls (Peaucelle et al., 2011; Kozioł et al., 2017). By effecting pectin degradation, PGs are hypothesized to cause a decrease in cell wall stiffness (Figures 3A,B), which is supported by the identified role of PGX2 in regulating the mechanical properties of stems (Xiao et al., 2017). PGs might also function in cell expansion by modulating the directionality of expansion. A mutant *Arabidopsis* line resistant to the drug cobtorin can suppress the cell-swelling phenotype induced by cobtorin, due to the overexpression of a PG gene in this line. Exogenously supplied PGs can also



rescue this phenotype in *Arabidopsis*, restoring the deposition of cellulose microfibrils in parallel with cortical microtubules with the help of PME (Yoneda et al., 2010). Exactly how PGs might influence expansion direction requires further study.

It is noteworthy that PGs might participate in cell and tissue patterning. PME5, which works upstream of PGs, plays a key role in controlling phyllotactic patterning (Peaucelle et al., 2008). Changes in the degree of methylesterification of pectins can enhance cell wall loosening, which underlies organ initiation (Peaucelle et al., 2011). One PME inhibitor protein, PME13, can suppress morphogenesis in inflorescence meristems (Peaucelle et al., 2008). The deformed pollen grains that result from abnormal intine formation in *BcMF2*, *BcMF9* antisense lines indicate that PGs might also function in determining cell morphology during pollen development. However, whether PGs act in organ patterning through regulating cell wall mechanics directly or by changing wall integrity signaling remains unknown.

Biological Functions of Polygalacturonase Genes in Plant Development and Fruit Ripening

The functions of PG genes, which have been identified by mutants or transgenic lines (Table 1) or have in some cases been inferred from their expression patterns (Supplementary Table 1), reveal their significant roles in both vegetative and reproductive development in plants. Ubiquitously expressed

PG genes have been identified as functioning in multiple processes during development, which might be important for plant survival. For instance, the expression of *Arabidopsis* *POLYGALACTURONASE INVOLVED IN EXPANSION1* (*PGX1*) is detected in seedlings, roots, leaves, and flowers (Xiao et al., 2014). Overexpression of this PG leads to significant elongation of etiolated hypocotyls and enhanced rosette leaf expansion. Its functions in cell and tissue expansion are supported by the opposite phenotypes in a *pgx1* knockout mutant. In flower development, abnormal angles between flower primordia and extra petals are found in both overexpression lines and mutants, which indicates that this PG might function in floral organ patterning (Xiao et al., 2014). Overexpressing *Arabidopsis* *POLYGALACTURONASE INVOLVED IN EXPANSION2* (*PGX2*) leads to faster flowering and enhanced stem lignification with enhanced tensile stiffness but lower bending stiffness, as well as enhanced etiolated hypocotyl length and rosette leaf area (Xiao et al., 2017). *PGX2* represents the first PG gene that has been found to regulate stem lignification. *Arabidopsis* *POLYGALACTURONASE INVOLVED IN EXPANSION3* (*PGX3*) also functions in rosette expansion and root elongation (Rui et al., 2017). Overexpressing this PG gene also causes larger stomatal pores in cotyledons, whereas its knockout mutant has the opposite phenotype. Without influencing stomatal dimensions in true leaves, *PGX3* asymmetrically affects mature stomatal

opening and closure (Rui et al., 2017). Interestingly, despite their common functions in rosette leaf expansion, *PGX1*, *PGX2*, and *PGX3* are not closely related in the phylogeny of Arabidopsis PGs (McCarthy et al., 2014).

How PG genes influence pollen development and fertilization has been a recent research priority. The expression of PG genes in tapeta, pollen grains, stigmas and pollinated pistils implies their role in tapetum degradation, pollen maturation, pollen tube growth, and pollination, as evinced by their functional characterization (**Table 1**). Suppressing the expression of *QRT3* in Arabidopsis and *BcMF6* in Chinese cabbage interferes with microspore separation after the tetrad stage (Rhee et al., 2003; Zhang et al., 2008). In addition, the latter mutant has smaller floral organs and a lower pollen germination rate caused by the disruption of microspore maturation (Zhang et al., 2008). RNA antisense lines of Chinese cabbage *BcMF2* and *BcMF9* show disturbed development of the pollen wall intine layer and of the pollen tube wall that is the continuation of the intine layer (Huang et al., 2009a,b). The downregulation of *BcMF2* causes pollen deformity and balloon-like swelling in the pollen tube tip, along with premature tapetum formation (Huang et al., 2009a). Downregulating *BcMF9* influences pollen wall exine layer formation as well (Huang et al., 2009b). When a soybean PG is heterologously overexpressed in Arabidopsis, inflorescence mortality is over 50%, and siliques and seeds significantly decrease in number (Wang et al., 2016). *NIMNA*, which plays a role in early embryo cells and suspensor elongation, is the only Arabidopsis PG gene that has been identified as functioning in embryo development (Babu et al., 2013).

Early studies examined the relationship between PGs and fruit development and mainly focused on the activity of PG proteins and their responses to exogenous stimuli. In recent studies, fruit-related PG gene functions have gradually attracted more attention, especially in species with edible fruit. The suppression of *FaPG1* leads to significant increase in strawberry (*Fragaria ananassa*) fruit firmness (Pose et al., 2013). A similar change is also found in *PG1* suppression lines of apple (*Malus × domestica*), with an enhancement of cell adhesion (Atkinson et al., 2012). Despite *MdPG1* showing fruit-specific expression, its overexpression lines display various novel phenotypes in vegetative growth, such as the silvery colored leaves, earlier leaf shedding, and abnormal stomatal structure (Atkinson et al., 2002). Its function in organ abscission is also supported by heterologous expression studies in Arabidopsis. Overexpressing *MdPG1* in Arabidopsis can drive early silique dehiscence, whereas the siliques of *MdPG1* antisense lines do not open normally, possibly due to the downregulation of endogenous PGs (Li et al., 2013). In non-softening fruit species like pepper (*Capsicum annuum*), PG genes perform crucial functions as well. A point mutation in one pepper PG gene (CA10g18920) in wild type downregulates its expression by generating a premature stop code in its 3' splicing site. This results in lower water-soluble pectin levels compared with a *soft flesh* mutant, preventing aberrant fruit softening (Kim S. et al., 2014).

Abscission and dehiscence help to remove old, impaired and infected tissues as well as release leaves, floral organs, pollen or

seeds, and are of great value in maintaining normal growth and reproduction. PG genes have been found to play a pivotal role in this process (**Supplementary Table 1D**). Mutation of tomato *PS-2* in intron recognition splice sites leads to non-dehiscent anthers and functional sterility (Gorguet et al., 2009). Similarly, an Arabidopsis *adpg1 adpg2 qrt2* triple mutant is delayed in anther dehiscence (Ogawa et al., 2009). This phenotype is not found in double mutants or single mutants, indicating that anther dehiscence is co-regulated by these three PG genes. However, non-dehiscent siliques are found in single mutants of *adpg1*, *adpg2* and more prevalently in an *adpg1 adpg2* double mutant. *ADPG1* and *ADPG2* potentially function in silique dehiscence by hydrolyzing only a small amount of pectin (Ogawa et al., 2009).

Polygalacturonase genes also participate in biotic and abiotic stress responses in plants. These findings have been nicely summarized by Sénéchal et al. (2014). When treating strawberry fruit with resistance inducers like chitosan, benzothiadiazole or a mixture of calcium and organic acids, the expression of PG genes changes significantly (Landi et al., 2014). In the same study, when stimulated by abiotic stresses, the expression of PG genes is up-regulated, which might elicit responses to environmental stresses (Liu et al., 2014).

REGULATORY NETWORKS OF POLYGALACTURONASE GENES AND ENCODED PROTEINS

Regulation of Polygalacturonase Gene Expression

The expression patterns of PG genes are regulated by various factors such as hormones, transcription factors, and environmental factors. These factors can function separately or be combined to form regulatory networks. Based on reported gene regulation in processes like lateral root initiation, root cap detachment, fruit ripening, and organ abscission, we summarize the regulatory network of PG genes in cell separation in **Figure 4**.

Lateral root initiation requires the primordium to emerge successively through the endodermal layer, cortical layer, and epidermal layer. PGs likely act to separate cells in the cortical and epidermal layers, contributing to the elongation of the primordium (Swarup et al., 2008; Kumpf et al., 2013). During this stage, a regulatory network (**Figure 4**) is formed centering on the auxin signal pathway and connecting with transcription factors, receptor-like kinases and OGs. Auxin influx carrier LIKE-AUXIN3 (*LAX3*) and its downstream auxin-inducing transcription factor LATERAL ORGAN BOUNDARIES DOMAIN18 (*LBD18*) indirectly increases the expression of *AtPG10* (Swarup et al., 2008; Kumpf et al., 2013; Lee et al., 2015). INFLORESCENCE DEFICIENT IN ABSCISSION (*IDA*) is the key regulator of cell separation, the expression of which is induced by AUXIN RESPONSE FACTOR7. After receiving a signal from *IDA*, the receptor-like kinase *HAESA* promotes the expression of *ADPG2* in lateral root development (Kumpf et al., 2013). However, whether *HAESA* regulates *ADPG2* directly has

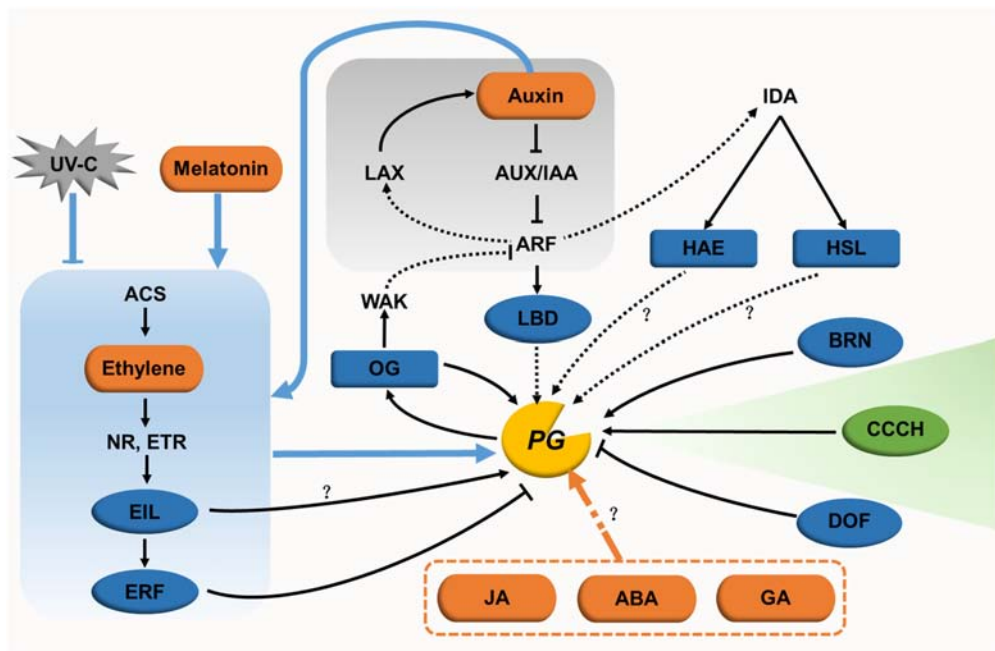


FIGURE 4 | Regulatory networks for PGs in cell expansion and separation. In cell expansion, only one regulator (green oval) has been identified. The regulatory network for this cell dynamic processes remain unclear. In cell separation, hormones (orange ovals) and their signaling pathways (with gray or blue background), transcription factors (blue ovals), other endogenous factors (blue rectangles), and the exogenous factor UV-C (gray polygon) form a genetic regulatory network for PGs. The direct or indirect action of some regulators needs to be further studied (black dotted lines). The regulatory pathways of JA, ABA, and GA in relation to PGs also need to be characterized further.

not been reported. OGs, the product of pectin degradation by PGs, might in turn control PG gene expression (Moscatiello et al., 2006). Exogenous OGs inhibit lateral root formation by suppressing the expression of auxin synthesis- and signaling-associated genes, such as *IAA5* and *SAUR16* (Savatin et al., 2011). Taking these results together, the interaction between auxin and PG-induced OGs is assumed to form a feedback regulation network (Ferrari et al., 2013). In this model, PG expression is induced by auxin, and the resulting PG proteins then release OGs in the apoplast. The accumulation of OGs in turn perturbs auxin responses. In root cap detachment, the transcription factor BEARSKIN1 directly binds to the *ROOT CAP POLYGALACTURONASE (RCPG)* promoter at a region containing a NAC-BINDING ELEMENT motif. The activation of *RCPG* accelerates cell detachment at the root cap (Kamiya et al., 2016).

In the process of fruit ripening, ethylene synthesis and signaling are crucial for regulating PG expression. Upon treatment with exogenous ethylene in the full bloom stage, expression of apple *PG1* is upregulated in the pericarp, leading to fruit softening (Tacken et al., 2010). The same expression pattern is found for the ethylene synthesis gene *ACC OXIDASE1*, which indicates that ethylene might regulate *PG1* expression. This idea was confirmed by the finding that EIN-LIKE3 transactivates the expression of *PG1* in transient assays (Tacken et al., 2010). Furthermore, ETHYLENE RESPONSE FACTOR9 (ERF9) in papaya (*Carica papaya*), which functions downstream of EIN3-LIKE, represses the expression of *CpPG5* in fruit

ripening by specifically binding to its promoter at a GCC-BOX motif. Its repressive role relies on the ERF-ASSOCIATED AMPHIPHILIC REPRESSION motif within ERF9 (Fu et al., 2016). Auxin functions in this process in concert with ethylene. With naphthalene acetic acid (NAA) treatment, the effect of an ethylene repressor is enhanced in downregulating the expression of ethylene-independent *MdPG2* in the cortex and abscission layer, and ethylene-dependent *MdPG1* in the abscission layer as well. However, NAA conversely upregulates the expression of *MdPG1* in the cortex, which results in the increased accumulation of ethylene in this layer (Li and Yuan, 2008). Apart from auxin, UV-C and melatonin also interact with ethylene in regulating PG genes. Treating postharvest tomato fruits with UV-C decreases the synthesis of ethylene, and suppresses the expression of PG genes together with other cell wall modifying genes during fruit softening (Bu et al., 2013). These data reveal that the regulatory mode of UV-C may be through the action of ethylene. Melatonin influences ethylene signaling in postharvest tomato ripening by promoting the expression of ethylene related genes (such as *NR*, *ETR4*, *EILs*, and *ERF2*), causing the upregulation of PG genes (Sun et al., 2015). ABA and GA also take part in this regulatory network. ABA treated tomato is downregulated in the expression of *SIPG* and other genes encoding wall-modifying enzymes, showing a delay in fruit softening (Sun et al., 2012). However, strawberry fruit treated with ABA showed slightly increased *FaPG1* expression without influencing its protein activity (Villarreal et al., 2009). GA affects flower blooming and fruit maturity in grapes (*Vitis labrusca* × *V. vinifera*) with two

230-fold upregulated PG genes (Cheng et al., 2015). However, the regulatory pathways for these two hormones need to be studied further.

During organ abscission, jasmonic acid (JA), transcription factors, and receptor-like kinases are involved in regulating PG expression. Microarray transcriptome data of a JA mutant treated with JA show that the expression of *ADPG1*, *ADPG2*, and *QRT2* are upregulated 10-fold upon JA treatment (Mandaokar et al., 2006). GUS activity of the *QRT2* promoter is detected in the anther of a JA mutant treated with methyl jasmonate (Ogawa et al., 2009). Taken together with the delayed dehiscence of anthers and siliques in the *adpg1 adpg2 qrt2* mutant, the non-dehiscent anther of the JA mutant may be caused by decreased PG gene expression (Ogawa et al., 2009). Additionally, the transcription of *ADPG2* in the AZ might be regulated directly via the binding of the transcription factor FOR DNA BINDING WITH ONE FINGER4.7 to the AAAG *cis*-regulatory element of its promoter (Wei et al., 2010). During lateral root emergence, both HAESA and HAESA-LIKE2 act downstream of IDA. These two kinases promote the expression of *ADPG2* in floral organ AZ via regulatory mechanisms that are currently unclear (González-Carranza et al., 2012).

Surprisingly little is known about the regulation of PG expression in cell expansion. Only a plant-specific tandem CCH zinc-finger C3H14 in Arabidopsis has been demonstrated to affect cell elongation by binding to the AU-RICH ELEMENT motif of the *ADPG1* promoter and activating its transcription (Kim W.C. et al., 2014). The roles of hormones, other transcription factors and environmental factors in regulating PG-mediated cell expansion need to be studied further.

Regulation of Polygalacturonase Biochemical Activities

External factors, other proteins, and the apoplastic microenvironment can all act as regulators of PG biochemical activities. PG activities can be controlled by temperature, pressure, exogenous polyamine and expansin (Crelrier et al., 2001; Fachin et al., 2002; Martinez-Tellez et al., 2002; Karakurt and Huber, 2003; Yuan et al., 2011; Nardi et al., 2013). UV-C, which was mentioned above as a transcriptional inhibitor, delays tomato fruit softening by negatively affecting PG activities (Barka et al., 2000). Hormones can also influence PG activities. During strawberry fruit ripening, exogenous NAA and an inhibitor of ethylene perception (1-MCP) decrease the expression of *FaPG1* and the activity of FaPG1, resulting in the delay of ripening. However, this phenotype is not necessarily caused by the difference in gene expression, because treatment with GA₃ reduces the activity of FaPG1 without significantly changing gene expression (Villarreal et al., 2009).

The wall microenvironment influences PG activities mainly through cations and pH. Ca²⁺ deficiency in potato leaves leads to a dramatic increase in PG enzyme activity (Seling et al., 2000), and the inhibitory role of Ca²⁺ for PG activities was also supported by treating grape berries with 1 mM Ca²⁺ (Cabanne and Doneche, 2001). In mango (*Mangifera indica*) fruit, cations like Li⁺, Mg²⁺, and a high substrate

(polygalacturonic acid) concentration suppress PG activities, whereas Fe³⁺ activates PG activities (Prasanna et al., 2006). PGs show high activity within physiological pH ranges. Two PG isoforms isolated from avocado (*Persea americana*) fruit mesocarp show highest activity when the pH is 6.0 (Wakabayashi and Huber, 2001). The influence of pH on PG activity can be altered by intercellular cations. The optimal pH for PG activity in tomato is 4–4.5 with NaCl, and changes to 5–6 with KCl (Chun and Huber, 1998). pH and Ca²⁺ are also likely to play complex roles in pectin degradation by influencing the action of PME and the formation of PG substrates.

Posttranslational modifications and interacting proteins can also influence PG activities. Polygalacturonase inhibiting proteins (PGIPs) can specifically bind and inhibit PG activities. Studies on the PGIP-PG interaction have been well summarized before (Kalunke et al., 2015). The β subunit of a tomato fruit polygalacturonase inhibits the interaction between PG and pectins (Zheng et al., 1992). To figure out whether a posttranslational modification works in affecting PG activities, phenylalanine residues within the FxxY sequence in a β subunit were modified to α , β -didehydro-Phe (Sergeant et al., 2017). This modification promoted β PG binding to pectins. It also contributed to the interaction of β PG with the catalytic subunit of PG (Sergeant et al., 2017). In addition, an 8-kD non-specific lipid transfer protein, ACTIVATOR, which is contained in a PG1 multiprotein complex, might change the form of associated PG into a heat-stable and active one (Tomassen et al., 2007). N-terminal prosequences have been hypothesized to control PG activity through influencing protein folding or protein inactivation (Hadfield and Bennett, 1998; Torki et al., 2000), but in at least some cases, the N-terminal prosequence instead controls the secretion of PG to the cell wall, as demonstrated in *Pichia pastoris* and Arabidopsis (Dal et al., 2001; Babu et al., 2013). To explore the posttranslational modifications of PG activity, an advanced approach for producing custom-folded proteins is required for further studies.

CONCLUSION

Major advances in our knowledge of the expression patterns, biological functions, and mechanisms of action of PGs have deepened our understanding of their roles in vegetative growth, pollen development, fruit ripening, and organ abscission. Although PGs are not the only pectin hydrolases, these studies provide convincing evidence for their unique importance in plant development. Based on genome data and current classification methods, we have a much clearer knowledge of PG structural characteristics, evolution, and expression patterns. However, shortcomings exist for each classification system, which calls for a better system that combines convenient calculations of phylogenetic relationships with high confidence values when dealing with PG families at large scales. PG genes with diverse expression patterns in the same tissues need to have their biological functions pinpointed for a better understanding of their individual activities and retention mechanisms.

To illuminate the ways in which PGs function in developmental processes, a comprehensive regulatory network needs to be built based on well-studied pathways that regulate PG expression and activity. As a wider suite of dyes and antibodies become available for observing pectin dynamics (Mravec et al., 2014, 2017; Pattathil et al., 2015; Anderson, 2016; Rydahl et al., 2018; Voiniciuc et al., 2018), we can dive into the mechanisms of PG action in cell wall dynamics. These and other studies can help us gain insights into the role of pectins in plant cell wall construction and dynamics as well as their relationships with other cell wall components.

AUTHOR CONTRIBUTIONS

YgY, YjY, and YL wrote the manuscript. CA and JC revised the paper. All the authors read and approved the final version of the manuscript.

REFERENCES

- Abbott, D. W., and Boraston, A. B. (2007). The structural basis for exopolygalacturonase activity in a family 28 glycoside hydrolase. *J. Mol. Biol.* 368, 1215–1222. doi: 10.1016/j.jmb.2007.02.083
- Albersheim, P., Darvill, A. G., O'Neill, M. A., Schols, H. A., and Voragen, A. G. J. (1996). An hypothesis: the same six polysaccharides are components of the primary cell walls of all higher plants. *Biotechnol. Prog.* 14, 47–55. doi: 10.1016/S0921-0423(96)80245-0
- Allen, R. L., and Lonsdale, D. M. (1993). Molecular characterization of one of the maize polygalacturonase gene family members which are expressed during late pollen development. *Plant J.* 3, 261–271. doi: 10.1111/j.1365-313X.1993.tb00177.x
- Anderson, C. T. (2016). We be Jammin': an update on pectin biosynthesis, trafficking and dynamics. *J. Exp. Bot.* 67, 495–502. doi: 10.1093/jxb/erv501
- Aouali, N., Laporte, P., and Clement, C. (2001). Pectin secretion and distribution in the anther during pollen development in *Lilium*. *Planta* 213, 71–79. doi: 10.1007/s004250000469
- Atkinson, R. G., Schroder, R., Hallett, I. C., Cohen, D., and MacRae, E. A. (2002). Overexpression of polygalacturonase in transgenic apple trees leads to a range of novel phenotypes involving changes in cell adhesion. *Plant Physiol.* 129, 122–133. doi: 10.1104/pp.010986
- Atkinson, R. G., Sutherland, P. W., Johnston, S. L., Gunaseelan, K., Hallett, I. C., Mitra, D., et al. (2012). Down-regulation of *POLYGALACTURONASE1* alters firmness, tensile strength and water loss in apple (*Malus x domestica*) fruit. *BMC Plant Biol.* 12:129. doi: 10.1186/1471-2229-12-129
- Babu, Y., Musielak, T., Henschen, A., and Bayer, M. (2013). Suspensor length determines developmental progression of the embryo in Arabidopsis. *Plant Physiol.* 162, 1448–1458. doi: 10.1104/pp.113.217166
- Barka, E. A., Kalantari, S., Makhlof, J., and Arul, J. (2000). Impact of UV-C irradiation on the cell wall-degrading enzymes during ripening of tomato (*Lycopersicon esculentum* L.) fruit. *J. Agric. Food Chem.* 48, 667–671. doi: 10.1021/jf9906174
- Boyer, J. S. (2016). Enzyme-less growth in *Chara* and terrestrial plants. *Front. Plant Sci.* 7:866. doi: 10.3389/fpls.2016.00866
- Braybrook, S. A., and Joensson, H. (2016). Shifting foundations: the mechanical cell wall and development. *Curr. Opin. Plant Biol.* 29, 115–120. doi: 10.1016/j.pbi.2015.12.009
- Bu, J., Yu, Y., Aisikaer, G., and Ying, T. (2013). Postharvest UV-C irradiation inhibits the production of ethylene and the activity of cell wall-degrading enzymes during softening of tomato (*Lycopersicon esculentum* L.) fruit. *Postharvest Biol. Technol.* 86, 337–345. doi: 10.1016/j.postharvbio.2013.07.026
- Bussink, H. J. D., Buxton, F. P., and Visser, J. (1991). Expression and sequence comparison of the *Aspergillus niger* and *Aspergillus tubingensis* genes encoding polygalacturonase-II. *Curr. Genet.* 19, 467–474. doi: 10.1007/BF00312738

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Nos. 31471877 and 31501769) and the Natural Science Foundation of Zhejiang Province (Grant No. LQ16C150005). The contributions of CA to this work were supported as part of The Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the United States Department of Energy, Office of Science, Basic Energy Sciences (Award No. DE-SC0001090).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.01208/full#supplementary-material>

- Cabanne, C., and Doneche, B. (2001). Changes in polygalacturonase activity and calcium content during ripening of grape berries. *Am. J. Enol. Vitic.* 52, 331–335.
- Cheng, C., Jiao, C., Singer, S. D., Gao, M., Xu, X., Zhou, Y., et al. (2015). Gibberellin-induced changes in the transcriptome of grapevine (*Vitis labrusca* x *V. vinifera*) cv. Kyoho flowers. *BMC Genomics* 16:128. doi: 10.1186/s12864-015-1324-8
- Chiang, J. Y., Shu, S. W., Ko, C. W., and Wang, C. S. (2006). Biochemical characterization of a pollen-specific cDNA encoding polygalacturonase in *Lilium longiflorum*. *Plant Sci.* 170, 433–440. doi: 10.1016/j.plantsci.2005.09.010
- Choi, J. K., Lee, B. H., Chae, C. H., and Shin, W. (2004). Computer modeling of the rhamnogalacturonase - "Hairy" pectin complex. *Proteins* 55, 22–33. doi: 10.1002/prot.10434
- Chun, J. P., and Huber, D. J. (1998). Polygalacturonase-mediated solubilization and depolymerization of pectic polymers in tomato fruit cell walls - Regulation by pH and ionic conditions. *Plant Physiol.* 117, 1293–1299. doi: 10.1104/pp.117.4.1293
- Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell. Bio.* 6, 850–861. doi: 10.1038/nrm1746
- Costa, F., Peace, C. P., Stella, S., Serra, S., Musacchi, S., Bazzani, M., et al. (2010). QTL dynamics for fruit firmness and softening around an ethylene-dependent polygalacturonase gene in apple (*Malus x domestica* Borkh.). *J. Exp. Bot.* 61, 3029–3039. doi: 10.1093/jxb/erq130
- Crelier, S., Robert, M. C., Claude, J., and Juillerat, M. A. (2001). Tomato (*Lycopersicon esculentum*) pectin methylesterase and polygalacturonase behaviors regarding heat- and pressure-induced inactivation. *J. Agric. Food Chem.* 49, 5566–5575. doi: 10.1021/jf010202u
- Daher, F. B., and Braybrook, S. A. (2015). How to let go: pectin and plant cell adhesion. *Front. Plant Sci.* 6:523. doi: 10.3389/fpls.2015.00523
- Dal, F. D., Child, R., Svendsen, I., and Ulvskov, P. (2001). The cleavable N-terminal domain of plant endopolygalacturonases from clade B may be involved in a regulated secretion mechanism. *J. Biol. Chem.* 276, 35297–35304. doi: 10.1074/jbc.M102136200
- Damak, N., Abdeljalil, S., Taeib, N. H., and Gargouri, A. (2015). Cloning and genomic organization of a rhamnogalacturonase gene from locally isolated strain of *Aspergillus niger*. *Appl. Biochem. Biotechnol.* 176, 2314–2327. doi: 10.1007/s12010-015-1720-1
- Dardelle, F., Lehner, A., Ramdani, Y., Bardor, M., Lerouge, P., Driouich, A., et al. (2010). Biochemical and immunocytological characterizations of Arabidopsis pollen tube cell wall. *Plant Physiol.* 153, 1563–1576. doi: 10.1104/pp.110.158881
- Duan, W., Huang, Z., Song, X., Liu, T., Liu, H., Hou, X., et al. (2016). Comprehensive analysis of the polygalacturonase and pectin methylesterase genes in *Brassica rapa* shed light on their different evolutionary patterns. *Sci. Rep.* 6:25107. doi: 10.1038/srep25107
- Fachin, D., Van Loey, A., VanLoeyIndrawati, A., Ludikhuyze, L., and Hendrickx, M. (2002). Thermal and high-pressure inactivation of tomato polygalacturonase:

- a kinetic study. *J. Food Sci.* 67, 1610–1615. doi: 10.1111/j.1365-2621.2002.tb08692.x
- Ferrari, S., Savatin, D. V., Sicilia, F., Gramegna, G., Cervone, F., and De Lorenzo, G. (2013). Oligogalacturonides: plant damage-associated molecular patterns and regulators of growth and development. *Front. Plant Sci.* 4:49. doi: 10.3389/fpls.2013.00049
- Francis, K. E., Lam, S. Y., and Copenhaver, G. P. (2006). Separation of Arabidopsis pollen tetrads is regulated by *QUARTET1*, a pectin methylesterase gene. *Plant Physiol.* 142, 1004–1013. doi: 10.1104/pp.106.085274
- Fu, C. C., Han, Y. C., Qi, X. Y., Shan, W., Chen, J. Y., Lu, W. J., et al. (2016). Papaya CpERF9 acts as a transcriptional repressor of cell-wall-modifying genes *CpPME1/2* and *CpPG5* involved in fruit ripening. *Plant Cell Rep.* 35, 2341–2352. doi: 10.1007/s00299-016-2038-3
- González-Carranza, Z. H., Elliott, K. A., and Roberts, J. A. (2007). Expression of polygalacturonases and evidence to support their role during cell separation processes in *Arabidopsis thaliana*. *J. Exp. Bot.* 58, 3719–3730. doi: 10.1093/jxb/erm222
- González-Carranza, Z. H., Shahid, A. A., Zhang, L., Liu, Y., Ninsuwan, U., and Roberts, J. A. (2012). A novel approach to dissect the abscission process in *Arabidopsis*. *Plant Physiol.* 160, 1342–1356. doi: 10.1104/pp.112.205955
- Gorguet, B., Schipper, D., van Lammeren, A., Visser, R. G. F., and van Heusden, A. W. (2009). *ps-2*, the gene responsible for functional sterility in tomato, due to non-dehiscent anthers, is the result of a mutation in a novel polygalacturonase gene. *Theor. Appl. Genet.* 118, 1199–1209. doi: 10.1007/s00122-009-0974-9
- Gou, J. Y., Miller, L. M., Hou, G., Yu, X. H., Chen, X. Y., and Liu, C. J. (2012). Acetylase-mediated deacetylation of pectin impairs cell elongation, pollen germination, and plant reproduction. *Plant Cell* 24, 50–65. doi: 10.1105/tpc.111.092411
- Gu, C., Wang, L., Wang, W., Zhou, H., Ma, B., Zheng, H., et al. (2016). Copy number variation of a gene cluster encoding endopolygalacturonase mediates flesh texture and stone adhesion in peach. *J. Exp. Bot.* 67, 1993–2005. doi: 10.1093/jxb/erw021
- Hadfield, K. A., and Bennett, A. B. (1998). Polygalacturonases: many genes in search of a function. *Plant Physiol.* 117, 337–343. doi: 10.1104/pp.117.2.337
- Hadfield, K. A., Rose, J. K. C., Yaver, D. S., Berka, R. M., and Bennett, A. B. (1998). Polygalacturonase gene expression in ripe melon fruit supports a role for polygalacturonase in ripening-associated pectin disassembly. *Plant Physiol.* 117, 363–373. doi: 10.1104/pp.117.2.363
- Herron, S. R., Benen, J. A. E., Scavetta, R. D., Visser, J., and Jurnak, F. (2000). Structure and function of pectic enzymes: virulence factors of plant pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8762–8769. doi: 10.1073/pnas.97.16.8762
- Hocq, L., Pelloux, J., and Lefebvre, V. (2017). Connecting homogalacturonan-type pectin remodeling to acid growth. *Trends Plant Sci.* 22, 20–29. doi: 10.1016/j.tplants.2016.10.009
- Huang, L., Cao, J., Zhang, A., Ye, Y., Zhang, Y., and Liu, T. (2009a). The polygalacturonase gene *BcMF2* from *Brassica campestris* is associated with intine development. *J. Exp. Bot.* 60, 301–313. doi: 10.1093/jxb/ern295
- Huang, L., Ye, Y., Zhang, Y., Zhang, A., Liu, T., and Cao, J. (2009b). *BcMF9*, a novel polygalacturonase gene, is required for both *Brassica campestris* intine and exine formation. *Ann. Bot.* 104, 1339–1351. doi: 10.1093/aob/mcp244
- Innan, H., and Kondrashov, F. (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* 11, 97–108. doi: 10.1038/nrg2689
- Irshad, M., Canut, H., Borderies, G., Pont-Lezica, R., and Jamet, E. (2008). A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: confirmed actors and newcomers. *BMC Plant Biol.* 8:94. doi: 10.1186/1471-2229-8-94
- Jenkins, J., Mayans, O., and Pickersgill, R. (1998). Structure and evolution of parallel beta-helix proteins. *J. Struct. Biol.* 122, 236–246. doi: 10.1006/jsbi.1998.3985
- Jolie, R. P., Duvetter, T., Van Loey, A. M., and Hendrickx, M. E. (2010). Pectin methylesterase and its proteinaceous inhibitor: a review. *Carbohydr. Res.* 345, 2583–2595. doi: 10.1016/j.carres.2010.10.002
- Kalaitzis, P., Solomos, T., and Tucker, M. L. (1997). Three different polygalacturonases are expressed in tomato leaf and flower abscission, each with a different temporal expression pattern. *Plant Physiol.* 113, 1303–1308. doi: 10.1104/pp.113.4.1303
- Kalunke, R. M., Tundo, S., Benedetti, M., Cervone, F., De Lorenzo, G., and D'Ovidio, R. (2015). An update on polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein that protects crop plants against pathogens. *Front. Plant Sci.* 6:146. doi: 10.3389/fpls.2015.00146
- Kamiya, M., Higashio, S. Y., Isomoto, A., Kim, J. M., Seki, M., Miyashima, S., et al. (2016). Control of root cap maturation and cell detachment by BEARSKIN transcription factors in *Arabidopsis*. *Development* 143, 4063–4072. doi: 10.1242/dev.142331
- Karakurt, Y., and Huber, D. J. (2003). Activities of several membrane and cell-wall hydrolases, ethylene biosynthetic enzymes, and cell wall polyuronide degradation during low-temperature storage of intact and fresh-cut papaya (*Carica papaya*) fruit. *Postharvest Biol. Technol.* 28, 219–229. doi: 10.1016/S0925-5214(02)00177-1
- Kim, J., and Patterson, S. E. (2006). Expression divergence and functional redundancy of polygalacturonases in floral organ abscission. *Plant Signal. Behav.* 1, 281–283. doi: 10.4161/psb.1.6.3541
- Kim, J., Shiu, S. H., Thoma, S., Li, W. H., and Patterson, S. E. (2006). Patterns of expansion and expression divergence in the plant polygalacturonase gene family. *Genome Biol.* 7:R87. doi: 10.1186/gb-2006-7-9-r87
- Kim, J., Sundaresan, S., Philosoph-Hadas, S., Yang, R., Meir, S., and Tucker, M. L. (2015). Examination of the abscission-associated transcriptomes for soybean, tomato, and *Arabidopsis* highlights the conserved biosynthesis of an extensible extracellular matrix and boundary layer. *Front. Plant Sci.* 6:1109. doi: 10.3389/fpls.2015.01109
- Kim, S., Park, M., Yeom, S. I., Kim, Y. M., Lee, J. M., Lee, H. A., et al. (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nat. Genet.* 46, 270–278. doi: 10.1038/ng.2877
- Kim, W. C., Kim, J. Y., Ko, J. H., Kang, H., Kim, J., and Han, K. H. (2014). AtC3H14, a plant-specific tandem CCH zinc-finger protein, binds to its target mRNAs in a sequence-specific manner and affects cell elongation in *Arabidopsis thaliana*. *Plant J.* 80, 772–784. doi: 10.1111/tpj.12667
- Kluskens, L. D., van Alebeek, G., Walther, J., Voragen, A. G. J., de Vos, W. M., and van der Oost, J. (2005). Characterization and mode of action of an exopolygalacturonase from the hyperthermophilic bacterium *Thermotoga maritima*. *FEBS J.* 272, 5464–5473. doi: 10.1111/j.1742-4658.2005.04935.x
- Kozioł, A., Cybulska, J., Pieczywek, P. M., and Zdunek, A. (2017). Changes of pectin nanostructure and cell wall stiffness induced *in vitro* by pectinase. *Carbohydr. Polym.* 161, 197–207. doi: 10.1016/j.carbpol.2017.01.014
- Kumpf, R. P., Shi, C. L., Larrieu, A., Sto, I. M., Butenko, M. A., Peret, B., et al. (2013). Floral organ abscission peptide IDA and its HAE/HSL2 receptors control cell separation during lateral root emergence. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5235–5240. doi: 10.1073/pnas.1210835110
- Landi, L., Feliziani, E., and Romanazzi, G. (2014). Expression of defense genes in strawberry fruits treated with different resistance inducers. *J. Agric. Food Chem.* 62, 3047–3056. doi: 10.1021/jf404423x
- Lee, H. W., Cho, C., and Kim, J. (2015). Lateral organ boundaries Domain16 and 18 Act downstream of the AUXIN1 and LIKE-AUXIN3 auxin influx carriers to control lateral root development in *Arabidopsis*. *Plant Physiol.* 168, 1792–U1177. doi: 10.1104/pp.15.00578
- Lester, D. R., Sherman, W. B., and Atwell, B. J. (1996). Endopolygalacturonase and the Melting flesh (M) locus in peach. *J. Am. Soc. Hortic. Sci.* 121, 231–235.
- Li, J., and Yuan, R. (2008). NAA and ethylene regulate expression of genes related to ethylene biosynthesis, perception, and cell wall degradation during fruit abscission and ripening in 'Delicious' apples. *J. Plant Growth Regul.* 27, 283–295. doi: 10.1007/s00344-008-9055-6
- Li, J., Zhu, H., and Yuan, R. (2010). Profiling the expression of genes related to ethylene biosynthesis, ethylene perception, and cell wall degradation during fruit abscission and fruit ripening in apple. *J. Am. Soc. Hortic. Sci.* 135, 391–401.
- Li, M., Zhang, Y., Zhang, Z., Ji, X., Zhang, R., Liu, D., et al. (2013). Hypersensitive ethylene signaling and ZmDPG1 expression lead to fruit softening and dehiscence. *PLoS One* 8:e58745. doi: 10.1371/journal.pone.0058745
- Liang, Y., Yu, Y., Cui, J., Lyu, M., Xu, L., and Cao, J. (2016). A comparative analysis of the evolution, expression, and cis-regulatory element of polygalacturonase genes in grasses and dicots. *Funct. Integr. Genomics* 16, 641–656. doi: 10.1007/s10142-016-0503-2

- Liang, Y., Yu, Y., Shen, X., Dong, H., Lyu, M., Xu, L., et al. (2015). Dissecting the complex molecular evolution and expression of polygalacturonase gene family in *Brassica rapa* ssp. *chinensis*. *Plant Mol. Biol.* 89, 629–646. doi: 10.1007/s11103-015-0390-2
- Liu, H., Ma, Y., Chen, N., Guo, S., Liu, H., Guo, X., et al. (2014). Overexpression of stress-inducible *OsBURP16*, the beta subunit of polygalacturonase 1, decreases pectin content and cell adhesion and increases abiotic stress sensitivity in rice. *Plant Cell Environ.* 37, 1144–1158. doi: 10.1111/pce.12223
- Mandaokar, A., Thines, B., Shin, B., Lange, B. M., Choi, G., Koo, Y. J., et al. (2006). Transcriptional regulators of stamen development in *Arabidopsis* identified by transcriptional profiling. *Plant J.* 46, 984–1008. doi: 10.1111/j.1365-313X.2006.02756.x
- Marković, O., and Janeček, S. (2001). Pectin degrading glycoside hydrolases of family 28: sequence-structural features, specificities and evolution. *Protein Eng.* 14, 615–631. doi: 10.1093/protein/14.9.615
- Martinez-Tellez, M. A., Ramos-Clamont, M. G., Gardea, A. A., and Vargas-Arispuro, I. (2002). Effect of infiltrated polyamines on polygalacturonase activity and chilling injury responses in zucchini squash (*Cucurbita pepo* L.). *Biochem. Biophys. Res. Commun.* 295, 98–101. doi: 10.1016/S0006-291X(02)00631-9
- Mayans, O., Scott, M., Connerton, I., Gravesen, T., Benen, J., Visser, J., et al. (1997). Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate-binding clefts of pectin and pectate lyases. *Structure* 5, 677–689. doi: 10.1016/S0969-2126(97)00222-0
- Mbégué-A-Mbégué, D., Hubert, O., Baurens, F. C., Matsumoto, T., Chillet, M., Fils-Lycaon, B., et al. (2009). Expression patterns of cell wall-modifying genes from banana during fruit ripening and in relationship with finger drop. *J. Exp. Bot.* 60, 2021–2034. doi: 10.1093/jxb/erp079
- McCarthy, T. W., Der, J. P., Honaas, L. A., dePamphilis, C. W., and Anderson, C. T. (2014). Phylogenetic analysis of pectin-related gene families in *Physcomitrella patens* and nine other plant species yields evolutionary insights into cell walls. *BMC Plant Biol.* 14:79. doi: 10.1186/1471-2229-14-79
- Moscattello, R., Mariani, P., Sanders, D., and Maathuis, F. J. M. (2006). Transcriptional analysis of calcium-dependent and calcium-independent signalling pathways induced by oligogalacturonides. *J. Exp. Bot.* 57, 2847–2865. doi: 10.1093/jxb/erl043
- Mravec, J., Kračun, S. K., Rydahl, M. G., Westereng, B., Miart, F., Clausen, M. H., et al. (2014). Tracking developmentally regulated post-synthetic processing of homogalacturonan and chitin using reciprocal oligosaccharide probes. *Development* 141, 4841–4850. doi: 10.1242/dev.113365
- Mravec, J., Kračun, S. K., Rydahl, M. G., Westereng, B., Pontiggia, D., De Lorenzo, G., et al. (2017). An oligogalacturonide-derived molecular probe demonstrates the dynamics of calcium-mediated pectin complexation in cell walls of tip-growing structures. *Plant J.* 91, 534–546. doi: 10.1111/tpj.13574
- Mutter, M., Beldman, G., Pitson, S. M., Schols, H. A., and Voragen, A. G. J. (1998). Rhamnogalacturonan alpha-D-galactopyranosyluronohydrolase - An enzyme that specifically removes the terminal nonreducing galacturonosyl residue in rhamnogalacturonan regions of pectin. *Plant Physiol.* 117, 153–163. doi: 10.1104/pp.117.1.153
- Nakashima, J., Endo, S., and Fukuda, H. (2004). Immunocytochemical localization of polygalacturonase during tracheary element differentiation in *Zinnia elegans*. *Planta* 218, 729–739. doi: 10.1007/s00425-003-1167-4
- Nardi, C., Escudero, C., Villarreal, N., Martinez, G., and Marcos Cívolo, P. (2013). The carbohydrate-binding module of *Fragaria x ananassa* expansin 2 (CBM-FaExp2) binds to cell wall polysaccharides and decreases cell wall enzyme activities "in vitro". *J. Plant Res.* 126, 151–159. doi: 10.1007/s10265-012-0504-8
- Ogawa, M., Kay, P., Wilson, S., and Swain, S. M. (2009). ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE1 (ADPG1), ADPG2, and QUARTET2 are Polygalacturonases required for cell separation during reproductive development in *Arabidopsis*. *Plant Cell* 21, 216–233. doi: 10.1105/tpc.108.063768
- Palin, R., and Geitmann, A. (2012). The role of pectin in plant morphogenesis. *Biosystems* 109, 397–402. doi: 10.1016/j.biosystems.2012.04.006
- Park, K. C., Kwon, S. J., and Kim, N. S. (2010). Intron loss mediated structural dynamics and functional differentiation of the polygalacturonase gene family in land plants. *Genes Genomics* 32, 570–577. doi: 10.1007/s13258-010-0076-8
- Park, K. C., Kwon, S. J., Kim, P. H., Bureau, T., and Kim, N. S. (2008). Gene structure dynamics and divergence of the polygalacturonase gene family of plants and fungus. *Genome* 51, 30–40. doi: 10.1139/G07-093
- Pattathil, S., Hahn, M. G., Dale, B. E., and Chundawat, S. P. S. (2015). Insights into plant cell wall structure, architecture, and integrity using glycome profiling of native and AFEX (TM)-pre-treated biomass. *J. Exp. Bot.* 66, 4279–4294. doi: 10.1093/jxb/erv107
- Patterson, S. E., and Bleecker, A. B. (2004). Ethylene-dependent and -independent processes associated with floral organ abscission in *Arabidopsis*. *Plant Physiol.* 134, 194–203. doi: 10.1104/pp.103.028027
- Peaucelle, A., Braybrook, S. A., Le Guillou, L., Bron, E., Kuhlmeier, C., and Hoeft, H. (2011). Pectin-induced changes in cell wall mechanics underlie organ initiation in *Arabidopsis*. *Curr. Biol.* 21, 1720–1726. doi: 10.1016/j.cub.2011.08.057
- Peaucelle, A., Louvet, R., Johansen, J. N., Hoeft, H., Laufs, P., Pelloux, J., et al. (2008). *Arabidopsis* phyllotaxis is controlled by the methyl-esterification status of cell-wall pectins. *Curr. Biol.* 18, 1943–1948. doi: 10.1016/j.cub.2008.10.065
- Peaucelle, A., Wightman, R., and Hoeft, H. (2015). The control of growth symmetry breaking in the *Arabidopsis* hypocotyl. *Curr. Biol.* 25, 1746–1752. doi: 10.1016/j.cub.2015.06.032
- Pickersgill, R., Smith, D., Worboys, K., and Jenkins, J. (1998). Crystal structure of polygalacturonase from *Erwinia carotovora* ssp. *carotovora*. *J. Biol. Chem.* 273, 24660–24664.
- Pose, S., Paniagua, C., Cifuentes, M., Blanco-Portales, R., Quesada, M. A., and Mercado, J. A. (2013). Insights into the effects of polygalacturonase *FaPG1* gene silencing on pectin matrix disassembly, enhanced tissue integrity, and firmness in ripe strawberry fruits. *J. Exp. Bot.* 64, 3803–3815. doi: 10.1093/jxb/ert210
- Prasanna, V., Prabha, T. N., and Tharanathan, R. N. (2006). Multiple forms of polygalacturonase from mango (*Mangifera indica* L. cv *Alphonso*) fruit. *Food Chem.* 95, 30–36. doi: 10.1016/j.foodchem.2004.12.014
- Rao, M. N., Kembhavi, A. A., and Pant, A. (1996). Implication of tryptophan and histidine in the active site of endo-polygalacturonase from *Aspergillus ustus*: elucidation of the reaction mechanism. *Biochim. Biophys. Acta* 1296, 167–173.
- Rexovabekova, L. (1990). Evidence for the role of carboxyl groups in activity of endopolygalacturonase of *Aspergillus niger* chemical modification by carbodiimide reagent. *Collect. Czechoslov. Chem. Commun.* 55, 1389–1395. doi: 10.1135/cccc19901389
- Rhee, S. Y., Osborne, E., Poindexter, P. D., and Somerville, C. R. (2003). Microspore separation in the *quartet3* mutants of *Arabidopsis* is impaired by a defect in a developmentally regulated polygalacturonase required for pollen mother cell wall degradation. *Plant Physiol.* 133, 1170–1180. doi: 10.1104/pp.103.028266
- Ridley, B. L., O'Neill, M. A., and Mohnen, D. A. (2001). Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* 57, 929–967. doi: 10.1016/S0031-9422(01)00113-3
- Rodríguez-Llorente, I. D., Pérez-Hormaeche, J., El Mounadi, K., Dary, M., Caviedes, M. A., Cosson, V., et al. (2004). From pollen tubes to infection threads: recruitment of *Medicago* floral pectic genes for symbiosis. *Plant J.* 39, 587–598. doi: 10.1111/j.1365-313X.2004.02155.x
- Rui, Y., Xiao, C., Yi, H., Kandemir, B., Wang, J. Z., Puri, V. M., et al. (2017). POLYGALACTURONASE INVOLVED IN EXPANSION3 functions in seedling development, rosette growth, and stomatal dynamics in *Arabidopsis thaliana*. *Plant Cell* 29, 2413–2432. doi: 10.1105/tpc.17.00568
- Rydahl, M. G., Hansen, A. R., Kračun, S. K., and Mravec, J. (2018). Report on the current inventory of the toolbox for plant cell wall analysis: proteinaceous and small molecular probes. *Front. Plant Sci.* 9:581. doi: 10.3389/fpls.2018.00581
- Savatin, D. V., Ferrari, S., Sicilia, F., and De Lorenzo, G. (2011). Oligogalacturonide-auxin antagonism does not require posttranscriptional gene silencing or stabilization of auxin response repressors in *Arabidopsis*. *Plant Physiol.* 157, 1163–1174. doi: 10.1104/pp.111.184663
- Sekine, D., Munemura, I., Gao, M., Mitsuhashi, W., Toyomasu, T., and Murayama, H. (2006). Cloning of cDNAs encoding cell-wall hydrolases from pear (*Pyrus communis*) fruit and their involvement in fruit softening and development of melting texture. *Physiol. Plant.* 126, 163–174. doi: 10.1111/j.1399-3054.2006.00583.x
- Seling, S., Wissemeier, A. H., Cambier, P., and Van Cutsem, P. (2000). Calcium deficiency in potato (*Solanum tuberosum* ssp. *tuberosum*) leaves and its effects on the pectic composition of the apoplastic fluid. *Physiol. Plant.* 109, 44–50. doi: 10.1034/j.1399-3054.2000.100107.x

- Sénéchal, F., Wattier, C., Rusterucci, C., and Pelloux, J. (2014). Homogalacturonan-modifying enzymes: structure, expression, and roles in plants. *J. Exp. Bot.* 65, 5125–5160. doi: 10.1093/jxb/eru272
- Sergeant, K., Printz, B., Gutsch, A., Behr, M., Renaut, J., and Hausman, J. F. (2017). Dihydrophenylalanine, an abundant modification in the beta subunit of plant polygalacturonases. *PLoS One* 12:e0171990. doi: 10.1371/journal.pone.0171990
- Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., et al. (2004). Toward a systems approach to understanding plant-cell walls. *Science* 306, 2206–2211. doi: 10.1126/science.1102765
- Sun, L., Sun, Y., Zhang, M., Wang, L., Ren, J., Cui, M., et al. (2012). Suppression of 9-cis-epoxycarotenoid dioxygenase, which encodes a key enzyme in abscisic acid biosynthesis, alters fruit texture in transgenic tomato. *Plant Physiol.* 158, 283–298. doi: 10.1104/pp.111.186866
- Sun, Q., Zhang, N., Wang, J., Zhang, H., Li, D., Shi, J., et al. (2015). Melatonin promotes ripening and improves quality of tomato fruit during postharvest life. *J. Exp. Bot.* 66, 657–668. doi: 10.1093/jxb/eru332
- Swarup, K., Benkova, E., Swarup, R., Casimiro, I., Peret, B., Yang, Y., et al. (2008). The auxin influx carrier LAX3 promotes lateral root emergence. *Nat. Cell Biol.* 10, 946–954. doi: 10.1038/ncb1754
- Tacken, E., Ireland, H., Gunaseelan, K., Karunairetnam, S., Wang, D., Schultz, K., et al. (2010). The role of ethylene and cold temperature in the regulation of the apple POLYGALACTURONASE1 gene and fruit softening. *Plant Physiol.* 153, 294–305. doi: 10.1104/pp.109.151092
- Tebbutt, S. J., Rogers, H. J., and Lonsdale, D. M. (1994). Characterization of a tobacco gene encoding a pollen-specific polygalacturonase. *Plant Mol. Biol.* 25, 283–297. doi: 10.1007/BF00023244
- Tomassen, M. M. M., Barrett, D. M., van der Valk, H. C. P. M., and Woltering, E. J. (2007). Isolation and characterization of a tomato non-specific lipid transfer protein involved in polygalacturonase-mediated pectin degradation. *J. Exp. Bot.* 58, 1151–1160. doi: 10.1093/jxb/erl288
- Torki, M., Mandaron, P., Mache, R., and Falconet, D. (2000). Characterization of a ubiquitous expressed gene family encoding polygalacturonase in *Arabidopsis thaliana*. *Gene* 242, 427–436. doi: 10.1016/S0378-1119(99)00497-7
- Tu, T., Meng, K., Luo, H., Turunen, O., Zhang, L., Cheng, Y., et al. (2015). New insights into the role of T3 loop in determining catalytic efficiency of GH28 endo-polygalacturonases. *PLoS One* 10:e0135413. doi: 10.1371/journal.pone.0135413
- Tyler, L., Bragg, J. N., Wu, J., Yang, X., Tuskan, G. A., and Vogel, J. P. (2010). Annotation and comparative analysis of the glycoside hydrolase genes in *Brachypodium distachyon*. *BMC Genomics* 11:600. doi: 10.1186/1471-2164-11-600
- Villarreal, N. M., Martinez, G. A., and Civello, P. M. (2009). Influence of plant growth regulators on polygalacturonase expression in strawberry fruit. *Plant Sci.* 176, 749–757. doi: 10.1016/j.plantsci.2009.02.019
- Villarreal, N. M., Rosli, H. G., Martinez, G. A., and Civello, M. (2008). Polygalacturonase activity and expression of related genes during ripening of strawberry cultivars with contrasting fruit firmness. *Postharvest Biol. Technol.* 47, 141–150. doi: 10.1016/j.postharvbio.2007.06.011
- Voiniciuc, C., Pauly, M., and Usadel, B. (2018). Monitoring polysaccharide dynamics in the plant cell wall. *Plant Physiol.* 176, 2590–2600. doi: 10.1104/pp.17.01776
- Wakabayashi, K., and Huber, D. J. (2001). Purification and catalytic properties of polygalacturonase isoforms from ripe avocado (*Persea americana*) fruit mesocarp. *Physiol. Plant.* 113, 210–216. doi: 10.1034/j.1399-3054.2001.1130208.x
- Wakasa, Y., Kudo, H., Ishikawa, R., Akada, S., Senda, M., Niizeki, M., et al. (2006). Low expression of an endopolygalacturonase gene in apple fruit with long-term storage potential. *Postharvest Biol. Technol.* 39, 193–198. doi: 10.1016/j.postharvbio.2005.10.005
- Wang, F., Sun, X., Shi, X., Zhai, H., Tian, C., Kong, F., et al. (2016). A global analysis of the polygalacturonase gene family in soybean (*Glycine max*). *PLoS One* 11:e0163012. doi: 10.1371/journal.pone.0163012
- Wei, P. C., Tan, F., Gao, X. Q., Zhang, X. Q., Wang, G. Q., Xu, H., et al. (2010). Overexpression of AtDOF4.7, an Arabidopsis DOF family transcription factor, induces floral organ abscission deficiency in Arabidopsis. *Plant Physiol.* 153, 1031–1045. doi: 10.1104/pp.110.153247
- Willats, W. G. T., McCartney, L., Mackie, W., and Knox, J. P. (2001). Pectin: cell biology and prospects for functional analysis. *Plant Mol. Biol.* 47, 9–27. doi: 10.1007/978-94-010-0668-2_2
- Xiao, C., Barnes, W. J., Zamil, M. S., Yi, H., Puri, V. M., and Anderson, C. T. (2017). Activation tagging of Arabidopsis POLYGALACTURONASE INVOLVED IN EXPANSION2 promotes hypocotyl elongation, leaf expansion, stem lignification, mechanical stiffening, and lodging. *Plant J.* 89, 1159–1173. doi: 10.1111/tpj.13453
- Xiao, C., Somerville, C., and Anderson, C. T. (2014). POLYGALACTURONASE INVOLVED IN EXPANSION1 functions in cell Elongation and flower development in Arabidopsis. *Plant Cell* 26, 1018–1035. doi: 10.1105/tpc.114.123968
- Yang, Z. L., Liu, H. J., Wang, X. R., and Zeng, Q. Y. (2013). Molecular evolution and expression divergence of the *Populus* polygalacturonase supergene family shed light on the evolution of increasingly complex organs in plants. *New Phytol.* 197, 1353–1365. doi: 10.1007/s11103-015-0390-2
- Yoneda, A., Ito, T., Higaki, T., Kutsuna, N., Saito, T., Ishimizu, T., et al. (2010). Cobtorin target analysis reveals that pectin functions in the deposition of cellulose microfibrils in parallel with cortical microtubules. *Plant J.* 64, 657–667. doi: 10.1111/j.1365-313X.2010.04356.x
- Yu, Y., Liang, Y., Lv, M., Wu, J., Lu, G., and Cao, J. (2014). Genome-wide identification and characterization of polygalacturonase genes in *Cucumis sativus* and *Citrullus lanatus*. *Plant Physiol. Biochem.* 74, 263–275. doi: 10.1016/j.plaphy.2013.11.022
- Yuan, P., Meng, K., Huang, H., Shi, P., Luo, H., Yang, P., et al. (2011). A novel acidic and low-temperature-active endo-polygalacturonase from *Penicillium* sp CGMCC 1669 with potential for application in apple juice clarification. *Food Chem.* 129, 1369–1375. doi: 10.1016/j.foodchem.2011.05.065
- Zhang, Q., Huang, L., Liu, T., Yu, X., and Cao, J. (2008). Functional analysis of a pollen-expressed polygalacturonase gene BcMF6 in Chinese cabbage (*Brassica campestris* L. ssp *chinensis* Makino). *Plant Cell Rep.* 27, 1207–1215. doi: 10.1007/s00299-008-0541-x
- Zheng, L. S., Heupel, R. C., and Dellapenna, D. (1992). The beta-subunit of tomato fruit polygalacturonase isoenzyme-1 - Isolation, characterization, and identification of unique structural features. *Plant Cell* 4, 1147–1156. doi: 10.1105/tpc.4.9.1147

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Yang, Yu, Liang, Anderson and Cao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Current Models for Transcriptional Regulation of Secondary Cell Wall Biosynthesis in Grasses

Xiaolan Rao^{1,2*} and Richard A. Dixon^{1,2,3}

¹ BioDiscovery Institute and Department of Biological Sciences, University of North Texas, Denton, TX, United States, ² BioEnergy Science Center, United States Department of Energy, Oak Ridge, TN, United States, ³ Center for Bioenergy Innovation, United States Department of Energy, Oak Ridge, TN, United States

OPEN ACCESS

Edited by:

Shucai Wang,
Northeast Normal University, China

Reviewed by:

Anongpat Suttangkakul,
Kasetsart University, Thailand
Rui Shi,
North Carolina State University,
United States

*Correspondence:

Xiaolan Rao
xiaolan.rao@unt.edu

Specialty section:

This article was submitted to
Plant Physiology,
a section of the journal
Frontiers in Plant Science

Received: 18 November 2017

Accepted: 13 March 2018

Published: 04 April 2018

Citation:

Rao X and Dixon RA (2018) Current Models for Transcriptional Regulation of Secondary Cell Wall Biosynthesis in Grasses. *Front. Plant Sci.* 9:399. doi: 10.3389/fpls.2018.00399

Secondary cell walls mediate many crucial biological processes in plants including mechanical support, water and nutrient transport and stress management. They also provide an abundant resource of renewable feed, fiber, and fuel. The grass family contains the most important food, forage, and biofuel crops. Understanding the regulatory mechanism of secondary wall formation in grasses is necessary for exploiting these plants for agriculture and industry. Previous research has established a detailed model of the secondary wall regulatory network in the dicot model species *Arabidopsis thaliana*. Grasses, branching off from the dicot ancestor 140–150 million years ago, display distinct cell wall morphology and composition, suggesting potential for a different secondary wall regulation program from that established for dicots. Recently, combined application of molecular, genetic and bioinformatics approaches have revealed more transcription factors involved in secondary cell wall biosynthesis in grasses. Compared with the dicots, grasses exhibit a relatively conserved but nevertheless divergent transcriptional regulatory program to activate their secondary cell wall development and to coordinate secondary wall biosynthesis with other physiological processes.

Keywords: secondary cell wall, secondary cell wall regulation, transcription factor, grasses, lignin biosynthesis

INTRODUCTION

The plant cell wall is a structural layer located outside of the cell membrane that provides the physical strength to maintain cell shape against gravity (Taiz and Zeiger, 1998). There are two types of cell wall, primary and secondary. The primary cell wall is a thin layer with considerable flexibility for extension, and is formed in most plant cells. In contrast, the secondary cell wall is a thicker layer deposited between the primary cell wall and the cell membrane, and is formed in specialized types of cells such as tracheid/vessel elements and fibers (Taiz and Zeiger, 1998; Cosgrove and Jarvis, 2012). Secondary cell walls play a pivotal role during plant development and are involved in resistance to abiotic/biotic stresses (Houston et al., 2016). At the same time, cell wall recalcitrance, resulting in large part from the complex cross-linked matrix of the lignified secondary cell wall, is the major barrier in conversion of biomass into biofuel (Himmel et al., 2007; Pauly and Keegstra, 2010). Plants in the grass (Poaceae) family supply the most abundant, renewable sources of both nutrition and sustainable energy. Therefore, knowledge of grass secondary wall regulation

can be applied for genetic modification to improve the quality of food, forages and fuel crops that sustainably supply economic and ecological benefits.

Transcriptional regulation in secondary wall formation has been extensively elucidated in the dicot model species *Arabidopsis thaliana* (Zhong and Ye, 2015). However, details of the regulatory network in grass secondary wall formation are still under investigation. The emergence of secondary cell walls in plants occurred about 430 million years ago, as an adaptation for colonizing from ocean to dry land (Li and Chapple, 2010). Around 140–150 million years ago, monocots achieved divergence from the dicot ancestor (Chaw et al., 2004). Subsequently, particular classes of transcription factors (TFs) have expanded in monocot lineages, including the R2R3 MYB TF class to which many secondary wall-regulators belong (Rabinowicz et al., 1999; Zhao and Bartley, 2014). The evolutionary history suggests that grasses may share conservation of secondary cell wall regulation with dicots to some degree, but present their unique aspects. Recent evidence indirectly or directly supports this view. This review focuses on current advances in secondary cell wall regulation in grasses, and discusses the commonalities and the differences between grasses and dicots.

CROSSTALK BETWEEN SECONDARY WALL SYNTHESIS AND OTHER PHYSIOLOGICAL PROCESSES

Establishment of secondary cell walls is not an independent event, but involves crosstalk with other biological processes. First, secondary wall accumulation is determined by sugar levels in the plant controlled by light and the circadian clock (Rogers et al., 2005). Plants have to maintain a balance between carbon supply captured through photosynthesis and carbon assimilation, which converts carbon resources into cell wall polymers (Smith and Stitt, 2007; Loque et al., 2015). Second, secondary walls are deposited in specialized cells that have ceased growth and achieved their final cell shape (Cosgrove and Jarvis, 2012). The events of cell-cycle exit and cell wall remodeling occur at the initial stage of secondary wall formation through differential regulation of cell cycle controllers and wall-modifying enzymes, respectively (Goulao et al., 2011; Didi et al., 2015; Polyn et al., 2015). In the process of tracheary element differentiation, secondary cell wall biosynthesis is required to be tightly coupled with programmed cell death (PCD) (Groover and Jones, 1999; Ohashi-Ito et al., 2010). Third, a strictly coordinated biosynthetic program is observed among individual secondary wall components including cellulose, xylan and lignin, which leads to the proper assembly of the secondary wall. Finally, the lignin biosynthesis pathway shares common intermediates with other secondary metabolism pathways such as flavonoid biosynthesis (Dixon et al., 2013), allowing plants to recruit controllers to shift the metabolic flow upon demand (Bhargava et al., 2010). To achieve this coordination, plants have employed a limited number of TF

families to constitute a complex regulatory network that is capable of coordinating secondary cell wall biosynthesis with other physiological processes.

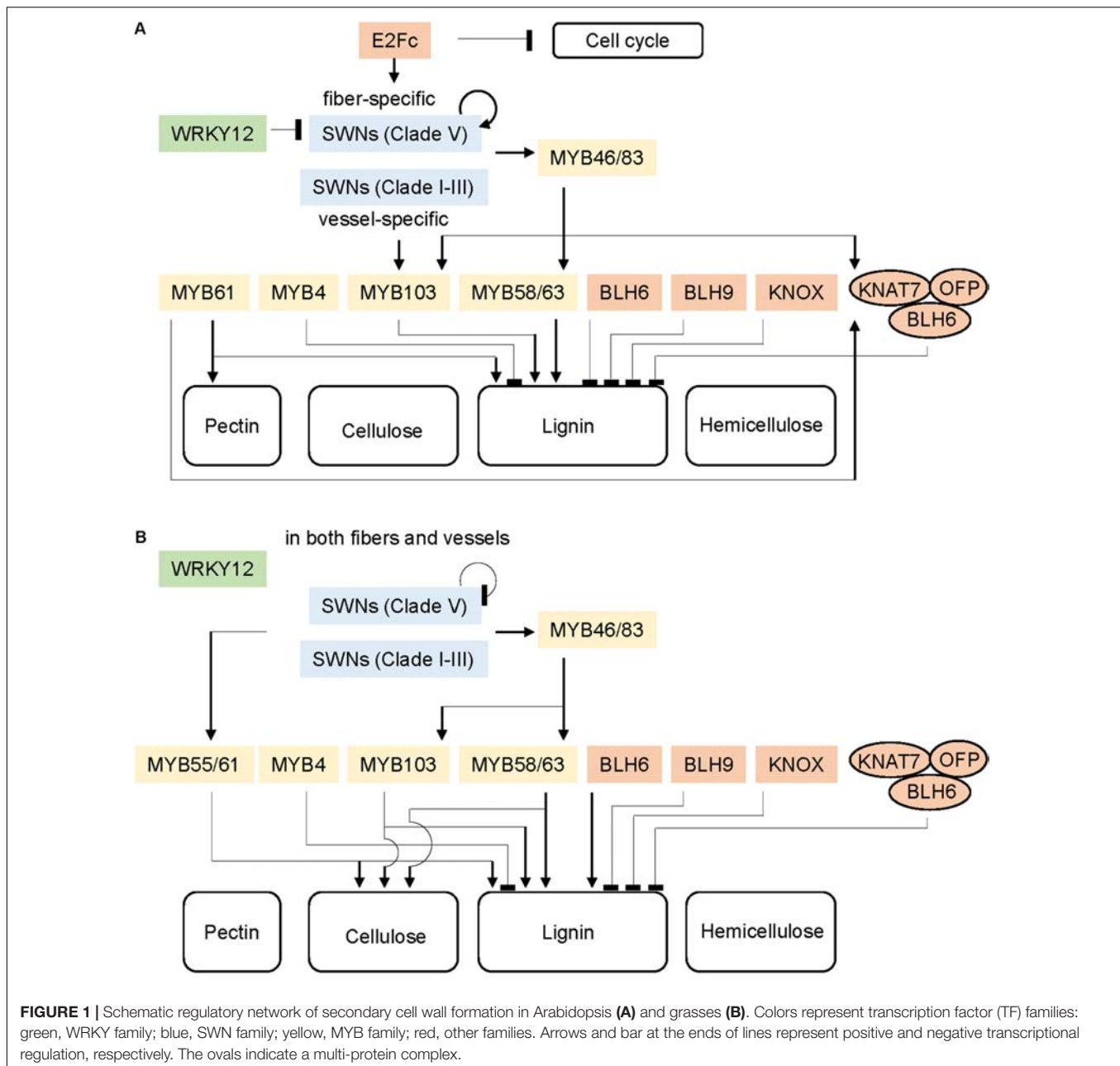
TFs INVOLVED IN GRASS SECONDARY WALL FORMATION

SWNs as Ancestral Master Switches for the Secondary Wall Program

A subgroup of NAC TFs, called secondary wall NACs (SWNs), function as top-level master switches for secondary cell wall biosynthesis. Diverse SWN orthologs exist in vascular plants, with first appearance in *S. moellendorffii* (Zhong et al., 2010a; Yao et al., 2012; Nakano et al., 2015). It has been considered that vascular plants may have employed these ancestral NACs via duplication for controlling secondary wall biosynthesis at the early stage of colonization of the land (Zhong et al., 2010a; Yao et al., 2012; Nakano et al., 2015).

Secondary wall NACs can be divided into four clades, according to their protein alignment (**Supplementary Figure S1**). *Arabidopsis* SWNs specifically expressed in vessels and fibers belong to clades I to III (called VNDs) and clade IV (called NST/SND), respectively (Zhong et al., 2010a). In *Arabidopsis*, AtVND6 and AtVND7 are responsible for determining tracheary element differentiation through controlling both secondary wall thickening and PCD in vessels (Ohashi-Ito et al., 2010; Yamaguchi et al., 2010), while AtNST1 and AtSND1 (also named as NST3) redundantly activate the whole secondary wall program in fiber cells (Zhong et al., 2006, 2007b; Mitsuda et al., 2007). The conserved function of SWN orthologs has been observed in many other dicots such as *Medicago* and poplar (Zhao et al., 2010; Zhong et al., 2010b; Wang et al., 2011), and in grasses including rice, maize, *Brachypodium*, and switchgrass (Zhong et al., 2011, 2015; Valdivia et al., 2013; Yoshida et al., 2013; Xiao et al., 2017). The exogenous overexpression of rice, maize and switchgrass SWNs in the *Arabidopsis nst1 snd1* double mutant can rescue the deficit of secondary wall development, and the endogenous overexpression of SWNs in rice, maize, and *Brachypodium* leads to secondary cell wall thickening and an upregulation of secondary wall-related genes (Zhong et al., 2011, 2015; Valdivia et al., 2013; Yoshida et al., 2013; Xiao et al., 2017). SWNs from rice, maize, *Brachypodium*, and switchgrass are capable of directly inducing the expression of secondary wall biosynthesis genes in *Arabidopsis* through binding to the SNBE (secondary wall NAC binding element) motif in the target gene's promoters (Zhong et al., 2006, 2011, 2015; Valdivia et al., 2013). Moreover, an upregulation of PCD genes was observed following endogenous/exogenous overexpression of SWNs in clades I, II and III, but not of SWNs in clade IV, in both *Arabidopsis* and grasses (Zhong et al., 2011, 2015; Valdivia et al., 2013).

Though highly conserved functions of SWNs are shared in vascular plants, some differences in expression pattern and regulatory mechanisms of SWNs have been detected in grasses and dicots. Unlike the differentiation of spatial expression in *Arabidopsis*, SWNs in all four clades display a similar expression



pattern in all the secondary wall-enriched cells including xylem vessels and cortical fibers in rice, maize, *Brachypodium*, and switchgrass (Zhong et al., 2011, 2015; Valdivia et al., 2013; Yoshida et al., 2013; Xiao et al., 2017). One explanation is that, in Arabidopsis, xylem fibers do not undergo cell death, as a result of the recruitment of SWNs in clade VI that activate secondary wall development but do not induce cell death, while the xylem vessel elements endure the coupled programs of secondary wall formation and cell death caused by SWNs in clades I, II, and III (Bollhoner et al., 2012). This may be not the case in grasses. Moreover, the regulation of SWNs displays different features among vascular plants. In the dicots Arabidopsis and *Medicago truncatula*, SND1 shows a relatively simple feedback- regulation

that can be auto-activated via binding to its own promoter and negatively regulated by its downstream MYB TFs (Wang et al., 2011). In wood development in *Populus trichocarpa*, a more complex regulation is apparent. Full-size PtrSND1 members self-activate their own genes as that in Arabidopsis, whereas splice variants from PtrSND1-A2 and PtrVND6-C1 reciprocally cross-inhibit the expression of all SWN members in clades I to III and clade IV, respectively, without auto-repression of their cognate TFs (Li Q. et al., 2012; Lin et al., 2017). However, in rice, the alternatively spliced form of OsSWN2, which lacks the transcriptional activation domain, may participate in a negative feedback loop to OsSWN1 and its cognate gene OsSWN2 (Yoshida et al., 2013). Taken together, these observations suggest

that, although grasses and dicots evolved from the last common ancestor to recruit SWNs as master switches in the secondary cell wall program, plants may utilize lineage-specific self-regulation of SWNs and different SWNs with functional specialization.

MYB Clades as Activators in Secondary Wall Accumulation

Secondary wall NACs service as master switches in secondary wall biosynthesis though directly regulating the transcriptional changes in secondary wall-structural genes and downstream TFs. In Arabidopsis, AtMYB46 and its paralog AtMYB83 are specifically expressed in both fibers and vessels, and redundantly activate secondary cell wall enhancement (Zhong et al., 2007a; McCarthy et al., 2009). The MYB46/83 orthologs in rice, maize, and switchgrass are capable of rescuing the defect in secondary cell wall-thickening in the Arabidopsis *myb46/83* double mutant (Zhong et al., 2011, 2015). Similar to AtMYB46, constitutive overexpression of ZmMYB46, OsMYB46, and PvMYB46 in Arabidopsis led to ectopic secondary wall deposition in stem and increased the content of cellulose, xylan and lignin, without activating the PCD genes (Zhong et al., 2011, 2015; Ko et al., 2012; Kim et al., 2013). Moreover, AtMYB46 and its ortholog PvMYB46 share a high similarity in activation efficiency on eight *cis*-acting elements [named the secondary wall MYB-responsive element (SMRE)] to induce the expression of target genes involved in secondary wall-related cellulose, xylan, and lignin biosynthesis (Ko et al., 2012; Zhong and Ye, 2012; Kim et al., 2013; Zhong et al., 2015), indicating the conservation of MYB46 function in grasses and Arabidopsis.

Two clades of MYBs, MYB58/63, and MYB42/85 (Supplementary Figure S2), are considered to be lower-level regulators of secondary wall biosynthesis, whose promoters can be bound by MYB46/83. In Arabidopsis, AtMYB58/63 and AtMYB42/85 are grouped as lignin-specific regulators because they show exclusive activation of all lignin biosynthesis genes (except AtF5H) (Zhou et al., 2009; Zhao and Dixon, 2011). Consistently, overexpression of OsMYB58/63 or OsMYB42/85 in rice leads to an elevated lignin content in the vascular bundles and sclerenchyma (Hirano et al., 2013b), and overexpression of SbMYB60 (the ortholog of AtMYB58/63) activates the expression of lignin biosynthesis genes and increases the lignin concentration in the biomass (Scully et al., 2016). Both results indicate the positive roles of these grass TFs in lignin accumulation. However, OsMYB58/63 triggers the additional expression of secondary wall-related cellulose synthase genes (Noda et al., 2015), and SbMYB60 overexpression affects the abundance of cellulose and xylan in the cell wall (Scully et al., 2016), neither of which effects are associated with AtMYB58/63 in Arabidopsis (Zhou et al., 2009). Interestingly, promoter analysis reveals that OsMYB58/63 and its Arabidopsis ortholog AtMYB58/63 proteins display a similar capacity for recognizing their binding sites (called AC-elements) (Zhou et al., 2009; Noda et al., 2015). AtMYB58/63 can activate the expression of secondary wall-related cellulose synthase genes in rice, but not in Arabidopsis (Noda et al., 2015). One explanation is a change in promoter elements during evolution. AC elements are found in

the promoter regions of lignin biosynthesis genes (except *F5H*) in Arabidopsis, but are absent in cellulose and xylan biosynthesis genes (Zhou et al., 2009; Zhao and Dixon, 2011). However, in rice, AC elements appear in the promoters of many secondary wall-related cellulose, xylan, and lignin biosynthesis genes (Zhou et al., 2009; Noda et al., 2015). Though rice and Arabidopsis MYB58/63 share commonalities of regulatory binding sites, the changed composition in *cis*-regulatory elements provides the basis for MYB58/63 to induce the biosynthesis program of all three secondary wall-components in rice but not in Arabidopsis.

Genes in the MYB55/61 and MYB103 clades (Supplementary Figure S2) are also positive regulators of secondary cell wall biosynthesis in Arabidopsis and grasses. The *atmyb61* mutant of Arabidopsis displayed fewer differentiated xylem vessels, and with reduced secondary wall-thickening, in the inflorescence stem (Newman et al., 2004; Romano et al., 2012). The target genes of AtMYB61 include a secondary wall-repressor *AtKNAT7*, a pectin methylesterase (*AtPME*) and *AtCCoAOMT7* (encoding the caffeoyl CoA 3-*O*-methyltransferase of monolignol biosynthesis) (Romano et al., 2012). In rice, the expression of OsMYB55/61 can be directly induced by OsSWN2 and OsSWN3 (Huang et al., 2015). OsMYB55/61 is capable of modulating lignin content in vascular bundles through activating lignin biosynthesis genes (at least *CAD2*) (Hirano et al., 2013b), and promoting secondary wall-related cellulose synthesis through binding to a GAMYB motif in the promoter region of *CESA* genes (Huang et al., 2015). OsMYB55/61 may contribute to the coordination of both cellulose and lignin biosynthesis in secondary wall formation.

AtMYB103, a direct transcriptional target of AtSWNs (SND1, NST1/2, and VND6/7) and AtMYB46/83, has been shown to interact with the promoter of a secondary wall-related cellulose synthesis gene *AtCESA8* in an Arabidopsis leaf protoplast transactivation system (Zhong et al., 2008; Ohashi-Ito et al., 2010; Yamaguchi et al., 2010). Interestingly, the Arabidopsis *atmyb103* mutant exhibits specific alteration of lignin composition via reduction of the expression of one lignin biosynthesis gene, *AtF5H*, but is not affected in the total lignin or cellulose content (Ohman et al., 2013). However, current evidence does not support the direct linkage of MYB103 and F5H in grasses. In rice, overexpression of OsMYB103 leads to increased cellulose content and enhanced secondary wall accumulation in sclerenchyma, whereas downregulating OsMYB103 results in decreased cellulose content and thinner cell walls in sclerenchyma, which together contribute to weakened mechanical strength of the culm (Hirano et al., 2013b; Yang et al., 2014; Ye et al., 2015). Microarray and transactivation experiments reveal that OsMYB103 significantly activates the expression of three secondary wall-related cellulose synthesis genes (*OsCESA4*, *OsCESA7*, and *OsCESA9*) and one secondary wall-related cellulose deposition gene (*OsBC1*) (Yang et al., 2014; Ye et al., 2015). Notably, a gibberellin (GA) signaling repressor (SLENDER RICE1, OsSLR1) shows physical interaction with OsMYB103 (Yang et al., 2014; Ye et al., 2015), and OsSWN2 and OsSWN3, which subsequently activate the expression of OsMYB55/61 (Huang et al., 2015). This suggests that

OsMYB55/61 and OsMYB103 may play a role in controlling GA-mediated secondary wall biosynthesis (Yang et al., 2014; Huang et al., 2015; Ye et al., 2015) (**Figure 1**).

A MYB Clade of Repressors of Secondary Wall Accumulation

Genes in the clade of MYB4/32 are proposed to be negative regulators of secondary wall biosynthesis in vascular plants (Zhao and Bartley, 2014) (**Supplementary Figure S2**). More accurately, the MYB4/32 clade should be considered as a controller that shifts the flux from the phenylpropanoid pathway to other metabolic pathways. In Arabidopsis, AtMYB4 suppresses the expression level of cinnamate-4-hydroxylase (*C4H*) and 4-coumarate:CoA ligase (*4CL*) genes, rather than other lignin biosynthesis genes, to control the accumulation of sinapate esters in response to ultraviolet-B (UV-B) irradiation (Jin et al., 2000). The AtMYB4 overexpressing Arabidopsis line has a decreased content of sinapate esters, with no effect on flavonoid composition (Jin et al., 2000). However, AtMYB7 and AtMYB32, two paralogs of AtMYB4, repress and induce genes involved in the flavonoid pathway, respectively (Preston et al., 2004; Fornalé et al., 2014); loss of function of AtMYB7 and AtMYB32 lead to notable induction of flavonoid content and alteration of pollen wall composition, respectively (Preston et al., 2004; Fornalé et al., 2014).

In contrast, AtMYB4 homologs in grasses have been observed to function in a more lineage-specific fashion for the regulation of lignin biosynthesis genes (Agarwal et al., 2016). In maize, ChIP-seq and coimmunoprecipitation (co-IP) assays revealed that ZmMYB11, ZmMYB31, and ZmMYB42 down-regulated different lignin biosynthesis genes, with the commonalities of *COMT* and *4CL2* (Vélez-Bermúdez et al., 2015). Exogenous overexpression of ZmMYB31 and ZmMYB42 in Arabidopsis redirected the phenylpropanoid flux by downregulating different lignin biosynthesis genes compared to maize (Fornalé et al., 2006, 2010; Sonbol et al., 2009; Vélez-Bermúdez et al., 2015). For example, ZmMYB31 and ZmMYB42 do not repress *ZmF5H* in maize, but do repress *AtF5H* in Arabidopsis, which causes decreased lignin S/G ratio (S, syringyl units; G, guaiacyl units) (Sonbol et al., 2009; Fornalé et al., 2010; Vélez-Bermúdez et al., 2015). Comparatively, overexpression of PvMYB4 in tobacco results in significantly reduced expression of 10 lignin biosynthesis genes leading to reduced lignin content and higher S/G ratio, whereas overexpressing PvMYB4 in switchgrass does not alter lignin composition (Shen et al., 2012). This suggests that grasses may utilize different regulatory mechanisms using MYB4/32 clade TFs to balance the flux between the lignin and flavonoid pathways.

Interestingly, although MYB4/32 homologs in grasses predominantly recognize a conserved domain in the promoter of target genes, they inhibit different phenylpropanoid genes within grass lineages (Shen et al., 2012; Vélez-Bermúdez et al., 2015; Agarwal et al., 2016). In maize, sorghum and rice leaves, MYB4/32 syntelogs share the common target of O-methyltransferase (*COMT1*), but display divergent binding

to the promoters of 4-coumarate-CoA ligase (*4CL2*), ferulate-5-hydroxylase (*F5H*), and caffeoyl shikimate esterase (*CSE*) (Agarwal et al., 2016). This suggests that genes in the MYB4/32 clade may have undergone sub-functionalization for fine-tuning of phenylpropanoid flux in some grass lineages (Agarwal et al., 2016).

WRKY12 as Repressor

WRKY12, a member of group IIc of the WRKY TF family (Rushton et al., 2010; Phukan et al., 2016), has been shown to control pith cell maintenance through repressing lignification in pith cell walls in dicots (Wang et al., 2010; Gallego-Giraldo et al., 2016; Yang et al., 2016). In Arabidopsis, *M. truncatula*, *Populus*, and alfalfa (*M. sativa*) a reduction of WRKY12 expression leads to an enhanced and/or ectopic deposition of secondary cell walls in the pith cells of the stem (Wang et al., 2010; Gallego-Giraldo et al., 2016; Yang et al., 2016). Similarly, alteration in secondary cell wall deposition is observed on down-regulation WRKY12 orthologs in switchgrass and maize (Gallego-Giraldo et al., 2016). This suggests a conserved function of WRKY12 as a repressor of secondary cell wall accumulation in grasses and dicots.

In *Medicago* and Arabidopsis, WRKY12 inhibits secondary wall formation though directly binding to the promoter of NST2, while the expression of WRKY12 is auto-repressed (Wang et al., 2010), which is a feature of WRKY signaling (Rushton et al., 2010; Phukan et al., 2016). In grasses, we suggest that WRKY12 may serve as a repressor in a similar way by down-regulating SWNs and auto-regulating itself.

KNOX, BEL, and OFP Groups of TFs Involved in Secondary Wall Accumulation in GA-Signaling and Organ Development

KNOX and BEL, two subclasses belonging to the TALE (Three Amino acid Loop Extension) homeodomain superclass, are the oldest TF groups diversely represented across the plant kingdom including green and red algae (Mukherjee et al., 2009). Members of class I KNOX genes from Arabidopsis, *Populus*, peach, maize, and switchgrass have been identified to be negative regulators of secondary wall deposition (Mukherjee et al., 2009; Hay and Tsiantis, 2010; Li E. et al., 2012; Townsley et al., 2013; Gong et al., 2014; Liu et al., 2014; Wuddineh et al., 2016). Overexpression of ZmKN1 in maize and tobacco significantly reduced the lignin content and altered lignin composition (Townsley et al., 2013). Partially similarly, a switchgrass PvKN1 (the ortholog of maize ZmKN1) overexpressing line displayed abnormal growth with a slightly reduced lignin content, and altered expression of some structural genes involved in cellulose, hemicellulose, and lignin biosynthesis (Wuddineh et al., 2016). Notably, ChIP-seq and qRT-PCR experiments revealed that ZmKN1 and PvKN1 can reduce the expression of the GA 20-oxidase (*GA20ox*, GA synthesis) gene while inducing the expression of GA 2-oxidase (*GA2ox*, GA catabolism), suggesting the roles of the TFs in modulation of GA signaling and maintenance of GA homeostasis (Bolduc et al., 2012; Wuddineh et al., 2016).

Arabidopsis AtKNAT7, a member of the class II KNOX gene family, and AtBLH6, a member of the BEL gene

family, are both also considered as repressors in secondary cell wall biosynthesis (Li E. et al., 2012; Liu et al., 2014). Furthermore, interactions between AtKNAT7/AtBLH6 and AtOPF1 and AtOPF4, two members of the OPF (OVATE FAMILY PROTEIN) family, result in heterodimeric complexes with enhanced activity to repress secondary wall thickening in the interfascicular fibers of inflorescence stems (Li E. et al., 2012; Liu et al., 2014; Liu and Douglas, 2015). As mentioned above, the expression of AtKNAT7 can be induced by the secondary wall-activators AtMYB46/83 and AtMYB61 (Zhong et al., 2008; Romano et al., 2012). This suggests that AtKNAT and the formation of the AtKNAT7-AtBLH6-OPFs multi-protein complex contribute to a negative feedback loop for fine tuning of secondary wall biosynthesis (Li E. et al., 2012; Liu and Douglas, 2015). In rice, overexpressing OsOPF2 causes disruption of vascular bundle arrangement in the stem and lower GA content, through alteration of gene expression associated with lignin biosynthesis, vascular development, and GA synthesis (Schmitz et al., 2015). In addition, yeast two-hybrid assays have proven the interactions between OsOPF2, OsKNAT7 and OsBLH6-like 1 and OsBLH6-like 2 (Schmitz et al., 2015). Considering that OPF is a land plant-specific TF family (Wang et al., 2016), it has been proposed that grasses and dicots have evolved OPF TFs which interact with KNOX and BEL members rooted from the last common ancestors with non-vascular plants, to control vascular development through suppression of GA and lignin biosynthesis (Figure 1).

In addition, the BEL-type homeodomain genes contribute to controlling lignin biosynthesis in replum development and seed shattering. AtBLH9 (also named as *REPLUMLESS*, RPL) is a key regulator for determining the orientation of stem growth (Bencivenga et al., 2016). The target genes of AtBLH9 identified by genome-wide ChIP-seq include *AtBGLU45* encoding stem-specific monolignol β -glucosidase (Chapelle et al., 2012) and the S-lignin biosynthesis-specific gene *AtF5H* (Bencivenga et al., 2016). Similarly, the high expression of OsSH5 (the homolog of AtBLH9 in rice) in rice pedicels inhibits the accumulation of lignin content by repressing the expression of lignin biosynthesis genes in the abscission zone (Yoon et al., 2014). This suggests that OsSH5 and AtBLH9 may play similar roles in repressing lignin biosynthesis during organ development.

Though a conserved function of members of the TALE and OPF families may be shared in grasses and dicots for the regulation of secondary wall accumulation, some differences are observed. Overexpression of AtBLH6 in Arabidopsis causes a reduction of secondary wall thickness in interfascicular fibers and a significant repression of stem growth (Liu et al., 2014). Interestingly, overexpressing OsBLH6, the third ortholog of Arabidopsis AtBLH6, in rice causes enhanced secondary wall-development in the stem but similar plant growth compared with the control; an OsBLH6 knock down line exhibits reduced lignin content, especially in the sclerenchyma in stems (Hirano et al., 2013b). The opposite direction of regulation by rice and Arabidopsis BLH6 orthologs in secondary wall accumulation suggests

that they have undergone functional specialization after gene duplication.

C2H2 Group TFs Involved in Secondary Wall Formation

Besides the NAC, MYB and TALE families, the C2H2 family is listed in TF families that have the most abundant members co-expressed with secondary cell wall structural genes in rice and Arabidopsis (Hirano et al., 2013a). One C2H2 member named OsIDD2 was proven to be a negative regulator of secondary wall formation (Huang et al., 2017). Overexpressing OsIDD2 in rice decreases the lignin content with a reduced expression of several structural genes involved in lignin, cellulose, and sucrose biosynthesis. The direct repression of OsCAD2 and OsCAD3 expression by OsIDD2 indicates its negative role in lignin accumulation (Huang et al., 2017).

The E2Fc Group of TFs Coupling Secondary Wall Initiation and Cell Cycle Exit

Secondary cell wall formation is coupled with cell cycle exit because secondary walls are deposited in the cells during the phase when growth stops and differentiation begins (Kwok and Wong, 2003). The inhibition of genes involved in cell division and the activation of genes involved in secondary cell wall biosynthesis occur at the same time in hormone-induced suspension cells of Arabidopsis and switchgrass (Pauwels et al., 2008; Rao et al., 2017). In Arabidopsis, E2Fc, a member of the E2F family, is considered to play a dual regulatory role in cell proliferation and secondary wall formation (del Pozo et al., 2002, 2006; Heckmann et al., 2011). E2Fc and its variants are capable of directly binding to the promoter of the centromere-specific histone in a cell cycle-dependent manner, and to the promoters of several secondary cell wall biosynthesis genes (Taylor-Teeple et al., 2015). E2Fc can activate the expression of VND7 in a dose-dependent manner (Taylor-Teeple et al., 2015), which further triggers a rapid cell death-program and secondary cell wall initiation in the tracheary element-differentiation process (Yamaguchi et al., 2010; Bollhoner et al., 2012). Four *E2F* genes, homologs of known cell cycle regulators in Arabidopsis, show tight coexpression with lignin biosynthesis genes in the time-course of secondary cell wall formation induced by the hormone brassinosteroid in switchgrass suspension cells (Rao et al., 2017). Their role in secondary cell wall formation and cell proliferation is worth exploring in the future.

Lineage-Specific TFs

Besides the examples mentioned above, additional functional divergences in cell wall regulation have been reported in monocots and dicots. For instance, homologous overexpression of the *SHN* gene in Arabidopsis (Aharoni et al., 2004; Shi et al., 2011), rice (Wang et al., 2012; Zhou et al., 2014), and switchgrass (Wuddineh et al., 2015) provides evidence for its function in wax biosynthesis. However, the heterologous expression of *AtSHN* in rice resulted in the downregulation of lignin biosynthesis genes

TABLE 1 | Summary of the commonalities and differences in transcriptional regulation of secondary wall formation in Arabidopsis and grasses.

TF group	Arabidopsis	Grasses	Reference
SWN	Binds to the SNBE motif; activates MYB46/83 and lower-level MYBs; induces cellulose, xylan, and lignin		Zhao et al., 2010; Zhong et al., 2010a,b, 2011, 2015; Wang et al., 2011; Valdivia et al., 2013; Yoshida et al., 2013; Xiao et al., 2017
	Some SWNs are expressed only in vessels or fibers	Expressed in both vessels and fibers	Zhong et al., 2010a, 2011, 2015; Valdivia et al., 2013; Yoshida et al., 2013; Xiao et al., 2017
	SND1 shows auto-activation	Rice OsSWN2 variant may negatively regulate OsSWN1 and OsSWN2	Wang et al., 2011; Yoshida et al., 2013
MYB46/83	Binds to the SMRE motif; activates lower-level MYBs; induces cellulose, xylan, and lignin		Zhong et al., 2007a, 2011, 2015; McCarthy et al., 2009; Ko et al., 2012; Kim et al., 2013
MYB58/63	Specifically induces lignin	Regulates both lignin and secondary wall-related cellulose in rice and sorghum	Zhou et al., 2009; Zhao and Dixon, 2011; Noda et al., 2015; Scully et al., 2016
MYB55/61	Activates genes involved in lignin biosynthesis (CCoAOMT7) and pectin modification (PME)	Activates lignin and cellulose biosynthesis in rice	Newman et al., 2004; Romano et al., 2012; Hirano et al., 2013b; Huang et al., 2015
MYB103	Alters lignin composition via regulation of AtF5H	No evidence to show specific regulation of F5H; increases total lignin and cellulose content in rice	Hirano et al., 2013b; Ohman et al., 2013; Yang et al., 2014; Ye et al., 2015
MYB4/32	MYB4 represses the lignin biosynthesis genes C4H and 4CL	Share common target of COMT but display lineage-specific suppression of other lignin biosynthesis genes in grasses	Jin et al., 2000; Sonbol et al., 2009; Fornalé et al., 2010; Shen et al., 2012; Vélez-Bermúdez et al., 2015; Agarwal et al., 2016
WRKY12	Represses lignification		Wang et al., 2010; Gallego-Giraldo et al., 2016; Yang et al., 2016
KNOX	Negatively regulates secondary wall deposition		Mukherjee et al., 2009; Hay and Tsiantis, 2010; Li E. et al., 2012; Townsley et al., 2013; Gong et al., 2014; Liu et al., 2014; Wuddineh et al., 2016
KNAT/BEL/OPF	May form KNAT-BLH-OPF multi-protein complex as a negative regulator of secondary wall development		Li E. et al., 2012; Liu et al., 2014; Liu and Douglas, 2015; Schmitz et al., 2015
BLH9	Represses lignin accumulation		Chapelle et al., 2012; Yoon et al., 2014; Bencivenga et al., 2016
BLH6	Represses secondary wall development	Induces secondary wall development in rice	Hirano et al., 2013b; Liu et al., 2014
C2H2	NA*	OsIDD2 represses secondary wall development in rice	Huang et al., 2017
E2Fc	Triggers cell proliferation and secondary wall formation	NA*	del Pozo et al., 2002, 2006; Heckmann et al., 2011; Taylor-Teeples et al., 2015

NA*, members of the C2H2 family in Arabidopsis and E2Fc in grasses show co-expression with secondary wall structural genes but have not been proven to be regulators of secondary wall formation (Hirano et al., 2013a; Rao et al., 2017).

and upregulation of cellulose synthesis genes, which led to a significant increase in cellulose and decrease in lignin content (Ambavaram et al., 2011). The discrepancy between homologous and heterologous expression of SHN may reflect the divergence between regulatory mechanisms of cell wall development in monocots and dicots.

Though sharing many TFs rooted from the last common ancestor, monocots and dicots have developed some lineage-specific TFs through gene duplication. For instance, AtMYB75 belongs to a dicot-specific group with orthologs found in poplar, but not in grasses (Zhao and Bartley, 2014). AtMYB75 is a master switch to control the shift from secondary wall formation to anthocyanin accumulation via repression of secondary wall-related cellulose synthase genes and lignin biosynthesis genes

and activation of the late anthocyanin biosynthetic genes in a light-dependent manner (Bhargava et al., 2010; Li et al., 2016). Monocots may have evolved other TFs participating in light-controlled secondary wall accumulation. Zhao and Bartley (2014) have identified several grass-specific TF clades. Candidates in these clades may have potential roles in secondary wall regulation in a grass-specific manner.

Recent tissue-specific and time-course transcriptome analyses from sorghum, *Miscanthus lutarioriparius*, and switchgrass have revealed 100s of TF genes whose expression is highly correlated with the dynamic process of lignification (Hu et al., 2017; Kebrom et al., 2017; Rao et al., 2017; Yan et al., 2017), providing more TF candidates in grasses for future functional identification.

CONCLUSION

According to current research, grasses and dicots share a conserved transcriptional regulatory network for secondary wall biosynthesis, nevertheless with many grass-specific features. The differences and commodities in the transcriptional networks for secondary cell wall regulation between *Arabidopsis* and grasses are summarized in **Figure 1** and **Table 1**, and details of the individual TFs are listed in **Supplementary Table S1**. The differences may be caused by changes in spatial expression of TFs, *cis*-regulatory element composition of structural genes and sub-functionalization after gene duplication. Considering the economic and ecological importance of the grass family, further research is needed to better understand the grass-specific transcriptional regulation of secondary cell wall development.

AUTHOR CONTRIBUTIONS

XR collected data from literature and wrote the manuscript. RD revised the article.

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
- Aharoni, A., Dixit, S., Jetter, R., Thoenes, E., van Arkel, G., and Pereira, A. (2004). The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in *Arabidopsis*. *Plant Cell* 16, 2463–2480. doi: 10.1105/tpc.104.022897
- 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
- Bencivenga, S., Serrano-Mislata, A., Bush, M., Fox, S., and Sablowski, R. (2016). Control of oriented tissue growth through repression of organ boundary genes promotes stem morphogenesis. *Dev. Cell* 39, 198–208. doi: 10.1016/j.devcel.2016.08.013
- Bhargava, A., Mansfield, S. D., Hall, H. C., Douglas, C. J., and Ellis, B. E. (2010). MYB75 functions in regulation of secondary cell wall formation in the *Arabidopsis* inflorescence stem. *Plant Physiol.* 154, 1428–1438. doi: 10.1104/pp.110.162735
- Bolduc, N., Yilmaz, A., Mejia-Guerra, M. K., Morohashi, K., O'Connor, D., Grotewold, E., et al. (2012). Unraveling the KNOTTED1 regulatory network in maize meristems. *Genes Dev.* 26, 1685–1690. doi: 10.1101/gad.193433.112
- Bollhoner, B., Prestele, J., and Tuominen, H. (2012). Xylem cell death: emerging understanding of regulation and function. *J. Exp. Bot.* 63, 1081–1094. doi: 10.1093/jxb/err438
- Chapelle, A., Morreel, K., Vanholme, R., Le-Bris, P., Morin, H., Lapierre, C., et al. (2012). Impact of the absence of stem-specific beta-glucosidases on lignin and monolignols. *Plant Physiol.* 160, 1204–1217. doi: 10.1104/pp.112.203364
- Chaw, S. M., Chang, C. C., Chen, H. L., and Li, W. H. (2004). Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *J. Mol. Evol.* 58, 424–441. doi: 10.1007/s00239-003-2564-9
- Cosgrove, D. J., and Jarvis, M. C. (2012). Comparative structure and biomechanics of plant primary and secondary cell walls. *Front. Plant Sci.* 3:204. doi: 10.3389/fpls.2012.00204
- del Pozo, J. C., Boniotti, M. B., and Gutierrez, C. (2002). *Arabidopsis* E2Fc functions in cell division and is degraded by the ubiquitin-SCF(AtSKP2) pathway in response to light. *Plant Cell* 14, 3057–3071. doi: 10.1105/tpc.006791
- del Pozo, J. C., Diaz-Trivino, S., Cisneros, N., and Gutierrez, C. (2006). The balance between cell division and endoreplication depends on E2FC-DPB, transcription factors regulated by the ubiquitin-SCF^{SKP2A} pathway in *Arabidopsis*. *Plant Cell* 18, 2224–2235. doi: 10.1105/tpc.105.039651
- Didi, V., Jackson, P., and Hejatko, J. (2015). Hormonal regulation of secondary cell wall formation. *J. Exp. Bot.* 66, 5015–5027. doi: 10.1093/jxb/erv222
- Dixon, R. A., Liu, C., and Jun, J. H. (2013). Metabolic engineering of anthocyanins and condensed tannins in plants. *Curr. Opin. Biotechnol.* 24, 329–335. doi: 10.1016/j.copbio.2012.07.004
- Fornalé, S., Lopez, E., Salazar-Henao, J. E., Fernandez-Nohales, P., Rigau, J., and Caparros-Ruiz, D. (2014). AtMYB7, a new player in the regulation of UV-screens in *Arabidopsis thaliana*. *Plant Cell Physiol.* 55, 507–516. doi: 10.1093/pcp/pct187
- Fornalé, 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.1365-3113.2010.04363.x
- Fornalé, S., Sonbol, F. M., Maes, T., Capellades, M., Puigdomenech, 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
- Gallego-Giraldo, L., Shadle, G., Shen, H., Barros-Rios, J., Fresquet Corrales, S., Wang, H., et al. (2016). Combining enhanced biomass density with reduced lignin level for improved forage quality. *Plant Biotechnol. J.* 14, 895–904. doi: 10.1111/pbi.12439
- Gong, S. Y., Huang, G. Q., Sun, X., Qin, L. X., Li, Y., Zhou, L., et al. (2014). Cotton KN1, encoding a class II KNOX transcription factor, is involved in regulation of fibre development. *J. Exp. Bot.* 65, 4133–4147. doi: 10.1093/jxb/eru182
- Goulao, L. F., Vieira-Silva, S., and Jackson, P. A. (2011). Association of hemicellulose- and pectin-modifying gene expression with *Eucalyptus globulus* secondary growth. *Plant Physiol. Biochem.* 49, 873–881. doi: 10.1016/j.plaphy.2011.02.020
- Groover, A., and Jones, A. M. (1999). Tracheary element differentiation uses a novel mechanism coordinating programmed cell death and secondary cell wall synthesis. *Plant Physiol.* 119, 375–384. doi: 10.1104/pp.119.2.375

FUNDING

This work was supported by the United States Department of Energy Bioenergy Sciences Center (BESC, grant # BER DE-AC05-00OR2727), through the Office of Biological and Environmental Research in the DOE Office of Science.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00399/full#supplementary-material>

FIGURE S1 | Phylogenetic analysis of SWNs from *Arabidopsis* and five grass species.

FIGURE S2 | Phylogenetic analysis of secondary wall-related MYBs from *Arabidopsis* and five grass species.

TABLE S1 | Transcription factors involved in grass secondary cell wall formation.

- Hay, A., and Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. *Development* 137, 3153–3165. doi: 10.1242/dev.030049
- Heckmann, S., Lermontova, I., Berckmans, B., De Veylder, L., Baumlein, H., and Schubert, I. (2011). The E2F transcription factor family regulates *CENH3* expression in *Arabidopsis thaliana*. *Plant J.* 68, 646–656. doi: 10.1111/j.1365-3113.2011.04715.x
- Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., et al. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315, 804–807. doi: 10.1126/science.1137016
- 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
- Houston, K., Tucker, M. R., Chowdhury, J., Shirley, N., and Little, A. (2016). The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. *Front. Plant Sci.* 7:984. doi: 10.3389/fpls.2016.00984
- Hu, R., Xu, Y., Yu, C., He, K., Tang, Q., Jia, C., et al. (2017). Transcriptome analysis of genes involved in secondary cell wall biosynthesis in developing internodes of *Miscanthus lutarioriparius*. *Sci. Rep.* 7:9034. doi: 10.1038/s41598-017-08690-8
- Huang, D., Wang, S., and Zhang, B. (2015). A gibberellin-mediated DELLA-NAC signaling cascade regulates cellulose synthesis in rice. *Plant Cell* 27, 1681–1696. doi: 10.1105/tpc.15.00015
- Huang, P., Yoshida, H., Yano, K., Kinoshita, S., Kawai, K., Koketsu, E., et al. (2017). OsIDD2, a zinc finger and INDETERMINATE DOMAIN protein, regulate secondary cell wall formation. *J. Integr. Plant Biol.* 60, 130–143. doi: 10.1111/jipb.12557
- 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
- Kebrom, T. H., McKinley, B., and Mullet, J. E. (2017). Dynamics of gene expression during development and expansion of vegetative stem internodes of bioenergy sorghum. *Biotechnol. Biofuels* 10:159. doi: 10.1186/s13068-017-0848-3
- Kim, W. C., Kim, J. Y., Ko, J. H., Kim, J., and Han, K. H. (2013). Transcription factor MYB46 is an obligate component of the transcriptional regulatory complex for functional expression of secondary wall-associated cellulose synthases in *Arabidopsis thaliana*. *J. Plant Physiol.* 170, 1374–1378. doi: 10.1016/j.jplph.2013.04.012
- Ko, J. H., Kim, W. C., Kim, J. Y., Ahn, S. J., and Han, K. H. (2012). MYB46-mediated transcriptional regulation of secondary wall biosynthesis. *Mol. Plant* 5, 961–963. doi: 10.1093/mp/sss076
- Kwok, A. C., and Wong, J. T. (2003). Cellulose synthesis is coupled to cell cycle progression at G1 in the dinoflagellate *Cryptothecodinium cohnii*. *Plant Physiol.* 131, 1681–1691. doi: 10.1104/pp.102.018945
- Li, E., Bhargava, A., Qiang, W., Friedmann, M. C., Forneris, N., Savidge, R. A., et al. (2012). The Class II KNOX gene *KNAT7* negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in *Populus*. *New Phytol.* 194, 102–115. doi: 10.1111/j.1469-8137.2011.04016.x
- Li, Q., Lin, Y. C., Sun, Y. H., Song, J., Chen, H., Zhang, X. H., et al. (2012). Splice variant of the SND1 transcription factor is a dominant negative of SND1 members and their regulation in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14699–14704. doi: 10.1073/pnas.1212977109
- Li, S., Wang, W., Gao, J., Yin, K., Wang, R., Wang, C., et al. (2016). MYB75 phosphorylation by MPK4 is required for light-induced anthocyanin accumulation in *Arabidopsis*. *Plant Cell* 28, 2866–2883. doi: 10.1105/tpc.16.00130
- Li, X., and Chapple, C. (2010). Understanding lignification: challenges beyond monolignol biosynthesis. *Plant Physiol.* 154, 449–452. doi: 10.1104/pp.110.162842
- Lin, Y. J., Chen, H., Li, Q., Li, W., Wang, J. P., Shi, R., et al. (2017). Reciprocal cross-regulation of VND and SND multigene TF families for wood formation in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9722–E9729. doi: 10.1073/pnas.1714422114
- Liu, Y., and Douglas, C. J. (2015). A role for OVATE FAMILY PROTEIN1 (OFP1) and OFP4 in a BLH6-KNAT7 multi-protein complex regulating secondary cell wall formation in *Arabidopsis thaliana*. *Plant Signal. Behav.* 10:e1033126. doi: 10.1080/15592324.2015.1033126
- Liu, Y., You, S., Taylor-Teeples, M., Li, W. L., Schuetz, M., Brady, S. M., et al. (2014). BEL1-LIKE HOMEODOMAIN6 and KNOTTED ARABIDOPSIS THALIANA7 interact and regulate secondary cell wall formation via repression of *REVOLUTA*. *Plant Cell* 26, 4843–4861. doi: 10.1105/tpc.114.128322
- Loque, D., Scheller, H. V., and Pauly, M. (2015). Engineering of plant cell walls for enhanced biofuel production. *Curr. Opin. Plant Biol.* 25, 151–161. doi: 10.1016/j.pbi.2015.05.018
- 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
- 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
- Mukherjee, K., Brocchieri, L., and Burglin, T. R. (2009). A comprehensive classification and evolutionary analysis of plant homeobox genes. *Mol. Biol. Evol.* 26, 2775–2794. doi: 10.1093/molbev/msp201
- Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N. A., and Ohtani, M. (2015). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* 6:288. doi: 10.3389/fpls.2015.00288
- 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.1365-3113.2003.01953.x
- Noda, S., Koshiba, T., Hattori, T., Yamaguchi, M., Suzuki, S., and Umezawa, T. (2015). The expression of a rice secondary wall-specific cellulose synthase gene, OsCesA7, is directly regulated by a rice transcription factor, OsMYB58/63. *Planta* 242, 589–600. doi: 10.1007/s00425-015-2343-z
- Ohashi-Ito, K., Oda, Y., and Fukuda, H. (2010). *Arabidopsis* VASCULAR-RELATED NAC-DOMAIN6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. *Plant Cell* 22, 3461–3473. doi: 10.1105/tpc.110.075036
- Ohman, D., Demedts, B., Kumar, M., Gerber, L., Gorzsas, A., Goeminne, G., et al. (2013). MYB103 is required for *FERULATE-5-HYDROXYLASE* expression and syringyl lignin biosynthesis in *Arabidopsis* stems. *Plant J.* 73, 63–76. doi: 10.1111/tpj.12018
- Pauly, M., and Keegstra, K. (2010). Plant cell wall polymers as precursors for biofuels. *Curr. Opin. Plant Biol.* 13, 305–312. doi: 10.1016/j.pbi.2009.12.009
- Pauwels, L., Morreel, K., De Witte, E., Lammertyn, F., Van Montagu, M., Boerjan, W., et al. (2008). Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1380–1385. doi: 10.1073/pnas.0711203105
- Phukan, U. J., Jeena, G. S., and Shukla, R. K. (2016). WRKY transcription factors: molecular regulation and stress responses in plants. *Front. Plant Sci.* 7:760. doi: 10.3389/fpls.2016.00760
- Polyn, S., Willems, A., and De Veylder, L. (2015). Cell cycle entry, maintenance, and exit during plant development. *Curr. Opin. Plant Biol.* 23, 1–7. doi: 10.1016/j.pbi.2014.09.012
- 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.1365-3113.2004.02280.x
- Rabinowicz, P. D., Braun, E. L., Wolfe, A. D., Bowen, B., and Grotewold, E. (1999). Maize R2R3 Myb genes: sequence analysis reveals amplification in the higher plants. *Genetics* 153, 427–444.
- Rao, X., Shen, H., Pattathil, S., Hahn, M. G., Gelineo-Albersheim, I., Mohnen, D., et al. (2017). Dynamic changes in transcriptome and cell wall composition

- underlying brassinosteroid-mediated lignification of switchgrass suspension cells. *Biotechnol. Biofuels* 10:266. doi: 10.1186/s13068-017-0954-2
- Rogers, L. A., Dubos, C., Cullis, I. F., Surman, C., Poole, M., Willment, J., et al. (2005). Light, the circadian clock, and sugar perception in the control of lignin biosynthesis. *J. Exp. Bot.* 56, 1651–1663. doi: 10.1093/jxb/eri162
- 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
- Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. (2010). WRKY transcription factors. *Trends Plant Sci.* 15, 247–258. doi: 10.1016/j.tplants.2010.02.006
- Schmitz, A. J., Begcy, K., Sarath, G., and Walia, H. (2015). Rice ovate family protein 2 (OFP2) alters hormonal homeostasis and vasculature development. *Plant Sci.* 241, 177–188. doi: 10.1016/j.plantsci.2015.10.011
- 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., 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
- Shi, J. X., Malitsky, S., De Oliveira, S., Branigan, C., Franke, R. B., Schreiber, L., et al. (2011). SHINE transcription factors act redundantly to pattern the archetypal surface of Arabidopsis flower organs. *PLoS Genet.* 7:e1001388. doi: 10.1371/journal.pgen.1001388
- Smith, A. M., and Stitt, M. (2007). Coordination of carbon supply and plant growth. *Plant Cell Environ.* 30, 1126–1149. doi: 10.1111/j.1365-3040.2007.01708.x
- 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
- Taiz, L., and Zeiger, E. (1998). *Cell Walls: Structure, Biogenesis, and Expansion*. Plant Physiology. Sunderland, MA: Sinauer Associates, 409–443.
- Taylor-Teeple, 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
- Townsley, B. T., Sinha, N. R., and Kang, J. (2013). KNOX1 genes regulate lignin deposition and composition in monocots and dicots. *Front. Plant Sci.* 4:121. doi: 10.3389/fpls.2013.00121
- Valdivia, E. R., Herrera, M. T., Gianzo, C., Fidalgo, J., Revilla, G., Zarra, I., et al. (2013). Regulation of secondary wall synthesis and cell death by NAC transcription factors in the monocot *Brachypodium distachyon*. *J. Exp. Bot.* 64, 1333–1343. doi: 10.1093/jxb/ers394
- Vélez-Bermúdez, I. C., Salazar-Henao, J. E., and Fornale, S. (2015). A MYB/ZML complex regulates wound-induced lignin genes in maize. *Plant Cell* 27, 3245–3259. doi: 10.1105/tpc.15.00545
- Wang, H., Avci, U., Nakashima, J., Hahn, M. G., Chen, F., and Dixon, R. A. (2010). Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22338–22343. doi: 10.1073/pnas.1016436107
- Wang, H., Zhao, Q., Chen, F., Wang, M., and Dixon, R. A. (2011). NAC domain function and transcriptional control of a secondary cell wall master switch. *Plant J.* 68, 1104–1114. doi: 10.1111/j.1365-3113X.2011.04764.x
- Wang, S., Chang, Y., and Ellis, B. (2016). Overview of OVATE FAMILY PROTEINS, a novel class of plant-specific growth regulators. *Front. Plant Sci.* 7:417. doi: 10.3389/fpls.2016.00417
- Wang, Y., Wan, L., Zhang, L., Zhang, Z., Zhang, H., Quan, R., et al. (2012). An ethylene response factor OsWRI responsive to drought stress transcriptionally activates wax synthesis related genes and increases wax production in rice. *Plant Mol. Biol.* 78, 275–288. doi: 10.1007/s11103-011-9861-2
- Wuddineh, W. A., Mazarei, M., Turner, G. B., Sykes, R. W., Decker, S. R., Davis, M. F., et al. (2015). Identification and molecular characterization of the switchgrass AP2/ERF transcription factor superfamily, and overexpression of PvERF001 for improvement of biomass characteristics for biofuel. *Front. Bioeng. Biotechnol.* 3:101. doi: 10.3389/fbioe.2015.00101
- Wuddineh, W. A., Mazarei, M., Zhang, J. Y., Turner, G. B., Sykes, R. W., Decker, S. R., et al. (2016). Identification and overexpression of a knotted1-like transcription factor in switchgrass (*Panicum virgatum* L.) for lignocellulosic feedstock improvement. *Front. Plant Sci.* 7:520. doi: 10.3389/fpls.2016.00520
- Xiao, W., Yang, Y., and Yu, J. (2017). ZmNST3 and ZmNST4 are master switches for secondary wall deposition in maize (*Zea mays* L.). *Plant Sci.* 266, 83–94. doi: 10.1016/j.plantsci.2017.03.012
- Yamaguchi, M., Goue, N., Igarashi, H., Ohtani, M., Nakano, Y., Mortimer, J. C., et al. (2010). VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant Physiol.* 153, 906–914. doi: 10.1104/pp.110.154013
- Yan, H., Zhang, A., Ye, Y., Xu, B., Chen, J., He, X., et al. (2017). Genome-wide survey of switchgrass NACs family provides new insights into motif and structure arrangements and reveals stress-related and tissue-specific NACs. *Sci. Rep.* 7:3056. doi: 10.1038/s41598-017-03435-z
- Yang, C., Li, D., Liu, X., Ji, C., Hao, L., Zhao, X., et al. (2014). OsMYB103L, an R2R3-MYB transcription factor, influences leaf rolling and mechanical strength in rice (*Oryza sativa* L.). *BMC Plant Biol.* 14:158. doi: 10.1186/1471-2229-14-158
- Yang, L., Zhao, X., Yang, F., Fan, D., Jiang, Y., and Luo, K. (2016). PtrWRKY19, a novel WRKY transcription factor, contributes to the regulation of pith secondary wall formation in *Populus trichocarpa*. *Sci. Rep.* 6:18643. doi: 10.1038/srep18643
- Yao, D., Wei, Q., Xu, W., Syrenne, R. D., Yuan, J. S., and Su, Z. (2012). Comparative genomic analysis of NAC transcriptional factors to dissect the regulatory mechanisms for cell wall biosynthesis. *BMC Bioinformatics* 13(Suppl. 15):S10. doi: 10.1186/1471-2105-13-s15-s10
- Ye, Y., Liu, B., Zhao, M., Wu, K., Cheng, W., Chen, X., et al. (2015). CEF1/OsMYB103L is involved in GA-mediated regulation of secondary wall biosynthesis in rice. *Plant Mol. Biol.* 89, 385–401. doi: 10.1007/s11103-015-0376-0
- Yoon, J., Cho, L. H., Kim, S. L., Choi, H., Koh, H. J., and An, G. (2014). The BEL1-type homeobox gene SH5 induces seed shattering by enhancing abscission-zone development and inhibiting lignin biosynthesis. *Plant J.* 79, 717–728. doi: 10.1111/tpj.12581
- Yoshida, K., Sakamoto, S., Kawai, T., Kobayashi, Y., Sato, K., Ichinose, Y., et al. (2013). Engineering the *Oryza sativa* cell wall with rice NAC transcription factors regulating secondary wall formation. *Front. Plant Sci.* 4:383. doi: 10.3389/fpls.2013.00383
- Zhao, K., and Bartley, L. E. (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. (2011). Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends Plant Sci.* 16, 227–233. doi: 10.1016/j.tplants.2010.12.005
- Zhao, Q., Gallego-Giraldo, L., Wang, H., Zeng, Y., Ding, S. Y., Chen, F., et al. (2010). An NAC transcription factor orchestrates multiple features of cell wall development in *Medicago truncatula*. *Plant J.* 63, 100–114. doi: 10.1111/j.1365-3113X.2010.04223.x
- Zhong, R., Demura, T., and Ye, Z. H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*. *Plant Cell* 18, 3158–3170. doi: 10.1105/tpc.106.047399
- 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. (2010a). Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends Plant Sci.* 15, 625–632. doi: 10.1016/j.tplants.2010.08.007

- Zhong, R., Lee, C., and Ye, Z. H. (2010b). 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. (2007a). 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., Richardson, E. A., and Ye, Z. H. (2007b). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis*. *Planta* 225, 1603–1611. doi: 10.1007/s00425-007-0498-y
- Zhong, R., and Ye, Z. H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.* 53, 368–380. doi: 10.1093/pcp/pcr185
- Zhong, R., and Ye, Z. H. (2015). Secondary cell walls: biosynthesis, pattern deposition and transcriptional regulation. *Plant Cell Physiol.* 56, 195–214. doi: 10.1093/pcp/pcu140
- 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, X., Jenks, M. A., Liu, J., Liu, A., Zhang, X., Xiang, J., et al. (2014). Overexpression of transcription factor OsWR2 regulates wax and cutin biosynthesis in rice and enhances its tolerance to water deficit. *Plant Mol. Biol. Report.* 32, 719–731. doi: 10.1007/s11105-013-0687-8

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Rao and Dixon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

Edited by:

Charles T. Anderson,
Pennsylvania State University,
United States

Reviewed by:

Rui Shi,
North Carolina State University,
United States
Hairong Wei,
Michigan Technological University,
United States
Miguel Blazquez,
Consejo Superior de Investigaciones
Científicas (CSIC), Spain

*Correspondence:

Hannele Tuominen
hannele.tuominen@umu.se

†Present Address:

Carolin Seyfferth,
Department of Plant Physiology, Umeå
Plant Science Centre, Umeå
University, Umeå, Sweden
Soile Jokipii-Lukkari,
Department of Agricultural Sciences,
Viikki Plant Science Centre, University
of Helsinki, Helsinki, Finland

‡These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Plant Physiology,
a section of the journal
Frontiers in Plant Science

Received: 23 October 2017

Accepted: 16 February 2018

Published: 14 March 2018

Citation:

Seyfferth C, Wessels B,
Jokipii-Lukkari S, Sundberg B,
Delhomme N, Felten J and
Tuominen H (2018) Ethylene-Related
Gene Expression Networks in Wood
Formation. *Front. Plant Sci.* 9:272.
doi: 10.3389/fpls.2018.00272

Ethylene-Related Gene Expression Networks in Wood Formation

**Carolin Seyfferth^{1††}, Bernard Wessels^{2†}, Soile Jokipii-Lukkari^{2†}, Björn Sundberg¹,
Nicolas Delhomme¹, Judith Felten¹ and Hannele Tuominen^{2*}**

¹ Department of Forest Genetics and Plant Physiology, Umeå Plant Science Centre, Swedish University of Agricultural Sciences, Umeå, Sweden, ² Department of Plant Physiology, Umeå Plant Science Centre, Umeå University, Umeå, Sweden

Thickening of tree stems is the result of secondary growth, accomplished by the meristematic activity of the vascular cambium. Secondary growth of the stem entails developmental cascades resulting in the formation of secondary phloem outwards and secondary xylem (i.e., wood) inwards of the stem. Signaling and transcriptional reprogramming by the phytohormone ethylene modifies cambial growth and cell differentiation, but the molecular link between ethylene and secondary growth remains unknown. We addressed this shortcoming by analyzing expression profiles and co-expression networks of ethylene pathway genes using the AspWood transcriptome database which covers all stages of secondary growth in aspen (*Populus tremula*) stems. ACC synthase expression suggests that the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) is synthesized during xylem expansion and xylem cell maturation. Ethylene-mediated transcriptional reprogramming occurs during all stages of secondary growth, as deduced from AspWood expression profiles of ethylene-responsive genes. A network centrality analysis of the AspWood dataset identified *EIN3D* and 11 *ERFs* as hubs. No overlap was found between the co-expressed genes of the *EIN3* and *ERF* hubs, suggesting target diversification and hence independent roles for these transcription factor families during normal wood formation. The *EIN3D* hub was part of a large co-expression gene module, which contained 16 transcription factors, among them several new candidates that have not been earlier connected to wood formation and a VND-INTERACTING 2 (VNI2) homolog. We experimentally demonstrated *Populus EIN3D* function in ethylene signaling in *Arabidopsis thaliana*. The *ERF* hubs *ERF118* and *ERF119* were connected on the basis of their expression pattern and gene co-expression module composition to xylem cell expansion and secondary cell wall formation, respectively. We hereby establish data resources for ethylene-responsive genes and potential targets for *EIN3D* and *ERF* transcription factors in *Populus* stem tissues, which can help to understand the range of ethylene targeted biological processes during secondary growth.

Keywords: ethylene signaling, secondary growth, wood development, co-expression network, *EIN3*, *ERF*

INTRODUCTION

Woody tissue serves as plant stabilizing material, in nutrient storage and distribution of water and minerals. It is produced by the activity of the vascular cambium which undergoes periclinal cell divisions to produce secondary xylem (i.e., “wood”) inwards and secondary phloem outwards in the stem (Mellerowicz et al., 2001; Zhang et al., 2014). Xylem cells undergo four main phases of differentiation to achieve functional specialization; cell division, cell expansion (elongation and radial enlargement), secondary cell wall thickening (involving deposition of cellulose, hemicellulose, cell wall proteins, and lignin) and programmed cell death. Recent advances in genome sequencing of both gymnosperm and angiosperm tree species provide novel tools to dissect the complex nature of these phases (Tuskan et al., 2006; Birol et al., 2013; Nystedt et al., 2013; Myburg et al., 2014; Salojärvi et al., 2017). Especially powerful are the transcriptomic and proteomic datasets that have been generated in aspen (*Populus tremula*) and Norway spruce (*Picea abies*) (Obudulu et al., 2016; Bygdell et al., 2017; Jokipii-Lukkari et al., 2017; Sundell et al., 2017). Their high spatial resolution within the woody tissues allows analyses in specific phases of wood formation, from cell division to cell death. Transcriptomic datasets from aspen and Norway spruce (*Picea abies*) are also easily accessible in the form of the AspWood and NorWood databases (Jokipii-Lukkari et al., 2017; Sundell et al., 2017).

The activity of the vascular cambium as well as secondary xylem differentiation is regulated by various plant hormones. Whilst cytokinins control cambial cell division activity, auxin is related to both cambial cell proliferation and xylem differentiation (for a recent review see Zhang et al., 2014). Also gibberellins and brassinosteroids control specific aspects of xylem differentiation (Eriksson et al., 2000; Caño-Delgado et al., 2004). The gaseous plant hormone ethylene stimulates cambial growth and induces typical features of reaction wood (such as G-layer formation) when applied exogenously (Brown and Leopold, 1973; Du and Yamamoto, 2007; Love et al., 2009; Felten et al., 2018). Genetic evidence supports the role of ethylene as a stimulator of cambial activity in *Arabidopsis thaliana* (Etchells et al., 2012) and in gravitropically stimulated *Populus* trees (Love et al., 2009). A number of ethylene biosynthetic enzymes and transcription factors (TFs) have been shown to be expressed in woody tissues of *Populus* trees (Andersson-Gunnerås et al., 2003; Vahala et al., 2013), but description of the whole ethylene pathway in this context is lacking.

Ethylene is the product of stepwise conversion of the amino acid methionine to S-adenosylmethionine (S-AdoMet) by S-AdoMet synthetase (SAM), then to 1-aminocyclopropane-1-carboxylate (ACC) by ACC synthase (ACS) and finally to ethylene by ACC oxidase (ACO) (Figure 1A; recently reviewed in Vanderstraeten and Van Der Straeten, 2017). Perception of ethylene occurs at the endoplasmic reticulum (ER) by a family of membrane-bound ethylene receptors (ETR). The ethylene receptors associate with the serine/threonine protein kinase CONSTITUTIVE TRIPLE RESPONSE 1 (CTR1) (reviewed in Merchante et al., 2013; Xu and Zhang, 2014). In the absence of ethylene, the ETR-CTR1 complex represses downstream ethylene

signaling, while in the presence of ethylene ETR dissociates from CTR1 leading to activation of ethylene signaling by the downstream component ETHYLENE INSENSITIVE 2 (EIN2). Hereby, EIN2 is cleaved (Wen et al., 2012) and the C-terminus of EIN2 (EIN2-C) prevents proteasome-mediated degradation of the TFs ETHYLENE INSENSITIVE 3 (EIN3) and EIL1 (EIN3-like 1) by suppressing function of the two F-box proteins EIN3-BINDING F-BOX (EBF1) and EBF2 (Li et al., 2015; Merchante et al., 2015). An additional signal transduction mechanism is created by relocalization of EIN2-C into the nucleus and interaction with ENAP1 (EIN2 nuclear associated protein 1), which triggers EIN3-mediated transcriptional reprogramming due to EIN2-C regulated histone acetylation (Zhang et al., 2016, 2017). Although multiple EIN3-like proteins exist in *A. thaliana*, knock-out of *EIN3* alone is sufficient to perturb ethylene signaling (Chao et al., 1997). A second layer of transcriptional regulation of ethylene-responsive genes consists of TFs belonging to the large family of ETHYLENE RESPONSE FACTORS (ERFs) (Licausi et al., 2013; Müller and Munné-Bosch, 2015). Most of our current knowledge about the ethylene pathway comes from *A. thaliana*. In woody species like *Populus*, only ERFs are well described (Vahala et al., 2013; Wang et al., 2014; Yao et al., 2017). Better understanding of the composition, function and molecular regulation of the ethylene pathway components and downstream gene targets in woody species is needed to reveal the role of ethylene in wood formation.

Here, we aimed to elucidate the ethylene pathway in *Populus* trees by identifying the ethylene signaling components and putative downstream targets, and assaying their expression during various stages of wood formation. For this purpose, we utilized two datasets: first, the recently published AspWood database, encompassing a near-tissue level transcriptome data from the vascular cambium and its derivative tissues (Sundell et al., 2017). Exploring this resource enabled us to reveal how ethylene biosynthesis and signaling components are expressed during cambial growth and normal wood formation. In addition, we used RNA-Seq libraries obtained from wild type and ethylene-insensitive transgenic *Populus tremula* × *tremuloides* trees treated with the ethylene precursor ACC (Felten et al., 2018), thus representing rapid transcriptomic changes upon elevated ethylene signaling. This latter dataset allowed us to identify potential ethylene regulated target genes and to pinpoint the phases where ethylene-mediated transcriptional reprogramming occurs during normal wood formation. Since *Populus tremula* and *Populus tremula* × *tremuloides* were used for the AspWood and ACC-dataset, respectively, we decided to map the libraries from the latter one to the *Populus tremula* genome. Our analyses showed that the expression of ACC synthases is under strict spatial control in woody tissues, suggesting that ACC is synthesized specifically during xylem cell expansion and late maturation. Ethylene signaling, however, seems to occur in a more ubiquitous manner. A centrality analysis of the gene network led to the identification of key genes (so-called “hubs”) in the wood transcriptome, among them *EIN3D* and 11 ERFs. The majority of the hubs have not been previously connected to wood formation. Altogether, our study provides a comprehensive data resource on the expression of ethylene pathway genes and

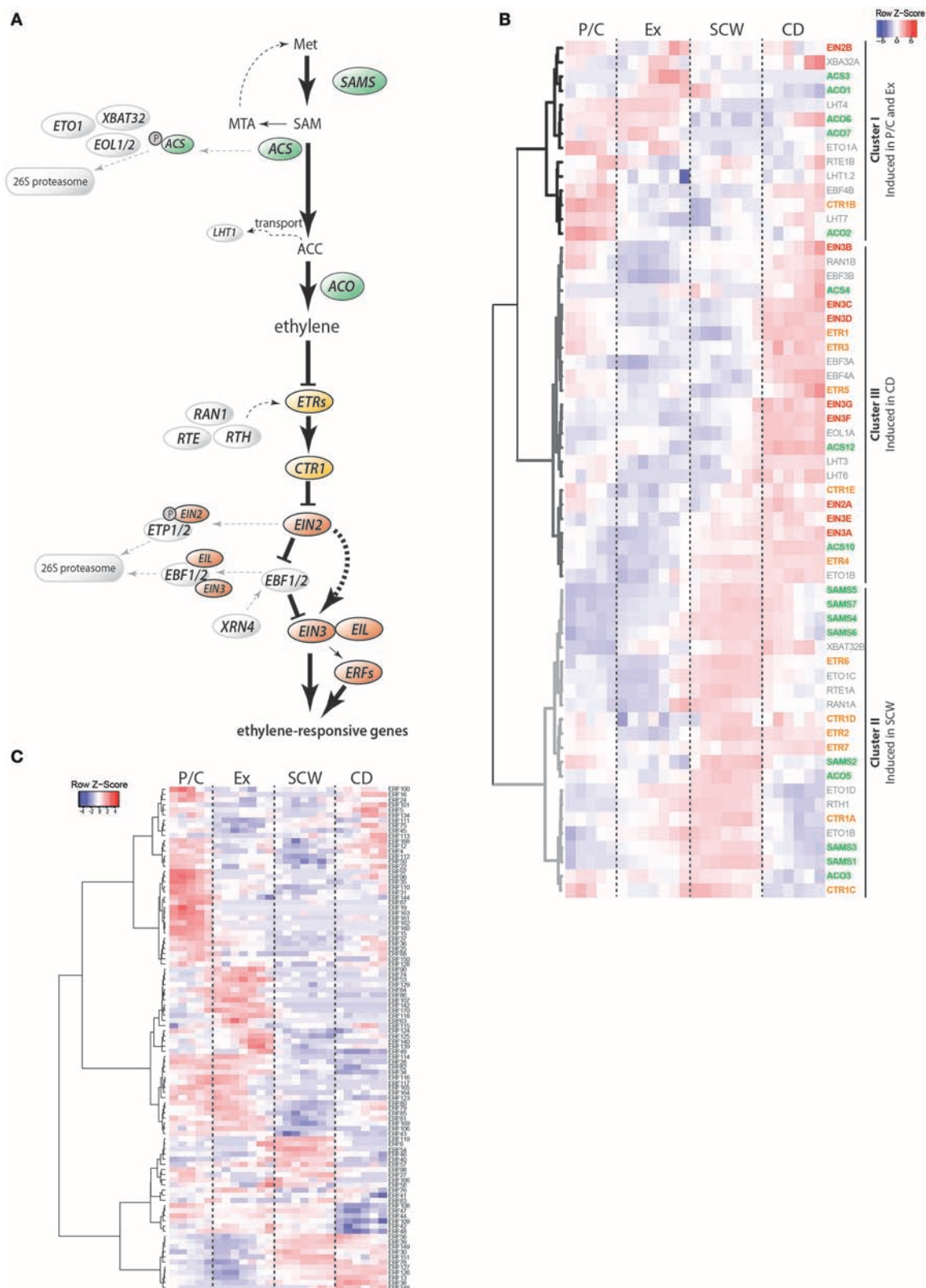


FIGURE 1 | Stem expression pattern of ethylene pathway genes. **(A)** Schematic representation of the current model of ethylene biosynthesis and signaling. Ethylene is the end product of subsequent conversion from methionine by S-AdoMet synthetases (SAMS), ACC synthetases (ACSs) and ACC oxidases (ACOs).

(Continued)

FIGURE 1 | Proteasome-mediated degradation of ACSs and transport of ACC are two mechanisms for plants to adjust cellular ethylene levels. Perception of ethylene is achieved through ER membrane-localized receptors (ETRs) that repress ethylene signaling in the absence of ethylene. Receptor activity is modulated by RAN1, RTE1 and RTH. In the absence of ethylene, ETRs prevent signaling and transcriptional reprogramming by activating the receptor-associated kinase CTR1. CTR1 controls activation of EIN2 in a phosphorylation-dependent manner. Once ethylene is bound to the receptors, the inhibitory activity of CTR1 is blocked, leading to cleavage of the EIN2 C-terminus which is now able to prevent proteasome-mediated degradation of the transcription factors EIN3 and EIL1. Downstream of EIN3, other transcription factors such as ethylene response factors (ERFs), contribute to transcriptional reprogramming of ethylene-responsive genes. Proteins are divided into biosynthesis (green), receptors and signaling (orange), transcriptional reprogramming (red) and regulatory (gray) components. Met, methionine; SAM, S-adenosylmethionine; MTA, Methylthioadenosine; ACC, 1-Aminocyclopropane-1-carboxylic acid; SAMS, S-adenosylmethionine synthetase; ACS, ACC synthase; MPK, Mitogen-activated protein kinase; ETO, Ethylene overproducer; EOL, Ethylene overproducer like; XBAT, XB3 ortholog; LHT, Lysine histidine transporter; ACO, ACC oxidase; ETR, Ethylene receptor; CTR, Constitutive triple response; RTE, Reversion-to-ethylene sensitivity; RTH, RTE-homolog; RAN, Resistance to antagonist; EIN, Ethylene insensitive; EIL, Ethylene-insensitive like; ETP, EIN2 targeting protein; EBF, EIN3-binding F-Box protein; XRN, Exoribonuclease; ERF, Ethylene response factor; P/C, Phloem/Cambium; Ex, Expanding xylem; SCW, Secondary cell wall formation; CD, Cell death **(B)** Heatmap showing the expression profiles of ethylene pathway genes in AspWood. Expression values are scaled per gene; red indicates gene expression higher than the associated cluster's average expression, while blue indicates gene expression lower than the average. Non-scaled expression values are shown in **Supplementary Figure S3** and listed in **Supplementary Table S2**. **(C)** Expression profiles of *ERFs* in AspWood. Scaled expression values are shown (non-scaled expression values are listed in **Supplementary Table S2**).

potential ethylene-regulated targets in *Populus*, and clues on their functional significance in wood formation.

RESULTS AND DISCUSSION

Populus Genome Displays Unique Features Within the Ethylene Pathway Gene Families

We performed phylogenetic analyses to identify homologs within the ethylene pathway gene families in the *Populus trichocarpa* (*Pt*), *Picea abies* (*Pa*) and *Arabidopsis thaliana* (*At*) genomes (**Supplementary Figure S1**, **Supplementary Table S1**). *P. trichocarpa* was used as it represents the *de facto* reference genome within the highly conserved *Populus* genus, while *Picea abies* was included to strengthen the identification of wood-related gene family members. For almost all gene families in the ethylene pathway, except the ethylene receptor family, *P. abies* genes clustered apart from *P. trichocarpa* and *A. thaliana*, as expected due to its evolutionary ancestry. Our analyses further revealed that all gene families in the ethylene pathway are expanded in the *P. trichocarpa* genome, in accordance with the rather recent whole genome duplication (Tuskan et al., 2006).

We observed expansion of the *CTR1* family in the *P. trichocarpa* genome, with five copies present compared to a single copy in the *A. thaliana* and spruce genomes (**Supplementary Figures S1E**, **S2**). Apart from three full-length *CTR1* isoforms, we found additional *Populus* genes encoding for either the N- or the C-terminus alone, *CTR1D* (*Potri.016095700*) and *CTR1E* (*Potri.016095800*), respectively (**Supplementary Figure S2**). In the absence of ethylene, the *CTR1* N-terminus binds to the receptors, while the C-terminus containing the kinase domain, prevents transcription of ethylene-regulated genes by triggering degradation of EIN2 (Clark et al., 1998; Huang et al., 2003; Ju et al., 2012). Mutations in either the N- or C-terminus of *CTR1* inhibits its suppressor function, indicating that both are essential to prevent constitutive ethylene signaling (Huang et al., 2003). A close relative of *P. trichocarpa*, *Salix purpurea*, encodes two genes that span the full-length *CTR1* sequence (*SapurV1A.0208s0240*; *SapurV1A.0208s0220*) and a third one (*SapurV1A.3019s0020*) which, similar to the *Populus CTR1E*, mainly covers the

C-terminus (**Supplementary Figure S2**). The functional relevance of this *CTR1* bifurcation and its presence in other species needs further investigation.

Spatial Expression Pattern of ACC Synthases Suggests That ACC Synthesis Occurs During Xylem Cell Expansion and Maturation

To identify sites of ethylene biosynthesis and ethylene-mediated transcriptional reprogramming, we surveyed transcript abundances for ethylene pathway genes using AspWood (**Supplementary Figure S3**, **Figure 1B**). AspWood represents a high-spatial-resolution transcriptomic database generated from RNA-Seq analysis of a longitudinal cryosection series from the stem, encompassing phloem/cambium differentiation (P/C), xylem cell expansion (Ex), secondary cell wall (SCW) formation and programmed cell death (CD) (Sundell et al., 2017). Hierarchical clustering of the ethylene pathway genes in AspWood (**Figure 1B**) identified three main clusters: Cluster I includes genes that are highest expressed in the phloem, cambium and expanding xylem; Cluster II includes genes that show highest expression during SCW formation; and Cluster III, comprised of genes that are induced during xylem maturation when xylem cells undergo CD.

Four of the ten *Populus* ACS genes (**Supplementary Figure S1**), *ACS3*, *ACS4*, *ACS10*, and *ACS12*, show expression profiles in AspWood (**Figure 1B**). Based on protein sequence similarity, the *Populus* *ACS3* and *ACS4* are most similar to the enzymatically active *A. thaliana* ACSs (Yamagami et al., 2003), while the *Populus* *ACS10* and *ACS12* are homologous to the enzymatically inactive ACSs in *A. thaliana* (*ACS10* and *ACS12*) (**Supplementary Figure S1B**). The *Populus* *ACS3* and *ACS4* had their highest expression in expanding xylem and during late xylem maturation (late CD zone), respectively (**Supplementary Figure S3B**, **Supplementary Table S2**). Genes encoding the potentially enzymatically inactive *Populus* ACSs, *ACS10*, and *ACS12*, had an overall high expression throughout the woody tissues (**Supplementary Figure S3B**). Similarly, expression analysis of *Pinus taeda* (*Pit*) xylem scrapings revealed low expression of *PitACS1* while the *PitACS1s* splice variant, lacking the tyrosine residue conserved in the active ACSs,

was expressed more than three-fold in comparison (Barnes et al., 2008). The tomato homolog of the enzymatically inactive *A. thaliana* ACS12 also showed highest expression within the ACS family in xylem tissues (Vanderstraeten and Van Der Straeten, 2017). Taken together, the expression of genes encoding the potentially enzymatically active *Populus* ACSs in the zones of xylem expansion and late maturation support these as the sites of ACC biosynthesis in the woody tissues. The constitutive expression of the genes encoding the potentially inactive ACSs support function of also these enzymes during wood formation.

ACC is distributed within the plant by transporters such as the recently identified LYSINE HISTIDINE TRANSPORTER 1 (LHT1) (Shin et al., 2015). Seven *LHT* genes were identified in *Populus*, with peaks in their expression in the P/C, Ex, and CD zones (Figure 1B). Thus, it may be that ACC is synthesized by the activity of the *Populus* ACS3 and ACS4 in the expanding xylem and late CD, from where it is transported to the other zones by the *LHT* transporters. This would be in line with the previously reported ubiquitous lateral distribution pattern of ACC throughout the woody tissues of *Populus* stems (Andersson-Gunnerås et al., 2003). We also observed that even though the six different *Populus* ACOs have very different expression patterns, they together provide ACO transcripts across all zones of wood formation (Supplementary Figures S1, S3C). From this follows that ethylene is potentially synthesized at any stage of xylem differentiation. Taken together, our results support that ACC biosynthesis occurs in the stages of xylem expansion and xylem maturation. Future work is needed to elucidate whether there are mechanisms, such as transport of ACC, which would allow lateral transport of ACC and hence production of ethylene in a ubiquitous manner during wood formation.

Ethylene Signaling Components and TFs Are Expressed Throughout All Phases of Wood Formation

We analyzed whether ethylene receptors and the downstream signaling components (see Supplementary Figure S1) show specific expression patterns during wood formation by investigating their expression in the AspWood database. The different isoforms of both *ETR* and *CTR1* families were highly expressed during either SCW formation (*ETR6*, *CTR1A*, *CTR1C*) or CD (*ETR1*, *ETR3*, *ETR5*), or during both (*ETR2*, *ETR4*, *ETR7*, *CTR1D*, *CTR1E*) (Figure 1B; Cluster II and III). *CTR1B* forms an exception as it is highest expressed in P/C and late CD. As the simultaneous presence of ethylene receptors and *CTR1* acts to suppress ethylene signaling it is possible, on the basis of their predominant expression during SCW formation, that downstream ethylene signaling is suppressed in this phase of xylem differentiation during normal wood formation.

Ethylene perception at the ER membrane is linked to gene regulation by the joint activity of the EIN2, EBF and EIN3 family members. The *Populus* EIN2 genes were constitutively expressed during secondary growth in the AspWood database (Supplementary Figure S3F, Supplementary Table S2). In contrast, the EBF F-box protein encoding genes showed high expression in P/C, a slight increase during SCW formation and

the highest expression during xylem maturation (CD zone) (Figure 1B, Supplementary Figure S3H). The EIN3 family in *Populus* consists of seven genes (Supplementary Figure S1G) which all had, similar to the EBFs, their highest expression in P/C and CD (Figure 1B, Supplementary Figure S3G). Co-expression of the EIN3 and EBF genes is consistent with the reported EIN3-mediated induction of *EBF2* expression in *A. thaliana* (Konishi and Yanagisawa, 2008). The high expression of the *Populus* EIN3s in P/C and during xylem maturation supports enhanced ethylene signaling during these phases of normal wood formation. However, even though EIN3 expression levels have been causally related to ethylene function in *A. thaliana* (Zhong et al., 2009), EIN3 is also regulated through posttranscriptional regulation (Binder et al., 2007). Functional studies are therefore needed to unequivocally demonstrate the function of the *Populus* EIN3 proteins in wood formation.

Populus ERFs are divided into 10 subgroups according to protein sequence similarity to *A. thaliana* ERFs (Vahala et al., 2013). Of the 170 ERFs of *P. trichocarpa*, 98 ERFs are expressed in AspWood (Figure 1B). The *Populus* ERFs cluster in groups with zone-specific peaks in their expression profile [e.g., *ERF87* (P/C), *ERF118* (Ex), *ERF119* (SCW), *ERF57* (Ex + SCW), or *ERF126* (SCW + CD)] (Figure 1C). Misregulated expression patterns and/or altered expression levels of ERFs have been shown to affect tree growth (Vahala et al., 2013). An important task for the future is, however, to identify which part of the ERF gene family is related to ethylene signaling and what biological processes are targeted by the various ERFs.

Ethylene-Mediated Transcriptional Reprogramming Occurs in All Developmental Zones

Parts of the genes in the ethylene pathway have been shown to respond to ethylene itself on a transcriptional level, creating both negative and positive feedback loops (Konishi and Yanagisawa, 2008; Shakeel et al., 2015; Prescott et al., 2016). The rapid effect of exogenous ACC (after 10 h) on ethylene pathway genes was recently analyzed in *Populus* stems (Felten et al., 2018). In this study, ACC was applied to *in vitro* grown wild type (T89) and two ethylene-insensitive (*pro35S::etr1-1* and *proLMX5::etr1-1*) *Populus* lines which expressed the *A. thaliana* dominant negative mutant allele *etr1-1* (Love et al., 2009). We mapped the sequencing data that originated from hybrid aspen (*P. tremula* × *tremuloides*) and that was presented in Felten et al. (2018) to the *P. tremula* genome. In accordance with expression data presented in Felten et al. (2018), ACC treatment induced expression of *ACO2*, *ETR1*, and *ETR2*, *CTR1* and *EIN3C* in an ethylene signaling-dependent manner (in wild type but not in any of the transgenic lines; Figure 2A). Transcriptional regulation of *ETRs* and *CTR1* upon stimulation of ethylene signaling suggests the existence of regulatory feedback loops in the pathway to achieve either increased sensitivity toward ethylene or suppression of the incoming signal depending on the stoichiometric balance between the level of ethylene and the abundance of these negative regulators of ethylene signaling. Consistent with results from Felten et al. (2018), expression

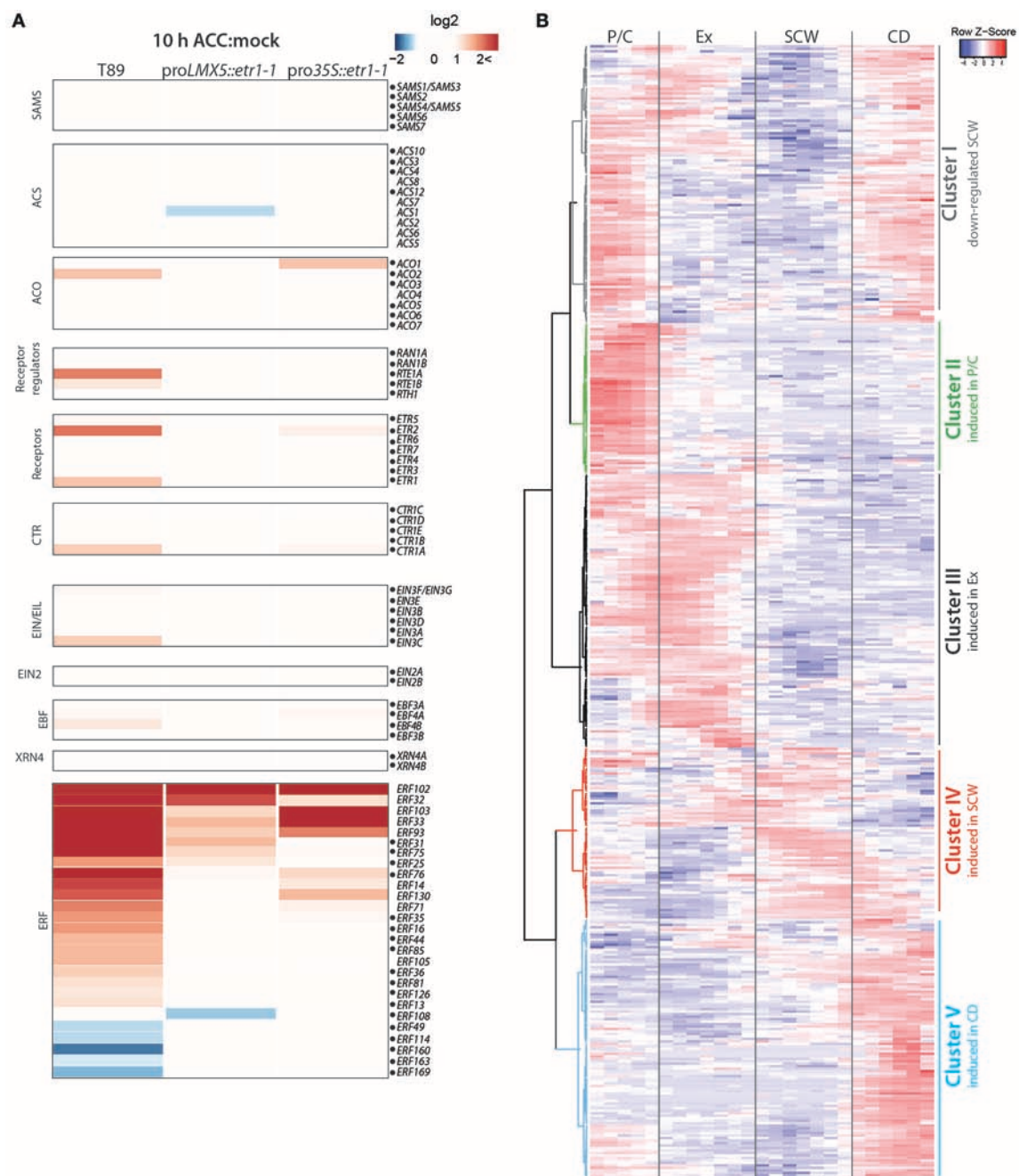


FIGURE 2 | Transcriptional regulation of ethylene-responsive genes occurs in all developmental zones. **(A)** Expression of ethylene pathway genes in *Populus* trees treated with either water (mock) or 100 μ M ACC for 10 h. RNA-Seq data was retrieved from Felten et al. (2018). Heatmap shows log2-values of gene induction/suppression between mock and ACC treatment in wild type (T89) and two ethylene-insensitive trees (*proLMX5::etr1-1* and *pro35S::etr1-1*). RNA-Seq reads were mapped to the *P. tremula* genome (corresponding *P. trichocarpa* gene identities are listed in **Supplementary Table S3**). Dots mark genes that also have expression profiles in AspWood. In case of the large *ERF* family we chose to only present *ERFs* with at least a two-fold expression change between mock and ACC treatment in any of the three genotypes. **(B)** Hierarchical clustering and scaled expression profiles of ethylene-responsive genes from AspWood ($n = 458$; **Supplementary Table S3**). Ethylene-responsive genes were defined as genes that were at least two-fold differentially expressed in wild type (T89) upon ACC treatment compared to mock treatment ($q < 0.01$). Gene IDs and expression values in response to ACC and mock treatment are listed in **Supplementary Table S3**.

of 27 *ERFs* (approximately 16% of all *ERFs*) was significantly altered under enhanced ethylene signaling in response to ACC application.

To study the role of ethylene-mediated transcriptional reprogramming for each phase of wood formation, we extracted stem expression profiles for ACC-responsive genes from the

AspWood database (**Figure 2B**). As a criterion for such genes (from now on called “ethylene-responsive genes”) a two-fold ACC-triggered gene induction/suppression in the ethylene-insensitive trees compared to wild type was used. We observed that 84% of all ethylene-responsive genes have expression profiles in AspWood (**Supplementary Table S3**). They can be divided on the basis of hierarchical clustering into five clusters that show enhanced transcription in one particular phase (Clusters II–V) or in several phases (Cluster I) of wood formation (**Figure 2B**). Cluster I, which includes genes with low expression during SCW formation, is enriched with EamA-like transporters, GTPase activating proteins, NAC TFs and peroxidases (see **Supplementary Table S3**). Among these genes we identified a homolog of *ANAC074* which in *A. thaliana* has been linked to SCW thickening, especially in fibers (Ko et al., 2007). Expression of Cluster II genes showed a peak in the P/C and encode for example ABC transporters, cytochrome P450 proteins, transferases and WRKY and GRAS TF family members. Expression of Cluster III genes was highest during xylem expansion and encode for proteins involved in transcription (e.g., members of the MYC, ERF, WRKY, and MYB TF families) but also several cell wall modifying enzymes (e.g., pectin lyases and xyloglucan endotransglucosylases). Cluster IV, showing the highest gene expression during SCW formation, contained the smallest number of ethylene-responsive genes. These did not show significant enrichment of any protein classes, but contained for instance several immunity-associated genes (e.g., *MAPKs*, *WRKY28*) and peroxidase, UDP-glycosyltransferase85A2 and tetratricopeptide repeat protein genes. Cluster V was comprised of genes encoding transporter and stress-associated proteins (e.g., cytochrome P450, peroxidases) with highest expression in the CD zone. On the basis of this data, we suggest that ethylene-mediated transcriptional reprogramming occurs in all developmental zones. This is in line with the known effects of ethylene on cambial activity (occurring in zone P/C), xylem cell expansion (occurring in the expansion zone) and xylem maturation (Love et al., 2009; Felten et al., 2018).

***EIN3D* and Several *ERFs* Are Hubs in the Wood Transcriptome**

Several TF families fulfill central roles during secondary growth and wood formation (Duval et al., 2014; Liu et al., 2015; Taylor-Teeples et al., 2015; Sakamoto et al., 2016). To study the importance of the *Populus* *EIN3s* and *ERFs* in these processes, we extracted network centrality/connectivity parameters of each annotated gene (“node”) in the AspWood dataset, including the following three parameters: betweenness (BTW), closeness (CI), and degree (using a default correlation threshold of five) (Sundell et al., 2017). The degree represents the amount of direct correlations of a gene and thus serves as an indicator of the number of other genes with the same (positive correlation) or opposite (negative correlation) expression profile. The BTW and CI instead give information about the importance of a gene for structure and organization within the network. The BTW serves as a measure of how connected the gene of interest is in the network; i.e., how often a gene is part of the shortest

path between other genes. As highly co-expressed genes are more likely to be co-regulated, the BTW serves as an indicator for the involvement of a gene of interest as a regulator of transcription of gene subsets within the network. The CI can be used as proxy for the distance of a gene of interest to other genes in the network and is calculated as the reciprocal sum of the distances that need to be taken to connect one gene to the others. Nodes (in our case genes) can be divided into four main categories: “center nodes,” highly connected within the network, indicated by a high BTW and CI and a high degree; “connecting nodes,” low degree but high BTW and CI indicating their role in connecting subsets of the network; “monopole nodes,” high BTW, but low CI and degree, typically the only connection between several genes in a small gene module; and “edge nodes,” low BTW and degree, which often are poorly connected genes with little contribution to the network structure. Of particular interest are highly connected genes (“hubs”) as these are likely to reveal important regulators of developmental switches. Hubs are characterized in general by a high BTW but not necessarily a high CI or degree (Valente et al., 2008). We decided to only focus on the top 20% genes characterized in AspWood ($n = 2,777$), ranked according to their BTW, and defined them as hubs in the current study (**Supplementary Table S4**). GO term analysis revealed that these hubs are enriched with transporters, potentially involved in membrane organization, ion trafficking and nutrient exchange. SNPs (single nucleotide polymorphisms) in five hubs (*Potri.001G096900*, *Potri.001G080400*, *Potri.005G237900*, *Potri.008G112300*, *Potri.018G127100*) were previously shown to associate with holocellulose and syringyl lignin content in *Populus* (Porth et al., 2013), supporting the importance of our selected hubs in the gene network underlying wood formation. Among our selected hubs we identified 221 TFs belonging to, among others, the NAC, MYB, C2H2, bZIP, bHLH, ERF, and TALE TF families (**Supplementary Table S4**). These hub TFs also include homologs of TFs with known functions during SCW formation, such as *SND2* (*SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN 2*; Hussey et al., 2011), *MYB46* (McCarthy et al., 2009; Ko et al., 2014) and *HDG11* (*HOMEODOMAIN GLABROUS 11*; Xu et al., 2014). Furthermore, we identified 21 *Populus* homologs of *A. thaliana* TFs that have been shown to bind to promoters of cellulose, xylan and lignin biosynthesis genes (**Supplementary Table S4**; Taylor-Teeples et al., 2015). One of them, PtMYB3 (*Potri.001G267300*), was also shown to bind to promoters of SCW-related genes in *Populus* (Zhong et al., 2013), again validating our selected hubs’ importance in wood formation.

In this network analysis, *EIN3D* and 11 *ERFs* were found among the hubs (**Figure 3**). Since our aim was to identify TFs that function as central nodes inside the wood transcriptome, we focused on those with a high connectivity. Taken into account their BTW, but also CI and degree, we focused on eight *ERFs* (*ERF27*, *ERF49*, *ERF75*, *ERF76*, *ERF83*, *ERF87*, *ERF118*, and *ERF119*) for further analysis. Based on their network centrality parameters, *ERF27*, *ERF87*, *ERF118*, *ERF119*, and *EIN3D* were highly connected with other genes in the network, suggesting potential roles as master regulators of gene expression during wood formation. *ERF49*, *ERF75*, *ERF76*, and *ERF83* might

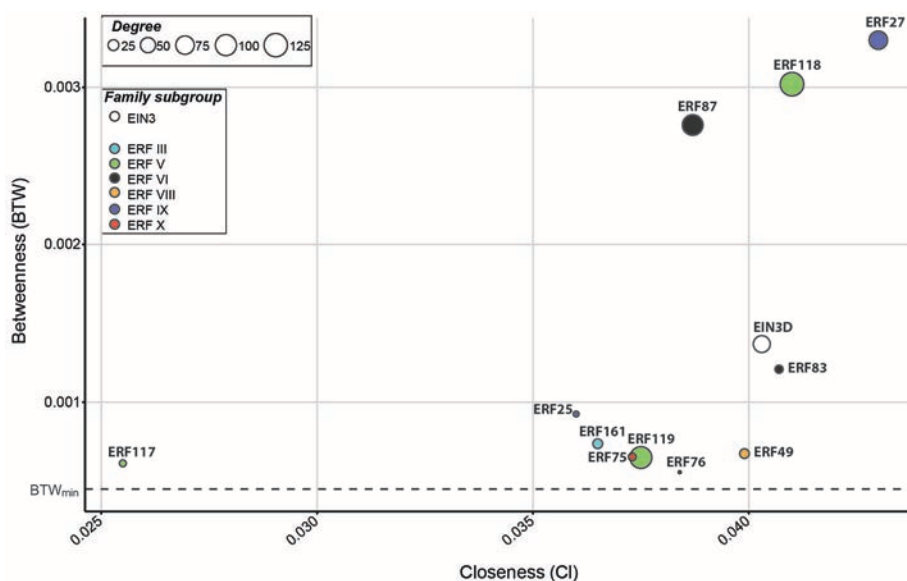


FIGURE 3 | EIN3D and ERFs function as hubs in the wood transcriptome. The graph shows betweenness (BTW) and closeness (Ci) values of *EIN3* and *ERFs* expressed in AspWood (defined in Sundell et al., 2017). Circle size corresponds to their degree. All genes with expression profiles in AspWood were ranked according to their BTW, and the top 20% genes ($n = 2,777$) were defined as hub genes (listed in **Supplementary Table S4**). Dotted lines mark the minimum BTW value (0.000447) calculated for a hub gene.

instead function as mediators of gene expression between diverse gene subsets. Interestingly, only *ERF49*, *ERF75*, *ERF76* were responsive to ACC treatment (**Figure 2A**). Furthermore, *ERF75* carries the EIN3-binding motif (TEIL motif) in its 2 kb promoter, suggesting that it acts directly downstream of EIN3 in ethylene signaling (Felten et al., 2018). In conclusion, the network centrality analysis resulted in identification of *EIN3D* and eleven *ERFs* as *Populus* TFs that are likely to control transcriptional changes during secondary growth. This is supported by recent finding on an *ERF118* homolog in *P. simonii* × *nigra* (*PsnSHN2*) in controlling the expression of SCW-related TFs and modulating secondary cell wall properties (Liu et al., 2017). Functional evidence exists also for *Populus* *ERF76* in connection to abiotic stress response (Yao et al., 2016, 2017). To our knowledge, these are however the only functionally characterized EIN3/ERF hubs so far, highlighting the need for future functional studies of these master regulators and their downstream targets.

The *EIN3D* Co-expression Gene Module Reveals Several TFs as Potential Novel Regulators of Wood Formation

Although the seven *Populus* *EIN3s* had very similar spatial expression patterns in AspWood, only *EIN3D* passed our selection for hubs in the AspWood dataset. This apparent functional diversification prompted us to study the structure of the co-expression network for the different *Populus* *EIN3* TFs. Using a co-expression cutoff of five (in accordance with the network analysis used in Sundell et al., 2017), we identified three distinct gene modules that contained genes co-expressed with either *EIN3C* and *EIN3D*, *EIN3F*, or *EIN3B* (**Figure 4A**).

Among these three gene modules, we found three genes that showed ethylene-dependent transcriptional regulation upon ACC treatment (marked with a cross in **Figure 4A**), indicating that EIN3s contribute to the transcriptional regulation of ethylene-responsive genes during wood formation. Furthermore, all *EIN3* gene modules were enriched with hubs (marked as orange nodes in **Figure 4B**), indicating co-expression and thus a potential regulatory function of *EIN3B*, *EIN3C/EIN3D*, and even *EIN3F* on other key genes during wood formation (**Supplementary Table S5**). Surprisingly, no *ERFs* were present in any of the *EIN3* gene modules, suggesting that the expression of *ERFs* might not require EIN3s during normal wood formation. It is possible that the control of *ERFs* through EIN3 becomes more prominent under stress, when ethylene levels increase. 14 of the 27 ACC-regulated *ERFs* carried the TEIL motif in their 2 kb promoter indicating a potential regulation of at least part of the *ERFs* through EIN3s upon high ethylene levels (Felten et al., 2018). Also other studies on EIN3-mediated regulation of *ERFs* support the idea of EIN3-mediated *ERF* regulation under stress conditions or exogenous ACC/ethylene application (for recent examples see Chang et al., 2013; Quan et al., 2017).

The largest gene module consists of 64 nodes including *EIN3D*. Representatives of eight different TF families, including EIN3 itself, NAC, bZIP, MYB-related, TALE, WOX, GRAS, C2H2, and C3H, were present in the *EIN3D* gene module (**Supplementary Table S5**). Among them we found a homolog of VND-INTERACTING 2 (VNI2; AT5G13180) which in *A. thaliana* interacts with VND7 and functions to suppress xylem vessel formation (Yamaguchi et al., 2010) and also NARS1/NAC2 (AT3G15510) shown to be involved in SCW development of seed coat epidermal cells (Voiniciuc et al.,

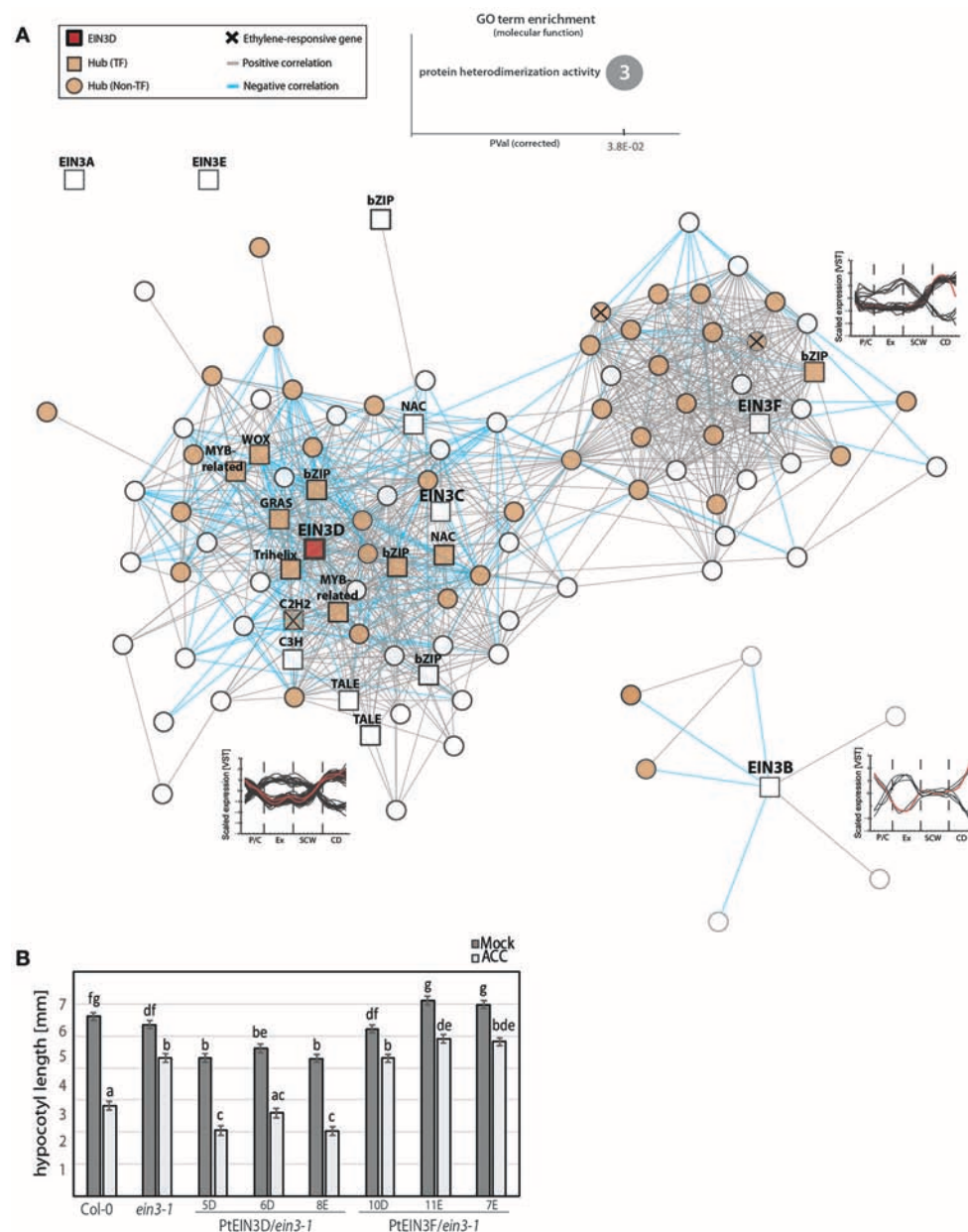


FIGURE 4 | EIN3 isoforms are co-expressed with TFs involved in transcriptional regulation and secondary cell wall biosynthesis. **(A)** Co-expression gene modules (threshold of five) of *EIN3B*, *EIN3C*, *EIN3D*, and *EIN3F*. TFs are shown as squares, while all *non-TFs* are represented by circles. Hubs are indicated by orange filling. Positive correlation between two genes is represented by gray lines (edges), negative correlations are indicated by blue edges. Ethylene-responsive genes shown in **Figure 2C** are marked with a cross (see **Supplementary Table S5**). Line graphs show scaled expression profiles of genes in each module during wood formation, with red lines depicting the expression pattern of each *EIN3* isoform. GO term enrichment was performed with *Arabidopsis* homologs. Number in the circle indicates number of genes belonging to this GO term found among all 87 *EIN3* co-expressed genes. **(B)** Estimate of hypocotyl length [in mm] with standard error, calculated using a linear effect model. Vernalized seeds were germinated in the dark for 72 h in the presence/absence of 10 μ M ACC. Hypocotyl length was determined from 29 to 44 biological replicates (depending on genotype, treatment and experiment) obtained from two independent experiments. Raw data, including number of biological replicates per line, treatment and experiment can be extracted from **Supplementary Table S8**.

2015). Comparing the *EIN3* co-expressed genes to a publicly available dataset of direct targets of *A. thaliana* EIN3 (Chang et al., 2013), we found ten shared genes of which eight belong to the *EIN3D* gene module (**Supplementary Table S5**). This result suggests that EIN3D directly controls at least part

of its co-expressed genes, including the homolog of *VNI2* (*Potri.003G166500*). Although protein interaction studies with EIN3 and other TFs are still elusive, this result suggests that EIN3D might act upstream or together with VNI2 during wood formation.

In order to assess *Populus* EIN3D functionality in ethylene signaling, we tested its capacity to complement the *A. thaliana* ethylene-insensitive *ein3-1* mutant by expressing it under the control of the 35S promoter (**Figure 4B**, **Supplementary Table S8**). We also included EIN3F as a representative of the three *P. trichocarpa* EIN3 genes that clustered together with *AtEIL3* which cannot complement *ein3-1* (**Supplementary Figure S1G**; Chao et al., 1997). Complementation was assessed using the triple response of dark-grown (etiolated) *A. thaliana* seedlings as the phenotypic output (Guzmán and Ecker, 1990) in three transgenic lines for each *Populus* EIN3 gene. External application of ACC or ethylene results in shortening and thickening of the hypocotyl and root as well as exaggerated curving of the apical hook (Stepanova and Alonso, 2009), which is diminished in *ein3-1*. Overexpression of EIN3D showed consistent ability to complement *ein3-1*, as demonstrated by the effect of ACC in reducing hypocotyl length of the transgenic lines to a length similar or smaller than in the wild type. Overexpression of EIN3F did not complement *ein3-1*. Hence, our data supports that the *Populus* EIN3D can rescue the loss of function of *A. thaliana* EIN3 and is therefore functional in the process of ethylene controlled hypocotyl elongation. However, the function of EIN3D during wood formation remains to be elucidated.

ERF Gene Modules Comprise Genes Associated With SCW Biogenesis

We selected eight out of 11 ERFs that were found as hubs for further analysis (ERF27, ERF49, ERF75, ERF76, ERF83, ERF87, ERF118, ERF119). We investigated co-expression gene modules of these hubs to elucidate their role in wood formation (**Figure 5**). GO term enrichment analysis with the *A. thaliana* homologs of the co-expressed genes showed a significant enrichment of genes associated to SCW biogenesis (e.g., genes encoding for lignan and xylan biosynthesis). Similar to our results, the AP2/ERF *Ii049* from *Isatis indigotica* was recently linked to lignan biosynthesis (Ma et al., 2017). We also observed that each of these ERFs is part of a gene module which contains several other hubs, further supporting their central function during wood formation.

The *Populus* ERF27 was part of the largest gene module containing co-regulated genes with strong expression in the phloem and during SCW formation or, *vice versa*, low during these phases and high in the cambium-xylem expansion phase and during xylem maturation (**Figure 5**). The module contained a few genes that were connected to carbohydrate household, such as *UDP-glycosyltransferase88A1* and *SUC2* (*ARABIDOPSIS THALIANA* SUCROSE-PROTON SYMPORTER 2), but also several oxidative stress- and salt stress-related genes. Interestingly, the expression of ERF27 correlated negatively with the expression of a homolog of *A. thaliana* HOMEBOX GENE 8 (*AtHB8*; **Supplementary Table S6**). *AtHB8* has been shown to control procambial cell specification in *A. thaliana* leaves (Donner et al., 2009), and overproduction of *AtHB8* stimulates cambial cell proliferation and xylem differentiation (Baima et al., 2001), thus linking ERF27 to processes occurring in the vascular cambium and differentiating xylem.

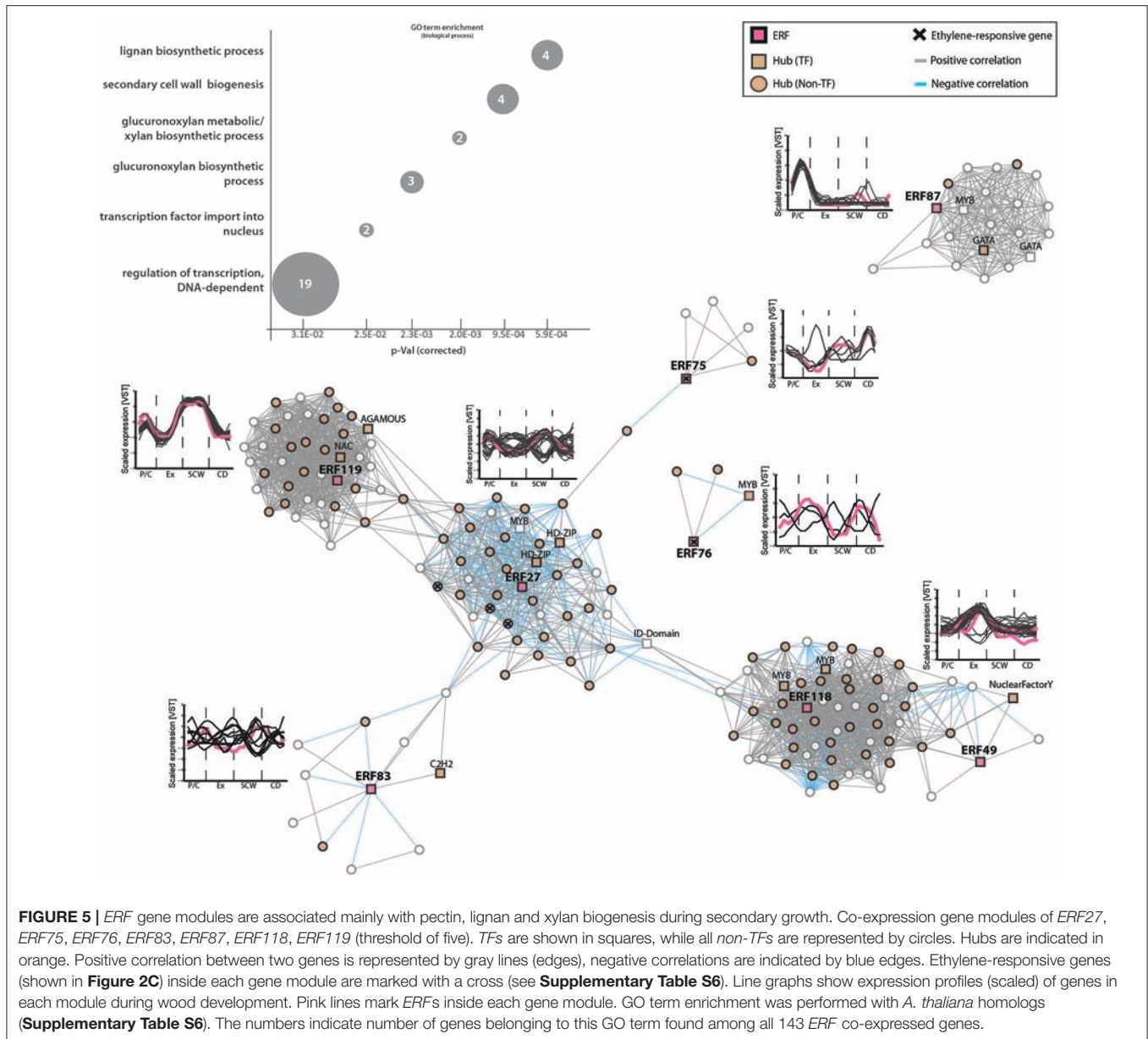
The *Populus* ERF118 and ERF49 were part of one gene module but not directly connected. All genes in this module had their highest expression in the xylem expansion phase. The module was enriched in genes encoding primary cell wall modifying enzymes, such as a xyloglucan endotransglycosylase (*Potri.013G005700*), a pectin methylesterase (*Potri.002G202500*) and a pectin lyase (*Potri.014G117100*) (**Supplementary Figure S6**). Hence we propose that genes in the ERF118/ERF49 module are associated with primary wall modifications during xylem cell expansion. ERF119 on the other hand was strongly induced during SCW formation, and the ERF119 co-expressed genes were associated for instance with lignan [pinorensinol reductase (*Potri.003G100200*)] and xylan biosynthesis (galacturonosyltransferases *Potri.001G416800* and *Potri.011G132600*).

In accordance with the EIN3 modules lacking ERFs, the ERF modules did not contain any EIN3 genes either (**Supplementary Table S6**), further supporting the hypothesis that EIN3s might only control ERFs that are responsive to high ethylene levels. In addition, we did not identify shared nodes between the EIN3 and ERF gene modules (**Supplementary Table S5**), supporting diverse functions for these TF families during wood formation.

Ethylene-Responsive TFs Are Co-expressed With Immune Response Genes During Wood Development

Only two *Populus* ERFs among the selected ERF hubs (ERF75 and ERF76) were ethylene responsive (**Figure 5**). In order to identify additional molecular players that connect ethylene signaling and wood formation, we enlarged our network analysis and included all TFs among the ethylene-responsive genes shown in **Figure 2B**, independent of their TF family background (**Figure 6**). We found three bZIP, three NAC, two C2HC, one Dof, one Myb-related, and one WRKY as hubs which displayed ethylene signaling-dependent expression in response to ACC (**Figure 6A**). One of the ethylene-responsive NAC TFs is a homolog of ANAC012 (*Potri.002G037100*), which is a negative regulator of secondary wall deposition in xylem fibers (Ko et al., 2007). A second NAC is a homolog of ANAC047 that has been shown to act downstream of the EIN2-EIN3 signaling cascade during leaf senescence (Kim et al., 2014). While the TEIL motif was not found in the promoter region (2 kb upstream the start codon) of the *Populus* homolog of ANAC012 nor ANAC047, promoter regions of two other ethylene-responsive TFs [ERF75 and ANAC100 (*Potri.017G086200*)] did contain a TEIL motif (Felten et al., 2018). These results support EIN3-mediated transcriptional regulation of a subset of the ethylene-responsive TFs and thus their potential function in wood formation in an ethylene-dependent manner.

Expression of eight ethylene-responsive TF hubs was co-regulated, with highest expression in phloem/cambium cells and during CD (gene module 1; **Figure 6B**; **Supplementary Table S7**). GO term enrichment tests indicated that the *A. thaliana* homologs of co-expressed genes are



involved in immune responses against chitin (for example WRKY-encoding *Potri.006G109100*, *Potri.015G099200*), phytohormone-mediated responses (jasmonic acid response; auxin-related indole glucosinolate biosynthesis), transcriptional regulation and zinc ion homeostasis. Jasmonates have previously been connected to secondary growth in *A. thaliana* (Sehr et al., 2010). Also, treatment with jasmonates to aspen plantlets induced the formation of tyloses, which are occlusions of xylem vessels that serve as a barrier against pathogens (Leśniewska et al., 2017). This study also showed that combined exogenous application of ACC and jasmonic acid triggered tyloses formation in an ethylene signaling-dependent manner. In wild type trees, ACC treatment led to suppression of all *bZIP*, *Dof* and one *NAC* TFs, suggesting that these are negative regulators of ethylene-mediated transcriptional reprogramming.

Indeed, one *bZIP* (*Potri.005G243400*, gene module 2) and the *Dof* TF (*Potri.004G038800*; gene module 5) showed mainly negative correlation with other hubs in their co-expression gene modules. The *bZIP* TF is a homolog to the *A. thaliana* FD, a flowering associated regulator (Abe et al., 2005), and its expression showed a low sharp peak during SCW formation. The importance of gene regulation by these ethylene-regulated TFs for wood development remains unanswered. However, we found several shared nodes between gene module 1 and the *EIN3D*-associated gene module (Supplementary Table S7), encoding for example for 6-BETA-TUBULIN (*Potri.009G067100*) or a homolog of mannanase regulator *AtBZIP44* (Iglesias-Fernández et al., 2013), pointing toward a link between *EIN3D* and the ethylene-regulated TFs on transcriptional regulation during secondary growth.

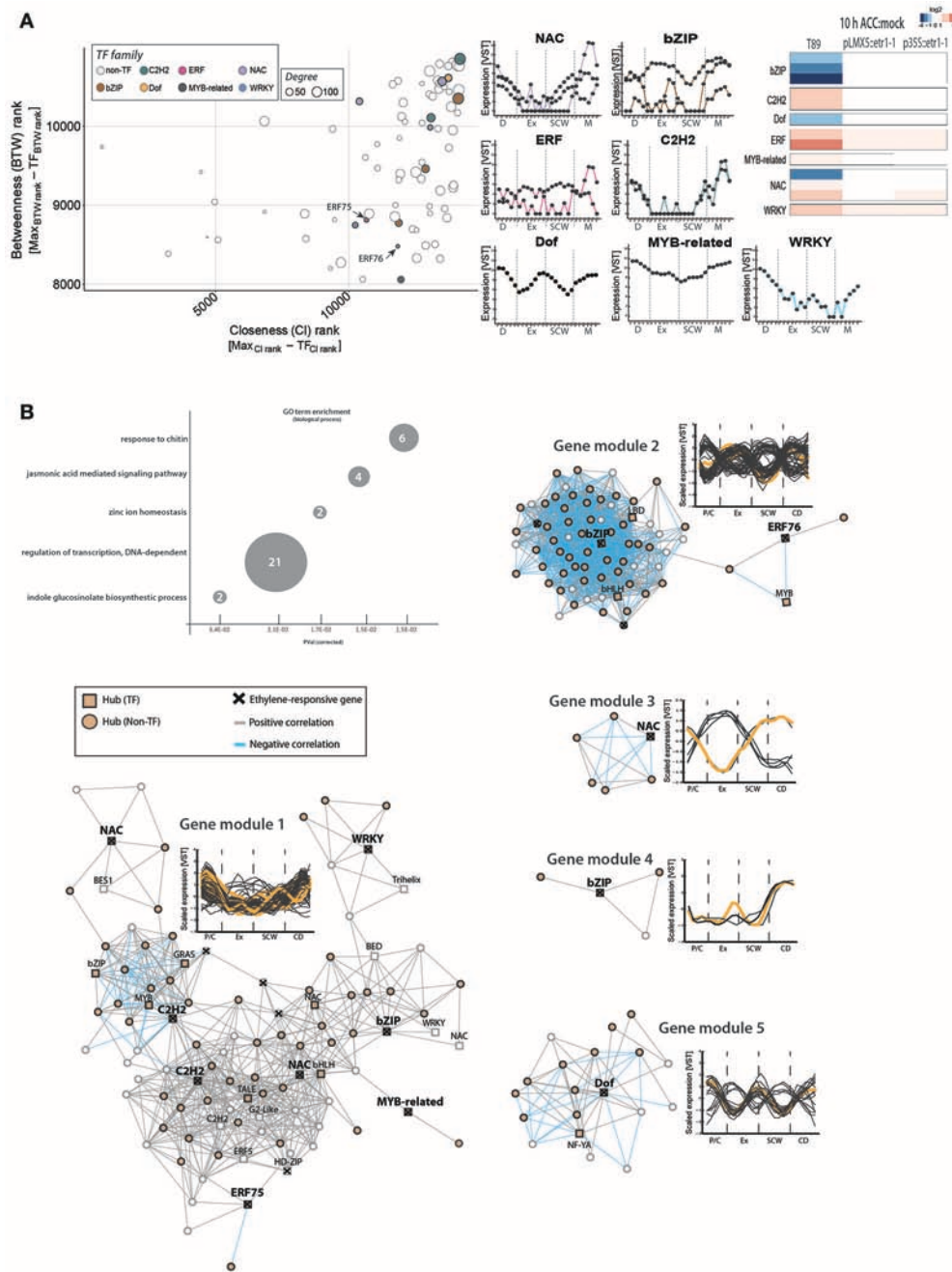


FIGURE 6 | Expression of ethylene-responsive hub TFs is associated with immunity-related phytohormone signaling during wood formation. **(A)** Gene centrality parameters and expression profiles of ethylene-responsive hub TFs. ACC-triggered expression values in wild type and two ethylene-insensitive trees were obtained as described in **Figure 2**. **(B)** Co-expression gene modules (threshold five) of ethylene-responsive TFs. Ethylene-responsive genes inside each gene module are marked with a cross (see **Supplementary Table S7**). GO term enrichment was performed with *A. thaliana* homologs. Number indicate number of genes belonging to this GO term found among all 150 co-expressed genes.

CONCLUSION

Unraveling transcriptional regulation mechanisms by phytohormones is key for understanding numerous aspects of the plant life cycle. Our study identified the homologs of

ethylene biosynthesis and signaling components in *Populus* and their expression profiles during secondary growth. Co-expression network analysis of the wood transcriptome revealed a plethora of biological processes, such as lignan, xylan and pectin biosynthesis, which were transcriptionally associated with

EIN3D and several *ERFs*. *EIN3D* and 11 *ERFs* were identified as hub TFs in a tissue-specific manner. Notably, we identified *EIN3D* and its' co-regulated TFs as potential transcriptional master switches during wood formation, and *ERF118* and *ERF119* having a potential role in regulating xylem expansion and SCW formation, respectively. Upcoming research projects focusing on the function and downstream targets of these TFs are predicted to significantly broaden our understanding of the role of ethylene in wood formation and to highlight possibilities to utilize ethylene pathway genes in forest biotechnology and tree breeding practices.

MATERIALS AND METHODS

Phylogenetic Analyses of Genes Putatively Related to Ethylene Signaling and Biosynthesis

The gene family members of ethylene receptor, *CTR1*, *EIN2*, *EIN3*, and *ERF* genes of *A. thaliana*, *P. trichocarpa* and *P. abies* were extracted from the Plant Genome Integrative Explorer resource (PlantGenIE.org; Sundell et al., 2015; Gene Family tab). In order to find homologous gene models putatively placed in separate gene families and *P. abies* full-length transcripts represented by low quality gene models, BLAST searches were also performed against PlantGenIE genome and transcriptome databases. Subsequently, the identified gene models and transcripts were aligned in Clustal Omega with default parameters (Sievers et al., 2011) and aberrant sequences were removed. The phylogenetic trees were created using the Galaxy platform (Goecks et al., 2010) hosted at PlantGenIE.org. The Galaxy workflow utilized the MUSCLE v3.8.31 program (maximum number of iterations: 16) for multiple alignment, and PhyML 3.1 (substitution model: WAG, aLRT test: SH-like, tree topology search operation: Nearest Neighbor Interchange) and Tree Vector programs for building and drawing phylogenetic trees, respectively.

Cloning and Mutant Generation

In order to ascertain whether *Populus* homologs of *A. thaliana* *EIN3* could rescue the triple response of the *A. thaliana* *ein3-1* mutant, genomic DNA was extracted from *P. trichocarpa* leaves using the E.Z.N.A. kit (Omega). Full length *Populus EIN3D* and *EIN3F* was amplified (primers listed in **Supplementary Table S9**) and cloned into pENTR/D-TOPO (pENTR Directional TOPO Cloning Kits, Invitrogen) and thereafter transferred to pB7GW2D through Gateway Recombination (LR Clonase II, Invitrogen). The resulting plasmids (p35S:*PtEIN3*) were transformed into *Agrobacterium tumefaciens* GV3101 (pMP90) followed by transformation of *ein3-1* through floral dip (Clough and Bent, 1998). Homozygous lines were selected based on Basta resistance (50 μ M). The triple response assay was performed with three homozygous lines.

Triple Response Assay

Transgenic lines were germinated for 72 h in the dark on Murashige-Skoog media that was either supplemented with 10 μ M ACC or without (Alonso-Stepanova Laboratory

Protocols)¹. Hypocotyl and root lengths were measured for each seedling in the presence and absence of ACC (root length not shown, but phenotypes are consistent in both tissues). Mean, standard error and *p*-values were calculated from 29 up to 44 biological replicates (depending on genotype and treatment) using a linear effect model (lme function in *limma*), with genotype and treatment as fixed effects. The *multcompView* package was used to assign significance letters using a *p*-value cutoff 0.01. **Supplementary Table S8** includes all raw data, number of replicates per genotype, treatment and experiment and output (mean, standard error, *p*-values and significance letters) calculated from the linear model.

RT-qPCR Analysis

Transgene expression was quantified from three pools of five seedlings harvested from control plates. Total RNA was isolated according to instructions using the AurumTM Total RNA Mini Kit (Bio-Rad) and a DNase treatment was performed using the Ambion[®] DNA-freeTM DNA Removal kit (Thermo Fisher Scientific). RNA was quantified and cDNA was synthesized using the iScript cDNA Synthesis Kit. Real-time quantitative PCR (qPCR) was performed on five times diluted cDNA template using a Bio-RAD CFX96 Real Time System with SYBR[®] Green Mastermix (Bio-Rad) and 5 pmol concentrated primers. PCR conditions were as follows; 3 min initial denaturation at 95°C, followed by 39 cycles of denaturation at 95°C for 10 s, primer annealing at 58°C for 10 s, and a 20 s extension step at 72°C. *AtEF1 α* (*At5G60390*) was used as a reference gene. Δ Ct values were calculated by subtracting average *EF1 α* Ct value from corresponding Ct value. Relative expression levels were calculated using the formula $2^{-\Delta\text{Ct}}$ for each sample (**Supplementary Figure S5**). Primers are listed in **Supplementary Table S9**.

RNA-SEQ Data Analysis

Description of tree growth conditions, experimental setup for treatment with ACC and RNA extraction and library preparation procedure can be found in Felten et al. (2018). Briefly, RNA was extracted from whole stems from wild type (T89) and two ethylene-insensitive trees (pLMX5::*etr1-1* and p35S::*etr1-1*). *In vitro*-grown plants were allowed to grow until a height of approximately 8 cm. For the treatment, 100 μ M ACC or water was applied on top of the medium. Stem material was pooled from six plants per treatment and genotype 10 h after application of either ACC or water. Frozen stems were ground and used for RNA extraction using the CTAB method and lithium chloride precipitation. DNA was removed using *DNAfree*TM (Ambion), left-over RNA was cleaned using the Qiagen MinElute kit and sent for library generation and paired-end Illumina sequencing to SciLifeLab (Science for Life laboratory, Stockholm, Sweden). The sequencing data is available from the European Nucleotide Archive under the accession number ERP012528. Quality assessment of raw sequence data, including removal of ribosomal RNAs and sequencing adapters and quality trimming

¹Alonso-Stepanova Laboratory Protocols. Available online at: http://www4.ncsu.edu/~jmalonso/Alonso-Stepanova_ACC.html (Accessed Sept 20, 2017).

of sequences was performed as described in Felten et al. (2018). Read pairs that passed the quality assessment were mapped to the latest *P. tremula* genome sequence retrieved from “PopGenIE” (www.popgenie.org). We chose the *P. tremula* genome since its genetic background is most similar to hybrid aspen (*P. tremula* × *tremuloides*; Hamzeh and Dayanandan, 2004), which was the species used for the ACC application experiment. Reads were transformed into a count per gene per library using HTSeq (Anders et al., 2015). Statistical data analysis was performed in R (version 3.2.2) using *EdgeR*. First, reads with less than 10 counts in at least one library were excluded from the dataset. Gene counts were normalized based on a calculated normalization factor (function `calcNormFactors` in the R package *edgeR*). Count data was log-transformed (`voom` function in R package *limma*) to obtain log2 counts per million. The `lmfit` function in the *limma* package was used to fit a mixed linear effect model (with genotype:treatment as fixed effects and biological replicate as random effect) to the log2 gene expression values and variance shrinkage was applied using the `eBayes` function in *limma* before calculating *p*-values. An FDR (false discovery rate) adjusted *p*-value (*q*-value) cutoff of 0.01 was used to extract differentially expressed genes. In order to compare the effect of ACC application in the ethylene-insensitive trees to wild type trees, we set a cutoff of two-fold ACC-triggered expression difference between wild type and both ethylene-insensitive trees and defined the genes that passed this criterion as “ethylene-responsive genes.” Heatmaps were generated with the *pheatmap* package in R.

Co-expression Network Analysis

Obtained RNA-Seq reads from stem sections of four trees were aligned to the *P. trichocarpa* genome and normalized using a variance stabilizing transformation (VST). Genes with a VST > 3 in at least two samples from at least three out of four trees were considered as differentially regulated. All samples (stem sections) were clustered using Euclidean distance and all genes were scaled and clustered using Pearson correlation. The co-expression network was performed using mutual information (MI) and context likelihood of relatedness (CLR) algorithm. The co-expression network is purely based on gene expression profiles irrespective of the cell type. Selection of hubs in the transcriptome during secondary growth was performed on the basis of the BTW rank obtained for each gene in AspWood (described in Sundell et al., 2017). The BTW rank was calculated as follows: first the betweenness was calculated (=number of cases in which a node lies on the shortest path between all pairs of other nodes) and afterwards calculated values were sorted in ascending order (highest betweenness value was associated to 1). The Cl rank (=reciprocal of the sum of distances to all other nodes) was calculated in the same way. All genes present in AspWood were ranked according to their BTW rank and the top 20% genes (*n* = 2,777) were defined as “hubs.” TFs among the hubs were selected according to their characterization in AspWood and presented in Figure 3, Supplementary Figure S4 with their centrality parameters (BTW, Cl and degree). In accordance to Sundell et al. (2017), all presented gene co-expression networks (Figures 4–6) are generated using a co-expression threshold of

five. Composition of each gene module was analyzed using a build-in function of the AspWood database.

GO Term Analysis

GO terms (in this case significantly enriched PFAMs) that were used to describe clusters of ethylene-responsive genes in Figure 2B were extracted from AspWood (listed in Supplementary Table S3). Cluster-based gene selection was done according to their expression profile in AspWood (defined by Clusters A–H in Aspwood). GO term analysis shown in Figures 4–6 were performed with *A. thaliana* homologs. Significantly enriched GO terms were extracted from “AtGenie” (<http://atgenie.org/enrichment>) using a *p*-value cutoff of 0.05 (listed in Supplementary Tables S5–S7).

AUTHOR CONTRIBUTIONS

CS performed co-expression network analysis and all bioinformatic analyses. BW performed the complementation experiment. SJ-L performed the phylogenetic analyses. ND helped with the RNA-Seq analysis. BS contributed to the complementation experiment. CS, BW, JE, and HT analyzed and discussed the data. The manuscript was written by CS and HT with contributions from all coauthors.

ACKNOWLEDGMENTS

The authors would like to thank Yuan Ma for the generation of transgenic plants, and Emma Hörnblad for performing initial screens to select positive lines. This work was supported by funds from the Kempe foundation (grants SMK-1649 and SMK-1533) and The Swedish Research Council Formas (grant 213-2011-1148). Computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00272/full#supplementary-material>

Supplementary Figure S1 | Phylogenetic analysis of genes involved in ethylene biosynthesis (A–C), perception and signaling (D,E), transcriptional reprogramming (F,G) and regulation (H) in *A. thaliana* (At; brown) and the two woody species *P. trichocarpa* (Pt; black) and *P. abies* (Pa; gray). Exact gene identities can be extracted from Supplementary Table S1.

Supplementary Figure S2 | Protein alignment of CTR1 isoform(s) of *A. thaliana* (At), *P. trichocarpa* (Potri), *P. abies* (MA), *Salix purpurea* (Sapur). Kinase domain is labeled according to Huang et al. (2003).

Supplementary Figure S3 | Expression pattern of analysis of genes involved in ethylene biosynthesis (A–C), perception and signaling (D,E), transcriptional reprogramming (F,G) and regulation (H) during secondary growth in *P. trichocarpa*. Gene expression data was extracted from the AspWood database (Sundell et al., 2017). Each data point represents expression values in cryosections of zones labeled as P/C, Phloem/Cambium cells; Ex, Expanding xylem; SCW, Secondary cell wall formation; CD, Cell death. Raw data used for this figure can be found in Supplementary Table S2.

Supplementary Figure S4 | TF hubs from the AspWood network. Centrality parameters of TF hubs (BTW, CI and degree, listed in **Supplementary Table S4**). Colors indicate TF family. The plant TF database (v4.0) was used to identify TF families. Families that are not present in top 20% genes are: ARR-B, BBR-BPC, BES1, CAMTA, E2F, FAR1, GeBP, HB-PHD, HB-others, HRT-like, LFY, LSD, NF-XI, NZZ/SPL, RAW, SAP, VOZ, Whirly.

Supplementary Figure S5 | Transgene expression levels in 35S::PTEIN3D/*ein3-1* and 35S::PTEIN3F/*ein3-1*. Each bar represents the mean of three pools of seedlings \pm SD, with each pool consisting of five seedlings per tested line.

Supplementary Table S1 | Gene identities and abbreviations used for phylogenetic analysis of gene families in the ethylene pathway.

Supplementary Table S2 | Expression values for genes in the ethylene pathway during secondary growth.

Supplementary Table S3 | Expression values of ethylene pathway and ethylene-responsive genes upon ACC treatment.

Supplementary Table S4 | Identified hubs from “AspWood” co-expression network.

Supplementary Table S5 | Structure and assembly of EIN3 co-expression networks.

Supplementary Table S6 | Structure and assembly of ERF co-expression networks.

Supplementary Table S7 | Structure and assembly of co-expression networks from ethylene-responsive TFs.

Supplementary Table S8 | Complementation of *ein3-1* mutants by PTEIN3D.

Supplementary Table S9 | Primers used in this study.

REFERENCES

- Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., et al. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* 309, 1052–1056. doi: 10.1126/science.1115983
- Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638
- Andersson-Gunnerås, S., Hellgren, J. M., Björklund, S., Regan, S., Moritz, T., and Sundberg, B. (2003). Asymmetric expression of a poplar ACC oxidase controls ethylene production during gravitational induction of tension wood. *Plant J.* 34, 339–349. doi: 10.1046/j.1365-3113X.2003.01727.x
- Baima, S., Possenti, M., Matteucci, A., Wisman, E., Altamura, M. M., Ruberti, I., et al. (2001). The Arabidopsis ATHB-8 HD-Zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiol.* 126, 643–655. doi: 10.1104/pp.126.2.643
- Barnes, J. R., Lorenz, W. W., and Dean, J. F. D. (2008). Characterization of a 1-aminocyclopropane-1-carboxylate synthase gene from loblolly pine (*Pinus taeda* L.). *Gene* 413, 18–31. doi: 10.1016/j.gene.2007.12.024
- Binder, B. M., Walker, J. M., Gagne, J. M., Emborg, T. J., Hemmann, G., Blecker, A. B., et al. (2007). The Arabidopsis EIN3 binding F-Box proteins EBF1 and EBF2 have distinct but overlapping roles in ethylene signaling. *Plant Cell* 19, 509–523. doi: 10.1105/tpc.106.048140
- Birol, I., Raymond, A., Jackman, S. D., Pleasance, S., Coope, R., Taylor, G. A., et al. (2013). Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinform. Oxf. Engl.* 29, 1492–1497. doi: 10.1093/bioinformatics/btt178
- Brown, K. M., and Leopold, A. C. (1973). Ethylene and the regulation of growth in pine. *Can. J. For. Res.* 3, 143–145. doi: 10.1139/x73-019
- Bygdell, J., Srivastava, V., Obudulu, O., Srivastava, M. K., Nilsson, R., Sundberg, B., et al. (2017). Protein expression in tension wood formation monitored at high tissue resolution in Populus. *J. Exp. Bot.* 68, 3405–3417. doi: 10.1093/jxb/erx186
- Caño-Delgado, A., Yin, Y., Yu, C., Vafeados, D., Mora-García, S., Cheng, J.-C., et al. (2004). BRL1 and BRL3 are novel brassinosteroid receptors that function in vascular differentiation in Arabidopsis. *Dev. Camb. Engl.* 131, 5341–5351. doi: 10.1242/dev.01403
- Chang, K. N., Zhong, S., Weirauch, M. T., Hon, G., Pelizzola, M., Li, H., et al. (2013). Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in Arabidopsis. *eLife* 2:e00675. doi: 10.7554/eLife.00675
- Chao, Q., Rothenberg, M., Solano, R., Roman, G., Terzaghi, W., and Ecker, J. R. (1997). Activation of the ethylene gas response pathway in Arabidopsis by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. *Cell* 89, 1133–1144.
- Clark, K. L., Larsen, P. B., Wang, X., and Chang, C. (1998). Association of the Arabidopsis CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5401–5406.
- Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for agrobacterium-mediated transformation of arabidopsis thaliana. *Plant J. Mol. Biol.* 16, 735–743.
- Donner, T. J., Sherr, I., and Scarpella, E. (2009). Regulation of preprocambial cell state acquisition by auxin signaling in Arabidopsis leaves. *Dev. Camb. Engl.* 136, 3235–3246. doi: 10.1242/dev.037028
- Du, S., and Yamamoto, F. (2007). An overview of the biology of reaction wood formation. *J. Integr. Plant Biol.* 49, 131–143. doi: 10.1111/j.1744-7909.2007.00427.x
- Duval, I., Lachance, D., Giguère, I., Bomal, C., Morency, M.-J., Pelletier, G., et al. (2014). Large-scale screening of transcription factor–promoter interactions in spruce reveals a transcriptional network involved in vascular development. *J. Exp. Bot.* 65, 2319–2333. doi: 10.1093/jxb/eru116
- Eriksson, M. E., Israelsson, M., Olsson, O., and Moritz, T. (2000). Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nat. Biotechnol.* 18, 784–788. doi: 10.1038/77355
- Etchells, J. P., Provost, C. M., and Turner, S. R. (2012). Plant vascular cell division is maintained by an interaction between PXY and ethylene signalling. *PLoS Genet.* 8:e1002997. doi: 10.1371/journal.pgen.1002997
- Felten, J., Vahala, J., Love, J., Gorzsas, A., Ruggeberg, M., Delhomme, N., et al. (2018). Ethylene signaling induces gelatinous layers with typical features of tension wood in hybrid aspen. *New Phytologist*. doi: 10.1111/nph.15078
- Goecks, J., Nekrutenko, A., Taylor, J., and Galaxy Team (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* 11:R86. doi: 10.1186/gb-2010-11-8-r86
- Guzmán, P., and Ecker, J. R. (1990). Exploiting the triple response of Arabidopsis to identify ethylene-related mutants. *Plant Cell* 2, 513–523.
- Hamzeh, M., and Dayanandan, S. (2004). Phylogeny of Populus (*Salicaceae*) based on nucleotide sequences of chloroplast TRNT-TRNF region and nuclear rDNA. *Am. J. Bot.* 91, 1398–1408. doi: 10.3732/ajb.91.9.1398
- Huang, Y., Li, H., Hutchison, C. E., Laskey, J., and Kieber, J. J. (2003). Biochemical and functional analysis of CTR1, a protein kinase that negatively regulates ethylene signaling in Arabidopsis. *Plant J.* 33, 221–233. doi: 10.1046/j.1365-3113X.2003.01620.x
- Hussey, S., Mizrahi, E., Berger, D., and Myburg, A. (2011). The role of SND2 in the regulation of Arabidopsis fibre secondary cell wall formation. *BMC Proc.* 5:P114. doi: 10.1186/1753-6561-5-S7-P114
- Iglesias-Fernández, R., Barrero-Sicilia, C., Carrillo-Barral, N., Oñate-Sánchez, L., and Carbonero, P. (2013). Arabidopsis thaliana bZIP44: a transcription factor affecting seed germination and expression of the mannanase-encoding gene AtMAN7. *Plant J.* 74, 767–780. doi: 10.1111/tpj.12162
- Jokipii-Lukkari, S., Sundell, D., Nilsson, O., Hvidsten, T. R., Street, N. R., and Tuominen, H. (2017). NorWood: a gene expression resource for evo-devo studies of conifer wood development. *New Phytol.* 216, 482–494. doi: 10.1111/nph.14458
- Ju, C., Yoon, G. M., Shemansky, J. M., Lin, D. Y., Ying, Z. I., Chang, J., et al. (2012). CTR1 phosphorylates the central regulator EIN2 to control ethylene hormone signaling from the ER membrane to the nucleus in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19486–19491. doi: 10.1073/pnas.1214848109
- Kim, H. J., Hong, S. H., Kim, Y. W., Lee, I. H., Jun, J. H., Phee, B.-K., et al. (2014). Gene regulatory cascade of senescence-associated NAC transcription factors

- activated by ETHYLENE-INSENSITIVE2-mediated leaf senescence signalling in Arabidopsis. *J. Exp. Bot.* 65, 4023–4036. doi: 10.1093/jxb/eru112
- Ko, J.-H., Jeon, H.-W., Kim, W.-C., Kim, J.-Y., and Han, K.-H. (2014). The MYB46/MYB83-mediated transcriptional regulatory programme is a gatekeeper of secondary wall biosynthesis. *Ann. Bot.* 114, 1099–1107. doi: 10.1093/aob/mcu126
- Ko, J.-H., Yang, S. H., Park, A. H., Lerouxel, O., and Han, K.-H. (2007). ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in Arabidopsis thaliana. *Plant J.* 50, 1035–1048. doi: 10.1111/j.1365-313X.2007.03109.x
- Konishi, M., and Yanagisawa, S. (2008). Ethylene signaling in Arabidopsis involves feedback regulation via the elaborate control of EBF2 expression by EIN3. *Plant J. Cell Mol. Biol.* 55, 821–831. doi: 10.1111/j.1365-313X.2008.03551.x
- Leśniewska, J., Öhman, D., Krzesłowska, M., Kushwah, S., Barciszewska-Pacac, M., Kleczkowski, L. A., et al. (2017). Defense responses in aspen with altered pectin methylesterase activity reveal the hormonal inducers of tyloses. *Plant Physiol.* 173, 1409–1419. doi: 10.1104/pp.16.01443
- Li, W., Ma, M., Feng, Y., Li, H., Wang, Y., Ma, Y., et al. (2015). EIN2-directed translational regulation of ethylene signaling in Arabidopsis. *Cell* 163, 670–683. doi: 10.1016/j.cell.2015.09.037
- Licausi, F., Ohme-Takagi, M., and Perata, P. (2013). APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. *New Phytol.* 199, 639–649. doi: 10.1111/nph.12291
- Liu, L., Ramsay, T., Zinkgraf, M., Sundell, D., Street, N. R., Filkov, V., et al. (2015). A resource for characterizing genome-wide binding and putative target genes of transcription factors expressed during secondary growth and wood formation in Populus. *Plant J. Cell Mol. Biol.* 82, 887–898. doi: 10.1111/tpj.12850
- Liu, Y., Wei, M., Hou, C., Lu, T., Liu, L., Wei, H., et al. (2017). Functional characterization of populus PsnSHN2 in coordinated regulation of secondary wall components in tobacco. *Sci. Rep.* 7:42. doi: 10.1038/s41598-017-00093-z
- Love, J., Björklund, S., Vahala, J., Hertzberg, M., Kangasjärvi, J., and Sundberg, B. (2009). Ethylene is an endogenous stimulator of cell division in the cambial meristem of Populus. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5984–5989. doi: 10.1073/pnas.0811660106
- Ma, R., Xiao, Y., Lv, Z., Tan, H., Chen, R., Li, Q., et al. (2017). AP2/ERF Transcription Factor, I049, positively regulates lignan biosynthesis in isatis indigotica through activating salicylic acid signaling and Lignan/Lignin pathway genes. *Front. Plant Sci.* 8:1361. doi: 10.3389/fpls.2017.01361
- 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
- Mellerowicz, E. J., Baucher, M., Sundberg, B., and Boerjan, W. (2001). Unravelling cell wall formation in the woody dicot stem. *Plant Mol. Biol.* 47, 239–274. doi: 10.1023/A:1010699919325
- Merchante, C., Alonso, J. M., and Stepanova, A. N. (2013). Ethylene signaling: simple ligand, complex regulation. *Curr. Opin. Plant Biol.* 16, 554–560. doi: 10.1016/j.pbi.2013.08.001
- Merchante, C., Brumos, J., Yun, J., Hu, Q., Spencer, K. R., Enriquez, P., et al. (2015). Gene-specific translation regulation mediated by the hormone-signaling molecule EIN2. *Cell* 163, 684–697. doi: 10.1016/j.cell.2015.09.036
- Müller, M., and Munné-Bosch, S. (2015). Ethylene response factors: a key regulatory hub in hormone and stress signaling. *Plant Physiol.* 169, 32–41. doi: 10.1104/pp.15.00677
- Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of *Eucalyptus grandis*. *Nature* 510, 356–362. doi: 10.1038/nature13308
- Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211
- Obudulu, O., Bygdell, J., Sundberg, B., Moritz, T., Hvidsten, T. R., Trygg, J., et al. (2016). Quantitative proteomics reveals protein profiles underlying major transitions in aspen wood development. *BMC Genomics* 17:119. doi: 10.1186/s12864-016-2458-z
- Porth, I., Klapšte, J., Skyba, O., Hannemann, J., McKown, A. D., Guy, R. D., et al. (2013). Genome-wide association mapping for wood characteristics in Populus identifies an array of candidate single nucleotide polymorphisms. *New Phytol.* 200, 710–726. doi: 10.1111/nph.12422
- Prescott, A. M., McCollough, F. W., Eldreth, B. L., Binder, B. M., and Abel, S. M. (2016). Analysis of network topologies underlying ethylene growth response kinetics. *Front. Plant Sci.* 7:1308. doi: 10.3389/fpls.2016.01308
- Quan, R., Wang, J., Yang, D., Zhang, H., Zhang, Z., and Huang, R. (2017). EIN3 and SOS2 synergistically modulate plant salt tolerance. *Sci. Rep.* 7:44637. doi: 10.1038/srep44637
- Sakamoto, S., Takata, N., Oshima, Y., Yoshida, K., Taniguchi, T., and Mitsuda, N. (2016). Wood reinforcement of poplar by rice NAC transcription factor. *Sci. Rep.* 6:19925. doi: 10.1038/srep19925
- Salojärvi, J., Smolander, O.-P., Nieminen, K., Rajaraman, S., Safronov, O., Safdari, P., et al. (2017). Genome sequencing and population genomic analyses provide insights into the adaptive landscape of silver birch. *Nat. Genet.* 49, 904–912. doi: 10.1038/ng.3862
- Sehr, E. M., Agusti, J., Lehner, R., Farmer, E. E., Schwarz, M., and Greb, T. (2010). Analysis of secondary growth in the Arabidopsis shoot reveals a positive role of jasmonate signalling in cambium formation. *Plant J. Cell Mol. Biol.* 63, 811–822. doi: 10.1111/j.1365-313X.2010.04283.x
- Shakeel, S. N., Gao, Z., Amir, M., Chen, Y.-F., Rai, M. I., Haq, N. U., et al. (2015). Ethylene regulates levels of ethylene receptor/CTR1 signaling complexes in Arabidopsis thaliana. *J. Biol. Chem.* 290, 12415–12424. doi: 10.1074/jbc.M115.652503
- Shin, K., Lee, S., Song, W.-Y., Lee, R.-A., Lee, I., Ha, K., et al. (2015). Genetic Identification of ACC-RESISTANT2 reveals involvement of LYSINE HISTIDINE TRANSPORTER1 in the uptake of 1-Aminocyclopropane-1-Carboxylic Acid in Arabidopsis thaliana. *Plant Cell Physiol.* 56, 572–582. doi: 10.1093/pcp/pcu201
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75
- Stepanova, A. N., and Alonso, J. M. (2009). Ethylene signaling and response: where different regulatory modules meet. *Curr. Opin. Plant Biol.* 12, 548–555. doi: 10.1016/j.pbi.2009.07.009
- Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y.-C., Sjödin, A., et al. (2015). The plant genome integrative explorer resource: plantgenie.org. *New Phytol.* 208, 1149–1156. doi: 10.1111/nph.13557
- Sundell, D., Street, N. R., Kumar, M., Mellerowicz, E. J., Kucukoglu, M., Johnsson, C., et al. (2017). AspWood: high-spatial-resolution transcriptome profiles reveal uncharacterized modularity of wood formation in Populus tremula. *Plant Cell* 29, 1585–1604. doi: 10.1105/tpc.17.00153
- Taylor-Teeple, 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
- Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691
- Vahala, J., Felten, J., Love, J., Gorzsás, A., Gerber, L., Lamminmäki, A., et al. (2013). A genome-wide screen for ethylene-induced ethylene response factors (ERFs) in hybrid aspen stem identifies ERF genes that modify stem growth and wood properties. *New Phytol.* 200, 511–522. doi: 10.1111/nph.12386
- Valente, T. W., Coronges, K., Lakon, C., and Costenbader, E. (2008). How correlated are network centrality measures? *Connect. Tor.* 28, 16–26.
- Vanderstraeten, L., and Van Der Straeten, D. (2017). Accumulation and Transport of 1-Aminocyclopropane-1-Carboxylic Acid (ACC) in plants: current status, considerations for future research and agronomic applications. *Front. Plant Sci.* 8:38. doi: 10.3389/fpls.2017.00038
- Voiniciuc, C., Yang, B., Schmidt, M. H.-W., Günl, M., and Usadel, B. (2015). Starting to Gel: how Arabidopsis seed coat epidermal cells produce specialized secondary cell walls. *Int. J. Mol. Sci.* 16, 3452–3473. doi: 10.3390/ijms16023452
- Wang, S., Yao, W., Wei, H., Jiang, T., and Zhou, B. (2014). Expression patterns of ERF genes underlying abiotic stresses in Di-haploid *Populus simonii* × *P. nigra*. *Sci. World J.* 2014:745091. doi: 10.1155/2014/745091
- Wen, X., Zhang, C., Ji, Y., Zhao, Q., He, W., An, F., et al. (2012). Activation of ethylene signaling is mediated by nuclear translocation of the cleaved EIN2 carboxyl terminus. *Cell Res.* 22, 1613–1616. doi: 10.1038/cr.2012.145

- Xu, J., and Zhang, S. (2014). Regulation of ethylene biosynthesis and signaling by protein kinases and phosphatases. *Mol. Plant* 7, 939–942. doi: 10.1093/mp/ssu059
- Xu, P., Cai, X.-T., Wang, Y., Xing, L., Chen, Q., and Xiang, C.-B. (2014). HDG11 upregulates cell-wall-loosening protein genes to promote root elongation in Arabidopsis. *J. Exp. Bot.* 65, 4285–4295. doi: 10.1093/jxb/eru202
- Yamagami, T., Tsuchisaka, A., Yamada, K., Haddon, W. F., Harden, L. A., and Theologis, A. (2003). Biochemical Diversity among the 1-Amino-cyclopropane-1-Carboxylate synthase isozymes encoded by the Arabidopsis gene family. *J. Biol. Chem.* 278, 49102–49112. doi: 10.1074/jbc.M308297200
- 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
- Yao, W., Wang, L., Zhou, B., Wang, S., Li, R., and Jiang, T. (2016). Over-expression of poplar transcription factor ERF76 gene confers salt tolerance in transgenic tobacco. *J. Plant Physiol.* 198, 23–31. doi: 10.1016/j.jplph.2016.03.015
- Yao, W., Zhang, X., Zhou, B., Zhao, K., Li, R., and Jiang, T. (2017). Expression Pattern of ERF gene family under multiple abiotic stresses in *Populus simonii* × *P. nigra*. *Front. Plant Sci.* 8:181. doi: 10.3389/fpls.2017.00181
- Zhang, F., Qi, B., Wang, L., Zhao, B., Rode, S., Riggan, N. D., et al. (2016). EIN2-dependent regulation of acetylation of histone H3K14 and non-canonical histone H3K23 in ethylene signalling. *Nat. Commun.* 7:13018. doi: 10.1038/ncomms13018
- Zhang, F., Wang, L., Qi, B., Zhao, B., Ko, E. E., Riggan, N. D., et al. (2017). EIN2 mediates direct regulation of histone acetylation in the ethylene response. *Proc. Natl. Acad. Sci. U.S.A.* 114, 10274–10279. doi: 10.1073/pnas.1707937114
- Zhang, J., Nieminen, K., Serra, J. A. A., and Helariutta, Y. (2014). The formation of wood and its control. *Curr. Opin. Plant Biol.* 17, 56–63. doi: 10.1016/j.pbi.2013.11.003
- Zhong, R., McCarthy, R. L., Haghighat, M., and Ye, Z.-H. (2013). The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PLoS ONE* 8:e69219. doi: 10.1371/journal.pone.0069219
- Zhong, S., Zhao, M., Shi, T., Shi, H., An, F., Zhao, Q., et al. (2009). EIN3/EIL1 cooperate with PIF1 to prevent photo-oxidation and to promote greening of Arabidopsis seedlings. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21431–21436. doi: 10.1073/pnas.0907670106

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer MB declared a past co-authorship with one of the authors HT to the handling Editor.

Copyright © 2018 Seyfferth, Wessels, Jokipii-Lukkari, Sundberg, Delhomme, Felten and Tuominen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership