BIOFUELS AND BIOENERGY

EDITED BY: Jin Zhang, Jaime Barros and Meng-Zhu Lu

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88966-392-7
DOI 10.3389/978-2-88966-392-7

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

BIOFUELS AND BIOENERGY

Topic Editors:
Jin Zhang, Zhejiang Agriculture & Forestry University, China
Jaime Barros, University of North Texas, United States
Meng-Zhu Lu, Zhejiang Agriculture & Forestry University, China

Citation: Zhang, J., Barros, J., Lu, M.-Z., eds. (2021). Biofuels and Bioenergy. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-392-7

Table of Contents

frontiers in Plant Science

EDITORIAL
published: 30 November 2020
doi: 10.3389/fpls.2020.621380

Editorial: Biofuels and Bioenergy

Jin Zhang[1,2]*, *Jaime Barros-Rios*[3] *and Mengzhu Lu*[1]

[1] State Key Laboratory of Subtropical Silviculture, School of Forestry and Biotechnology, Zhejiang A&F University, Hangzhou, China, [2] Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [3] Department of Biological Sciences, BioDiscovery Institute, University of North Texas, Denton, TX, United States

Keywords: biofuels, bioenergy, biomass, plant cell wall, bioengineering

Editorial on the Research Topic

Biofuels and Bioenergy

Plant cell wall biosynthesis is critical for plant growth and development, and a major global carbon sink in the biosphere. Plant secondary cell walls includes three main biological polymers: cellulose, hemicellulose, and lignin (Pauly and Keegstra, 2010). In the past half century, the research on the biosynthesis and metabolism of these three biopolymers has attracted the attention of researchers (Li et al., 2014; Zeng et al., 2014; Kumar et al., 2016; Xie et al., 2018; Polko and Kieber, 2019; Vanholme et al., 2019). In this Research Topic, we aim to gather knowledge about current advances in the field of plant biotechnology for bioenergy production. We have integrated several key components such as technology optimization for biorefineries, genetic modification to improve bioenergy crops, and gene discovery for biofuel and bioenergy production. Together, this knowledge will support the promotion of sustainable biofuels and bioproducts to contribute to the bioeconomy on a global scale.

This Research Topic, Biofuels and Bioenergy, including eight original research articles, four review articles, and one perspective paper, covers the following three topics.

OPEN ACCESS

Edited and reviewed by:
Maurice Bosch,
Aberystwyth University,
United Kingdom

***Correspondence:**
Jin Zhang
zhangj@zafu.edu.cn

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *26 October 2020*
Accepted: *09 November 2020*
Published: *30 November 2020*

Citation:
Zhang J, Barros-Rios J and Lu M (2020) Editorial: Biofuels and Bioenergy. Front. Plant Sci. 11:621380. doi: 10.3389/fpls.2020.621380

TECHNOLOGY OPTIMIZATION FOR BIOFUELS AND BIOENERGY

In the past few decades, cell wall engineering has been a key method to reduce biomass recalcitrance by altering the structure of lignin (Sattler and Funnell-Harris, 2013). Perturbations of genes in the lignin biosynthetic pathway lead to major structural changes, allowing the design of easily controlled biomass structures to produce biofuels and chemicals (Zhang et al., 2019; Fan et al., 2020). In this Research Topic, Kim et al. reported an integrated strategy combining biomass genetic engineering with a pretreatment using a bio-derived deep eutectic solvent (DES). Kim et al. strategically expressed a bacterial hydroxycinnamoyl-CoA hydratase-lyase (HCHL) gene using the lignifying tissue-specific *IRX5* promoter in *Arabidopsis* to reduce the degree of lignin polymerization via incorporation of side-chain-truncated monomers in lignin polymer ends. After pretreatment with the lignin-derived DES at mild conditions, the *IRX5:HCHL-1* transgenic *Arabidopsis* yielded higher levels of fermentable sugars. This approach could support the development of new raw materials for the optimization of the production process in biorefineries.

As one of the main components of plant cell walls, cellulose is the most abundant biopolymer on the planet (Mizrachi et al., 2012). Therefore, understanding the structure of cellulose can provide a basis for improving bioenergy plants by altering the cell wall components. Previous studies using atomic force microscopy (AFM) showed the intricate details of cellulose microfibril arrangement in plant cell walls (Kirby et al., 1996; Ding et al., 2014). However, the image quality is insufficient to resolve the dimension and the cross-sectional shape of cellulose microfibrils. In this Research Topic, Song et al. used the advantages of AFM imaging under aqueous condition and a combined effort of sample preparation, pre-selection of tips with 1 nm radius and systematic adjustment of

imaging parameters to optimize image quality, and analyzed the images of primary and secondary cell walls at the sub-nanometer scale.

When compared to the model plant *Arabidopsis*, many species used as bioenergy crops are recalcitrant to both transformation and *in vitro* regeneration (Clifton-Brown et al., 2019). As a fast-growing species, willow (*Salix* spp. L.) has attracted attention due to its potential as a raw material for bioenergy and biofuel production (Nordborg et al., 2018). In this Research Topic, Gomes et al. proposed a reproducible, rapid, and highly efficient *Agrobacterium rhizogenes*-mediated hairy root transformation system for *S. purpurea*. The transformed *Salix* hairy roots can be easily and quickly selected by fluorescent markers. This transformation system provides an effective method of gene function validation for improving bioenergy plants.

In addition to those traditional plants used for bioenergy, more and more plant species are being tested as potential raw materials for bioenergy and biofuel applications. White mustard (*Sinapis alba* L.) is an annual plant of the family Brassicaceae that originates from the Mediterranean region. Mitrović et al. summarized the white mustard seed oil as a promising feedstock for biodiesel production and discussed its fuel properties and performance.

GENETIC MODIFICATION TO IMPROVE BIOENERGY CROPS

The ideal bioenergy crop should contain a high proportion of C6-sugars in polysaccharides, and genetic modification of carbohydrate active enzymes can reduce the recalcitrance of the cell wall. In our Research Topic, Brandon and Scheller reviewed progress in the use of a variety of dominant genetic engineering strategies to improve biomass for bioenergy conversion and modulate polysaccharide biosynthesis. Wang et al. expressed a *Hypocrea jecorina* acetyl xylan esterase (*HjAXE*), a member of carbohydrate esterase (CE) family 5, driven by a wood-specific promoter (*PtGT43B* promoter) in hybrid poplar. AX was predicted to deacetylate polymeric xylan in the vicinity of cellulose due to the presence of a cellulose-binding module. Expression of *HjAXE* leads to reduced xylan acetylation and ∼30% increased glucose yields in enzymatic saccharification of wood without pretreatment, indicating that the CE5 family could be used as a source of enzymes to reduce biomass recalcitrance in poplar. In addition, Debra-Maceluch et al. evaluated the field performance of several acetylation reduced poplar lines including *HjAXE*, *AnAXE1* (*Aspergillus niger AXE1*, a CE1 family member) and *RWA* (*REDUCED WALL ACETYLATION*), driven by either a constitutive 35S promoter or a wood-specific promoter. This provides data from early field trials for evaluating genetic modification strategies and assessing the potential pros and cons of using genetic modified crops compared to non-genetic modified commercial crops. In this Research Topic, Mazarei et al. reported that genetic manipulation of a *Panicum virgatum* folylpolyglutamate synthetase gene (*PvFPGS1*) in the one-carbon pathway in switchgrass could lead to improved biofuel production without negatively impacting plant growth and biomass yield.

Moreover, genetic modification of microorganisms used for photo-bioreactors could enhance the efficiency of fermentation. Velmurugan and Incharoensakdi engineered the *Synechococcus elongatus* PCC7942 by knockout ADP-glucose pyrophosphorylase gene *glgC* and insertion of the two pyruvate decarboxylase genes *pdc-adh* from two different microorganisms; and increased the ethanol synthesis in the system. This provide another strategy to enhance biofuel products in the fermentation stage.

GENE DISCOVERY FOR BIOFUEL AND BIOENERGY

Although the transcriptional regulation model of the secondary cell wall has been established in plants in the past 10 years (Zhong et al., 2010; Zhang et al., 2018a), exploring new genes related to the formation of secondary cell wall and their post-transcriptional regulation is still an important issue. In the current Research Topic, Zhang J. et al. summarized recent progress on the transcriptional and post-transcriptional regulatory model of lignin biosynthesis pathway genes in the woody model plant *Populus*. For post-translational protein modification, Sulis and Wang review the phosphorylation, ubiquitination, glycosylation, and S-nitrosylation of transcription factors in monolignols biosynthetic enzymes, and peroxidases in the regulation of lignin biosynthesis and polymerization.

With the rapid development of next-generation sequencing technologies, RNA-Seq data provides a rich resource for gene expression and data mining. Zhang L. et al. analyzed the structural differences of developing xylem in five black poplar cultivars and profiled the transcriptome-wide gene expression. Based on a weighted gene co-expression network analysis, they identified a set of promising candidate regulators for genetic engineering to improve feedstock and enhance biofuel conversion.

Genome-wide association study (GWAS) is an effective strategy using "unstructured" populations to identify significant trait associations with small numbers of candidate genes with high resolution (Porth et al., 2013; Zhang et al., 2018b). In this Research Topic, Chhetri et al. used 869 unrelated *Populus trichocarpa* genotypes from a common garden and tested the association of 25 wood anatomical phenotypic traits and 9 multi-trait combinations with 6.741 million single nucleotide polymorphisms (SNPs). Combining with network-based Lines of Evidence (LOE) method, they identified GWAS hits that are strong gene candidates for experimental validation by integrating data from multiple sources. This study provides insights into the type of genes controlling wood anatomical traits for biotechnological approaches toward optimizing wood traits for biofuel production.

CONCLUDING REMARKS

The 13 articles on this Research Topic only covered a small portion of the current progress in the field of plant biotechnology

for bioenergy production. We hope that this collection of research articles contributes with valuable information for researchers and practitioners in the biofuel and bioenergy industry chain. As more researchers work collaboratively on this field, the output will accelerate the development of new bioenergy technologies, and ultimately promote the use of sustainable biofuels in the future bioeconomy.

AUTHOR CONTRIBUTIONS

JZ wrote the draft. JB-R and ML made editing and approved the final version for publication. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank all authors who submitted their work for this Research Topic as well as the invaluable help of reviewers in manuscript evaluation and the support of professional editorial staff at Frontiers. JZ was funded by the Zhejiang A&F University Research and Development Fund Talent Startup Project. ML was funded by the Zhejiang Provincial Department of Science and Technology Project (2016C02056-5-5). JB-R acknowledges a Center for Bioenergy Innovation (CBI)-Early Career Research Fellowship received as Visiting Fellow at Oak Ridge National Laboratory, during which period this Research Topic was compiled.

REFERENCES

Clifton-Brown, J., Harfouche, A., Casler, M. D., Dylan Jones, H., Macalpine, W. J., Murphy-Bokern, D., et al. (2019). Breeding progress and preparedness for mass-scale deployment of perennial lignocellulosic biomass crops switchgrass, miscanthus, willow and poplar. *Gcb Bioenergy* 11, 118–151. doi: 10.1111/gcbb.12566

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

Fan, D., Li, C., Fan, C., Hu, J., Li, J., Yao, S., et al. (2020). MicroRNA6443-mediated regulation of FERULATE 5-HYDROXYLASE gene alters lignin composition and enhances saccharification in *Populus tomentosa*. *New Phytol.* 226, 410–425. doi: 10.1111/nph.16379

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

Kumar, M., Campbell, L., and Turner, S. (2016). Secondary cell walls: biosynthesis and manipulation. *J. Exp. Bot.* 67, 515–531. doi: 10.1093/jxb/erv533

Li, Q., Song, J., Peng, S., Wang, J. P., Qu, G. Z., Sederoff, R. R., et al. (2014). Plant biotechnology for lignocellulosic biofuel production. *Plant Biotechnol. J.* 12,1174–1192. doi: 10.1111/pbi.12273

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

Nordborg, M., Berndes, G., Dimitriou, I., Henriksson, A., Mola-Yudego, B., and Rosenqvist, H. (2018). Energy analysis of willow production for bioenergy in Sweden. *Renew. Sustain. Energy Rev.* 93, 473–482. doi: 10.1016/j.rser.2018.05.045

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

Polko, J. K., and Kieber, J. J. (2019). The regulation of cellulose biosynthesis in plants. *Plant Cell* 31, 282–296. doi: 10.1105/tpc.18.00760

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

Sattler, S., and Funnell-Harris, D. (2013). Modifying lignin to improve bioenergy feedstocks: strengthening the barrier against pathogens? *Front. Plant Sci.* 4:70. doi: 10.3389/fpls.2013.00070

Vanholme, R., De Meester, B., Ralph, J., and Boerjan, W. (2019). Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.* 56, 230–239. doi: 10.1016/j.copbio.2019.02.018

Xie, M., Zhang, J., Tschaplinski, T. J., Tuskan, G. A., 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

Zeng, Y., Zhao, S., Yang, S., and Ding, S.-Y. (2014). Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Curr. Opin. Biotechnol.* 27, 38–45. doi: 10.1016/j.copbio.2013.09.008

Zhang, J., Li, M., Bryan, A. C., Yoo, C. G., Rottmann, W., Winkeler, K. A., et al. (2019). Overexpression of a serine hydroxymethyltransferase increases biomass production and reduces recalcitrance in the bioenergy crop Populus. *Sustain. Energy Fuels* 3, 195–207. doi: 10.1039/C8SE00471D

Zhang, J., Xie, M., Tuskan, G. A., Muchero, W., and Chen, J.-G. (2018a). 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

Zhang, J., Yang, Y., Zheng, K., Xie, M., Feng, K., Jawdy, S. S., et al. (2018b). 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.* 220, 502–516. doi: 10.1111/nph.15297

Zhong, R., Lee, C., and Ye, Z.-H. (2010). 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

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.

frontiers in Plant Science

PERSPECTIVE
published: 11 November 2019
doi: 10.3389/fpls.2019.01427

Hairy Root Transformation: A Useful Tool to Explore Gene Function and Expression in *Salix* spp. Recalcitrant to Transformation

*Carolina Gomes[1], Annabelle Dupas[2], Andrea Pagano[1,3], Jacqueline Grima-Pettenati[2] and Jorge Almiro P. Paiva[1]**

[1] Department of Integrative Plant Biology, Institute of Plant Genetics, Polish Academy of Sciences, Poznan, Poland,
[2] LRSV, Laboratoire de Recherche en Sciences Végétales, UPS, CNRS, Université Toulouse 3, Castanet Tolosan, France,
[3] Department of Biology and Biotechnology "L. Spallanzani", University of Pavia, Pavia, Italy

Willow (*Salix* spp. L.) species are fast-growing trees and shrubs that have attracted emergent attention for their potential as feedstocks for bioenergy and biofuel production, as well as for pharmaceutical and phytoremediation applications. This economic and environmental potential has propelled the creation of several genetic and genomic resources for *Salix* spp. Furthermore, the recent availability of an annotated genome for *Salix purpurea* has pinpointed novel candidate genes underlying economically relevant traits. However, functional studies have been stalled by the lack of rapid and efficient coupled regeneration-transformation systems for *Salix purpurea* and *Salix* spp. in general. In this report, we describe a fast and highly efficient hairy root transformation protocol for *S. purpurea.* It was effective for different explant sources and *S. purpurea* genotypes, with efficiencies between 63.4% and 98.7%, and the screening of the transformed hairy roots was easily carried out using the fluorescent marker DsRed. To test the applicability of this hairy root transformation system for gene functional analysis, we transformed hairy roots with the vector pGWAY-*SpDRM2*, where the gene *SpDRM2* encoding a putative Domain Rearranged Methyltransferase (DRM) was placed under the control of the CaMV 35S constitutive promoter. Indeed, the transgenic hairy roots obtained exhibited significantly increased expression of *SpDRM2 as* compared to controls, demonstrating that this protocol is suitable for the medium/high-throughput functional characterization of candidate genes in *S. purpurea and* other recalcitrant *Salix* spp.

Keywords: *Salix purpurea*, willow, domains rearranged methyltransferase 2 (DRM2), *Agrobacterium rhizogenes*-mediated transformation, pGWAY-0

OPEN ACCESS

Edited by:
Jin Zhang,
Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Zhenzhen Qiao,
Oak Ridge National Laboratory, United States
Aleksandra Krolicka,
University of Gdansk, Poland

***Correspondence:**
Jorge Almiro P. Paiva
jpai@igr.poznan.pl

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *20 August 2019*
Accepted: *15 October 2019*
Published: *11 November 2019*

Citation:
Gomes C, Dupas A, Pagano A, Grima-Pettenati J and Paiva JAP (2019) Hairy Root Transformation: A Useful Tool to Explore Gene Function and Expression in Salix spp. Recalcitrant to Transformation. Front. Plant Sci. 10:1427. doi: 10.3389/fpls.2019.01427

INTRODUCTION

Salix spp. L. (willows) are very diverse, comprising more than 400 identified species spread over a wide variety of natural habitats (Sulima et al., 2018). Willows display a high morphological diversity, occurring in the growth forms of trees, shrubs, or subshrubs. Shrub willows (*Vetrix* sub-genus) are ideal biomass feedstocks for bioenergy and biofuel applications given the ease of vegetative propagation, fast growth in short-rotation coppices (SRC) and high biomass yields (Kuzovkina et al., 2008; Karp et al., 2011). Some *Salix* spp. can also be used in phytoremediation strategies as they are characterized by physiological adaptations and ecological resilience, rendering them particularly suitable for the clean-up

of environmental contaminants (Kuzovkina and Quigley, 2005; Zalesny and Bauer, 2007). Besides, willow species such as *Salix purpurea* L. (purple willow) have great potential for the production of natural alternatives to synthetic aspirin, as their bark is a source of salicylic glycosides (SGs) (Sulima et al., 2017).

In the last few years, the development of genetic and genomics tools for *Salix* spp., together with extensive phenotyping efforts, have significantly extended our knowledge on the factors involved in trait determination and phenotypic adaptation in willows (Hanley and Karp, 2014; Sulima et al., 2017; Fabio and Smart, 2018; Sulima et al., 2018; Gouker et al., 2019). Furthermore, the recent availability of the *S. purpurea* genome (https://phytozome.jgi.doe.gov) coupled to transcriptomic studies (Carlson et al., 2017; Yanitch et al., 2017) allowed this species to become a model for the *Salix* genus. However, the functional characterization of new candidate genes has been hindered by the lack of rapid and efficient regeneration and transformation protocols for *S. purpurea* and *Salix* spp. in general.

Salix spp. are recalcitrant to both transformation and *in vitro* regeneration. There are few reports describing the *in vitro* regeneration of *Salix* plants (Grönroos et al., 1989; Stoehr et al., 1989; Lyyra et al., 2006), but only one reported the regeneration of a significant number of plantlets (Stoehr et al., 1989). Early attempts to regenerate transgenic willow species were also proven ineffective, as no shoots were regenerated from transformed *calli* (Vahala et al., 1989, Vahala et al.,1993). Yang et al. (2013) developed a coupled regeneration-transformation system for *Salix matsudana* Koidz using the embryo apical region of mature seeds as initial explant. Shoots were regenerated directly from cotyledonary nodes. The average transformation frequency was low, approximately 7%, and moreover, this method required a laborious screening of transformants, given the chimeric nature of transgenic plants produced. More recently, Guan et al. (2018) developed an *Agrobacterium tumefaciens*-mediated genetic transformation system using leaf-based *calli* of *S. mongolica* as explants for transformation. Differentiation of adventitious buds and rooting of plantlets was accomplished by adding different ratios of 2,4-dichlorophenoxyacetic acid (2,4-D), 6-benzyl aminopurine (BA), and naphthaleneacetic acid (NAA) into Murashige and Skoog (MS) medium. The *Agrobacterium*-mediated integration of the *β*-glucuronidase (*gus*) gene into *S. mongolica* genome of five transgenic lines was confirmed by Southern blot (Guan et al., 2018), but no data on transformation efficiency were provided.

A. rhizogenes-mediated hairy root transformation systems are particularly useful for species recalcitrant to transformation by *A. tumefaciens*, since much higher transformation efficiencies are obtained, and shorter transformation periods are attainable in comparison to *A. tumefaciens*-mediated transformation systems. Hairy root transformation systems have been previously established in woody species such as *Prunus* spp. (Bosselut et al., 2011), *Populus* spp. (Yoshida et al., 2015), *Eucalyptus grandis* (Plasencia et al., 2016), *Camelia sinensis* (Alagarsamy et al., 2018), and *Dryas* spp. (Billault-Penneteau et al., 2019). In poplar, hairy roots were used to study the role of the transcription factor *MYB182* on the regulation of proanthocyanidin and anthocyanin biosynthesis (Yoshida et al., 2015). In *Eucalyptus*, hairy roots were shown to be a suitable system for the functional characterization of genes involved in lignin biosynthesis, such as the *Eucalyptus cinnamoyl-CoA reductase1* (*EgCCR1*) (Plasencia et al., 2016). To our best knowledge, in *Salix* spp., the induction of hairy roots has only been described in *Salix alba* L. (Hauth and Beiderbeck, 1992) in a report dating back to the early nineties. However, the aim of this study was solely to improve root biomass by the production of *A. rhizogenes*-induced hairy roots, as the normal roots cultures of *Salix alba* presented slow growth rates. Given the nature of this study, no confirmation of putatively transformed hairy root lines was done. Therefore, there is still a need to develop a hairy root transformation protocol that would allow the rapid characterization of gene function in *Salix* spp.

Here, we propose a reproducible, rapid and highly efficient *A. rhizogenes*-mediated hairy root transformation system for *S. purpurea* (**Figure 1**; details in **Supplementary Material S1**). In this method, the transformed hairy roots are detectable by fluorescent markers, allowing an easy and fast selection of transgenic roots. Our results suggest that this transformation system can potentially be applied to different genotypes, enabling gene functional studies in selected *S. purpurea* genotypes.

Hairy Roots: A Highly Efficient Strategy for Purple Willow Transformation

Different hairy root transformation protocols were tested to check the effect of explant age, plant culture media, and genotype on the transformation efficiency. We initially tested four transformation protocols combining one genotype at two developmental stages and two culture media (**Supplementary Table S2.1**—Experiment A). DsRed fluorescence was easily detected in neo-formed *calli* at wounding sites and in emerging roots in average 7 days after inoculation (dpi), i.e. (**Figure 1**). The transformation efficiencies obtained 21 dpi were higher for two-week-old *S. purpurea* plantlets grown in MS30 (reaching 83.33%) with the other three tested conditions presenting lower efficiencies (between 63.41% and 67.86%). To validate the results of the first experiment and to check if the developed protocol could be applicable to different *S. purpurea* genotypes, a second transformation experiment (**Supplementary Table S2.1**—Experiment B) was performed comparing two non-related genotypes, ELB3/6 and ELB2/5, using two-week-old *in vitro* clonal lines of *S. purpurea* grown in MS30, as this was shown to be the best condition to maximize the transformation efficiency. In this second experiment, the transformation efficiencies were slightly higher in both genotypes (above 86%), showing that the developed protocol is reproducible, rapid, and highly efficient (**Supplementary Table S2.1**—Experiment B). Moreover, we observed that isolated transformed hairy roots were able to continue to grow in solid media, and composite plants could be grown in hydroponic culture, allowing the maintenance and multiplication of composite plants and hairy roots for further experiments.

Application of Hairy Roots Transformation of *S. purpurea* for Gene Functional Analysis

To confirm that this transformation approach was suitable for functional gene characterization and gene function hypothesis-testing, we transformed roots with *SpDRM2* (SapurV1A.0571s0130), a gene encoding a putative DNA methyltransferase with rearranged catalytic domains (For

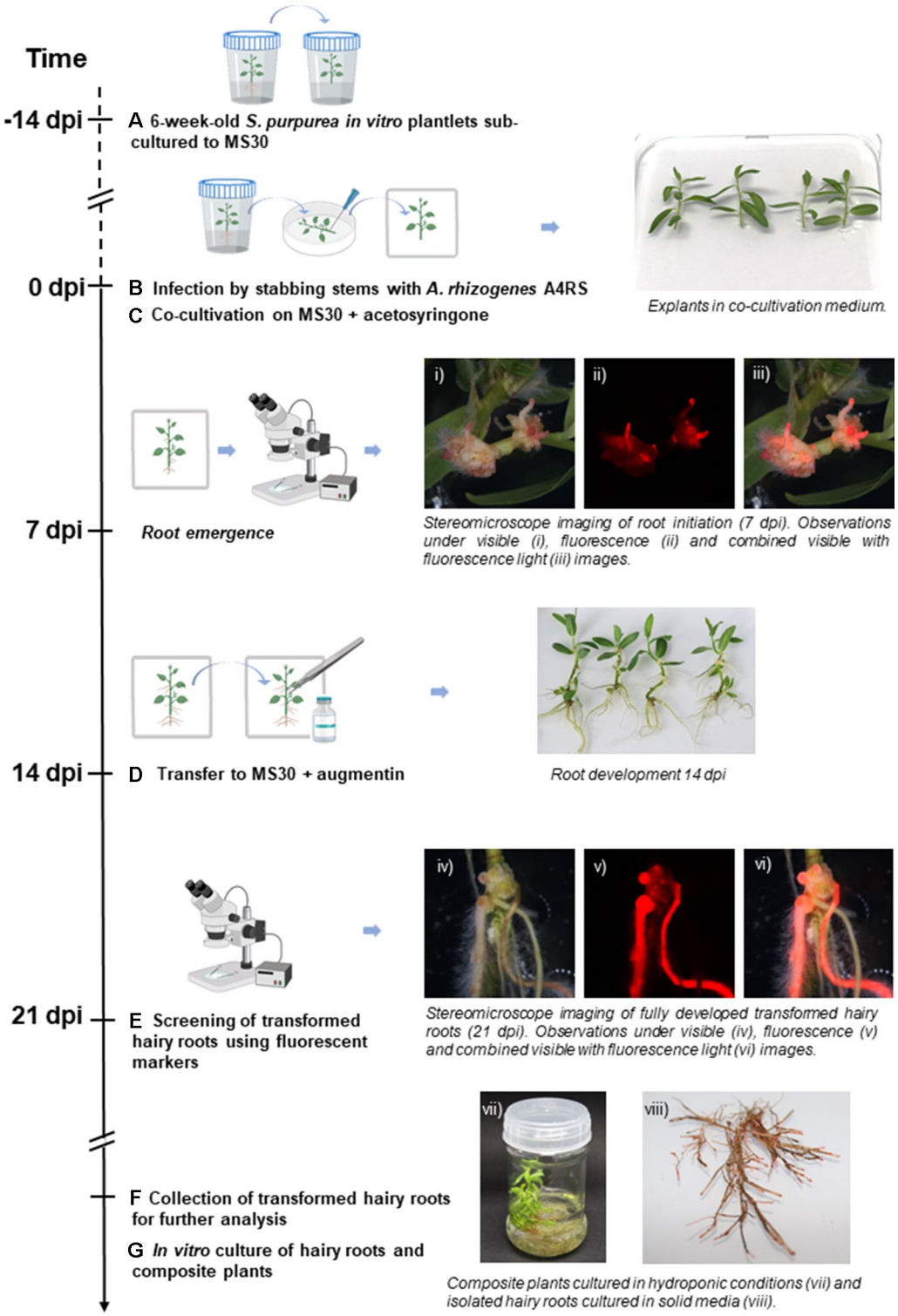

FIGURE 1 | Workflow of the hairy root transformation of *Salix purpurea in vitro* plantlets by A4RS harboring a DsRed-based binary vector. Emerging DsRed fluorescent roots are already detected 7 days post inoculation (dpi). **(A)** 6-week-old *S. purpurea in vitro* plantlets are sub-cultured to MS medium with full strength macroelements and 30 g L^{-1} sucrose (MS30). **(B)** 14-day-old *S. purpurea in vitro* plantlets were infected by stabbing the stem with a needle swabbed with *Agrobacterium rhizogenes*. **(C)** Infected plants were co-cultivated with *A. rhizogenes* for 14 days on MS30 supplemented with acetosyringone under dim light. **(D)** Plants were transferred to MS30 medium supplemented with Augmentin. **(E)** Generated hairy roots were examined at 21 dpi under a stereo fluorescence microscope. **(F)** Co-transformed roots were excised and collected for further analysis. **(G)** *In vitro* culture of hairy roots and composite plants. Figure adapted from Plasencia et al. (2016), introducing the protocol specificities of hairy roots transformation for *S. purpurea*.

details on cloning of SpDRM2 see **Supplementary Data S1**, **Supplementary Table S2.2**, and **Supplementary Figure S1**). This gene is an ortholog of *DRM2* of *Arabidopsis thaliana* (AT5G14620), where it was shown to encode an enzyme involved in *de novo* DNA methylation and gene silencing (Cao and Jacobsen, 2002). Transgenic hairy roots potentially containing the construct pGWAY-*SpDRM2* were screened using DsRed and transformation efficiencies were comparable to those obtained with pGWAY-0 (98.75% for 82 analyzed plants). Gene expression analysis by RT-qPCR confirmed the overexpression of *SpDRM2* in pools of hairy roots transformed with pGWAY-*SpDRM2*, collected from three different composite plants (**Figure 2**) (**Supplementary Data S1**, and **Supplementary Table S.2.3**). Comparing wild-type roots to roots transformed with pGWAY-0 (empty vector), no significant difference was observed in the transcript level of *SpDRM2*, indicating that the endogenous gene expression was not affected by the transformation process. In contrast, the expression level was significantly increased by 2.9-fold to 6.9-fold in *SpDRM2*-overexpressing lines as compared to wild-type (For details on cloning of SpDRM2 and RT-qPCR analysis of transformed hairy roots check S1- Detailed Material and methods, **Supplementary Table S2.2** and **Supplementary Table S2.3**). Based on the results obtained, we expect that this transformation method will be a valuable tool for the medium-/high-throughput functional characterization of candidate genes in *S. purpurea*.

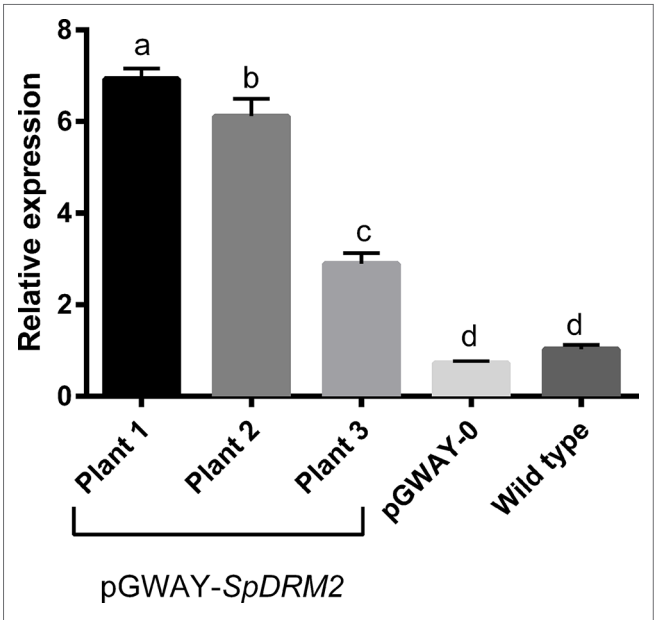

FIGURE 2 | Relative expression of SpDRM2 in transformed pGWAY-SpDRM2 hairy roots collected in three composite plants, in hairy roots transformed with pGWAY-0 (empty vectors) and wild-type roots. Relative expression is expressed in arbitrary units, by comparing with the expression in wild-type roots. Bars indicate the average of relative expression for pGWAY-DRM2 hairy roots (n = 3; technical replicates) and for pGWAY—0 hairy roots or wild type roots (n = 9 3 biological × 3 technical replicates). Standard error of mean is also indicated above the bars. Different letters above the bars indicate significant differences between conditions (Tukey's comparison test P-value < 0.05).

DISCUSSION

Current *A. tumefaciens*-mediated transformation procedures of willows (Vahala et al., 1989; Yang et al., 2013; Guan et al., 2018) still limit functional studies in *Salix* spp. As they are laborious, time-consuming, genotype-dependent, and hindered by low transformation efficiencies. Furthermore, no stable transformation protocol is currently available for *S. viminalis* and *S. purpurea*, two relevant willow species for the production of biofuel and pharmaceutical compounds, respectively. Thus, the use of *A. rhizogenes* to induce transformed hairy roots represents a novel approach in *Salix*, as these species are recalcitrant to regeneration and transformation with *A. tumefaciens*. Previous reports suggest that hairy roots can be induced on a wide range of woody species, using *in vitro* (Alpizar et al., 2006; Bosselut et al., 2011; Yoshida et al., 2015; Plasencia et al., 2016) and *in planta* strategies (Alagarsamy et al., 2018), and from different explant types, e.g. seedlings, roots, stems, and leaves. In the present study, *A. rhizogenes* strain A4RS harboring the vector PGWAY-0 (Plasencia-Casavedall, 2015; Plasencia et al., 2016) was proven highly effective at inducing transgenic hairy roots in *S. purpurea* in explants with different ages and different genotypes. Transgenic hairy roots were easily detectable using the fluorescent marker DsRed, with emerging transformed roots appearing in average 7 days after *A. rhizogenes* inoculation. The fluorescent marker DsRed facilitates the identification of transgenic roots (which can be selected in the absence or in addition to antibiotic selection), allowing the non-destructive and precocious identification of transgenic roots (Limpens et al, 2004; Chabaud et al., 2006; Meng et al., 2019). High efficiency of hairy root transformation (ranging from 63% to 98%) was achieved for all tested conditions. The hairy root transformation efficiencies obtained in this report are similar to those reported for other woody species (Alpizar et al., 2006; Bosselut et al., 2011; Plasencia et al., 2016; Alagarsamy et al., 2018), legumes (Estrada-Navarrete et al., 2007; Aggarwal et al., 2018), and tomato (Ho-Plágaro et al., 2018). However, most available protocols use seedlings as explant for hairy root induction while in this report we used *in vitro*-cultured plantlets of *S. purpurea* as explants. The use of *in vitro*-cultured clonal lines has the advantage of not requiring the sterilization and germination of seeds prior to transformation, which can be limiting steps in the protocol in case of seed contamination, poor seed germination, and limited seedling growth (Alagarsamy et al., 2018). Contrastingly, *in vitro*-cultured clonal lines only require simple maintenance procedures and can be an almost continuously available source of explants for hairy root induction experiments. Besides, the use of clonal lines allows testing the expression of target genes in a plant material with identical genetic background.

Using the developed protocol (**Figure 1**, and **Supplementary Material S1**) we were able to achieve very high hairy root co-transformation efficiencies for both tested *S. purpurea* genotypes, suggesting that this hairy root transformation protocol can potentially be transferred to other *S. purpurea* genotypes. Nonetheless, more experiments should be done to test the susceptibility to *A. rhizogenes* in different *S. purpurea* genotypes and other willows species. A main limitation of this protocol is that it does not allow the transformation of above ground plant tissues other than hairy roots. Still, this can also represent an advantage when investigating root-shoot interactions in composite plant systems, to study, for example, the transport

of small regulatory and signaling molecules (e.g. sRNAs, peptides and metabolites) between roots and aerial components of the plant. Transgenic hairy roots and composite plant systems can be used to study resistance against different biotic (Mellor et al., 2012; Xue et al., 2017) and abiotic stresses (Kajikawa et al., 2010; An et al., 2017; Du et al., 2018; Li et al., 2019), mycorrhizal associations, and root symbioses (Clemow et al., 2011; Indrasumunar et al., 2015; Billault-Penneteau et al., 2019). Furthermore, the use of the CRISPR/Cas9 technology with hairy root transformation has emerged as an efficient tool for plant genome editing and gene functional studies (Ron et al., 2014; Wang et al., 2016; Du et al., 2018), thus opening a whole new range of applications for *Salix* spp., such as the manipulation of target biosynthetic pathways through multiplexed genome editing. Indeed, as each single hairy root represents a single transformation event and can continue to grow autonomously, the system can be particularly useful as a medium-/high-throughput tool for functional analysis and biotechnological applications. The overexpression of *SpDRM2* in hairy roots transformed with the pGWAY-*SpDRM2* construct demonstrated the potential of *S. purpurea* hairy root transformation as a homologous, versatile, and efficient system that will enable the rapid validation of candidate genes and gene mining in *S. purpurea* and other recalcitrant willow species.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

JG-P and JP conceived the study and designed the experiments. CG, AD, AP, and JP performed the experiments and analysed the results. CG, AP, JG-P, and JP wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported NCN project Sonata-bis UMO-2015 /18/E/NZ2/00694 (PurpleWalls Project). CG acknowledges a PhD scholarship in the frame of PurpleWalls Project and the Short-Term Scientific Mission (STSM) grant in the frame of COST Action 15223. AP acknowledges the Erasmus+ program of the European Union and the IUSS (University School for Advanced Studies) of Pavia. JP acknowledges his research contract in the frame of EU-FP7-ERAChairs-PilotCAll-2013 project "Biotalent—The creation of the Department of Integrative Plant Biology" (FP7-REGPOT-621321) and the Polish financial sources for education in the years 2015–2019 allocated to an international co-financed project.

ACKNOWLEDGEMENTS

The authors thank Paweł Sulima and Jerzy A. Przyborowski (University of Warmia and Mazury in Olsztyn, Poland) for providing the plant material for this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01427/full#supplementary-material

REFERENCES

Aggarwal, P. R., Nag, P., Choudhary, P., Chakraborty, N., and Chakraborty, S. (2018). Genotype-independent Agrobacterium rhizogenes-mediated root transformation of chickpea: a rapid and efficient method for reverse genetics studies. *Plant Methods* 14, 55. doi: 10.1186/s13007-018-0315-6

Alagarsamy, K., Shamala, L. F., and Wei, S. (2018). Protocol: high-efficiency in-planta Agrobacterium-mediated transgenic hairy root induction of Camellia sinensis var. sinensis. *Plant Methods* 14, 17. doi: 10.1186/s13007-018-0285-8

Alpizar, E., Dechamp, E., Espeout, S., Royer, M., Lecouls, A. C., Nicole, M., et al. (2006). Efficient production of Agrobacterium rhizogenes-transformed roots and composite plants for studying gene expression in coffee roots. *Plant Cell Rep.* 25, 959–967. doi: 10.1007/s00299-006-0159-9

An, J., Hu, Z., Che, B., Chen, H., Yu, B., and Cai, W. (2017). Heterologous Expression of panax ginseng PgTIP1 Confers enhanced salt tolerance of Soybean Cotyledon hairy roots, composite, and whole plants. *Front. Plant Sci.* 8, 1232. doi: 10.3389/fpls.2017.01232

Billault-Penneteau, B., Sandré, A., Folgmann, J., Parniske, M., and Pawlowski, K. (2019). Dryas as a model for studying the root symbioses of the rosaceae. *Front. Plant Sci.* 10, 661. doi: 10.3389/fpls.2019.00661

Bosselut, N., Van Ghelder, C., Claverie, M., Voisin, R., Onesto, J.-P., Rosso, M.-N., et al. (2011). Agrobacterium rhizogenes-mediated transformation of Prunus as an alternative for gene functional analysis in hairy-roots and composite plants. *Plant Cell Rep.* 30, 1313–1326. doi: 10.1007/s00299-011-1043-9

Cao, X., and Jacobsen, S. E. (2002). Role of the Arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr. Biol.* 12, 1138–1144. doi: 10.1016/S0960-9822(02)00925-9

Carlson, C. H., Choi, Y., Chan, A. P., Serapiglia, M. J., Town, C. D., and Smart, L. B. (2017). Dominance and Sexual Dimorphism Pervade the Salix purpurea L. *Transcriptome. Genome Biol. Evol.* 9, 2377–2394. doi: 10.1093/gbe/evx174

Chabaud, M., Boisson-Dernier, A., Zhang, J., Taylor, C. G., Yu, O., Barker, D., et al. (2006). Agrobacterium rhizogenes-mediated root transformation. In: *The Medicago truncatula handbook*. U. Mathesius, E. P. Journet, L. W. Sumner (eds). Noble Research Institute (United States). http://www.noble.org/MedicagoHandbook/.

Clemow, S. R., Clairmont, L., Madsen, L. H., and Guinel, F. C. (2011). Reproducible hairy root transformation and spot-inoculation methods to study root symbioses of pea. *Plant Methods* 7, 46. doi: 10.1186/1746-4811-7-46

Du, Y.-T., Zhao, M.-J., Wang, C.-T., Gao, Y., Wang, Y.-X., Liu, Y.-W., et al. (2018). Identification and characterization of GmMYB118 responses to drought and salt stress. *BMC Plant Biol.* 18, 320. doi: 10.1186/s12870-018-1551-7

Estrada-Navarrete, G., Alvarado-Affantranger, X., Olivares, J.-E., Guillén, G., Díaz-Camino, C., Campos, F., et al. (2007). Fast, efficient and reproducible genetic transformation of Phaseolus spp. by Agrobacterium rhizogenes. *Nat. Protoc.* 2, 1819–1824. doi: 10.1038/nprot.2007.259

Fabio, E. S., and Smart, L. B. (2018). Differential growth response to fertilization of ten elite shrub willow (Salix spp.) bioenergy cultivars. *Trees - Struct. Funct.* 32, 1061–1072. doi: 10.1007/s00468-018-1695-y

Franche, C., Diouf, D., Le, Q. V., Bogusz, D., N'Diaye, A., Gherbi, H., et al. (1997). Genetic transformation of the actinorhizal tree Allocasuarina

verticillata by Agrobacterium tumefaciens. *Plant J.* 11, 897–904. doi: 10.1046/j.1365-313X.1997.11040897.x

Gouker, F. E., DiFazio, S. P., Bubner, B., Zander, M., and Smart, L. B. (2019). Genetic diversity and population structure of native, naturalized, and cultivated Salix purpurea. *Tree Genet. Genomes* 15, 47. doi: 10.1007/s11295-019-1359-0

Grönroos, L., Von Arnold, S., and Eriksson, T. (1989). Callus Production and Somatic Embryogenesis from Floral Explants of Basket Willow (Salix viminalis L.). *J. Plant Physiol.* 134, 558–566. doi: 10.1016/S0176-1617(89)80147-6

Guan, Q., He, M., Ma, H., Liao, X., Wang, Z., and Liu, S. (2018). Construction of genetic transformation system of Salix mongolica: in vitro leaf-based callus induction, adventitious buds differentiation, and plant regeneration. *Plant Cell Tissue Organ Cult.* 132, 213–217. doi: 10.1007/s11240-017-1265-9

Hanley, S. J., and Karp, A. (2014). Genetic strategies for dissecting complex traits in biomass willows (Salix spp.). *Tree Physiol.* 34, 1167–1180. doi: 10.1093/treephys/tpt089

Hauth, S., and Beiderbeck, R. (1992). In vitro culture of Agrobacterium rhizogenes-induced hairy roots of Salix alba L. *Silvae Genet.* 41, 46–48.

Ho-Plágaro, T., Huertas, R., Tamayo-Navarrete, M. I., Ocampo, J. A., and García-Garrido, J. M. (2018). An improved method for Agrobacterium rhizogenes-mediated transformation of tomato suitable for the study of arbuscular mycorrhizal symbiosis. *Plant Methods* 14, 34. doi: 10.1186/s13007-018-0304-9

Ieamkhang, S., and Chatchawankanphanich, O. (2005). Augmentin® as an alternative antibiotic for growth suppression of Agrobacterium for tomato (Lycopersicon esculentum) transformation. *Plant Cell. Tissue Organ Cult.* 82, 213–220. doi: 10.1007/s11240-005-0416-6

Indrasumunar, A., Wilde, J., Hayashi, S., Li, D., and Gresshoff, P. M. (2015). Functional analysis of duplicated Symbiosis Receptor Kinase (SymRK) genes during nodulation and mycorrhizal infection in soybean (Glycine max). *J. Plant Physiol.* 176, 157–168. doi: 10.1016/j.jplph.2015.01.002

Jouanin, L., Tourneur, J., Tourneur, C., and Casse-Delbart, F. (1986). Restriction maps and homologies of the three plasmids of Agrobacterium rhizogenes strain A4. *Plasmid* 16, 124–134. doi: 10.1016/0147-619X(86)90071-5

Kajikawa, M., Morikawa, K., Abe, Y., Yokota, A., and Akashi, K. (2010). Establishment of a transgenic hairy root system in wild and domesticated watermelon (Citrullus lanatus) for studying root vigor under drought. *Plant Cell Rep.* 29, 771–778. doi: 10.1007/s00299-010-0863-3

Karp, A., Hanley, S. J., Trybush, S. O., Macalpine, W., Pei, M., and Shield, I. (2011). Genetic Improvement of Willow for Bioenergy and Biofuels. *J. Integr. Plant Biol.* 53, 151–165. doi: 10.1111/j.1744-7909.2010.01015.x

Kuzovkina, Y. A., and Quigley, M. F. (2005). Willows Beyond Wetlands: Uses of Salix L. *Species Environ. Projects. Water Air Soil Pollut.* 162, 183–204. doi: 10.1007/s11270-005-6272-5

Kuzovkina, Y. A., Weih, M., Romero, M. A., Charles, J., Hust, S., McIvor, I., et al. (2008). "Salix: Botany and Global Horticulture," in *Horticultural Reviews* (Hoboken, NJ, USA: John Wiley & Sons, Inc.), 447–489. doi: 10.1002/9780470380147.ch8

Li, B., Liu, Y., Cui, X.-Y., Fu, J.-D., Zhou, Y.-B., Zheng, W.-J., et al. (2019). Genome-Wide Characterization and Expression Analysis of Soybean TGA Transcription Factors Identified a Novel TGA Gene Involved in Drought and Salt Tolerance. *Front. Plant Sci.* 10, 549. doi: 10.3389/fpls.2019.00549

Limpens, E., Ramos, J., Franken, C., Raz, V., Compaan, B., Franssen, H., et al. (2004). RNA interference in Agrobacterium rhizogenes-transformed roots of Arabidopsis and Medicago truncatula, J. Exp. *Botany* 55, 983–992. doi: 10.1093/jxb/erh122

Lyyra, S., Lima, A., and Merkle, S. A. (2006). In vitro regeneration of Salix nigra from adventitious shoots. *Tree Physiol.* 26, 969–975. doi: 10.1093/treephys/26.7.969

Mellor, K. E., Hoffman, A. M., and Timko, M. P. (2012). Use of ex vitro composite plants to study the interaction of cowpea (Vigna unguiculata L.) with the root parasitic angiosperm Striga gesnerioides. *Plant Methods* 8, 22. doi: 10.1186/1746-4811-8-22

Meng, D., Yang, Q., Dong, B., Song, Z., Niu, L., Wang, L., et al. (2019). Development of an efficient root transgenic system for pigeon pea and its application to other important economically plants. *Plant Biotechnol. J.* 17, 1804–1813. doi: 10.1111/pbi.13101

Nagel, R., Elliott, A., Masel, A., Birch, R. G., and Manners, J. M. (1990). Electroporation of binary Ti plasmid vector into Agrobacterium tumefaciens and Agrobacterium rhizogenes. *FEMS Microbiol. Lett.* 67, 325–328. doi: 10.1111/j.1574-6968.1990.tb04041.x

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, 45e– 45. doi: 10.1093/nar/29.9.e45

Plasencia, A., Soler, M., Dupas, A., Ladouce, N., Silva-Martins, G., Martinez, Y., et al. (2016). Eucalyptus hairy roots, a fast, efficient and versatile tool to explore function and expression of genes involved in wood formation. *Plant Biotechnol. J.* 14, 1381–1393. doi: 10.1111/pbi.12502

Plasencia-Casavedall, A. (2015). "Transcriptional regulation of wood formation in eucalyptus: Role of MYBtranscription factors and protein-protein interactions," in *Dissertation in Vegetal Biology.* Toulouse, France: Université Paul Sabatier -Toulouse III

Ron, M., Kajala, K., Pauluzzi, G., Wang, D., Reynoso, M. A., Zumstein, K., et al. (2014). Hairy Root Transformation Using Agrobacterium rhizogenes as a Tool for Exploring Cell Type-Specific Gene Expression and Function Using Tomato as a Model. *Plant Physiol.* 166, 455–469. doi: 10.1104/pp.114.239392

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

Stoehr, M. U., Cai, M., and Zsuffa, L. (1989). In vitro plant regeneration via callus culture of mature Salix exigua. *Can. J. For. Res.* 19, 1634–1638. doi: 10.1139/x89-247

Sulima, P., Krauze-Baranowska, M., and Przyborowski, J. A. (2017). Variations in the chemical composition and content of salicylic glycosides in the bark of Salix purpurea from natural locations and their significance for breeding. *Fitoterapia* 118, 118–125. doi: 10.1016/J.FITOTE.2017.03.005

Sulima, P., Prinz, K., and Przyborowski, J. (2018). Genetic Diversity and Genetic Relationships of Purple Willow (Salix purpurea L.) from Natural Locations. *Int. J. Mol. Sci.* 19, 105. doi: 10.3390/ijms19010105

Vahala, T., Eriksson, T., Tillberg, E., and Nicander, B. (1993). Expression of a cytokinin synthesis gene from Agrobacterium tumefaciens T-DNA in basket willow (Salix viminalis). *Physiol. Plant* 88, 439–445. doi: 10.1111/j.1399-3054.1993.tb01357.x

Vahala, T., Stabel, P., and Eriksson, T. (1989). Genetic transformation of willows (Salix spp.) by Agrobacterium tumefaciens. *Plant Cell Rep.* 8, 55–58. doi: 10.1007/BF00716837

Wang, L., Wang, L., Tan, Q., Fan, Q., Zhu, H., Hong, Z., et al. (2016). Efficient Inactivation of Symbiotic Nitrogen Fixation Related Genes in Lotus japonicus Using CRISPR-Cas9. *Front. Plant Sci.* 7, 1333. doi: 10.3389/fpls.2016.01333

Xue, R., Wu, X., Wang, Y., Zhuang, Y., Chen, J., Wu, J., et al. (2017). Hairy root transgene expression analysis of a secretory peroxidase (PvPOX1) from common bean infected by Fusarium wilt. *Plant Sci.* 260, 1–7. doi: 10.1016/j.plantsci.2017.03.011

Yang, J., Yi, J., Yang, C., and Li, C. (2013). Agrobacterium tumefaciens-mediated genetic transformation of Salix matsudana Koidz. *using mature seeds. Tree Physiol.* 33, 628–639. doi: 10.1093/treephys/tpt038

Yanitch, A., Brereton, N. J. B., Gonzalez, E., Labrecque, M., Joly, S., and Pitre, F. E. (2017). Transcriptomic Response of Purple Willow (Salix purpurea) to Arsenic Stress. *Front. Plant Sci.* 8, 1115. doi: 10.3389/FPLS.2017.01115

Yoshida, K., Ma, D., and Constabel, C. P. (2015). The MYB182 Protein Down-Regulates Proanthocyanidin and Anthocyanin Biosynthesis in Poplar by Repressing Both Structural and Regulatory Flavonoid Genes. *Plant Physiol.* 167, 693–710. doi: 10.1104/pp.114.253674

Zalesny, R. S., and Bauer, E. O. (2007). Selecting and Utilizing Populus and Salix for Landfill Covers: Implications for Leachate Irrigation. *Int. J. Phytorem.* 9, 497–511. doi: 10.1080/15226510701709689

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 29 January 2020
doi: 10.3389/fpls.2019.01774

Deep Eutectic Solvent Pretreatment of Transgenic Biomass With Increased C_6C_1 Lignin Monomers

Kwang Ho Kim[1,2*], *Yunxuan Wang*[3], *Masatsugu Takada*[2], *Aymerick Eudes*[4,5], *Chang Geun Yoo*[3], *Chang Soo Kim*[1] *and Jack Saddler*[2]

[1] Clean Energy Research Center, Korea Institute of Science and Technology, Seoul, South Korea, [2] Department of Wood Science, University of British Columbia, Vancouver, BC, Canada, [3] Department of Paper and Bioprocess Engineering, State University of New York College of Environmental Science and Forestry, Syracuse, NY, United States, [4] Feedstocks Division, Joint BioEnergy Institute, Emeryville, CA, United States, [5] Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States

The complex and heterogeneous polyphenolic structure of lignin confers recalcitrance to plant cell walls and challenges biomass processing for agroindustrial applications. Recently, significant efforts have been made to alter lignin composition to overcome its inherent intractability. In this work, to overcome technical difficulties related to biomass recalcitrance, we report an integrated strategy combining biomass genetic engineering with a pretreatment using a bio-derived deep eutectic solvent (DES). In particular, we employed biomass from an *Arabidopsis* line that expressed a bacterial hydroxycinnamoyl-CoA hydratase-lyase (HCHL) in lignifying tissues, which results in the accumulation of unusual C_6C_1 lignin monomers and a slight decrease in lignin molecular weight. The transgenic biomass was pretreated with renewable DES that can be synthesized from lignin-derived phenols. Biomass from the HCHL plant line containing C_6C_1 monomers showed increased pretreatment efficiency and released more fermentable sugars up to 34% compared to WT biomass. The enhanced biomass saccharification of the HCHL line is likely due to a reduction of lignin recalcitrance caused by the overproduction of C_6C_1 aromatics that act as degree of polymerization (DP) reducers and higher chemical reactivity of lignin structures with such C_6C_1 aromatics. Overall, our findings demonstrate that strategic plant genetic engineering, along with renewable DES pretreatment, could enable the development of sustainable biorefinery.

Keywords: lignocellulosic biomass, lignin engineering, hydroxycinnamoyl-CoA hydratase-lyase, bioenergy, saccharification

OPEN ACCESS

Edited by:
Jin Zhang,
Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Rajeev Kumar,
University of California, Riverside, Unites States
Kewei Zhang,
Zhejiang Normal University, China

****Correspondence:***
Kwang Ho Kim
kwanghokim@kist.re.kr

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *18 October 2019*
Accepted: *19 December 2019*
Published: *29 January 2020*

Citation:
Kim KH, Wang Y, Takada M, Eudes A, Yoo CG, Kim CS and Saddler J (2020) Deep Eutectic Solvent Pretreatment of Transgenic Biomass With Increased C_6C_1 Lignin Monomers. Front. Plant Sci. 10:1774. doi: 10.3389/fpls.2019.01774

INTRODUCTION

Growing concerns over global warming and our over-dependence on fossil resources have forced society to demand sustainable and green products (Isikgor and Becer, 2015). Lignocellulosic biomass presents a promising source of renewable carbon, holding enormous potential for the production of chemicals and fuels. In order to utilize biomass as a source of renewable energy and chemicals, there have been significant efforts to develop an efficient process for biomass conversion

(Demirbaş, 2001). Among several technologies developed within the biorefinery concept, the production of second-generation biofuels (e.g., bioethanol) from lignocellulosic biomass is well established and close to commercialization. Despite the successful demonstration of biomass conversion to biofuels, the production of cellulosic biofuels still encounters several technical challenges for achieving a sustainable energy future (Dale et al., 2014; Maity, 2015; Kim and Kim, 2018).

Lignin, an essential component of biomass that provides mechanical strength for upright growth and acts as a physical barrier against pathogens (Boudet, 2007), is often blamed for conferring recalcitrance to biomass against processing and utilization. In the typical biological conversion process, effective removal of lignin is crucial to maximize the utilization of carbohydrates to produce fuels and building block chemicals (Ding et al., 2012; Ragauskas et al., 2014). However, considering its complex and heterogeneous structure, hydrophobic character, and other intractable properties, lignin is one of the most challenging biomaterials to handle (Himmel, 2009).

During the last decades, researchers have developed several routes to overcome lignin-associated recalcitrance. For example, the lignin-first approach was introduced to extract the reactive lignin at the early stage of biomass fractionation, providing opportunities for the use of both lignin and carbohydrates (Renders et al., 2017). In contrast to the conventional carbohydrates-oriented pretreatment technologies of biomass, this new biorefinery scheme offers potential valorization routes for both lignin fractions and residual carbohydrates. While the above strategy focuses on the development of processes to overcome the technical difficulties related to lignin, another effort to reduce the biomass recalcitrance through genetic engineering has shown its effectiveness (Chen and Dixon, 2007; Bhagia et al., 2016; Thomas et al., 2017).

Previously, biomass cell wall engineering was directed to strategically reduce total lignin content by downregulating one or more enzymes in the monolignol pathway, which includes cinnamate 4-hydroxylase (*C4H*) (Schilmiller et al., 2009), 4-coumarate-CoA ligase (*4CL*) (Xu et al., 2011), and cinnamoyl-CoA reductase (*CCR*) (Chabannes et al., 2001). Although the decrease in the amount of lignin was proven to improve processing efficiency, this method often involves an agronomic penalty (Bonawitz and Chapple, 2013). Moreover, modification of lignin monomeric composition leading to structural modifications has been extensively studied to make biomass more amenable to processing without compromising biomass yield. In-planta expression of a bacterial 3-dehydroshikimate dehydratase resulted in the higher deposition of H-units and lower amounts of G- and S-units, resulting in more than a two-fold improvement in saccharification efficiency (Eudes et al., 2015). Incorporation of chemically labile ester linkages (zip-lignin) in the lignin backbone was proven to enhance biomass pretreatment efficiency (Wilkerson et al., 2014; Kim et al., 2017). Previous work also reported a strategy for the overproduction of uncommon lignin monomers through in-planta expression of a bacterial hydroxycinnamoyl-CoA hydratase-lyase (HCHL) in biomass (Mitra et al., 2002; Eudes et al., 2012). **Figure 1** describes the enzymatic reactions catalyzed by HCHL which is expressed in lignifying tissues of engineered plants. HCHL cleaves the side-chain of coumaroyl-CoA and feruloyl-CoA, resulting in an increased amount of unusual C_6C_1 end-groups in lignin, including 4-hydroxybenzaldehyde, vanillin, syringylaldehyde, and 4-hydroxybenzoic acid (Eudes et al., 2012).

Bio-derived deep eutectic solvents (DESs) have gained considerable attention because of their potential uses in biomass pretreatment and processing. DESs are mixtures of compounds formed by strong intermolecular hydrogen bonds, resulting in a lower melting point than that of any individual component (Dai et al., 2013). As an alternative to organic solvents for biomass pretreatment, DESs exhibit promising solvent properties including high dissolution capability, low vapor pressure, tunability, stabilization of carbohydrates through hydrogen-bond interactions, and compatibility with certain microorganisms (Vigier et al., 2015). Recently, DESs synthesized from lignin-derivable phenolic compounds were found to be effective in lignin removal, and thus resulted in enhanced biomass saccharification efficacy (Kim et al., 2018). Also, renewable DESs prepared from phenolic aldehydes (e.g., vanillin and 4-hydroxybenzaldehyde) with choline chloride (ChCl), integrated with the use of low-recalcitrant engineered biomass *via* down-regulation of cinnamyl alcohol dehydrogenase (CAD), demonstrated the potential of developing a closed-loop biorefinery process (Kim et al., 2019).

In this work, we employed, as a raw material for DES-assisted pretreatment, the biomass from a previously described plant genetic engineering approach that increases non-conventional C_6C_1 monomers in lignin *via* HCHL expression. We also demonstrate the use of renewable DES for pretreating such biomass designed for improved processability.

MATERIALS AND METHODS

Biomass Material

Arabidopsis thaliana (ecotype Columbia, Col-0) wild type (WT) and line *IRX5:HCHL*-1 (HCHL) were used in this study (Eudes et al., 2012). Previously, five transgenic lines that express HCHL were isolated and characterized, of which two showed reduced biomass and/or height. Line *IRX5:HCHL*-1and one with no yield penalty was selected for in-depth characterization and grown on a larger scale for further research (Eudes et al., 2012). Seeds were germinated directly on soil. Growing conditions were: 14 h light/day at 100 μmol/m^2/s, 22°C, and 55% humidity. For analyses, stems from mature senesced dried plants harvested without siliques and leaves were ball-milled to a fine powder using a Mixer Mill MM 400 (Retsch Inc., Newtown, PA) and stainless steel balls for 2 min.

Bio-Derived Deep Eutectic Solvent Preparation

All chemicals used for DES synthesis were purchased from Sigma-Aldrich (St. Louis, MO) and used without further

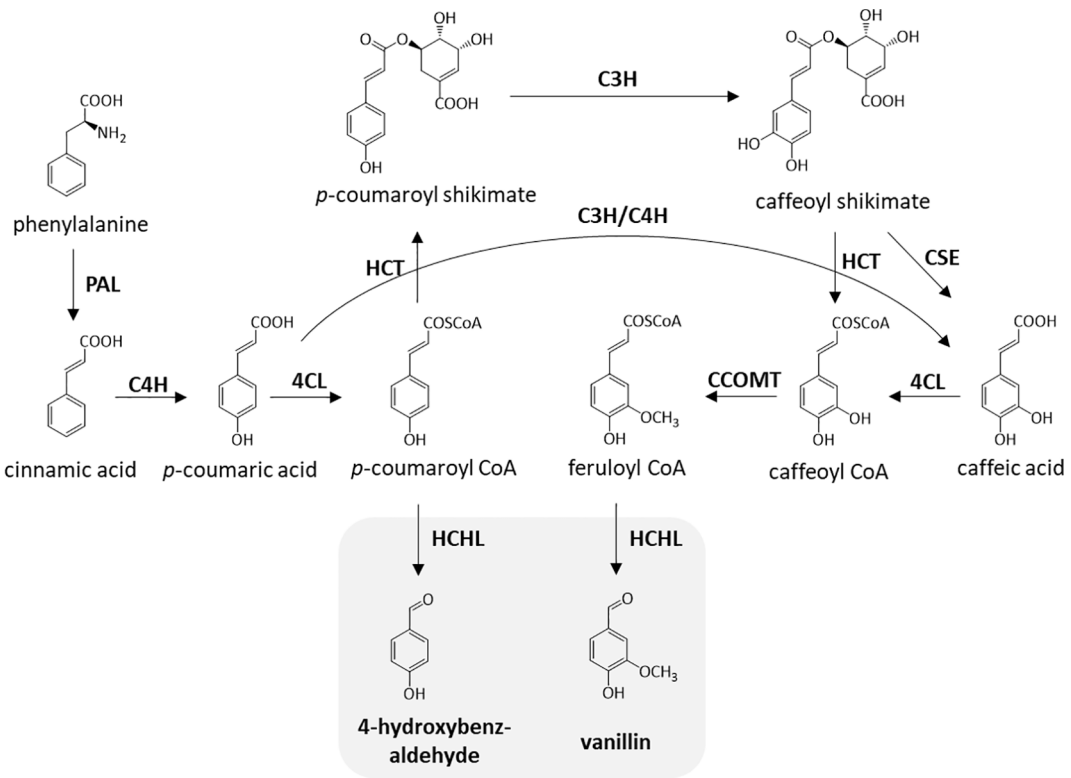

FIGURE 1 | Enzymatic steps catalyzed by hydroxycinnamoyl-CoA hydratase-lyase (HCHL). The two products of HCHL activity (in gray) are converted into several other C6C1 aromatics in plant tissues. *PAL*, phenylalanine ammonia lyase; *C4H*, cinnamate 4-hydroxylase; *4CL*, 4-coumarate-CoA ligase; *HCT*, hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase; *C3H,* coumarate 3-hydroxylase; *CSE*, caffeoyl shikimate esterase; *CCOMT,* caffeoyl-CoA *O*-methyltransferase.

purification. DES used in this work was synthesized using ChCl (purity: ≥98%) and vanillin (VAN, purity: 99%) as a hydrogen-bond acceptor and donor, respectively. ChCl and VAN were mixed at 1:2 molar ratio, and then heated to 100°C with continuous stirring until a homogeneous liquid was formed (Kim et al., 2018; Kim et al., 2019). The prepared DES (i.e., ChCl-VAN) was stored in a vacuum oven before use.

Pretreatment and Enzymatic Saccharification

For biomass pretreatment, 5 wt% biomass solution was prepared by mixing 0.2 g of biomass (WT and HCHL line) and 3.8 g ChCl-VAN in a 20 ml pressure tube (Ace Glass, Vineland, NJ). The tube was placed to the preheated oil bath at 80°C. A mild pretreatment temperature used in this work to examine the effect of engineered biomass on pretreatment efficiency (Kim et al., 2019). After 3 h, the pretreated slurry was transferred to a 15 ml centrifuge tube and washed with 50 ml (5 × 10 ml) ethanol and deionized water mixture (2:1, v/v) to separate lignin and remove any residual DES (Kim et al., 2019).

Enzymatic hydrolysis of the pretreated biomass was carried out in 5 ml of 50 mM citrate buffer supplemented with Cellic® CTec3 (266.3 mg/ml) and HTec (234.8 mg/ml) from Novozymes (9:1 ratio) at 10 mg protein per gram solid biomass (solid loading ≈2 wt%). The enzymatic saccharification was performed in an Incubator-Genie™ rotating hybridization oven (Scientific Industries, Bohemia, NY) at 50°C for 72 h.

Topochemical Analysis

Sections of intact stems of mature plants (5 cm from bottom) of unpretreated and DES-pretreated were gradually dehydrated with ethanol-water mixtures with increasing ethanol concentration up to 99.5%. After the ethanol was replaced by acetone, the samples were embedded in the epoxy resin. Ultrathin sections of 80 nm were prepared from embedded samples with a diamond knife mounted on ultramicrotome. Lignin was selectively stained with 1% $KMnO_4$ and the stained section was mounted on copper grids, and observed with a Hitachi H7600 transmission electron microscope (TEM) at an acceleration voltage of 80 keV.

Analytical Methods

Compositional analysis of raw materials and pretreated samples was conducted according to the National Renewable Energy Laboratory protocol (Sluiter et al., 2008), and the results are given in **Table 1**.

After pretreatment and saccharification, the hydrolysate was filtered through a 0.45 μm polytetrafluoroethylene (PTFE)

TABLE 1 | Compositional analysis of wild-type (WT) and hydroxycinnamoyl-CoA hydratase-lyase (HCHL) *Arabidopsis* line.

	WT	HCHL
Glucan, wt%	28.9 (0.3)	28.7 (1.7)
Xylan, wt%	11.8 (0.0)	12.3 (0.6)
Lignin, wt%	21.9 (0.4)	19.6 (0.6)
Extractives, wt%	22.3 (0.3)	29.9 (2.1)*

*The values in parentheses are standard deviations (statistical significance from WT *P < 0.05).*

syringe filter (VWR, Radnor, PA). The amount of glucose and xylose released was quantified using a YL 9100 high performance liquid chromatography (Young-Lin) equipped with a refractive index detector and a Bio-Rad Aminex HPX-87H ion-exchange column. A solution of 5 mM H_2SO_4 was used as a mobile phase at a constant flow rate of 0.6 ml/min, and the temperature of column compartment was maintained at 60°C.

Structural information of lignin was analyzed by two-dimensional (2D) ^{1}H-^{13}C heteronuclear single-quantum coherence (HSQC) nuclear magnetic resonance (NMR) spectroscopy. Before the NMR analysis, cellulolytic enzyme lignin was isolated from each line as described elsewhere (Yoo et al., 2016). The prepared lignin dissolved with DMSO-d_6 was placed in a 5 mm NMR tube. The NMR spectra were acquired using a Bruker Avance III HD 800 MHz NMR spectrometer equipped with a TCI Cryoprobe with the following acquisition parameters: spectra width 12 ppm in F2 (^{1}H) dimension with 1024 time of domain (acquisition time, 53.2 ms), 200 ppm in F1 (^{13}C) dimension with 512 time of domain (acquisition time, 6.4 ms), a 1.2-s relaxation delay, and 32 scans. Assignment of the HSQC spectra is described elsewhere (Kim and Ralph, 2010; Eudes et al., 2012).

Gel permeation chromatography (GPC) was used to determine the molecular weight distribution of the lignin from the WT control and HCHL transgenic. GPC measurement was conducted using a Waters 2489 GPC system equipped with a UV detector at 270 nm of wavelength. The eluent used for the analysis was tetrahydrofuran, and the three Waters Styragel columns (HR0.5, HR3, and HR4e) were used. All samples were acetylated using 1.0 ml acetic anhydride/pyridine (1:1, v/v) at room temperature for at least 24 h prior to GPC analysis.

Computational Analysis

Density functional theory (DFT) based computational analysis was performed to gain a mechanistic understanding of structurally altered lignin molecules. It is noted, considering the complex structure of lignin, that lignin model compounds with a representative interunit linkage are typically employed for computational studies (Sun et al., 2016). **Figure 2** shows five representative β-O-4 dimeric model compounds from typical lignin (1) and from the C_6C_1-enriched lignin of the HCHL line (2, 3, 4, and 5) tested for the mechanistic study. The geometry optimizations of model compounds were conducted using DFT with the B3LYP and the 6-31+G(2d,2p) basis set. Frequency calculations were also carried out to verify that the optimized structures corresponded to energy minima. In this study, the electrophilicity index, a global reactivity descriptor, was calculated using the electronegativity and the chemical hardness to compare the chemical reactivity of each model structure (Shi et al., 2016; Kim et al., 2019).

Statistical Analysis

Analysis of variance (ANOVA) was conducted using R (R Foundation for Statistical Computing, Vienna, Austria) to test the null hypothesis of no statistical difference in saccharification yield (glucose and xylose) between the WT and transgenic HCHL *Arabidopsis*. Three independent tests were conducted and each sample was analyzed. The null hypothesis was rejected at the 0.05 level.

RESULTS

Structural Analysis of Lignin in Wild-Type and Hydroxycinnamoyl-Coa Hydratase-Lyase *Arabidopsis*

The isolated lignins were subjected to HSQC NMR analysis to elucidate the structural differences between *Arabidopsis* WT and the HCHL line. The HSQC spectra, respective peak assignments, and lignin substructures of WT and transgenic biomass are shown in **Figure 3**. The HSQC spectra were divided broadly into aromatic (δ_H/δ_C 6.0–8.0/90–150) and aliphatic (δ_H/δ_C 3.0–5.5/50–90) regions. In the aromatic regions of the WT lignin,

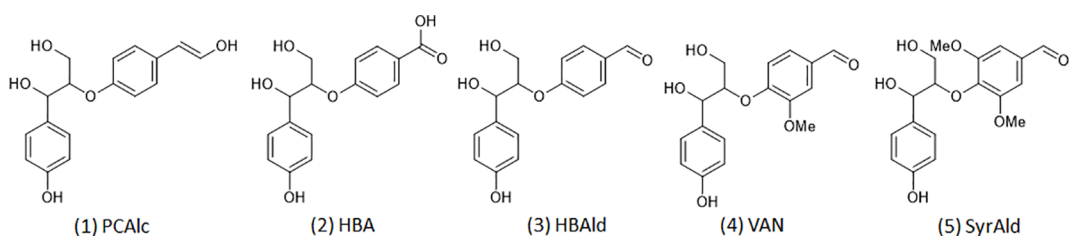

FIGURE 2 | Dimeric model compounds: 1) 4-hydroxylphenylglycerol-β-*p*-coumaryl alcohol ether (PCAlc), 2) 4-hydroxylphenylglycerol-β-4-oxybenzoic acid ether (HBA), 3) 4-hydroxylphenylglycerol-β-4-oxybenzaldehyde ether (HBAld), 4) 4-hydroxylphenylglycerol-β-vanillin ether (VAN), and 5) 4-hydroxylphenylglycerol-β-syringaldehyde ether (SyrAld).

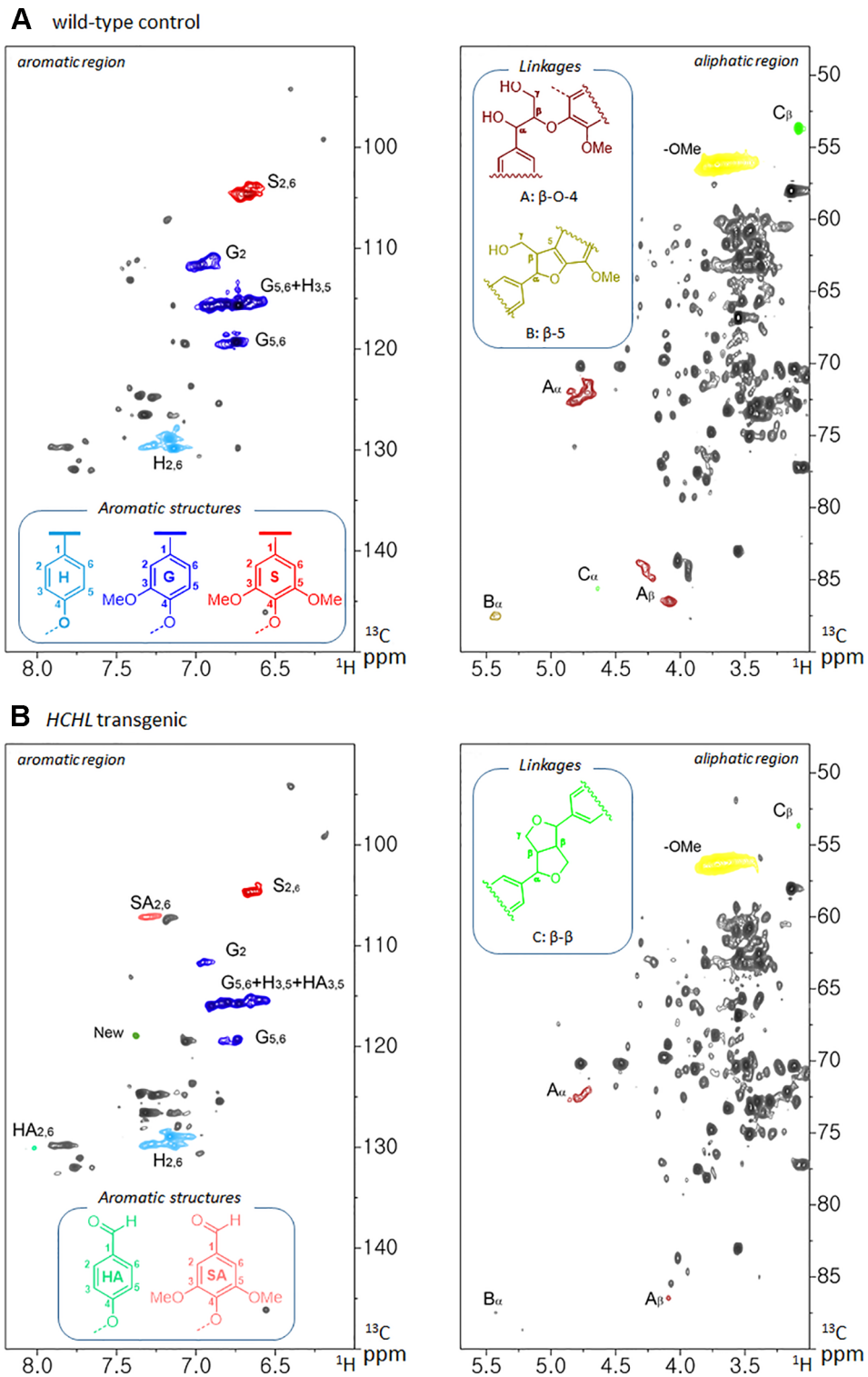

FIGURE 3 | Two-dimensional ^{1}H-^{13}C heteronuclear single-quantum coherence (2D HSQC) nuclear magnetic resonance (NMR) spectra of isolated lignins from *Arabidopsis* wild-type (WT) **(A)** and hydroxycinnamoyl-CoA hydratase-lyase (HCHL) line **(B)** with main structures of lignin subunits (H, *p*-hydroxyphenyl; G, guaiacyl; S, syringyl; A, β-O-4; B, β-5; and C, β-β).

typical lignin subunits, including *p*-hydroxyphenyl (**H**), guaiacyl (**G**), and syringyl (**S**) were observed. In the lignin isolated from the HCHL line, signal intensities of guaiacyl (**G**) and syringyl (**S**) units were relatively lower. Instead, the aromatic regions of the lignin obtained from the HCHL line had several new correlations compared with WT. For example, the H/C correlations from the oxidized S-unit (**SA**) at 7.3/107.2 ppm and oxidized H-unit (**HA**) at 8.0/130.1 ppm were observed, which is in agreement with the previous study (Eudes et al., 2012). It is also noted that the relative S/G levels of both lignins were similar (0.17 for the WT and 0.19 for the HCHL line), although the engineered lignin exhibited some structural changes.

The aliphatic sidechain region provides information about the structural types and distribution of interunit linkages of the lignin fraction (Kim and Ralph, 2010). HSQC spectra of the WT and engineered lignin showed correlations corresponding to β-ether (β-O-4, substructure **A**), phenylcoumaran (β-5, substructure **B**), and resinol (β-β, substructure **C**) units. As shown in the figure, both lignins isolated from WT and HCHL line exhibited the typical lignin sidechain hydrogen and carbon resonances, although the signal intensities of the engineered lignin were marginally lower than that from WT lignin.

The molecular weight distribution of isolated lignins was measured, and the results are presented in **Table 2**. As seen, the molecular weight of lignins from WT and HCHL line are marginally different (2,853 Da for WT and 2,973 Da for HCHL). Together with the data from biomass compositional analysis, this result suggests that the expression of the *HCHL* gene did not significantly affect lignin content, although some non-traditional side-chain-truncated units were observed.

Biomass Pretreatment Using Bio-Derived Deep Eutectic Solvent and Enzymatic Saccharification

For biomass pretreatment, the DES (ChCl-VAN) was synthesized at a 1:2 molar ratio. As previously reported, the eutectic point of ChCl-VAN was found to be 55–60°C, which is significantly lower than that of ChCl (302°C) and VAN (81–83°C) (Kim et al., 2019). The pretreatment of WT and transgenic *Arabidopsis* with ChCl-VAN was conducted at 80°C for 3 h. It is noted that the pretreatment conditions used in this work are relatively milder compared with other conventional pretreatment methods (e.g., dilute acid- or organosolv pretreatment) that are typically conducted between 100 and 250°C with catalyst (Wyman et al., 2005; Agbor et al., 2011; Wyman et al., 2013). Indeed, it was hypothesized that reduced biomass recalcitrance due to the incorporation of side-chain-truncated monomers in lignin of engineered plants would facilitate the release of sugars under low-severity pretreatment conditions. After DES pretreatment, 58.0% (WT) and 49.0% (HCHL) of the biomass was recovered. Compositional analysis of the pretreated biomass was conducted and the amount of glucan, xylan, and lignin was as follows; 40.7% (WT) and 40.6% (HCHL) for glucan, 13.2% (WT) and 13.3% (HCHL) for xylan, and 25.0% (WT) and 18.2% (HCHL) for lignin. It also revealed that lignin removal was found to be 33.6%, and 54.7% from WT and HCHL line, respectively. The pretreated solid was enzymatically hydrolyzed for 72 h at 50°C.

Figure 4 illustrates the sugar yields after ChCl-VAN pretreatment of biomass from the WT and HCHL lines followed by enzymatic saccharification. Glucose release for WT was 102 μg/mg untreated biomass after ChCl-VAN pretreatment, whereas the amount of glucose released from the engineered biomass was significantly higher (+ 34.3%) and reached 137 μg/mg untreated biomass. This clearly indicates that a strategic expression of the HCHL gene in biomass reduced lignin-associated recalcitrance and resulted in enhanced cell wall digestibility. Similarly, xylose release from the HCHL line was 37 μg/mg untreated biomass, which is 68% higher than that from the WT. The enhanced saccharification was also reported from HCHL-engineered *Arabidopsis* (+78% reducing sugars after hot water pretreatment at 120°C, +31% reducing sugars after 1.2% H_2SO_4 pretreatment at 120°C, and +71% reducing sugars after 0.25% NaOH pretreatment at 120°C) (Eudes et al., 2012). Although more comprehensive research is required to optimize DES-assisted pretreatment to maximize the sugar release, this work demonstrates that pretreatment using lignin-derived DES is effective on lignocellulosic biomass.

Cell Wall Ultrastructural Morphology

Electron microscopic analysis is widely used to obtain high-resolution information on the lignin distribution in

TABLE 2 | Molecular weight distribution of isolated lignins from wild-type (WT) and hydroxycinnamoyl-CoA hydratase-lyase (HCHL) transgenic lines.

	WT	HCHL transgenic
Mw	2853	2973
PDI	3.1	3.3

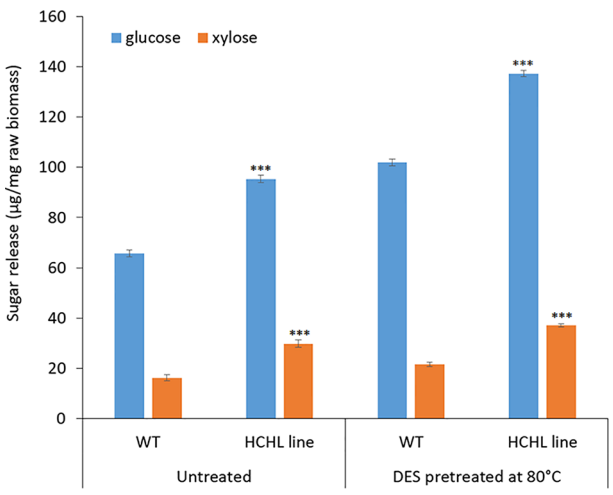

FIGURE 4 | Sugar yield (mean ± standard deviation of three independent tests) from the wild type (WT) and hydroxycinnamoyl-CoA hydratase-lyase (HCHL) lines before and after bio-derived deep eutectic solvent (DES) pretreatment at 80°C (statistical significance from WT ***$P < 0.001$).

biomass cell walls, which can be visualized by staining with potassium permanganate ($KMnO_4$) (Koch and Schmitt, 2013). **Figure 5** shows the TEM micrographs of intact stems from WT (**Figure 5A**) and HCHL lines (**Figure 5D**) before and after DES pretreatment at 80°C, in which lignin is selectively stained with 1% $KMnO_4$ and darker areas indicate higher lignin concentration.

Regarding the untreated samples, the different layers of fiber and vessel cell walls from both WT and HCHL lines exhibit general staining, including the middle lamella (ML), primary wall (P), and secondary wall (S_1, S_2, and S_3 layers). Particularly, the middle lamella portion is highly reactive to the staining which indicates higher lignin concentration. Previously, it was reported that the HCHL transgenic has a cell wall structure similar to that of WT using light microscopy and stainings such as toluidine blue, Maule, and phloroglucinol-HCl (Eudes et al., 2012). After the DES pretreatment, TEM images of cell wall structures showed little difference for both stem samples. The middle lamella, primary wall, and secondary wall layers are clearly observed for both genotypes, and the middle lamella portion contains higher lignin content than the secondary wall. Considering that the middle lamella lignin is relatively tougher to be removed compared to that in the primary and secondary cell walls (Terashima and Fukushima, 1988; Doherty et al., 2010), the lignin removal after DES pretreatment likely occurs in the primary and secondary cell walls.

Interestingly, some parts of thin wall cells such as in parenchyma cells were compressed and delamination zones were observed for both samples after DES-assisted pretreatment (arrow in **Figure 5B**). This is quite surprising because common DESs have negligible vapor pressure, and the pretreatment conditions used in this work are relatively mild. It is speculated that the cell wall delamination during the pretreatment is likely due to lignin dissolution in the DES, resulting in a loosened fiber cell wall with delamination. Changes in cell wall architecture by DES pretreatment, such as delamination, and possibly an increase in the porosity are main factors for enhanced digestibility (Donohoe et al., 2011). Taken together, the TEM images did not show any distinct differences between unpretreated WT and HCHL lines at the cell wall ultrastructure scale, implying that the enhanced efficiency of pretreatment and digestibility in the case of HCHL-engineered plants should be discussed at the molecular scale.

Computational Study

As observed above, the strategic introduction of the HCHL gene into lignifying tissue reduced biomass recalcitrance *via* incorporation of side-chain-truncated C_6C_1 lignin monomers, which led to enhanced saccharification yields. It was hypothesized that incorporation of C_6C_1 monomers makes lignin structure chemically more amenable. To validate this hypothesis, we employed the DFT-based computational study to compute a kinetic quantity of different lignin structures (Socha et al., 2014). Model dimers with aryl-ether linkage found in the typical WT lignin structure and engineered line were used to understand chemical reactivity. Normal β-O-4 structure containing *p*-coumaryl alcohol end group and those with side-chain-truncated monomers (C_6C_1: HBA, HBAld, and SyrAld)

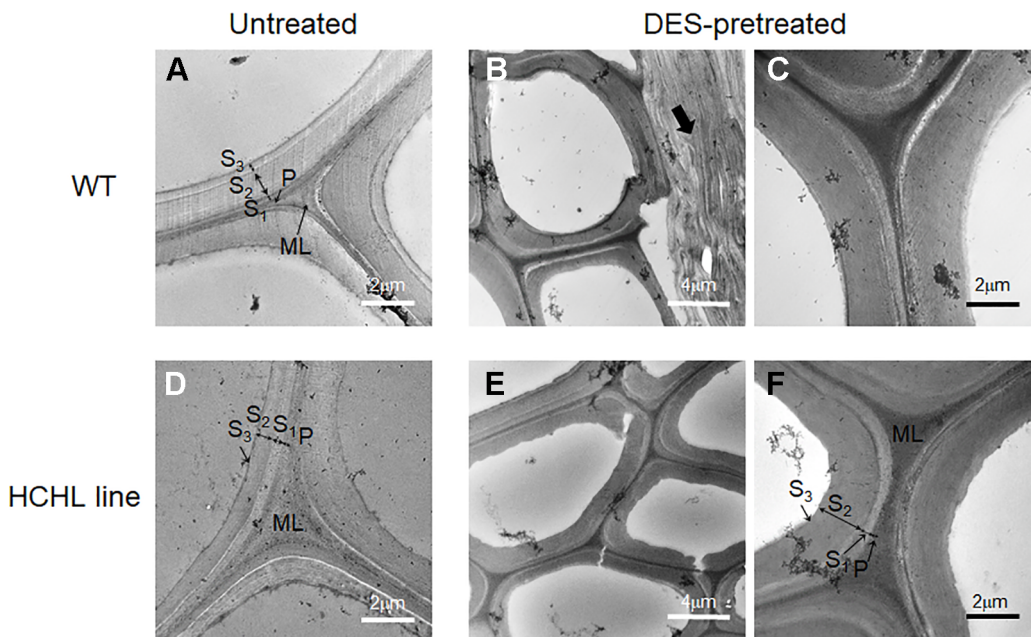

FIGURE 5 | TEM micrographs of cross-sections of wild type (WT) **(A–C)** and hydroxycinnamoyl-CoA hydratase-lyase (HCHL) lines **(D–F)** before **(A ,D)** and after ChCl-VAN pretreatment at 80°C **(B**, **C**, **E**, and **F)**.

were computed to find the optimized geometry. **Figure 6** depicts the optimized structures and the C_β-O bond length in the four lignin model dimers. In this work, the C_β-O bond length was particularly investigated because it is the most frequent linkage found in the lignin macromolecule and also the weakest linkage that is easily cleaved under heated reaction (Kim et al., 2011). As shown in **Figure 6**, the DFT calculations show that the bond length of C_β-O in PCAlc dimer is 1.4260 Å. In the case of model dimers with a shorter C_1 side chain, the dissociating bond length noticeably elongated. For example, the C_β-O bond length increased to 1.4297, 1.4340, 1.4426, and 1.4425 Å with HBA, HBAld, VAN, and SyrAld, respectively. This result indicates that it requires a lower bond dissociation energy for the ether linkage when the C_6C_1 monomers are incorporated into lignin structure.

The electrophilicity index was also calculated to assess the chemical reactivity of dimeric model compounds, and the results are given in **Figure 7**. The electrophilicity index of common lignin dimeric structure with *p*-coumaryl alcohol end group (PCAlc) was calculated to be 1.30 eV. Interestingly, the values of four dimeric model compounds with a shortened C_1 side-chain monomer end group were considerably higher than that of a conventional lignin structure, representing 1.56 (+20.2%), 1.86 (+43.1%), 1.89 (+45.3%), and 1.83 (+40.8%) with HBA, HBAld, VAN, and SyrAld, respectively. It is evident that lignin dimers with C_6C_1 monomers have a higher electrophilicity index, suggesting that these structures are chemically more reactive than the typical β-O-4 structure (Shi et al., 2016). The results from DFT computational study support the hypothesis that lignin structure in the HCHL line is more susceptible to biomass pretreatment.

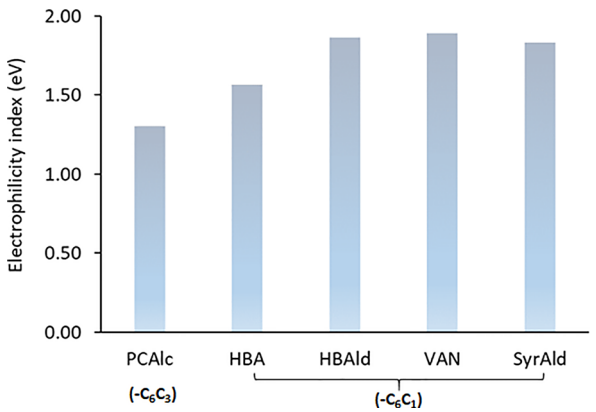

FIGURE 7 | Electrophilicity index of dimeric model compounds.

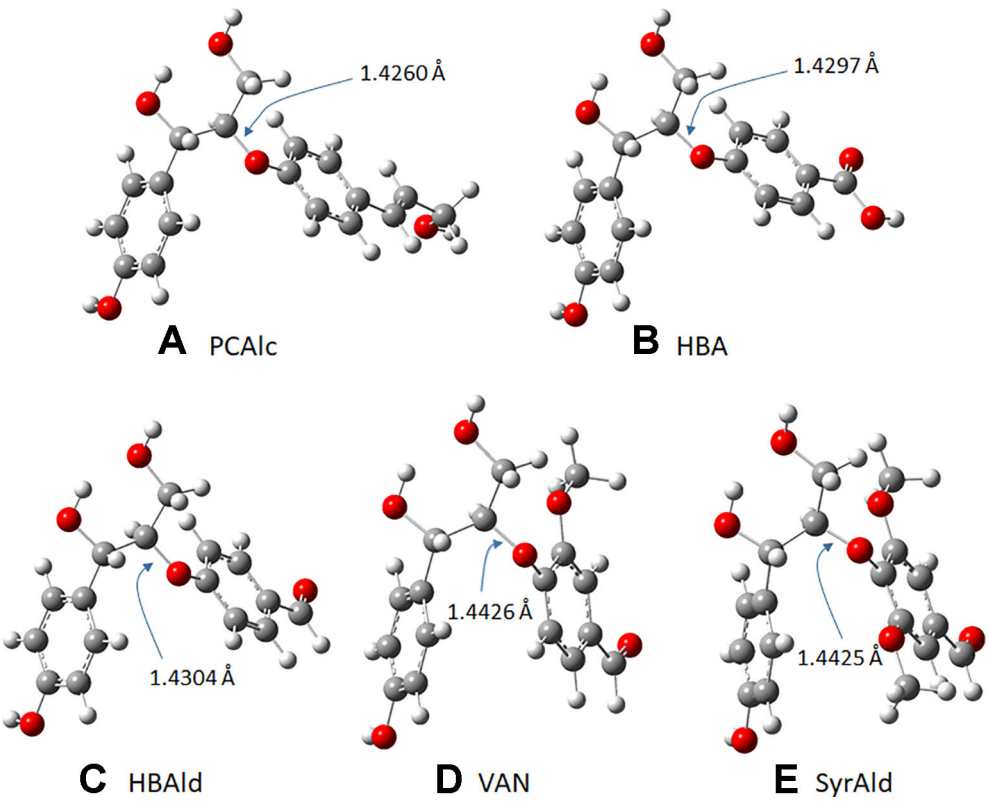

FIGURE 6 | (A–E) Optimized geometry of five lignin model dimers and dissociating bond (C_β-O) length.

DISCUSSION

Throughout the last decades, cell wall engineering has been a critical approach to reduce biomass recalcitrance by altering lignin structures. Perturbations of genes on the biosynthesis pathway of lignin result in significant structural changes, which allows designing readily tractable biomass structures for the production of biofuels and chemicals (Ralph et al., 2019). The strategic expression of HCHL in lignifying tissues of *Arabidopsis* resulted in the overproduction of side-chain-truncated (C_6C_1) lignin monomers incorporated in lignin as end-groups (Eudes et al., 2012). 2D HSQC NMR analysis confirms that isolated lignin from HCHL transgenic biomass contains a significant amount of oxidized C_6C_1 units. Considering the similar lignin content with a slight difference in molecular weight of isolated lignin, as well as comparable ultrastructural morphologies between the WT and HCHL lines, we propose that engineering the biosynthesis of C_6C_1 monomers in-planta is an effective approach to modify lignin without any agronomic penalty.

In this work, VAN, an oxidized C_6C_1 unit found in HCHL plants, was used to synthesize ChCl-VAN, a renewable DES that was previously shown to be effective for biomass pretreatment (Kim et al., 2018). Furthermore, integration of lignin-derived DES with HCHL-engineered biomass offers opportunities to move closer towards achieving a closed-loop biorefinery (Kim et al., 2019). The pretreatment of WT and HCHL biomass using ChCl-VAN showed that lignin removal from transgenic biomass was higher, resulting in enhanced saccharification efficiencies. Considering that lignin content in both samples is quite similar, the improvement of biomass digestibility for the HCHL line is associated with 1) significant structural alteration in lignin, and 2) the higher reactivity of short-side chain lignin monomers.

Regarding the chemical properties of lignin derived from HCHL plants, DFT-based calculation revealed higher reactivity for structures corresponding to C_6C_1 monomers linked to lignin. Moreover, lignin with side-chain-truncated monomers requires less energy for the cleavage of β-aryl ether bond, which is typically necessary for lignin fractionation during biomass pretreatment. Previously, lignins rich in H-units were computationally found to be more reactive than other types of lignins (G- and S-unit) (Shi et al., 2016), and transgenic biomass consisting of H-lignin yielded higher sugars upon saccharification (Bonawitz et al., 2014). As discussed above, the more reactive nature of C_6C_1 monomers with aldehyde or carboxylic acid functionalities that form lignin end-groups contributed to the increased pretreatment efficiency followed by higher saccharification yield.

It is noted that although both empirical and computational results undoubtedly support the hypothesis, more work is necessary to gain a far-reaching insight of structural modification resulting from the expression of HCHL. However, the pretreatment of transgenic biomass using a bio-derivable DES offers opportunities to operate reactors with reduced amount of energy and chemicals, which is highly desirable in developing a sustainable bioeconomy.

CONCLUSIONS

Plant cell wall engineering has been developed for the production of biofuels and renewable chemicals. In this study, we strategically expressed the HCHL gene in *Arabidopsis* to reduce the degree of lignin polymerization *via* incorporation of side-chain-truncated monomers in lignin polymer ends. The transgenic *Arabidopsis* yielded higher levels of fermentable sugars compared to WT plants when pretreated with the lignin-derived DES at mild conditions. The results of this work clearly indicate that interfering with the lignin biosynthetic pathway has the potential to improve the conversion of biomass into biofuels and other intermediate products. Together with the development of tailor-made biomass that is more amenable to chemical processes, biomass pretreatment using a renewable DES could make the biofuel industry more economically feasible in the future.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this article are not publicly available. Requests to access the datasets should be directed to KHK, kwanghokim@kist.re.kr.

AUTHOR CONTRIBUTIONS

KHK conceived and designed the experiments. KHK, YW, MT, and AE performed the experiments. KHK, CGY, CSK, and JS analyzed the data of the experiments. KHK, MT, AE, and CGY drafted manuscript. All authors read and approved the final manuscript.

FUNDING

This work is supported by the Korea Institute of Science and Technology–The University of British Columbia Biorefinery on-site laboratory project. NMR analysis is supported by SUNY ESF and NIH Shared Instrumentation Grant 1S10OD012254. This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

REFERENCES

Agbor, V. B., Cicek, N., Sparling, R., Berlin, A., and Levin, D. B. (2011). Biomass pretreatment: fundamentals toward application. *Biotechnol. Adv.* 29:6, 675–685. doi: 10.1016/j.biotechadv.2011.05.005

Bhagia, S., Muchero, W., Kumar, R., Tuskan, G. A., and Wyman, C. E. (2016). Natural genetic variability reduces recalcitrance in poplar. *Biotechnol. biofuels* 9:(1), 106. doi: 10.1186/s13068-016-0521-2

Bonawitz, N. D., and Chapple, C. (2013). Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Curr. Opin. Biotechnol.* 24:(2), 336–343. doi: 10.1016/j.copbio.2012.11.004

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:(7500), 376–380. doi: 10.1038/nature13084

Boudet, A. M. (2007). Evolution and current status of research in phenolic compounds. *Phytochemistry* 68:(22–24), 2722–2735. doi: 10.1016/j.phytochem.2007.06.012

Chabannes, M., Barakate, A., Lapierre, C., Marita, J. M., Ralph, J., Pean, M., et al. (2001). Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants. *Plant J.* 28:(3), 257–270. doi: 10.1046/j.1365-313X.2001.01140.x

Chen, F., and Dixon, R. A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.* 25:7, 759. doi: 10.1038/nbt1316

Dai, Y. T., van Spronsen, J., Witkamp, G. J., Verpoorte, R., and Choi, Y. H. (2013). Natural deep eutectic solvents as new potential media for green technology. *Anal. Chim. Acta* 766, 61–68. doi: 10.1016/j.aca.2012.12.019

Dale, B. E., Anderson, J. E., Brown, R. C., Csonka, S., Dale, V. H., Herwick, G., et al. (2014). Take a closer look: biofuels can support environmental, economic and social goals. *Environ. Sci. Technol.* 48:(13), 7200–7203. doi: 10.1021/es5025433

Demirbaş, A. (2001). Biomass resource facilities and biomass conversion processing for fuels and chemicals. *Energy Convers. Manage.* 42:(11), 1357–1378. doi: 10.1016/S0196-8904(00)00137-0

Ding, S.-Y., Liu, Y.-S., Zeng, Y., Himmel, M. E., Baker, J. O., and Bayer, E. A. (2012). How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? *Science* 338:(6110), 1055–1060. doi: 10.1126/science.1227491

Doherty, T. V., Mora-Pale, M., Foley, S. E., Linhardt, R. J., and Dordick, J. S. (2010). Ionic liquid solvent properties as predictors of lignocellulose pretreatment efficacy. *Green Chem.* 12:11, 1967–1975. doi: 10.1039/c0gc00206b

Donohoe, B. S., Vinzant, T. B., Elander, R. T., Pallapolu, V. R., Lee, Y. Y., Garlock, R. J., et al. (2011). Surface and ultrastructural characterization of raw and pretreated switchgrass. *Bioresour. Technol.* 102:(24), 11097–11104. doi: 10.1016/j.biortech.2011.03.092

Eudes, A., George, A., Mukerjee, P., Kim, J. S., Pollet, B., Benke, P. I., et al. (2012). Biosynthesis and incorporation of side-chain-truncated lignin monomers to reduce lignin polymerization and enhance saccharification. *Plant Biotechnol. J.* 10:(5), 609–620. doi: 10.1111/j.1467-7652.2012.00692.x

Eudes, A., Sathitsuksanoh, N., Baidoo, E. E. K., George, A., Liang, Y., Yang, F., et al. (2015). Expression of a bacterial 3-dehydroshikimate dehydratase reduces lignin content and improves biomass saccharification efficiency. *Plant Biotechnol. J.* 13:(9), 1241–1250. doi: 10.1111/pbi.12310

Himmel, M. E. (2009). *Biomass recalcitrance: deconstructing the plant cell wall for bioenergy* (Oxford: Wiley-Blackwell).

Isikgor, F. H., and Becer, C. R. (2015). Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polym. Chem.* 6:(25), 4497–4559. doi: 10.1039/c5py00263j

Kim, K. H., and Kim, C. S. (2018). Recent Efforts to Prevent Undesirable Reactions From Fractionation to Depolymerization of Lignin: Toward Maximizing the Value From Lignin. *Front. Energy Res.* 6, 92. doi: 10.3389/fenrg.2018.00092

Kim, H., and Ralph, J. (2010). Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d(6)/pyridine-d(5). *Org. Biomol. Chem.* 8:(3), 576–591. doi: 10.1039/b916070a

Kim, S., Chmely, S. C., Nimos, M. R., Bomble, Y. J., Foust, T. D., Paton, R. S., et al. (2011). Computational study of bond dissociation enthalpies for a large range of native and modified lignins. *J. Phys. Chem. Lett.* 2:(22), 2846–2852. doi: 10.1021/jz201182w

Kim, K. H., Dutta, T., Ralph, J., Mansfield, S. D., Simmons, B. A., and Singh, S. (2017). Impact of lignin polymer backbone esters on ionic liquid pretreatment of poplar. *Biotechnol. Biofuels* 10. doi: 10.1186/s13068-017-0784-2

Kim, K. H., Dutta, T., Sun, J., Simmons, B., and Singh, S. (2018). Biomass pretreatment using deep eutectic solvents from lignin derived phenols. *Green Chem.* 20:(4), 809–815. doi: 10.1039/c7gc03029k

Kim, K. H., Eudes, A., Jeong, K., Yoo, C. G., Kim, C. S., and Ragauskas, A. (2019). Integration of renewable deep eutectic solvents with engineered biomass to achieve a closed-loop biorefinery. *Proc. Nat. Acad. Sci. U.S.A.* 116:(28), 13816–13824. doi: 10.1073/pnas.1904636116

Koch, G., and Schmitt, U. (2013). "Topochemical and electron microscopic analyses on the lignification of individual cell wall layers during wood formation and secondary changes" in *Cellular aspects of wood formation*. Ed. J. Fromm (Heidelberg: Springer), 41–69. doi: 10.1007/978-3-642-36491-4_2

Maity, S. K. (2015). Opportunities, recent trends and challenges of integrated biorefinery: Part I. *Renewable Sustainable Energy Rev.* 43, 1427–1445. doi: 10.1016/j.rser.2014.11.092

Mitra, A., Mayer, M. J., Mellon, F. A., Michael, A. J., Narbard, A., Parr, A. J., et al. (2002). 4-Hydroxycinnamoyl-CoA hydratase/lyase, an enzyme of phenylpropanoid cleavage from Pseudomonas causes formation of C-6-C-1 acid and alcohol glucose conjugates when expressed in hairy roots of Datura stramonium L. *Planta* 215:(1), 79–89. doi: 10.1007/s00425-001-0712-2

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:(6185), 1246843. doi: 10.1126/science.1246843

Ralph, J., Lapierre, C., and Boerjan, W. (2019). Lignin structure and its engineering. *Curr. Opin. Biotechnol.* 56, 240–249. doi: 10.1016/j.copbio.2019.02.019

Renders, T., Van den Bosch, S., Koelewijn, S. F., Schutyser, W., and Sels, B. F. (2017). Lignin-first biomass fractionation: the advent of active stabilisation strategies. *Energy Environ. Sci.* 10:(7), 1551–1557. doi: 10.1039/c7ee01298e

Schilmiller, A. L., Stout, J., Weng, J. K., Humphreys, J., Ruegger, M. O., and Chapple, C. (2009). Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis. *Plant J.* 60:(5), 771–782. doi: 10.1111/j.1365-313X.2009.03996.x

Shi, J., Pattathil, S., Parthasarathi, R., Anderson, N. A., Kim, J. I., Venketachalam, S., et al. (2016). Impact of engineered lignin composition on biomass recalcitrance and ionic liquid pretreatment efficiency. *Green Chem.* 18:(18), 4884–4895. doi: 10.1039/c6gc01193d

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., et al. (2008). *"Determination of structural carbohydrates and lignin in biomass-laboratory analytical procedure (LAP)"* (Colorado, USA: National Renewable Energy Laboratory). Report number TP-510-42618).

Socha, A. M., Parthasarathi, R., Shi, J., Pattathil, S., Whyte, D., Bergeron, M., et al. (2014). Efficient biomass pretreatment using ionic liquids derived from lignin and hemicellulose. *Proc. Nat. Acad. Sci. U. S. A.* 111:(35), E3587–E3595. doi: 10.1073/pnas.1405685111

Sun, J., Dutta, T., Parthasarathi, R., Kim, K. H., Tolic, N., Chu, R. K., et al. (2016). Rapid room temperature solubilization and depolymerization of polymeric lignin at high loadings. *Green Chem.* 18:(22), 6012–6020. doi: 10.1039/c6gc02258h

Terashima, N., and Fukushima, K. (1988). Heterogeneity in Formation of Lignin.11. An Autoradiographic Study of the Heterogeneous Formation and Structure of Pine Lignin. *Wood Sci. Technol.* 22:(3), 259–270. doi: 10.1007/Bf00386021

Thomas, V. A., Kothari, N., Bhagia, S., Akinosho, H., Li, M., Pu, Y., et al. (2017). Comparative evaluation of Populus variants total sugar release and structural features following pretreatment and digestion by two distinct biological systems. *Biotechnol. Biofuels* 10:1, 292. doi: 10.1186/s13068-017-0975-x

Vigier, K. D., Chatel, G., and Jerome, F. (2015). Contribution of deep eutectic solvents for biomass processing: opportunities, challenges, and limitations. *Chemcatchem* 7:8, 1250–1260. doi: 10.1002/cctc.201500134

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:(6179), 90–93. doi: 10.1126/science.1250161

Wyman, C. E., Dale, B. E., Elander, R. T., Holtzapple, M., Ladisch, M. R., and Lee, Y. (2005). Comparative sugar recovery data from laboratory scale application of leading pretreatment technologies to corn stover. *Bioresour. Technol.* 96:(18), 2026–2032. doi: 10.1016/j.biortech.2005.01.018

Wyman, C. E., Dale, B. E., Balan, V., Elander, R. T., Holtzapple, M. T., Ramirez, R. S., et al. (2013). "Comparative performance of leading pretreatment technologies for biological conversion of corn stover, poplar wood, and switchgrass to sugars" in *Aqueous pretreatment plant biomass for biological and chemical conversion to fuels and chemicals*, Ed. C. E. Wyman (Chichester: Wiley), 239–259. doi: 10.1002/9780470975831.ch12

Xu, B., Escamilla-Treviño, L. L., Sathitsuksanoh, N., Shen, Z., Shen, H., Percival Zhang, Y. H., et al. (2011). Silencing of 4-coumarate: coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. *New Phytol.* 192:(3), 611–625. doi: 10.1111/j.1469-8137.2011.03830.x

Yoo, C. G., Pu, Y. Q., Li, M., and Ragauskas, A. J. (2016). Elucidating Structural Characteristics of Biomass using Solution-State 2D NMR with a Mixture of Deuterated Dimethylsulfoxide and Hexamethylphosphoramide. *Chemsuschem* 9:(10), 1090–1095. doi: 10.1002/cssc.201600135

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 February 2020
doi: 10.3389/fpls.2020.00074

Heterologous Expression of Ethanol Synthesis Pathway in Glycogen Deficient *Synechococcus elongatus* PCC 7942 Resulted in Enhanced Production of Ethanol and Exopolysaccharides

Rajendran Velmurugan[1] *and Aran Incharoensakdi*[1,2]*

[1] Cyanobacterial Biotechnology Laboratory, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok, Thailand, [2] Academy of Science, Royal Society of Thailand, Bangkok, Thailand

OPEN ACCESS

Edited by:
Jaime Barros-Rios,
University of North Texas,
United States

Reviewed by:
Alberto A. Iglesias,
National University of the Littoral,
Argentina
Rei Narikawa,
Shizuoka University, Japan

***Correspondence:**
Aran Incharoensakdi
aran.i@chula.ac.th

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: 18 September 2019
Accepted: 20 January 2020
Published: 14 February 2020

Citation:
Velmurugan R and Incharoensakdi A (2020) Heterologous Expression of Ethanol Synthesis Pathway in Glycogen Deficient Synechococcus elongatus PCC 7942 Resulted in Enhanced Production of Ethanol and Exopolysaccharides. Front. Plant Sci. 11:74. doi: 10.3389/fpls.2020.00074

In this study, the *Synechococcus elongatus* PCC 7942 (hereafter *S. elongatus*) was engineered by the *glgC* knockout as well as the insertion of the *pdc-adh* genes from two different microorganisms. The insertion of *pdc-adh* genes increased the ethanol synthesis with further improvement in the productivity upon the destruction of glycogen synthesis pathway and the supplementation of cofactor. The abolition of glycogen synthesis pathway led to a considerable increase of the engineered *S. elongatus* metabolites involved in the ethanol synthesis pathway. Moreover, the studies on cofactor addition highlighted the importance of Mg^{+2}, Zn^{+2}, thiamine pyrophosphate, and $NADP^{+}$ in ethanol synthesis. The yields of 3856 mg/L ethanol and 109.5 μg/10^8 cells exopolysaccharides were obtained in the engineered *S. elongatus* using a photo-bioreactor under optimized conditions. This enhanced production in ethanol and exopolysaccharides are attributed to the flux of carbon from glycogen synthesis pathway and proper availability of essential components.

Keywords: ***Synechococcus elongatus*, ethanol, metabolic engineering, cofactor, exopolysaccharides**

INTRODUCTION

The energy crisis and depletion of fossil fuel necessitate the demand toward alternative fuel production. As an alternative fuel, ethanol is already alleviating the dependency on fossil fuel, and it is known to be sustainable and environmentally friendly (Mofijur et al., 2016). The ethanol production from cyanobacteria is considered as a future fuel source due to the direct fuel molecule production from atmospheric CO_2 (Nozzi et al., 2013; Norena-Caro and Benton, 2018). Although the research on ethanol has reached up to the industrial level, there are certain factors limiting its commercialization (Su et al., 2017). In general, the engineering of ethanol synthesis pathway consisted of an insertion of two genes *pdc* and *adh* encoding pyruvate decarboxylase which converts pyruvate to acetaldehyde, and alcohol dehydrogenase which converts acetaldehyde to ethanol,

respectively (Deng and Coleman, 1999). The engineering of ethanol synthesis pathway has been carried out widely in two genera, such as *Synechocystis* and *Synechococcus*, in which the genes of interest and expression systems were varied (Deng and Coleman, 1999; Gao et al., 2012). The first case of ethanol production improvement was carried out in *Synechococcus* sp. PCC 7942, which expresses PDC and ADH from *Zymomonas mobilis* under the control of the *rbcLS* promoter (Deng and Coleman, 1999). Another attempt to redirect more carbon sources to ethanol synthesis by inactivating the glycogen and PHB synthesis pathways led to an improvement of the ethanol synthesis in *Synechocystis* sp. (Gao et al., 2012; Velmurugan and Incharoensakdi, 2020). Although the engineering strategies improved the ethanol synthesis, the production process is still facing problems associated with stress tolerance, adaptability, and productivity (Pade et al., 2017). As the wild type cyanobacterial strains have certain characteristics to grow under normal growth conditions, the engineered strain to produce certain products does not grow well (Dexter and Fu, 2009). Therefore, it is important to study the growth pattern under challenging environment imposed by the incorporation of synthetic pathways. The incorporation of ethanol synthesis pathway into a new host strain may cause insufficient availability of related cofactors (Choi and Park, 2016). On the other hand, several physiological factors can decrease the product yield, such as an increased ethanol concentration, electron imbalances, and membrane transport systems (Tian et al., 2013). Higher ethanol concentration is reported to trigger stress response in cyanobacteria (Tian et al., 2013). The *Synechococcus* holds its natural traits like possessing various unique components in cyanobacteria like exopolysaccharides, glycogen, and carotenoids, and these components act as an extracellular protecting agent and intracellular electron sink, respectively (Pereira et al., 2009; Velmurugan and Incharoensakdi, 2018). As one of the strains developed in this study is defective in glycogen synthesis, the demonstration of the changes of exopolysaccharides content in response to stress can be informative.

In this study, the ethanol synthesis pathway engineered *S. elongatus* was used to analyze the ethanol production and also to investigate the changes in intracellular and extracellular concentration of biomolecules. The contribution of cofactor has been evaluated to fulfill the ethanol synthesis pathway as the pathway is new to the host strain *S. elongatus*. Besides, the exopolysaccharides content in response to physiological condition imposed by pathway engineering was determined.

MATERIALS AND METHODS

Materials

All the standard chemicals used were purchased from Sigma-Aldrich (USA) and the nucleotide bases were purchased from Fermentas (Canada). Taq polymerase, ligase, and restriction enzymes (BamHI, BmtI, HindIII, KpnI, and NdeI) were purchased from Fermentas (Canada). The kits used for plasmid extraction and PCR purification were obtained from Geneaid Biotech Ltd. Oligonucleotide primers were designed with the help of primer3 online software, and the synthesis was performed by Pacific Science Co. Ltd, Thailand. The vectors such as pSyn_1, pGEM-T easy, and pUC4K were standard commercial products.

Strains and Cultivation Conditions

Synechococcus elongatus PCC 7942 (Pasteur Institute, France) was propagated on BG-11 agar medium (Rippka et al., 1979). The wild type and engineered strains were cultivated in 250 ml Erlenmeyer flasks containing 100 ml BG-11 medium (pH 7.5) under continuous illumination of 50 $\mu E/m^2/s$ at 28 ± 1°C with atmospheric CO_2 supplementation. The commercial microorganisms *Escherichia coli* DH5α and *Saccharomyces cerevisiae* (MTCC-170) were cultivated in LB medium and, yeast extract peptone and dextrose medium (YEPD), respectively.

Plasmid Construction

All the primers used and the strains constructed are presented in **Table 1** and **Figure 1**, respectively. The expression vector pAPX was constructed by the insertion of $NADP^+$ dependent alcohol dehydrogenase (*adh*: slr0942) into pSyn_1 vector under the control of *Psc* promoter. The gene *adh* was amplified with respective primers (0942F and 0942R) using the genomic DNA of *Synechocystis* as a template and the gene *pdc* was amplified with respective primers (*PdcF* and *PdcR*) using the genomic DNA of *S. cerevisiae* as a template (**Figure 1**). Plasmid pGK was constructed by inserting 1.3 kb ADP-glucose pyrophosphorylase (*glgC*) gene into pGEM-T easy vector. The *glgC* was amplified with respective primers (0603F and 0603R) using the genomic DNA of *S. elongatus* as a template. The inactivation of *glgC* was carried out by inserting a 0.92 kb kanamycin resistance cassette from pUC4K vector into BmtI site of pGK resulting in plasmid with kanamycin resistance.

TABLE 1 | Primers used for engineering of *S. elongatus* PCC 7942.

Gene	Primers	Oligonucleotide	PCR product length (bp)
adh	0942F	GTGGATCCGTGCAGAGTTTCAATAGG	990
	0942R	CGGGTACCTTAAATTTCATCCCATAGG	
pdc	PdcF	GCGGAAGCTTATGTCTGAAATTACTTTG	1698
	PdcR	GCGGATCCTTATTGCTTAGCGTTGGT	
glgC	0603F	TGGTACCGTGAAAAACGTGCTGGCGAT	1299
	0603R	GTCATATGTTAGATCACCGTGTTGTCGGG	
Km	KmF	GCAAGCTAGCAAGCCACGTTGTGTCTCA	932
	KmR	GCCAGCTAGCGATTAGAAAAACTCATCG	
Adh-RT-PCR	Adh-RTF	AACTTTGCAGGATTTGGGTCTA	297
	Adh-RTR	AGCAAGTCTGATTGTTGGAGGTA	
Pdc- RT-PCR	Pdc-RTF	CCAGCTTTCGTCACCCCAAT	265
	Pdc-RTR	CGAATTTCATTTGGACACCTGG	
16s	16s-RTF	CTTCGCGTTGCATCGAATTAAACCAC	368
	16s-RTR	GCGTGGGGCTCAACCTCATAC	
NSI	NS1F	GGCAGCTTGGAAGGGCG	1568
	NS1R	GGCGTTGCCAATATCAAGATTGC	
Promoter	pSyPF	CGGTCTGATCTTAGCGG	Not applicable

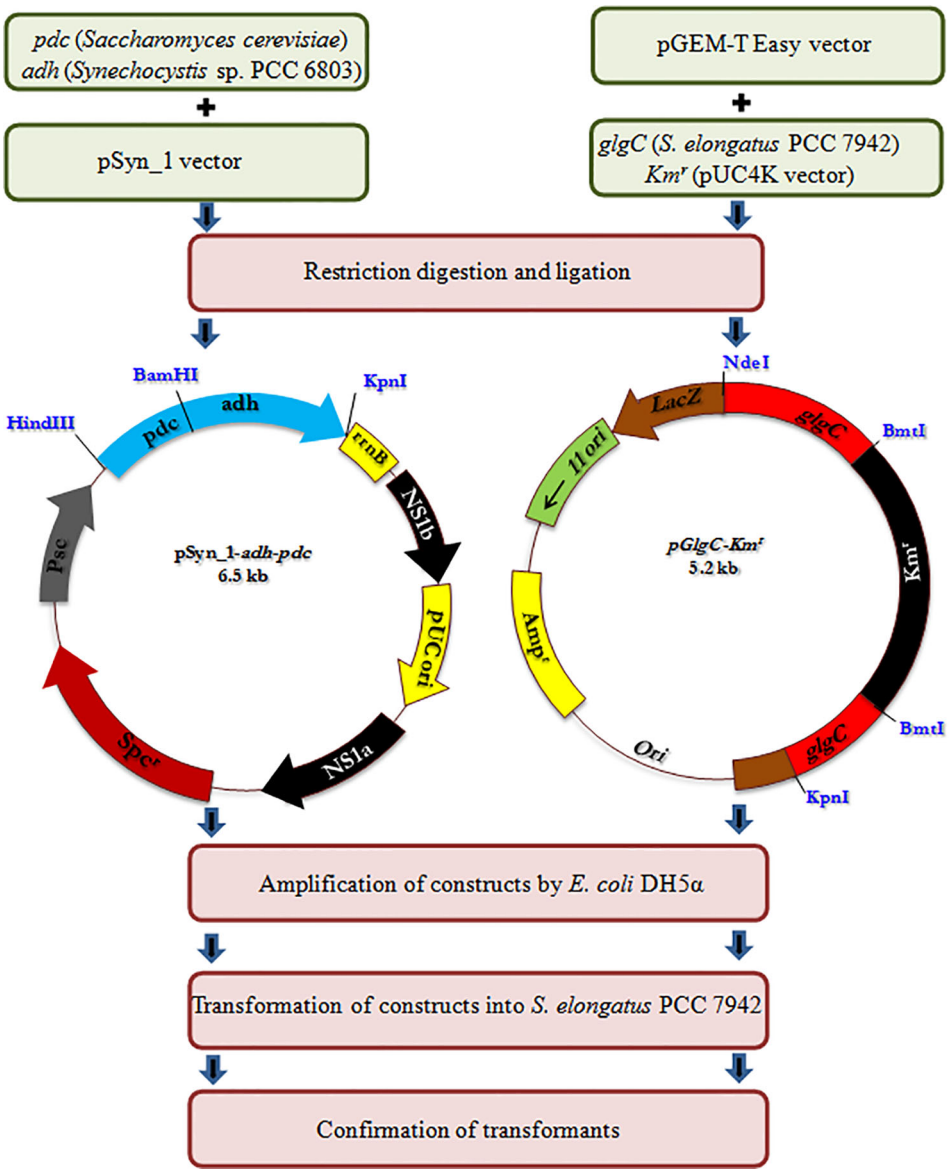

FIGURE 1 | Schematic representation of engineering of *S. elongatus*.

Construction of Engineered Strains

Escherichia coli DH5α strain was used for routine propagation of all plasmid constructs and was cultivated in LB medium containing respective antibiotics at appropriate concentration. The transformation of plasmid constructs into *E. coli* DH5α was performed by heat shock transformation and the amplified plasmids were extracted after the cultivation for overnight. The engineered *S. elongatus* was obtained by natural transformation. Briefly, *S. elongatus* was grown in BG-11 medium for 2 to 3 days until the cell density reached OD_{730} = 0.3. Cells were harvested by centrifugation and resuspended in fresh BG-11 medium (100 μL). The transformation was carried out by adding 5μL of appropriate plasmid DNA into the cell suspension. The mixture was incubated at 27°C under low light for 6 h before spreading on a plate containing appropriate antibiotics for natural transformation. After about 2 weeks of incubation, the single colonies were isolated and subcultured for at least seven generations. After full segregation was achieved, engineered strains were verified by forward and reverse primers of NS1 (NS1F and NS1R), *glgC* (0603F and 0603R), and antibiotic cassette (KmF and KmR), respectively (**Figure S1**). The engineered strains were maintained in BG-11 medium containing appropriate antibiotics such as ampicillin (30 μg/ml), kanamycin (50 μg/ml), and spectinomycin (30 μg/ml) (Kanwal et al., 2015).

Characterization of Wild-Type and Engineered Strains

The wild type and engineered *S. elongatus* were cultured in 250 ml Erlenmeyer flasks containing 100 ml BG11 medium (pH 7.5) with a continuous illumination of 100 μE/m^2/s at 28 ± 1°C and atmospheric CO_2 was supplemented. The initial cell density was 4×10^7 cells/ml. The optimization of cofactor concentration such as $MgSO_4.7H_2O$ (0, 200, 400, 600, and 800 μM), $ZnSO_4.7H_2O$ (0, 1, 2, 3, 4 μM), TPP (0, 50, 100, 150, and 200) and $NADP^+$ (0, 50, 100, 150, and 200) concentrations were optimized in 250-ml flask experiment as outlined above. The individual effect of cofactors on growth, exopolysaccharides, and ethanol content was analyzed in 5 L photo-bioreactor with the atmospheric CO_2 supplementation at the flow rate of 200 ml/min. Cells were harvested after 20 days of growth by centrifugation at 5,000*g* for 10 min at room temperature. To estimate the intracellular metabolites, cells were vigorously mixed with 500 μl of 70% methanol by vortex mixer. The mixture was incubated for 2 h at room temperature and then centrifuged at 6,000*g* for 10 min at 4°C. The supernatant was collected and dried in a vacuum evaporator at 40°C. The pellet left after drying was dissolved and mixed thoroughly in 250 μl of water and 50 μl of chloroform followed by centrifugation at 6,000*g* for 10 min. The uppermost water phase (200 μl) was collected, and filtered through a 0.45 μm Millipore filter before the analysis by high performance liquid chromatography (HPLC).

To estimate the exopolysaccharides content, the cell suspension obtained from 50-ml culture was made up to 10 ml using Milli-Q water (Millipore, USA) and vortexed five times for 30 s with 1 min intervals. The samples were centrifuged at 15,000*g* for 20 min under room temperature, and the supernatant was concentrated approximately 10 fold by evaporation at 60°C for 8 to 12 h. The exopolysaccharides in the concentrated liquid were precipitated by gradual addition of 3 volumes of cold ethanol and kept overnight at 4°C. After evaporation, the precipitate was washed with cold absolute ethanol, followed by centrifugation. The gel-like pellet was dialyzed against 5 volumes of MilliQ water at room temperature for 12 h. The samples were dissolved in HCl (2M final) and autoclaved at 110°C for 10 min. The concentrations of glucose, galactose, xylose, mannose, arabinose, and uronic acid were determined by HPLC (Panoff et al., 1998).

Transcript Analysis by RT-PCR

Total RNA was extracted by using the TRI Reagent (Molecular Research Center). After DNase (Fermentas) treatment, single-stranded cDNA was synthesized from 1 μg of total RNA with the SuperScript™ III First-Strand Synthesis Kit (Invitrogen) according to the manufacturer's instruction. RT-PCR using cDNA of the *adh*, *pdc*, and *16s* as a reference gene was performed using forward and reverse primers as described in **Table 1**. The PCR conditions consisted of: denaturation at 94°C for 5 min, followed by 20 cycles of 94°C for 30 s, annealing temperature of 55°C for 1 min and 72°C for 20 s, and then a final extension at 72°C for 3 min. The PCR product was analyzed using a 1.0% (w/v) agarose gel electrophoresis system (Kanwal et al., 2014).

Analytical Methods

Intracellular pigments of *S. elongatus* cell suspension were extracted by dimethylformamide. Chlorophyll *a* and carotenoid contents were determined according to the methods of Moran (1982) and Chamovitz et al. (1993), respectively. A Clark-type oxygen electrode was employed for oxygen evolution measurement (YSI 5300A, YSI Inc., USA) at 30°C (Jantaro et al., 2005). PHB contents were determined using HPLC (Shimadzu, Japan) equipped with InertSustain 3-μm C18 column (GL Sciences, Japan) and UV/Vis detector as described by Monshupanee and Incharoensakdi (2014). The estimation of glycogen was performed as previously described (Velmurugan and Incharoensakdi, 2016). The sugar (glucose, galactose, xylose, mannose, and arabinose) and ethanol contents were quantified using HPLC system equipped with refractive index detector (RID 10A, Shimadzu, Japan). Metabolic intermediates such as acetate, pyruvate, succinate, and uronic acid were quantified using HPLC equipped with UV/Vis detector (SPD-20A, Shimadzu, Japan). The components were separated in Phenomenex, Rezex ROA-Organic acid column (150 × 7.8 mm) with 5 mM H_2SO_4 as a mobile phase at a flow rate of 0.6 ml/min (Velmurugan and Incharoensakdi, 2017).

Enzyme Characterization

The intracellular protein was extracted from the cells after 10 and 20 days of cultivation. Cells were harvested from 50-ml liquid culture by centrifugation at 6,000*g* for 10 min and resuspended in 3 ml of 30 mM Tris-HCl, pH 8.0. The resuspended cells were lysed by ultrasonic treatment for 30 s with 60% power input and were repeated three times in a pre-chilled water bath. The lysate was centrifuged at 6000*g* for 5 min, and the supernatant was used for protein determination, ADH and PDC assays.

The protein concentration was determined by Bradford method using BSA as a standard protein (Bradford, 1976). Activities of ADH and PDC were measured by monitoring the increase and decrease in absorbance at 340 nm with utilization of 200 μM $NADP^+$ and NADPH, respectively. Briefly, the ADH activity was measured by adding the enzyme extract in 30 mM Tris (pH 8.0) buffer containing 200 μm $NADP^+$ and ethanol. The PDC activity was measured in 100 mM Tris (pH 7.5) buffer containing 200 μm NADPH, 0.1 mM $MgCl_2$, 0.1 mM thiamine pyrophosphate, and 10 mM pyruvate (Gao et al., 2012).

Statistical Analysis

The experiments were performed with three biological replicates, and the average values are reported. The average standard deviation values were calculated using the respective functions (AVERAGE, STDEV) available in Microsoft Excel. The significance of the results were analysed by two-tailed Student's t-test.

RESULTS

Growth of Wild Type and Engineered Strain

The growth pattern of wild type, *adh-pdc* overexpressing strain (▲APX) and *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) under photosynthetic growth condition is shown in **Figure 2A**. *The* two engineered strains (▲APX and ▲APX-ΔGK) had lower growth than the wild type strain, whereas the *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) showed the lowest growth of 6.04×10^8 cells/ml, suggesting that the glycogen plays an important role in the cell growth. In addition, the culture broth of the *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) showed slightly yellowish color while the wild type culture stayed greenish at least until 20 days. The insertion of ethanol synthesis pathway improved the chlorophyll *a* content and the disruption of glycogen synthesis further improved the chlorophyll *a* content of 3.50 μg/10^8 cells (**Figure 2A**). The increase in chlorophyll *a* content was observed concomitant with the decrease in glycogen content of both *adh-pdc* overexpressing strain (▲APX) and *adh-pdc* overexpressing strain containing *glgC* knockout (APX-ΔGK) (**Figure 2B**). The oxygen evolution rate was also increased in engineered strains in a similar manner to the increase of chlorophyll *a* with the highest rate of 1.8×10^2 μmol O_2/mg chl*a*·h observed in the overexpressing strain containing glycogen synthesis knockout (**Figure 2A**).The wild type strain contained 15.5% glycogen content which was slightly reduced to 13.4% in the strain engineered for enhanced ethanol production (▲APX) and drastically reduced to 0.5% in overexpressing strain containing glycogen synthesis knockout (▲APX -ΔGK) (**Figure 2B**). On the other hand, the lipid and protein contents were increased to 13.8 and 22% w/w respectively in ▲APX -ΔGK strain (**Figure 2B**).

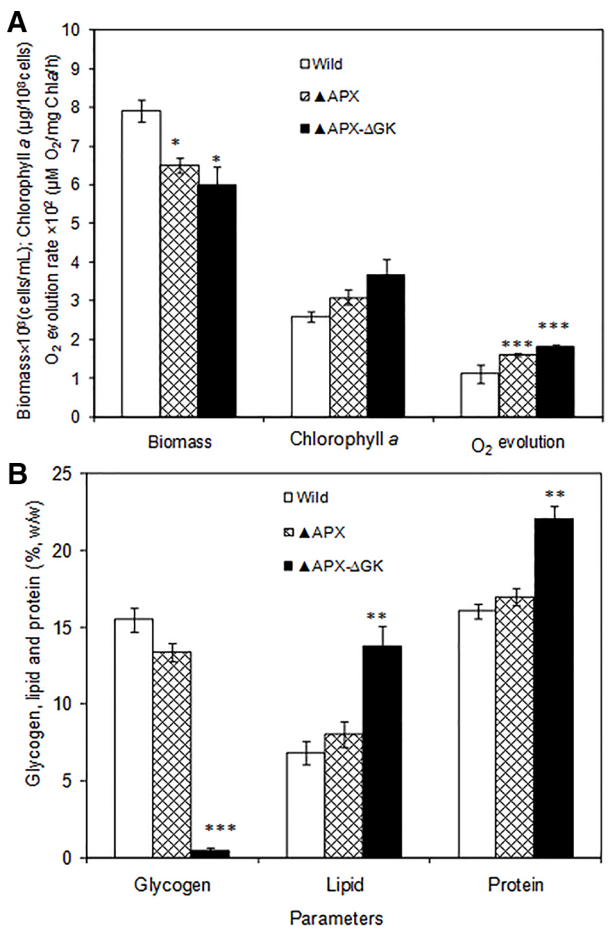

FIGURE 2 | (A) Biomass and chlorophyll *a* content and oxygen evolution rate and **(B)** glycogen, lipid and protein contents in wild type and engineered *S. elongatus* (conditions: 20 days; BG-11 medium). The values are mean ± standard deviation (n = 3), statistical significance (compared to wild type) are indicated by * for $p < 0.05$; ** for $p < 0.01$; *** for $p < 0.001$.

Metabolic Changes in Engineered Strains

The intracellular glucose, acetate, pyruvate, and succinate contents were drastically decreased in engineered strains as compared to those in the wild type (**Figure 3A**). Notably, the intracellular ethanol content was remarkably increased in the overexpressing strains. When comparing the *adh-pdc* overexpressing strain (▲APX) with *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX -ΔGK), the extracellular concentrations of glucose, acetate, pyruvate, and succinate were increased in glycogen synthesis knockout strain indicating the excessive amount of these three metabolites being excreted out when their intracellular contents exceeded the limit (**Figure 3B**). Remarkably, the ethanol production is highly improved from 919.7 mg/L to 2059.7 mg/L by the *glgC* knockout in *S. elongatus* (**Figure 3B**).

Expression Levels of Engineered Strains

The expression level of *adh* and *pdc* genes were quantified by analyzing their enzyme activities and the *adh* and *pdc* transcripts of total RNA extracted from *S. elongatus* cells (wild type, *adh-pdc* overexpressing; ▲APX, and *adh-pdc* overexpressing strain containing *glgC* knockout; ▲APX-ΔGK) (**Figures 4A, B**). Since multiple copies of chromosomes exist in S. *elongatus*, the *16s* gene was chosen as a reference gene. As shown in **Figure 4B**, the transcripts of both *adh* and *pdc* were not detected in the wild type while they were present in high amounts in the engineered strains, indicating the successful segregation of respective genes in the *S. elongatus*. The transcript level of *adh* was not much different for both *adh-pdc* overexpressing strain (▲APX) and *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK), whereas the latter showed a 1.5-fold increase in *pdc* transcript level. The maximum transcript level of 0.55 and 0.60 were observed for *adh* and *pdc* respectively in *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK). Subsequently, the activities of ADH and PDC in the wild type, ▲APX and ▲APX-ΔGK strains were examined (**Figure 4A**). In accord with the transcript levels, the enzyme activities were not observed in the wild-type, whereas the *adh-pdc* overexpressing strain (▲APX) and *adh-pdc* overexpressing strain containing

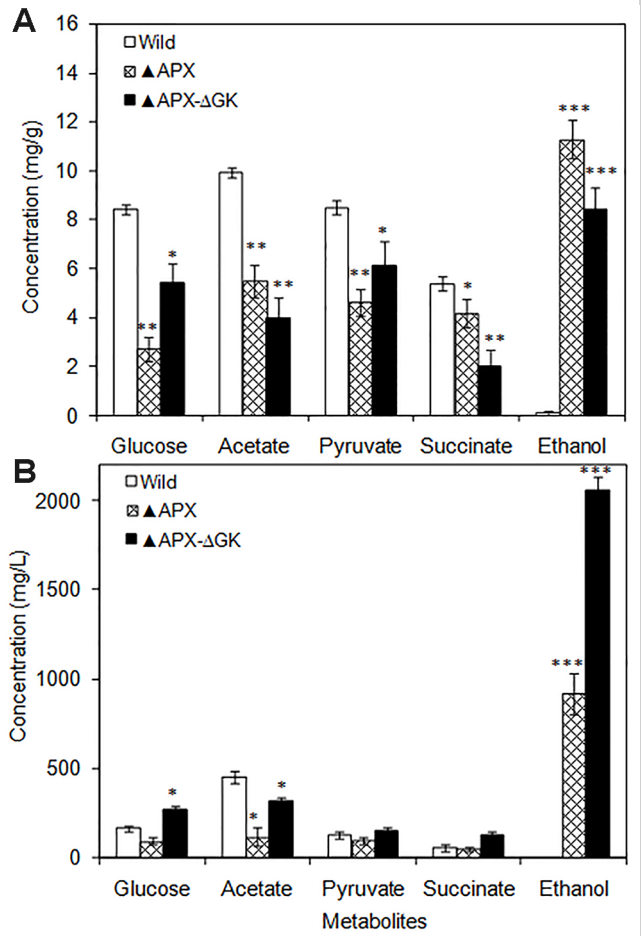

FIGURE 3 | (A) Intracellular and **(B)** extracellular metabolites of wild type and engineered *S. elongatus* (conditions: 20 days; BG-11 medium). Values are mean ± standard deviation (n = 3), statistical significant values (compared to wild type) are indicated by * for $p < 0.05$; ** for $p < 0.01$; *** for $p < 0.001$.

glgC knockout (▲APX-ΔGK) had considerable enzyme activities, in which the *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) produced the maximum ADH and PDC activities of 168 and 194 nmol/min/mg respectively.

Contribution of Cofactors in Ethanol Synthesis

To analyze the actual requirement of Mg^{+2} and Zn^{+2} for cell growth and ethanol production, the $MgSO_4{\cdot}7H_2O$ and $ZnSO_4{\cdot}7H_2O$ were supplemented by varying their concentrations. Under all Mg^{+2} concentrations tested, the growth of the wild type was highest followed by that of *adh-pdc* overexpressing strain (▲APX) and *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) respectively (**Figure 5A**). Nevertheless, the growth of all three strains was improved with an increase in the concentration of Mg^{+2} (**Figure 5A**). At 400 μM of Mg^{+2}, the *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) showed maximum ethanol concentration of 2643 mg/L (**Figure 5B**) with the cell count of 7.82×10^8 cells/ml (**Figure 5A**). Similar to the effect by Mg^{+2}, the Zn^{+2} also increased ethanol production upon an increase in the concentration of Zn^{+2} and the maximum ethanol concentration of 2943 mg/L was observed at 2 μM of Zn^{+2} with the cell count of 8.89×10^8 cells/ml in *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) (**Figures 5C, D**). The optimization of metals ($MgSO_4{\cdot}7H_2O$ and $ZnSO_4{\cdot}7H_2O$) concentrations together improved the ethanol and biomass production up to 1.43 and 1.48 folds, respectively, compared to that with normal BG-11 medium without metal supplementation.

The TPP and $NADP^+$ act as co-factors for PDC and ADH, respectively. Moreover, TPP is an essential cofactor for various enzymes such as transketolase, α-ketoglutarate dehydrogenase, pyruvate dehydrogenase, and α-keto acid dehydrogenase. The catalytic reactions mediated by these enzymes are the major source of ATP, NADPH, and ribose-5-phosphate. As a result of an increase in overall cellular energy, the biomass contents of the wild type and the engineered strains were increased upon TPP addition. The wild type strain had increased biomass with the supplementation of TPP, while the ethanol synthesis pathway engineered strains showed reduced biomass, even in the absence of glycogen synthesis (**Figure 6A**). The maximum biomass of 11.4×10^8 cells/ml was observed in the wild type at 150 μM TPP. The ethanol production was also increased by TPP significantly and reached the maximum at 150 μM TPP (**Figure 6B**). On the other hand, the supplementation of $NADP^+$ had beneficial effect on *S. elongatus* growth, as it is a major source of ATP synthesis (**Figure 6C**). The maximum ethanol concentration of 3857 mg/L was observed at 2 μM $NADP^+$, which is 1.3-fold and 1.9-fold higher than that in metal supplemented medium and normal BG-11 medium, respectively (**Figure 6D**).

Exopolysaccharides Accumulation in Response to Pathway Engineering and Cofactor Supplementation

As can be seen in **Figures 7A–C**, the biomass, exopolysaccharides, and ethanol concentration was varied upon cofactor supplementation in BG11 medium. The production pattern of exopolysaccharides and ethanol was very similar to each other. Although the wild type produced the exopolysaccharides at detectable level (25.5 μg/10^8 cells), the concentration of exopolysaccharides in engineered strain ▲APX was considerably improved (45.7 μg/10^8 cells). On the other hand, the *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) improved the exopolysaccharides content further to 63.9 μg/10^8 cells (**Figure 7B**). The major components of exopolysaccharides were analyzed to measure the alternative sugar monomer produced in place of the glycogen (**Table 2**). As expected, the contents of glucose, galactose, xylose, and mannose in exopolysaccharides were increased upon engineering the strain. When comparing the *adh-pdc* overexpressing strain (▲APX) with *adh-pdc* overexpressing

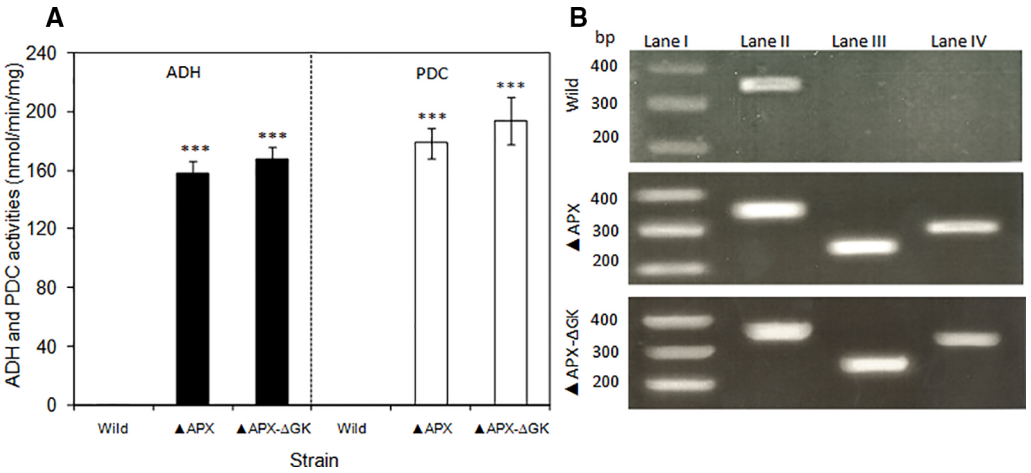

FIGURE 4 | (A) Activities of pyruvate decarboxylase and alcohol dehydrogenase of wild type and engineered *S. elongatus* and **(B)** electrophoresis of RT-PCR products of wild type and engineered *S. elongatus*. Lane I: gene marker, lane II: 16s, lane III: *pdc*, lane IV: *adh*. (conditions: BG-11 medium; 3 days for RT-PCR, 20 days for enzyme activity). Values are mean ± standard deviation (n = 3), statistical significant values (compared to wild type) are indicated by *** for $p < 0.001$.

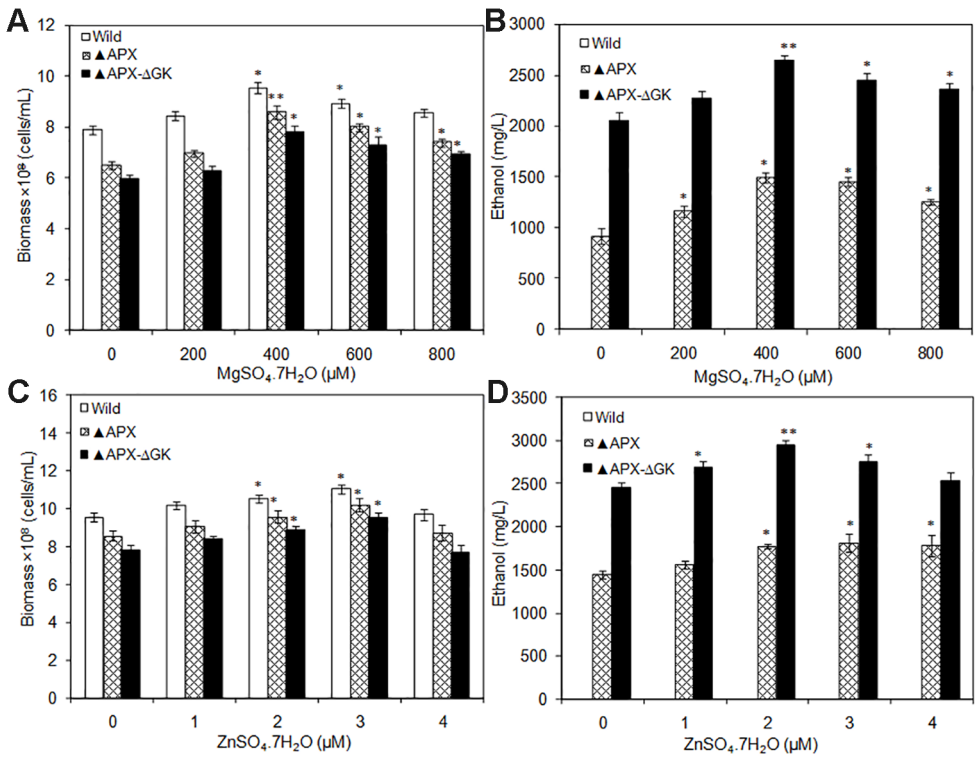

FIGURE 5 | (A, B) Influence of $MgSO_4.7H_2O$ addition on biomass and ethanol production (conditions: 20 days; BG-11 medium) and **(C, D)** Influence of $ZnSO_4.7H_2O$ addition on biomass and ethanol production (conditions: 20 days; BG-11 medium supplemented with 400 μM $MgSO_4.7H_2O$). Values are mean ± standard deviation (n = 3), statistical significant values (compared to no supplementation) are indicated by * for $p < 0.05$; ** for $p < 0.01$.

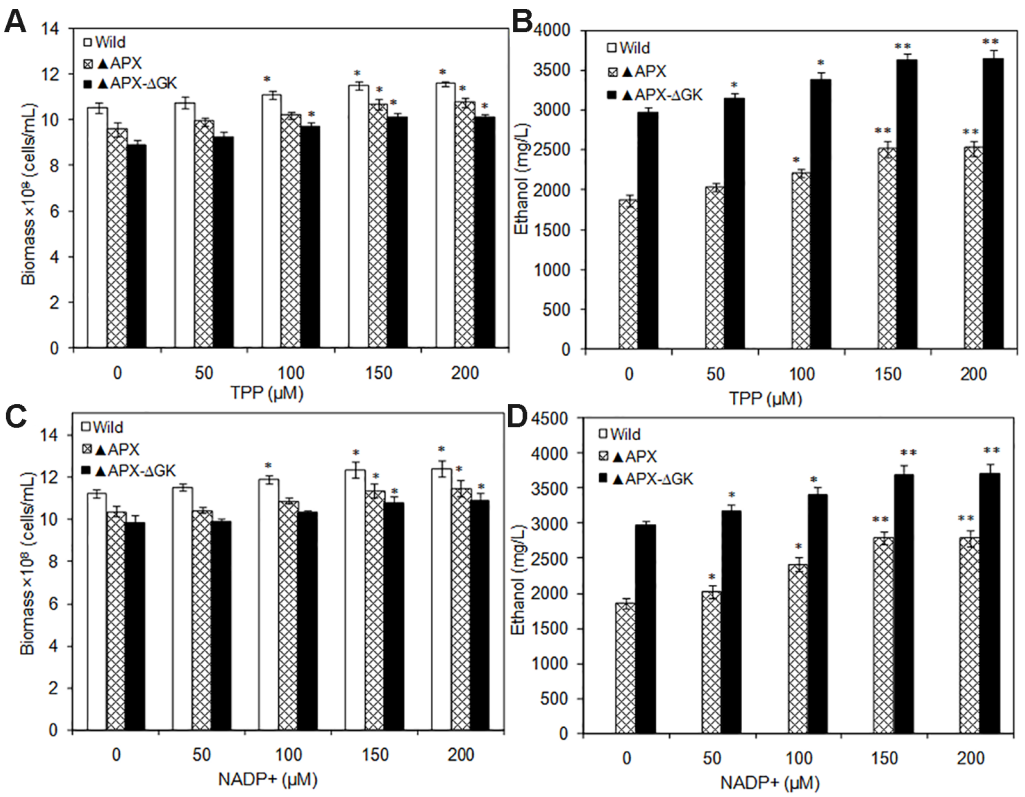

FIGURE 6 | (A, B) Influence of TPP addition on biomass and ethanol production (conditions: 20 days; BG-11 medium supplemented with 400 μM $MgSO_4.7H_2O$; 2 μM $ZnSO_4.7H_2O$), **(C, D)** Influence of $NADP^+$ addition on biomass and ethanol production (conditions: 20 days; BG-11 medium supplemented with 400 μM $MgSO_4.7H_2O$; 2 μM $ZnSO_4.7H_2O$ and 150 μM TPP). Values are mean ± standard deviation (n = 3), statistical significant values (compared to no supplementation) are indicated by * for $p < 0.05$; ** for $p < 0.01$.

strain containing *glgC* knockout (▲APX-ΔGK), the concentration of these sugars was higher in glycogen synthesis knockout strain which confirm the redirection of carbon sources from glycogen synthesis pathway to exopolysaccharides synthesis (**Table 2**). The contents were increased further upon the addition of co-factors. The maximum exopolysaccharides content of 109.5 μg/10^8 cells was observed in Zn^{+2} supplemented *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK).

DISCUSSION

Influence of Pathway Engineering in *S. elongatus*

The reduction in cell growth upon engineering of ethanol synthesis pathway in *S. elongatus* (▲APX) with amore reduction of growth when further engineered with glycogen synthesis knockout (▲APX-ΔGK) suggested the negative impact on cell growth upon engineering of the strain; moreover, it also highlighted the important role of glycogen in the cell growth (**Figure 2A**). The improvement in lipid and protein contents in the engineered strains occurred in concomitant with the decrease of glycogen content, suggesting the possible break down of glycogen as well as the re-direction of carbon flow towards the synthesis of lipid and protein rather than toward glycogen synthesis. The huge storage of carbon sources like glycogen in glycogen synthesis knockout strain makes it possible to provide the carbon substrate in the ethanol synthesis pathway, even though the glycogen synthesis pathway and ethanol synthesis pathway seem unlikely to be competitive. The increase in acetate (source of lipid biosynthesis), succinate (TCA cycle intermediate), pyruvate (source of ethanol synthesis pathway) concentration in *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK) also confirms the redirection of carbon sources towards ethanol synthesis (**Figure 3A**). On the other hand, the extracellular concentrations of these metabolites were also higher in *adh-pdc* overexpressing strain (▲APX) compared to that in *adh-pdc* overexpressing strain containing *glgC* knockout (▲APX-ΔGK), indicating the possible excretion of excess carbon sources by *S. elongatus* (**Figure 3B**). The excretion of succinate, pyruvate, and acetate has already been previously reported for *S. elongatus*

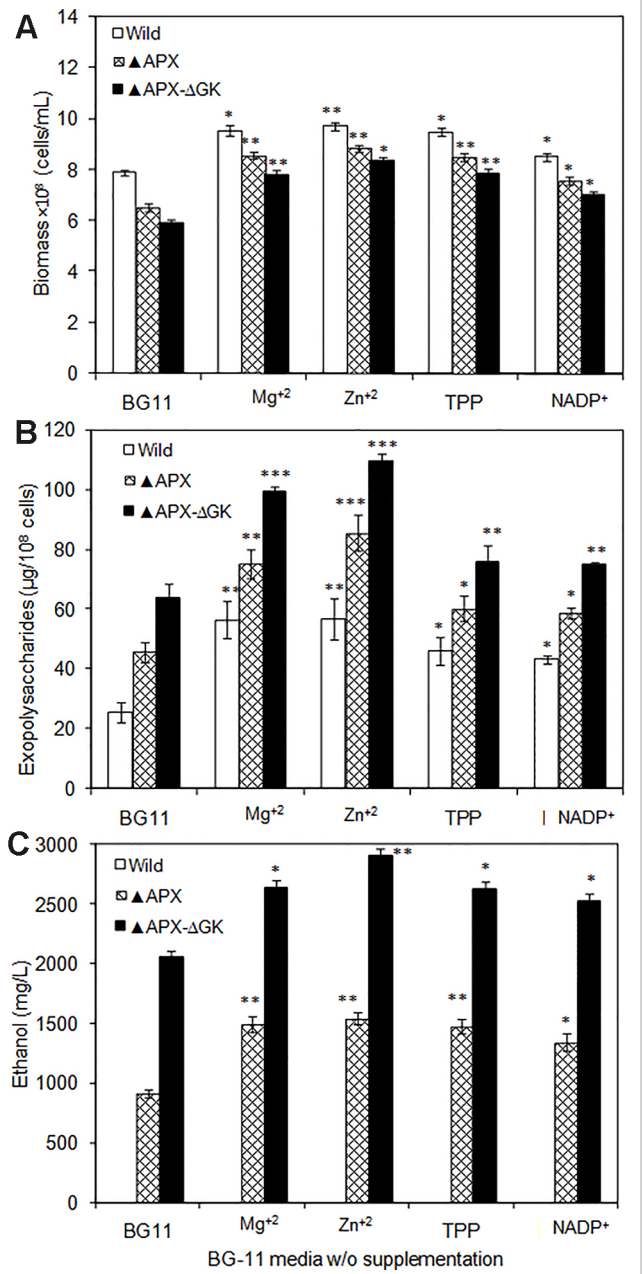

FIGURE 7 | (A) Biomass, **(B)** exopolysaccharides, and **(C)** ethanol production in wild type and engineered *S. elongatus* under various conditions (conditions: 20 days; BG-11 medium or BG-11 medium supplemented with 400 µM $MgSO_4 \cdot 7H_2O$ or 2 µM $ZnSO_4 \cdot 7H_2O$ or 150 µM TPP or 150 µM $NADP^+$). Values are mean ± standard deviation (n = 3), statistical significant values (compared to no supplementation) are indicated by * for $p < 0.05$; ** for $p < 0.01$; *** for $p < 0.001$.

(Hickman et al., 2013). The improvement in extracellular glucose, pyruvate, and succinate concentrations indicated the overproduction due to the lack of glycogen synthesis in the *glgC* knockout strain.

Cofactor Supplementation Improves Ethanol Production

S. elongatus has been engineered to produce ethanol, as it does not carry an effective natural ethanol synthesis pathway. The production of ethanol from *S. elongatus* has been initiated by the incorporation of *pdc* and *adh* from *Zymomonas mobilis* and produced maximum ethanol concentration of 5 mM (≈0.23 g/L) (Deng and Coleman, 1999). Recently, Kopka et al. (2017) engineered *pdc* from *Zymomonas mobilis* and *adh* from *Synechocystis* sp. PCC 6803 in *Synechococcus* sp. PCC 7002 and observed the production of 0.25% (v/v) ethanol. In general, Mg^{+2} and Zn^{+2} are the cofactors for PDC and ADH, respectively. The incorporation of respective genes (*pdc* and *adh*) into a new host might disturb the availability of cofactors for enzymes in ethanol synthesis pathway. Noticeably, both biomass and ethanol production were increased upon $MgSO_4 \cdot 7H_2O$ and $ZnSO_4 \cdot 7H_2O$ supplementation up to a certain concentration, which signify the requirement of more cofactors for ethanol synthesis pathway engineered strain. The results clearly indicate that both cofactors are important requirements of ethanol synthesis pathway, in which the Mg^{+2} is more effective for cell growth and ethanol production. It is noted that the wild type showed no ethanol production either with or without Mg^{+2} and Zn^{+2} addition (**Figures 5B, D**).

Similarly, the co-factors such as TPP and $NADP^+$ supplementation improved the biomass and ethanol production. This improvement is due to the contribution of those cofactors for ethanol production as well as other cellular functions. Recent computational modeling studies demonstrated that the intracellular ATP and NADPH concentrations could enhance the biofuel production (Shabestary and Hudson, 2016). In particular, the decrease in ATP to NADPH ratio was predicted to enhance the ethanol production (Shabestary and Hudson, 2016). According to this hypothesis, the increase in NADPH concentration can automatically decrease the ATP to NADPH ratio, which can further increase the ethanol production. Choi and Park (2016) increased the NADPH concentration by overexpressing *Zwf* gene (encoding glucose-6-phosphate dehydrogenase) in *Synechocystis* sp. PCC 6803 and observed an enhancement in ethanol production up to 0.59 g/L under autotrophic condition. Apart from pyruvate to ethanol conversion steps, NADPH is also reported to improve the Calvin cycle, TCA cycle, and acetyl CoA formation through oxidoreduction reactions in *Synechocystis* sp. PCC 6803 (Hasunuma et al., 2014); thereby improving the ethanol production. However, as the NADPH dependent alcohol dehydrogenase gene is heterologously expressed in *S. elongatus*, the demand for NADPH might be increased upon the insertion of a particular gene. Obviously, the experimental results also showed that the $NADP^+$ addition increases ethanol production to 1345 mg/L (**Figure 7**), which is higher than that of the NADPH engineered strain (0.59 g/L) (Choi and Park, 2016). In conclusion, the cofactors TPP and $NADP^+$ are necessary factors for the improvement of ethanol synthesis in engineered *S. elongatus*. The extracellular addition of such cofactors are not economically feasible; however, the

TABLE 2 | Composition of exopolysaccharides in wild type and engineered strains.

Strain	Glucose (μg/ml)	Galactose (μg/ml)	Xylose (μg/ml)	Mannose (μg/ml)	Arabinose (μg/ml)	Uronic acid (μg/ml)
			BG-11			
Wild type	30 ± 1.6	42 ± 1.5	20 ± 2.2	27 ± 0.6	54 ± 2.8	29 ± 0.3
▲APX	58 ± 2.3***	54 ± 6.1***	37 ± 3.3***	40 ± 2.5***	65 ± 1.3***	43 ± 1.5***
▲APX-ΔGK	87 ± 6.4***	81 ± 2.1***	58 ± 0.9***	47 ± 2.9***	76 ± 2.8***	30 ± 1.0
			Mg^{+2}			
Wild type	103 ± 3.8	109 ± 3.4	59 ± 1.7	76 ± 4.9	106 ± 3.7	81 ± 3.5
▲APX	105 ± 4.7	104 ± 2.7	72 ± 2.6***	114 ± 6.1***	135 ± 5.1***	109 ± 4.8***
▲APX-ΔGK	142 ± 2.4***	123 ± 1.1**	93 ± 4.1***	125 ± 4.0***	187 ± 4.6***	104 ± 2.6***
			Zn^{+2}			
Wild type	85 ± 2.5	78 ± 5.2	56 ± 3.3	71 ± 3.5	112 ± 3.9	83 ± 4.1
▲APX	106 ± 6.7***	113 ± 3.2***	93 ± 4.1***	119 ± 5.0***	122 ± 3.8**	95 ± 3.5*
▲APX-ΔGK	136 ± 4.3***	146 ± 4.8***	106 ± 4.6***	131 ± 4.1***	149 ± 4.2***	101 ± 4.0**
			TPP			
Wild type	77 ± 3.9	84 ± 2.5	50 ± 1.3	68 ± 2.9	121 ± 5.7	76 ± 3.6
▲APX	97 ± 4.7**	100 ± 4.1**	87 ± 2.8***	88 ± 3.3**	119 ± 3.8	76 ± 2.7
▲APX-ΔGK	145 ± 2.2***	143 ± 5.9***	115 ± 3.9***	85 ± 3.4**	135 ± 3.3*	58 ± 1.4*
			$NADP^{+}$			
Wild type	71 ± 4.1	81 ± 3.6	51 ± 1.4	62 ± 3.8	95 ± 5.3	60 ± 1.1
▲APX	90 ± 3.0**	93 ± 2.4*	67 ± 4.6*	83 ± 1.4***	104 ± 4.1	82 ± 2.9**
▲APX-ΔGK	128 ± 5.8***	126 ± 4.0***	106 ± 5.1***	88 ± 2.1***	121 ± 5.5**	59 ± 0.9

*Values are mean ± standard deviation (n = 3), statistical significant values (compared to wild type) are indicated by * for $p < 0.05$; ** for $p < 0.01$; *** for $p < 0.001$.*

overexpression of particular genes involved in TPP and $NADP^{+}$ synthesis can lead to an improved regulatory mechanism for re-directing the primary carbon sources into ethanol synthesis pathway.

Pathway Engineering and Cofactor Supplementation Induces Accumulation of Exopolysaccharides

In cyanobacteria, the exopolysaccharides are accumulated under various stress conditions as a putative physical protective mechanism of the cell (DePhilippis and Vincenzini, 1998; Li et al., 2001; Zeyons et al., 2009; Planchon et al., 2013). The present study is focused on exopolysaccharides production in response to engineering of ethanol synthesis pathway and cofactor supplementation. As presented in the results, the insertion of ethanol synthesis pathway improved exopolysaccharides production, which is due to the formation of a cover-shield in response to the stress condition imposed by ethanol synthesis. In addition, the glycogen synthesis pathway destructed strain showed further improvement in exopolysaccharides content, which might be due to the functioning of exopolysaccharides as an alternative to the glycogen. On the other hand, it is also possible that there is the diversion of most of the carbon sources of glycogen pathway to other relatively close pathways, such as xylose, arabinose, and galactose synthesis. The presence of uronic acids and pentoses (xylose, arabinose, and ribose) are the peculiar components of cyanobacteria which makes them negatively charged. This negatively charged characteristic of cyanobacterial surface usually shows a high affinity for metal cations and other positively charged molecules (DePhilippis and Vincenzini, 1998). Additionally, the presence of ribose, fructose, galactosamine, glucosamine, and in some cases, *N*-acetyl glucosamine, 2,3-*O*-methyl rhamnose, and 3-*O*-methyl glucose have been reported (De Philippis et al., 2011).

Irrespective of the strains used, the concentration of sugars were varied in which glucose, galactose, and xylose were in high concentration followed by mannose, arabinose, and uronic acid. Similar pattern was reported for *Anacystis nidulans* by Sangar and Dugan (1972) in which the glucose, galactose, and mannose were in the ratio of 60:14:20. On the other hand, in terms of production, the results of the present study showed lower exopolysaccharides content than that of *Synechocystis* sp. PCC 6803 (36 pg/cell) (Panoff et al., 1998), However, the results showing an increase in exopolysaccharides content upon cofactor supplementation (**Figure 7**) would be beneficial to further stimulate the exopolysaccharides production in *Synechocystis* sp. PCC 6803. The increase in exopolysaccharides along with ethanol concentration clearly indicates that the exopolysaccharides were accumulated in response to ethanol stress condition, which is an added advantage in biorefinery approach.

CONCLUSION

An integrative expression of a pyruvate decarboxylase *pdc* from *S. cerevisiae* and an alcohol dehydrogenase *adh* from *Synechocystis* worked well in *S. elongatus*. This *S. elongatus* overexpressing both *adh* and *pdc* showed drastic improvement in ethanol synthesis when its glycogen synthesis pathway was knocked out. The supplementation of co-factors such as Mg^{+2}, Zn^{+2}, TPP, and $NADP^{+}$ improved further the ethanol production. As a result of engineering, the engineered cells undergo various metabolic changes to cope with the changes in metabolic flux. On the other hand, as a stress induced cellular response, the engineered *S. elongatus* increased the content of exopolysaccharides. Altogether, the engineered *S. elongatus* produced the maximum ethanol yield of 3856 mg/L. In conclusion, the glycogen deficient *S. elongatus* produced exopolysaccharides as a stress response even at a low ethanol

concentration. Additionally, the results clearly showed that the improvement of co-factors availability can further promote the ethanol production in *S. elongatus*.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

RV and AI designed this research and wrote the manuscript. RV performed all experiments.

FUNDING

RV is thankful to the Graduate School and Faculty of Science, Chulalongkorn University (CU), for senior post-doctoral fellowship from Rachadaphiseksomphot Endowment Fund. AI acknowledges the research grants from CU on the Frontier Research Energy Cluster (CU-59-048-EN) and from Thailand Research Fund (IRG 5780008).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00074/full#supplementary-material

REFERENCES

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1006/abio.1976.9999

Chamovitz, D., Sandmann, G., and Hirschberg, J. (1993). Molecular and biochemical characterization of herbicide-resistant mutants of cyanobacteria reveals that phytoene desaturation is a rate-limiting step in carotenoid biosynthesis. *J. Biol. Chem.* 268, 17348–17353.

Choi, Y. N., and Park, J. M. (2016). Enhancing biomass and ethanol production by increasing NADPH production in *Synechocystis*sp.PCC 6803. *Bioresour. Technol.* 213, 54–57. doi: 10.1016/j.biortech.2016.02.056

De Philippis, R., Colica, G., and Micheletti, E. (2011). Exopolysaccharide-producing cyanobacteria in heavy metal removal from water: molecular basis and practical applicability of the biosorption process. *Appl. Microbiol. Biotechnol.* 92, 697–708. doi: 10.1007/s00253-011-3601-z

Deng, M. D., and Coleman, J. R. (1999). Ethanol synthesis by genetic engineering in cyanobacteria. *Appl. Environ. Microbiol.* 65, 523–528. doi: 10.1128/AEM.65.2.523-528.1999

DePhilippis, R., and Vincenzini, M. (1998). Exocellular polysaccharides from cyanobacteria and their possible applications. *FEMS Microbiol. Rev.* 22, 151–175. doi: 10.1111/j.1574-6976.1998.tb00365.x

Dexter, J., and Fu, P. (2009). Metabolic engineering of cyanobacteria for ethanol production. *Energy Environ. Sci.* 2, 857–864. doi: 10.1039/b811937f

Gao, Z., Zhao, H., Li, Z., Tana, X., and Lu, X. (2012). Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy Environ. Sci.* 5, 9857. doi: 10.1039/C2EE22675H

Hasunuma, T., Matsuda, M., Senga, Y., Aikawa, S., Toyoshima, M., Shimakawa, G., et al. (2014). Overexpression of flv3 improves photosynthesis in the cyanobacterium *Synechocystis* sp. PCC6803 by enhancement of alternative electron flow. *Biotechnol. Biofuels* 7, 493. doi: 10.1186/s13068-014-0183-x

Hickman, J. W., Kotovic, K. M., Miller, C., Warrener, P., Kaiser, B., Jurista, T., et al. (2013). Glycogen synthesis is a required component of the nitrogen stress response in *Synechococcus elongatus* PCC 7942. *Algal Res.* 2, 98–106. doi: 10.1016/j.algal.2013.01.008

Jantaro, S., Mulo, P., Jansen, T., Incharoensakdi, A., and Maenpaa, P. (2005). Effects of long-term ionic and osmotic stress conditions on photosynthesis in the cyanobacterium *Synechocystis* sp. PCC 6803. *Funct. Plant Biol.* 32, 807–815. doi: 10.1071/FP04219

Kanwal, S., Rastogi, R. P., and Incharoensakdi, A. (2014). Glutamate decarboxylase activity and gamma-aminobutyric acid content in *Synechocystis* sp. PCC 6803 under osmotic stress and different carbon sources. *J. Appl. Phycol.* 26, 2327–2333. doi: 10.1007/s10811-014-0259-9

Kanwal, S., Khetkorn, W., and Incharoensakdi, A. (2015). GABA accumulation in response to different nitrogenous compounds in unicellular cyanobacterium *Synechocystis* sp. PCC 6803. *Curr. Microbiol.* 70, 96–102. doi: 10.1007/s00284-014-0687-4

Kopka, J., Schmidt, S., Dethloff, F., Pade, N., Berendt, S., and Schottkowski, M. (2017). Systems analysis of ethanol production in the genetically engineered cyanobacterium*Synechococcus* sp. PCC 7002. *Biotechnol. Biofuels* 10, 56. doi: 10.1186/s13068-017-0741-0

Li, P., Harding, S. E., and Liu, Z. (2001). Cyanobacterial exopolysaccharides: their nature and potential biotechnological applications. *Biotechnol. Genet. Eng. Rev.* 18, 375–404. doi: 10.1080/02648725.2001.10648020

Mofijur, M., Rasul, M. G., Hyde, J., Azad, A. K., Mamat, R., and Bhuiya, M. M. K. (2016). Role of biofuel and their binary (diesel–biodiesel) and ternary (ethanol–biodiesel–diesel) blends on internal combustion engines emission reduction. *Renew. Sust. Energ. Rev.* 53, 265–278. doi: 10.1016/j.rser.2015.08.046

Monshupanee, T., and Incharoensakdi, A. (2014). Enhanced accumulation of glycogen, lipids and polyhydroxybutyrate under optimal nutrients and light intensities in the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Appl. Microbiol.* 116, 830–838. doi: 10.1111/jam.12409

Moran, R. (1982). Formulae for determination of chlorophyllous pigments extracted with N,N-dimethylformamide. *Plant Physiol.* 69, 1376–1381. doi: 10.1104/pp.69.6.1376

Norena-Caro, D., and Benton, M. G. (2018). Cyanobacteria as photoautotrophic biofactories of high-value chemicals. *J. CO_2 Util.* 28, 335–366. doi: 10.1016/j.jcou.2018.10.008

Nozzi, N. E., Oliver, J. W. K., and Atsumi, S. (2013). Cyanobacteria as a platform for biofuel production. *Front. Bioeng. Biotechnol.* 1, 7. doi: 10.3389/fbioe.2013.00007

Pade, N., Mikkat, S., and Hagemann, M. (2017). Ethanol, glycogen and glucosylglycerol represent competing carbon pools in ethanol-producing cells of *Synechocystis* sp. PCC 6803 under high-salt conditions. *Microbiology* 163, 300–307. doi: 10.1099/mic.0.000433

Panoff, J.-M., Priem, B., Morvan, H., and Joset, F. (1998). Sulphated exopolysaccharides produced by two unicellular strains of cyanobacteria, *Synechocystis* PCC 6803 and 6714. *Arch. Microbiol.* 150, 558– 563. doi: 10.1007/BF00408249

Pereira, S., Zille, A., Micheletti, E., Moradas-Ferreira, P., De Philippis, R., and Tamagnini, P. (2009). Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. *FEMS Microbiol. Rev.* 33, 917–941. doi: 10.1111/j.1574-6976.2009.00183.x

Planchon, M., Jittawuttipoka, T., Cassier-Chauvat, C., Guyot, F., Gelabert, A., and Benedetti, M. F. (2013). Exopolysaccharides protect *Synechocystis* against the deleterious effects of titanium dioxide nanoparticles in natural and artificial waters. *J. Colloid Interface Sci.* 405, 35–43. doi: 10.1016/j.jcis.2013.05.061

Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., and Stanier, R. Y. (1979). Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J. Gen. Microbiol.* 111, 1–61. doi: 10.1099/00221287-111-1-1

Sangar, V. K., and Dugan, P. R. (1972). Polysaccharide produced by *Anacystis nidulans*: its ecological implication. *Appl. Microbiol.* 24, 732–734. doi: 10.1128/AEM.24.5.732-734.1972

Shabestary, K., and Hudson, E. P. (2016). Computational metabolic engineering strategies for growth-coupled biofuel production by *Synechocystis*. *Metab. Eng. Commun.* 3, 216–226. doi: 10.1016/j.meteno.2016.07.003

Su, Y., Song, K., Zhang, P., Su, Y., Cheng, J., and Chen, X. (2017). Progress of microalgae biofuel's commercialization. *Renew. Sust. Energ. Rev.* 74, 402–411. doi: 10.1016/j.rser.2016.12.078

Tian, X., Chen, L., Wang, J., Qiao, J., and Zhang, W. (2013). Quantitative proteomics reveals dynamic responses of *Synechocystis* sp. PCC 6803 to next-generation biofuel butanol. *J. Proteome* 78, 326–345. doi: 10.1016/j.jprot.2012.10.002

Velmurugan, R., and Incharoensakdi, A. (2016). Potential of metal oxides in fractionation of Synechocystis sp. PCC 6803 biomass for biofuel production. *Algal Res.* 19, 96–103. doi: 10.1016/j.algal.2016.07.018

Velmurugan, R., and Incharoensakdi, A. (2017). Immobilization of α-amylase on metal nanoparticles mediated by xylan aldehyde improves hydrolysis of glycogen from *Synechocystis* sp. PCC 6803. *Fuel* 210, 334–342. doi: 10.1016/j.fuel.2017.08.073

Velmurugan, R., and Incharoensakdi, A. (2018). Disruption of polyhydroxybutyrate synthesis redirects carbon flow towards glycogen synthesis in *Synechocystis* sp. PCC 6803 overexpressing *glgC*/*glgA*. *Plant Cell Physiol.* 59, 2020–2029. doi: 10.1093/pcp/pcy121

Velmurugan, R., and Incharoensakdi, A. (2020). Co-cultivation of two engineered strains of *Synechocystis* sp. PCC 6803 results in improved bioethanol production. *Renew. Energy* 146, 1124–1133. doi: 10.1016/j.renene.2019.07.025

Zeyons, O., Thill, A., Chauvat, F., Menguy, N., Cassier-Chauvat, C., Orear, ,. C., et al. (2009). Direct and indirect CeO_2 nanoparticles toxicity for *Escherichia coli* and *Synechocystis*. *Nanotoxicol.* 3, 284–295. doi: 10.3109/17435390903305260

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.

frontiers in Plant Science

REVIEW
published: 11 March 2020
doi: 10.3389/fpls.2020.00282

Engineering of Bioenergy Crops: Dominant Genetic Approaches to Improve Polysaccharide Properties and Composition in Biomass

Andrew G. Brandon[1,2,3] *and Henrik V. Scheller*[1,2,3]*

[1] *Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA, United States,* [2] *Feedstocks Division, Joint BioEnergy Institute, Emeryville, CA, United States,* [3] *Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States*

Large-scale, sustainable production of lignocellulosic bioenergy from biomass will depend on a variety of dedicated bioenergy crops. Despite their great genetic diversity, prospective bioenergy crops share many similarities in the polysaccharide composition of their cell walls, and the changes needed to optimize them for conversion are largely universal. Therefore, biomass modification strategies that do not depend on genetic background or require mutant varieties are extremely valuable. Due to their preferential fermentation and conversion by microorganisms downstream, the ideal bioenergy crop should contain a high proportion of C6-sugars in polysaccharides like cellulose, callose, galactan, and mixed-linkage glucans. In addition, the biomass should be reduced in inhibitors of fermentation like pentoses and acetate. Finally, the overall complexity of the plant cell wall should be modified to reduce its recalcitrance to enzymatic deconstruction in ways that do no compromise plant health or come at a yield penalty. This review will focus on progress in the use of a variety of genetically dominant strategies to reach these ideals. Due to the breadth and volume of research in the field of lignin bioengineering, this review will instead focus on approaches to improve polysaccharide component plant biomass. Carbohydrate content can be dramatically increased by transgenic overexpression of enzymes involved in cell wall polysaccharide biosynthesis. Additionally, the recalcitrance of the cell wall can be reduced via the overexpression of native or non-native carbohydrate active enzymes like glycosyl hydrolases or carbohydrate esterases. Some research in this area has focused on engineering plants that accumulate cell wall-degrading enzymes that are sequestered to organelles or only active at very high temperatures. The rationale being that, in order to avoid potential negative effects of cell wall modification during plant growth, the enzymes could be activated post-harvest, and post-maturation of the cell wall. A potentially significant limitation of this approach is that at harvest, the cell wall is heavily lignified, making the substrates for these enzymes inaccessible and their activity ineffective. Therefore, this review will only include research employing enzymes that are at least partially active under the ambient conditions of plant growth and cell wall development.

Keywords: lignocellulosic biomass, dedicated bioenergy crops, genetic engineering, cellulose, hemicellulose, cell walls, carbohydrate active enzymes, polysaccharides

OPEN ACCESS

Edited by:
Jin Zhang,
Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Mitra Mazarei,
The University of Tennessee, Knoxville, United States
Ajaya K. Biswal,
University of Georgia, United States

***Correspondence:**
Henrik V. Scheller
hscheller@lbl.gov

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *12 December 2019*
Accepted: *25 February 2020*
Published: *11 March 2020*

Citation:
Brandon AG and Scheller HV (2020) Engineering of Bioenergy Crops: Dominant Genetic Approaches to Improve Polysaccharide Properties and Composition in Biomass. Front. Plant Sci. 11:282. doi: 10.3389/fpls.2020.00282

INTRODUCTION

Lignocellulosic plant biomass represents the largest renewable source of organic carbon on earth. Organic carbon that can be converted into a wide variety of compounds, including high-energy liquid fuel, thereby curbing our dependence on non-renewable sources and limiting the net production of carbon dioxide. The bulk of plant biomass is contained in the cell wall, specifically the thick secondary cell walls (SCW) of the vasculature and fiber tissues. The cell wall has evolved a highly complex and rigid structure to resist the mechanical forces of growth and protect the plant from various stresses. The most valuable component of the cell wall, from a bioenergy perspective, is the 6-carbon sugar glucose comprising the linear polysaccharide cellulose. Produced and extruded into the developing cell wall by cellulose synthase (CesA) complexes at the plasma membrane, individual cellulose chains coalesce through hydrogen bonding to form crystalline cellulose microfibrils. In all vascular plants, the SCW is composed of cellulose microfibrils embedded in a matrix of the aromatic polymer lignin. Hemicellulose like xylan coat the cellulose microfibrils and can form covalent linkages with cell wall proteins, lignin, and other hemicelluloses (Gírio et al., 2010; Scheller and Ulvskov, 2010; Meents et al., 2018). This natural complexity makes the deconstruction and recovery of usable sugars costly and resource intensive.

The prevailing sources of biofuels to date have been sucrose from crops like sugarcane or sugar beet and starch from corn. While much simpler and cheaper to process, their sustainability at a larger scale is dubious considering all are also major food and forage crops. The "food vs. fuel" competition for arable land could drive up the price of food and have negative socioeconomic impacts. Therefore, modern approaches to sustainable bioenergy emphasize the development of dedicated bioenergy crops that can be grown on marginal land (Himmel and Bayer, 2009; Cai et al., 2011). Ideally, these crops will be fast-growing perennials, producing the maximum biomass per unit land over multi-year cycles and minimizing nutrient input needs (Sanderson and Adler, 2008). The additional constraints of varying soil quality, water availability, and average temperature mean biomass productivity will vary depending on where the crops are grown and that no single engineered species can meet bioenergy demands around the world (Somerville et al., 2010; Chen and Peng, 2013). For example, in tropical and sub-tropical climates, elephant or Napier grass (*Pennisetum purpureum*) produces more biomass per hectare annually than any other vegetation. Grasses like *Miscanthus* × *giganteus*, sorghum, and switchgrass (*Panicum virgatum*), and tree species like poplar, aspen, and willow are capable of producing large amounts of biomass in temperate regions like Europe and the United States (Heaton et al., 2008; Guidi et al., 2013; Junior et al., 2016). While the massive bioethanol productivity of Brazil comes from sugarcane sucrose, the lignocellulosic biomass left over could additionally be engineered for better bioenergy conversion. Crassulacean acid metabolism plants like *Agave* spp. have the highest water use efficiency of all plants, making them attractive bioenergy feedstock crops for cultivation on the increasing percentage of the world's land area considered arid (<800 mm of rainfall per year) (United Nations Environment Programme, 2007; Borland et al., 2009).

The wide variety of potential feedstock species means efforts to improve their quality through biotechnology should be as broadly useful as possible. Research into the fundamental cell biology of plants and the organisms that degrade them has revealed the causes of biomass recalcitrance and a variety of approaches to reduce it. Significant improvements in biomass have been demonstrated by suppressing or eliminating the expression of various genes related to the biosynthesis of specific polysaccharides (Biswal et al., 2015, 2018; Loqué et al., 2015; Kalluri et al., 2016; Wang et al., 2016; Willis et al., 2016; Bhatia et al., 2017; Donev et al., 2018; Li et al., 2019b). However, the usefulness of the same approaches in a range of bioenergy crops can be limited by their comparative genetic complexity. Therefore, this review will focus on cell wall polysaccharide engineering strategies that act independent of genetic or genomic context, primarily via the overexpression of recombinant or native enzymes. Importantly, we include only enzymes that are active *in planta* concurrent with cell wall development. The engineering of plants to accumulate hyperthermophilic or inactive cell wall-degrading enzymes has been well reviewed in other works (Mir et al., 2014; Damm et al., 2016; Park et al., 2016).

MODULATION OF POLYSACCHARIDE BIOSYNTHESIS

Cellulose

Two major features of an "ideal" bioenergy crop are high cellulose content and a high ratio of C6:C5 sugar residues comprising the polysaccharide content. Increasing cellulose biosynthesis is an important goal for engineering efforts because cellulose consists entirely of the C6 sugar glucose and the genes involved in its biosynthesis are relatively well-understood and conserved among land plants. Secondary cell wall (SCW) cellulose, which accounts for the bulk of cellulosic biomass in bioenergy-relevant crops, is synthesized at the plasma membrane by three, non-redundant cellulose synthase (CesA) proteins CesA4, CesA7, and CesA8 (McFarlane et al., 2014). Overexpression of CesAs is, therefore, a logical approach to generating transgenic plants enriched in cellulose. However, attempts to overexpress SCW CesAs in aspen and barley resulted in co-suppression and decreased cellulose content (Joshi et al., 2011; Tan et al., 2015). Overexpression of SCW CesA4 and CesA6 driven by a maize ubiquitin promoter in switchgrass also resulted in decreased cellulose, even though co-suppression did not occur (Mazarei et al., 2018). See **Table 1** for a summary of studies discussed in this review. The reason for the decrease in cellulose in the study by Mazarei and coworkers is unclear, but the plants exhibited reduced growth. The study would suggest that careful consideration of promoters may be critical for the success of overexpressing SCW CesAs in crops. Greater success has been demonstrated in Arabidopsis (*Arabidopsis thaliana*) by overexpression of either of the primary cell wall (PCW) CesAs, *CesA2*, *CesA5* and *CesA6* (Hu et al., 2018). CesA2, CesA5, and CesA6 are each functional in a PCW cellulose

TABLE 1 | Summary of dominant approaches to modify biomass polysaccharide composition.

Engineered species	Transgene expressed	Effects on biomass composition and conversion	References
A. thaliana	*AtCesA2*, *AtCesA5*, *AtCesA6*	(+) 29–37% Cel	Hu et al., 2018
	GhCOBL9A	(+) 59% Cel	Niu et al., 2018
	AtGalS1::AtUGE2::AtURGT1	(+) 80% Gal	Gondolf et al., 2014; Aznar et al., 2018
	pSAG12:OsCSLF6	(+) non-Cellulosic Glc	Vega-Sánchez et al., 2015
	AtIRX10^{G283D}, *AtIRX10*^{E293Q}	(–) 39–55% Xyl	Brandon et al., 2019
	AtSBD123	(+) 76% FW; (+) 50% HC; (+) 30% Pec; (+) 100% non-crystalline Cel; (+) 28% IVD	Grisolia et al., 2017
	pEST:PcPL1	(+) 90–100% SE, post-induction	Tomassetti et al., 2015
	pSAG12:AnPGA2	(+) 50–100% SE, post-senescence	Tomassetti et al., 2015
	AtPMEI-2	(+) 50% SE; (+) 68% DW	Lionetti et al., 2010
Populus spp.	*GhSuSy*	(+) 2–6% Cel; (+) CrI	Coleman et al., 2009
	PdDUF266A	(+) 17–34% DW; (+) 37% Cel; (+) 13% Cel DP; (+) 38% SE	Yang et al., 2017a
	AnAXE1	(+) 26% SE	Pawar et al., 2016, 2017
	PdDUF231A	(+) 8–21% Cel; (–) 6–8% lignin	Yang et al., 2017b
	AaXEG2	(+) DW; (+) 81% SE	Kaida et al., 2009; Park et al., 2004
	PtPL1	(+) SE	Biswal et al., 2014
Tobacco	*PsnSuSy1/2*	(+) 18% Cel; (–) 28% lignin	Wei et al., 2015; Li et al., 2019a
	AcCel5	(+) 10–15% SE	Brunecky et al., 2011
	TrCel5	(–) FW; (–) Cel	Klose et al., 2015
	ClEXPA1/2	(+) FW; (+) 30–50% Cel	Wang et al., 2011
	AnPGA2	(+) 100% SE; (–) 50–84% FW	Lionetti et al., 2010
P. virgatum	*PvCesA4*, *PvCesA6*	(–) DW; (–) 6–33% Cel; (+) 2–12% Xyl	Mazarei et al., 2018
	OsAT10	(+) 40% Cel	Li et al., 2018
O. sativa	*OsSuSy3*	(+) 15–26% Cel; (+) 11–13% HC	Fan et al., 2017, 2019
	OsGH9B1, *OsGH9B3*	(+) 63% SE	Huang et al., 2019
	OsARAF1, *OsARAF3*	(–) 19–25% Ara; (–) 28–34% Cel	Sumiyoshi et al., 2013
	OsAT10	(+) 8–19% Cel; (+) 40% SE	Bartley et al., 2013
Saccharum spp.	*CsCESA*	(+) 31% Cel; (+) 28–39% SE; (+) 25% Suc; (+) 56% non-Cel Glc; (+) 22% Gal; (+) 53% GalA	Ndimande, 2013
F. arundinacea	*AnFAE*	(+) 10–14% IVD	de Buanafina et al., 2008, 2010; Morris et al., 2017

Strong, constitutive promoters were used unless otherwise indicated. Cel, cellulose; Ara, arabinose; Glc, glucose; Gal, galactose; GalA, galacturonic acid; Suc, sucrose; Xyl, xylose; HC, hemicellulose; Pec, pectin; FW, fresh weight; DW, dry weight; DP, degree of polymerization; CrI, cellulose crystallinity index; SE, saccharification efficiency; IVD, in vitro digestibility.

synthase complex including CesA1 and CesA3. Transgenic plants overexpressing one of the three genes had 29–37% increase in crystalline cellulose content. Expression of both *CesA1* and *CesA3* was significantly higher in transgenic lines, indicating the possible secretion and activity of PCW CesA complexes even during SCW development. Transgenic lines also had more xylan and a slight, but significant, increase in lignin, potentially counteracting the benefits of the increase in cellulose. However, enzymatic saccharification efficiency of the transgenic biomass was not reported and further work is required to determine the usefulness of PCW *CesA*-overexpression. A clever approach to increase SCW cellulose content that avoids the negative effects of co-suppression has been employed to generate transgenic sugarcane (*Saccharum* spp.). While plants, many bacteria and some fungi can produce cellulose, only one group of animals is known to do so: the marine invertebrates of subphylum *Tunicata*, otherwise known as sea squirts (Matthysse et al., 2004; Kimura and Itoh, 2007). A cellulose synthase cDNA from *Ciona savignyi* (*CsCesA*) was used to create transgenic sugarcane overexpressing a functional form of the protein (Ndimande, 2013). Use of this divergent gene sequence did not cause any co-suppression and the internode cellulose content was increased by up to 31%. Additionally, all tissues of *CsCesA*-overexpressing sugarcane lines had increased saccharification efficiency, with increases of 39% and 28% in young and mature internodes, respectively. Since sucrose is the primary source of bioenergy potential in sugarcane, total soluble sugars were also measured. Transgenic lines yielded up to 25% more, contrary to intuition that sucrose content would be depleted by increased conversion to UDP-glucose (UDP-Glc) for cellulose production. It could be that this depletion acts as a signal to source organs to increase production and/or transport of sucrose to sink organs. It has previously been demonstrated that induced depletion of sucrose from sugarcane stems increases photosynthetic productivity and

phloem loading from leaves (McCormick et al., 2009; Wang et al., 2013). In the non-cellulosic polysaccharide fraction of the cell wall, Ndimande (2013) also observed increased a 56% increase in glucose. 22% increase in galactose and a 53% increase in galacturonic acid. The author attributes the increase in glucose to increased biosynthesis of mixed-linkage glucan, which uses UDP-Glc as its substrate. An increased mixed-linkage glucan deposition driven by higher UDP-Glc has previously been reported in barley starch mutants (Christensen and Scheller, 2012). While sugarcane accumulates large amounts of soluble sucrose, the majority of the carbohydrate mass of the plant is still the cell wall, making it an interesting model for the study of carbon flux to cell wall biosynthesis.

Increased cellulose biosynthesis has been accomplished by overexpressing the enzyme responsible for producing its substrate. Sucrose synthase (SuSy) proteins catalyze the cleavage of sucrose to fructose and UDP-Glc, which is the sole substrate for the biosynthesis of glucans like cellulose, callose and mixed-linkage glucans. Evidence suggests that some SuSy isoforms interact directly with the CesA complex, channeling UDP-Glc directly to cellulose biosynthesis (**Figure 1**; Fujii et al., 2010; Stein and Granot, 2019). Overexpression of cotton (*Gossypium hirsutum*) *SuSy* in hybrid poplar (*Populus alba* × *grandidentata*) resulted in small (2–6%) increases in cellulose and increased cellulose crystallinity (CrI) compared to controls (Coleman et al., 2009). CrI reflects the degree of hydrogen bonding between individual cellulose chains and is a primary contributor to biomass recalcitrance by reducing the proportion of cellulose exposed to cellulolytic enzymes (Hall et al., 2010). Overexpression of the hybrid poplar (*Populus simonii* × *Populus nigra)* gene *PsnSuSy2* in tobacco led to 25% thicker cell walls, containing up to 18% more cellulose and decrease in lignin of up to 28%, when compared to controls. Additionally, the degree of cellulose crystallinity (CrI) was reduced by 9–11% (Wei et al., 2015; Li et al., 2019a). Overexpression of the endogenous *PvSUS1* in switchgrass resulted in up to 14% more biomass, but in contrast to the studies of tobacco and poplar, lignin was increased rather than decreased (Poovaiah et al., 2015). Cellulose crystallinity was not reported, but saccharification was reduced, perhaps due to the increased lignin. A similar approach in transgenic rice (*Oryza sativa*) lines overexpressing *OsSuSy3* under a SCW-specific or a constitutive promoter resulted in dramatic improvements in multiple bioenergy-relevant characteristics (Fan et al., 2017, 2019). The effects on cell wall composition were similar using either promoter. Total plant biomass was only slightly increased in transgenic lines, but microscopic analysis of the cell wall revealed a 68% increase in cell wall thickness compared to controls. The increase in cell wall thickness corresponded to a 15–26% increase in cellulose and 11–13% increase in hemicelluloses, while lignin content was not significantly changed. Cellulose from transgenic plants had a 7–10% reduction in CrI, but an increased cellulose degree of polymerization (DP) by 8–15%. DP, like CrI, is negatively correlated with saccharification efficiency. After pretreatment, transgenic lines yielded 13–23% total sugars, leading to a 20–49% greater ethanol yield and 22% higher conversion efficiency when compared to controls. The 2017 and 2019 studies distinguish themselves by investigating secondary effects of *OsSuSy3* overexpression on two important agronomic traits, respectively: lodging resistance and susceptibility to pathogen or insect attack. Lodging is a complex trait and crop susceptibility to lodging can dramatically reduce yield and increase the cost of harvesting. Lodging Index is a combination of several measurements taken from the plant to determine its physicomechanical strength and resistance to lodging. All four independent transgenic lines overexpressing *OsSuSy3* showed increases of 17–50% in Lodging Index compared to controls, suggesting they could grow robustly and be high-yielding in the field. In the more recent study by Fan and coworkers, *OsSuSy3* transgenic lines were observed to be less susceptible to a number of biotic stresses including bacterial blight, fungal rice blast, and herbivory by the brown planthopper. These resistances correlated with a significant increase in callose deposition upon initiation of infection or pest attack. Since both callose and cellulose are synthesized from UDP-Glc, it is possible that *OsSuSy3* overexpression could also stimulate callose biosynthesis (**Figure 1**). The β-(1,3)-Glc bonds in callose make it much less crystalline than cellulose, thus more amenable to saccharification. It would be intriguing to see if this accumulation of callose phenotype could be exploited, perhaps by triggering an immune response at senescence in order to rapidly accumulate low-recalcitrance callose before harvest. These works demonstrate that overexpression of a single gene involved in carbohydrate flux can simultaneously improve a variety of traits important in bioenergy crops.

In addition to the *CesA* genes, several other genes that are important for, or even critical to, cellulose biosynthesis have been identified based on the phenotypes of their loss-of-function mutants. Their gene products often lack structural similarity to CesA proteins or even to glycosyltransferases in general, so the roles they play in cellulose biosynthesis can be difficult to determine. Despite this, a few studies have demonstrated that the overexpression of such proteins can lead to dramatic improvements to biomass composition and recalcitrance. A specific example is the engineered overexpression of a gene, *Domain of Unknown Function-266A (DUF266A)*, in poplar (*Populus deltoides)* and Arabidopsis (Yang et al., 2017a). DUF266-containing proteins are only present in land plants and are categorized as "non-classifiable GT" (Lao et al., 2014), although they are distantly related to GT family 14. GT14s have been characterized as having a range of activities, from arabinogalactan biosynthesis in plants to protein *O*-glycosylation in animals (Cantarel et al., 2009; Hansen et al., 2012; Knoch et al., 2013). The rice gene *Brittle Culm 10 (OsBC10)* is the only previously characterized gene encoding a DUF266-containing protein. *OsBC10* encodes a Type II transmembrane Golgi protein and loss-of-function mutant plants were dwarfed with reductions in cellulose, and increases in xylose and lignin, indicating a role in cell wall biosynthesis (Zhou et al., 2009). The transgenic poplar lines overexpressing *PdDUF266A* generated by Yang et al. (2017a) had increased total biomass of 17–34% and cellulose content up to 37% greater than wild type. Their cellulose DP was also increased by 13%, although the CrI was not significantly altered. Total sugar release after enzymatic saccharification was

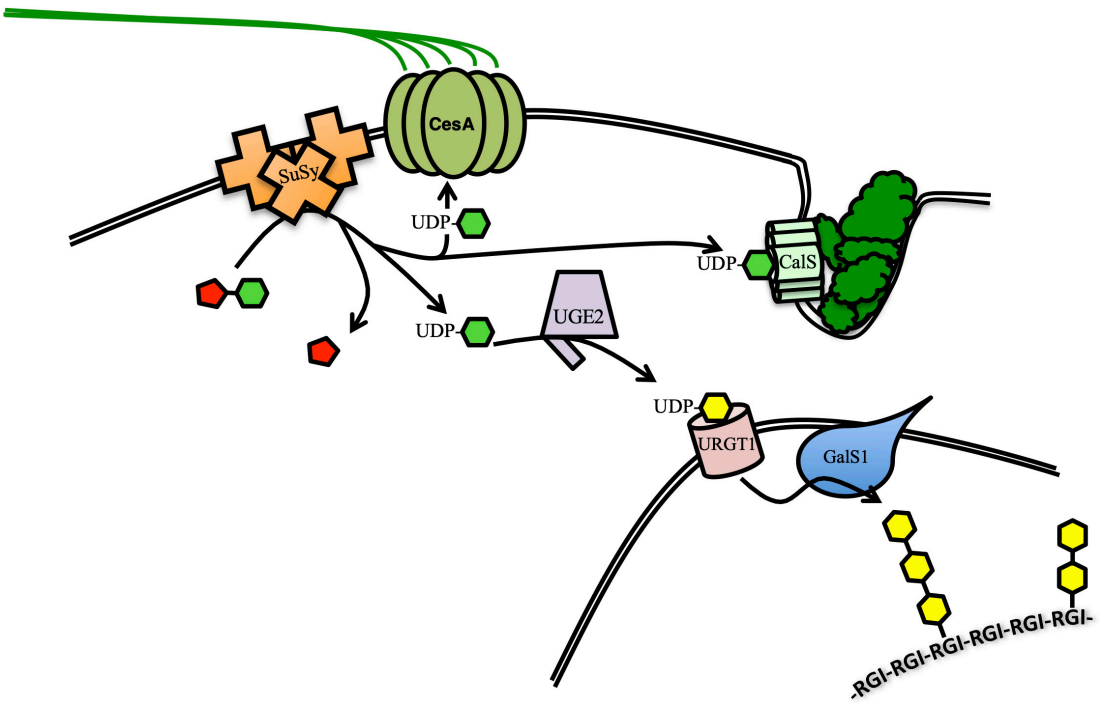

FIGURE 1 | Schematic representation of the role of sucrose synthase (SuSy) in increasing the biosynthesis of several C6-sugar polysaccharides. SuSy isoforms can be localized to the cytoplasm, or tightly associated with the plasma membrane, some interacting directly with CesA. All catalyze the reversible conversion of sucrose into fructose (red) and UDP-Glc (green). UDP-Glc is the substrate for cellulose synthesis by CesA complex or callose synthesis by CalS at the plasma membrane. UDP-Glc can also be converted to UDP-Gal (yellow) by UGE1, imported to the Golgi by URGT1, and used to synthesize the galactan side chains of pectic RGI.

increased in *PdDUF266A*-overexpressors by 38% compared to controls. These effects could be indirect, since several genes involved with SCW cellulose biosynthesis were found to be significantly upregulated. How a protein without a predicted function, residing in the Golgi, is able to alter gene expression is not currently understood. However, there is a more direct effect *PdDUF266A* overexpression may have had on cell wall composition (Yang et al., 2017a). When probing transverse section of the *osbc10* mutant, Zhou et al. (2009) found a dramatic decrease in signal using antibodies specific to arabinogalactan proteins. Compositional analysis revealed *osbc10* mutants to have a 72% reduction in cell wall arabinogalactan proteins compared to controls. While it was not measured in the study, it is possible that overexpression of *PdDUF266A* modified the arabinogalactan protein profile of the cell wall, which could cause alterations in the organization or interaction of various wall polysaccharides. PdDUF266A could also be involved directly in the glycosylation of CesA proteins or other proteins critical to cellulose synthesis like KORRIGAN and COBRA. The latter two proteins play roles in microfibril deposition and require posttranslational *N*-glycosylation to be fully functional (Roudier et al., 2005; Liebminger et al., 2013). Recently, a *COBRA-like* gene from cotton (*Gossypium hirsutum*), *GhCOBL9A*, was overexpressed in Arabidopsis, leading to dramatic increases in total biomass and cellulose content (Niu et al., 2018). The cell walls of cotton fibers consist almost entirely of cellulose and are an interesting model for high-level cellulose production. Of the 33 identified *COBL* genes in the cotton genome, *COBL9A* was expressed highly during SCW development and co-expressed with SCW *CesA* genes. COBL proteins are secreted and have a glycosylphosphatidylinositol anchor to the plasma membrane. COBLs also have a carbohydrate-binding domain (CBM) that preferentially binds to crystalline cellulose. COBLs are required for cellulose production, believed to direct the orderly deposition of nascent cellulose microfibrils but are not components of the cellulose synthase complex (Liu et al., 2013). Niu et al. (2018) found that overexpression of *GhCOBL9A* in Arabidopsis produced plants that grew taller and contained up to 59% more cellulose. The expression of three SCW *CesA* genes, *CesA4*, *CesA7* and *CesA8*, were also dramatically increased, approximately 10-fold higher than controls. Interestingly, examination of transverse stem sections of transgenic plants revealed increased cellulose and wall thickening, not only in fiber and vessel cells, but also in pith and parenchyma cells. These tissues normally contain only cells with a thin primary wall. Increasing cellulose production in these cell types or even engineering them to deposit a SCW would be an interesting strategy with potential to increase total cellulose content in bioenergy crops.

Hemicelluloses and Pectin

Although it is by far the most abundant, cellulose is not the only C6-sugar polysaccharide in the cell wall. Co-overexpression of *Galactan Synthase1* (*GalS1*) with the gene encoding an enzyme that provides its substrate, *UDP-Glc/UDP-Gal-4-Epimerase2*

(*UGE2*), in Arabidopsis increased galactose content in the cell wall by up to 80% (Gondolf et al., 2014). GalS1 transfers galactose to β-1,4-galactan side chains of the rhamnogalacturonan I (RGI) backbone domain of pectin. Overexpression of *UDP-Rha/UDP-Gal Transporter1 (URGT1)*, responsible for transport of UDP-Gal from the cytoplasm to the Golgi, in combination with *GalS1* and *UGE2* further boosted galactose content in the stems to four times the levels in wild-type plants (**Figure 1**; Aznar et al., 2018). In these plants the transgenes were driven by SCW-promoters and β-1,4-galactan accumulated in secondary walls where it is not normally found, except for specialized walls in gelatinous fibers. Another hexose polysaccharide found in the cell walls of grasses is mixed-linkage glucan. The accumulation of significant amounts of mixed-linkage glucans has been successfully engineered in Arabidopsis through overexpression of *Cellulose synthase-like F6* from rice *(OsCslF6*) (Vega-Sánchez et al., 2015). The cell walls of *OsCslF6*-overexpressor lines contained four times more non-cellulosic glucose. Additionally, the saccharification efficiency of plants producing mixed-linkage glucans was increased by 42% compared to wild type. The variety of linkages between glucose residues in mixed-linkage glucans make it much more amorphous and soluble than cellulose, and thus more amenable to saccharification. The greater solubility of mixed-linkage glucan means it is easily extracted from the biomass post-harvest, which in turn exposes more of the cellulose surface to hydrolytic enzymes. While high-level production of mixed-linkage glucans is, therefore, an attractive trait to engineer into bioenergy crops, the choice of promoter appears to be critical. Overexpression of *CslF6* in barley, Arabidopsis and tobacco with constitutive or SCW-specific promoters had severely adverse effects on plant growth (Burton et al., 2011; Vega-Sánchez et al., 2015). The successful outcomes of the study by Vega-Sanchez and coworkers depended on the use of the promoter of *A. thaliana SENESCENCE ASSOCIATED GENE-12* (*pAtSAG12*), a promoter active only during senescence.

In conjunction with increasing C6 sugars like glucose and galactose, a reduction in C5 sugars like xylose is also desirable in the cell walls of dedicated bioenergy crops. Mutants in xylan biosynthesis have been identified in model species, but genetic redundancy and greater genomic complexity in bioenergy crops make full knockouts or knockdowns difficult to generate (Brown et al., 2007, 2009; Peña et al., 2007; Wu et al., 2010; Lee et al., 2012; Chen et al., 2013; Mortimer et al., 2015). In all cases, significant decreases in xylan biosynthesis in mutants was accompanied by severely reduce growth. Recent work from our lab has successfully demonstrated a novel approach to reducing the amount of a specific polysaccharide with a protein-level, dominant knock-down of xylan biosynthesis (Brandon et al., 2019). The gene *Irregular Xylem 10 (IRX10)* and its partially redundant homolog *Irregular Xylem 10-like (IRX10-L)* encode GT47 enzymes that unambiguously exhibit xylan β-(1,4)-xylosyltransferase activity in recombinant systems (Jensen et al., 2014; Urbanowicz et al., 2014). However, other proteins, IRX9 and IRX14 in particular, play critical, likely structural roles in the functional xylan synthase complex (Wu et al., 2010; Ren et al., 2014). As of this publication, no crystal structure of a GT47 protein has been resolved, so the catalytic site of IRX10 is unknown. However, potentially important amino acid residues can be inferred to be involved in catalysis based on their very high degree of conservation in evolutionarily divergent species. Two of these residues, Gly-283 and Glu-293 drastically reduced or eliminated enzymatic activity when mutated. By overexpressing the mutated *IRX10* (*dnIRX10*) genes in wild-type Arabidopsis, the mutant isoforms ($IRX10^{G283D}$ and $IRX10^{E293Q}$) out-competed the native IRX10 for its place in the proposed xylan synthase complex. The stems of *dnIRX10*-overexpressing plants had reductions in xylose content of up to 55% compared to wild type. While xylan has detrimental effects on recalcitrance and biomass conversion, it is critical to the strength of vessel and fiber SCWs, and mutants in xylan biosynthesis exhibit severe growth defects due primarily to collapsed xylem vessels and impaired water and nutrient transport (Brown et al., 2005, 2009; Lee et al., 2007; Wu et al., 2010). The phenotypes of *dnIRX10*-overexpressing lines, unsurprisingly, mimicked those of xylan knockout mutants. Previous work in our lab has demonstrated that the growth phenotype of xylan biosynthetic mutants can be rescued by expression of a functional copy of the gene under the control of a vessel cell-specific promoter (Petersen et al., 2012). Therefore, much of the xylose reduction could be maintained without growth penalty if expression of the *dnIRX10* transgene was abrogated in vessel cells or if a strong, fiber-specific promoter were used.

POLYSACCHARIDE MODIFICATION

Cellulose

Biomass quality can be improved by transgenic expression of enzymes that modify the polysaccharides of the cell wall prior to its full maturation. Plants express various glycosyl hydrolases (GH) for building and remodeling the wall in many different tissues and stages of development (Barnes and Anderson, 2018). The GH9 β-1,4-endoglucanases in plants are hypothesized to play a role in cellulose remodeling and biosynthesis by cutting specifically between glucose residues of a cellulose chain. The GH9 subfamily B is distinguished by the absence of either a transmembrane domain or carbohydrate binding module (CBM) (Hayashi et al., 2005). Recent work in rice has demonstrated that overexpression of two native GH9Bs (*OsGH9B1* and *OsGH9B3*) dramatically improved biomass quality without significant growth or developmental defects (Huang et al., 2019). The cell wall composition was unchanged compared to the control and the plants grew normally. However, the transgenic lines exhibited a 18–23% decrease in cellulose DP and an 11–23% reduction in CrI. After pretreatment and enzymatic saccharification of the biomass, both *GH9B1* and *GH9B3*-overexpressors released 63% more reducing sugars than control lines. Since total cellulose content in transgenic lines was unchanged, the authors posit that the enzymes directly decreased DP and CrI by cleaving microfibrils, thereby increasing accessibility to cellulases by increasing the number of exposed cellulose ends (**Figure 2A**).

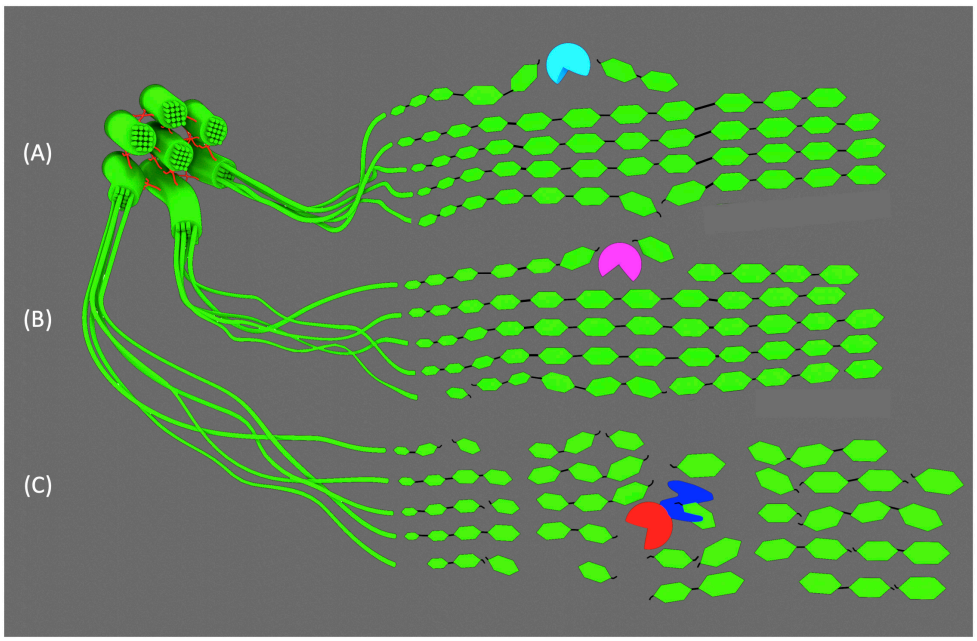

FIGURE 2 | GH cuts to cellulose microfibrils. **(A)** GH9B1 and GH9B3 and **(B)** thermophilic, CBM-truncated AcCel5A make relatively few cuts to superficial strands in the microfibril. **(C)** Mesophilic TrCel5A binds to cellulose via its CBM (blue), while the endoglucanase domain (red) makes cuts with significantly higher frequency due to its temperature optimum being similar to plant growth conditions.

Genes encoding GH enzymes from lignocellulose-degrading fungi can also be used to engineer bioenergy crops for heterologous expression. Work using the thermophilic endoglucanase gene from *Acidothermus cellulolyticus AcCel5A* suggests that even minimal cuts in cellulose chains can have dramatic effects on microfibril crystallinity, increasing enzymatic digestibility and sugar yield (Brunecky et al., 2011; Donohoe et al., 2017). AcCel5A is known to have high activity at high temperatures on synthetic substrates *in vitro,* but to be ineffective at degrading mature cell walls. Theorizing that *in planta* expression might be more effective, transgenic lines of maize and tobacco producing apoplast-targeted AcCel5A were generated. The high temperature optimum of the enzyme was a deliberate and important consideration, as it suggested activity would be low enough to avoid deleterious effects on normal plant growth. Indeed, both tobacco and maize plants grew normally and were less recalcitrant to bioconversion, yielding 10–15% more glucose from cellulose than untransformed plants under the same conditions (Brunecky et al., 2011). Importantly, adding recombinant AcCel5A to post-harvest cell wall material could not replicate these results, indicating that enzyme activity concurrent with cellulose production and deposition is key. In order to better understand the role of AcCel5A and build on prior work in maize and tobacco, the same endoglucanase gene was transformed into Arabidopsis to engineer overexpressor lines (**Figure 2B**; Donohoe et al., 2017). As in previous experiments using maize and tobacco, the composition of the cell wall was unchanged in transgenic plants and they grew normally. However, after closer inspection of the cell wall with scanning electron microscopy, large voids, pockets, and other structural irregularities were observed in cell walls of *AcCel5A*-overexpressors. Interestingly, these features mimicked the characteristics of electron micrographs of plant biomass after various chemical pretreatments. Thus, *in planta* expression of *AcCel5A* is able to mimic the role of pretreatment to increase the cellulose surface area accessible to hydrolytic enzymes. This presents the obvious benefit of likely reducing the chemical and enzyme input necessary downstream.

The choice of endoglucanase and the organism it is derived from is important. Overexpression and apoplast-targeting in tobacco of a similar GH from the mesophilic fungus *Trichoderma reesei*, TrCel5A, caused severe growth defects and a significant decrease in cellulose content (Klose et al., 2015). Presumably, TrCel5A would be more active than AcCel5A since *T. reesei* evolved in similar conditions to plants. Additionally, TrCel5A possesses a CBM that was truncated from the AcCel5A used by Brunecky et al. (2011). CBMs are known to facilitate hydrolytic activity and reduce CrI by physically disrupting the hydrogen bonds between cellulose chains (**Figure 2C**; Abramson et al., 2010). This disruptive effect could be exploited for the improvement of bioenergy crops, since CrI is strongly correlated with biomass recalcitrance. However, relatively few studies have been published exploring the engineered overexpression of CBMs, or CBM-containing proteins like expansins, to this end. Expansins are cell wall proteins that disrupt the intermolecular hydrogen bonds of cellulose and hemicelluloses, promoting the flexibility and extensibility of the primary cell wall. They play critical roles in primary growth by maintaining the delicate balance between turgor pressure and cell wall integrity that drive cell expansion (Cosgrove, 2000, 2005). However, some

expansins are specifically expressed in vessel and fiber cells during SCW development. Two such expansin genes from Chinese fir (*Cunninghamia lanceolata)* have been cloned and used to engineer overexpression lines in tobacco (Wang et al., 2011). Both *Expansin-A1 and Expansin-A2 (ClEXPA1 and ClEXPA2)* overexpressing lines grew taller and had thicker stems than wild type. The cell walls of xylem cells were 1.13 to 1.45 times thicker in transgenic plants and when the composition of stem cell wall material was analyzed, they contained 30–50% more cellulose. It is also interesting to explore the effects that CBMs alone can have on cell wall architecture. In Arabidopsis, the coding sequence of the CBM of *Starch Synthase III (SSIII)* was overexpressed and targeted to the cell wall (Grisolia et al., 2017). Though starch and cellulose are very different in structure, previous work indicated that the concatenated triplicate of CBMs (collectively referred to as SBD123) from SSIII had preferential affinity for the linear portions of the starch molecule. Thus, the rationale behind the work was that it may also bind to linear cellulose and modify the crystallinity of microfibrils. Like the tobacco plants expressing *ClEXPA* genes, Arabidopsis plants overexpressing *SBD123* grew significantly larger than untransformed lines. The average cell area was increased by 40% and dry biomass weight by 76%. Unlike the *ClEXPA*-expressing plants, Grisolia et al. (2017) observed a 27% reduction in cell wall thickness and similar amounts of cellulose in transgenic Arabidopsis stems. This suggests SBD123 loosened the components of the cell wall, stretching the cell wall to cover a greater cell volume. However, hemicellulose and pectin contents were significantly higher (50% and 30%, respectively) and dilute acid hydrolysis of cell wall material from transgenic plants yielded almost twice as much glucose. Dilute acid is generally insufficient to hydrolyze crystalline cellulose and there is relatively little xyloglucan in Arabidopsis stems. Thus, it is reasonable to conclude that the glucose released is derived from crystalline cellulose made amorphous by the disruptive effects of SBD123-cellulose interaction. Finally, an *in vitro* rumen digestibility assay determined biomass from transgenic plants was 28% easily digested than that of control plants. However, other attempts to engineer plants expressing CBMs have been less successful with variable results and sometimes negative effects on growth and cell wall recalcitrance (Safra-Dassa et al., 2006; Obembe et al., 2007; Keadtidumrongkul et al., 2017).

Xylan

Xylan is the most abundant hemicellulose in most bioenergy crops and contributes significantly to biomass recalcitrance by enveloping cellulose microfibrils and forming covalent linkages to lignin and other hemicelluloses in the cell wall. Thus, modifications to xylan are important engineering goals in the development of dedicated bioenergy crops. The xylan backbone is β-(1,4)-linked xylose residues, many of which can be mono- or di-acetylated, and is decorated to varying degrees with glucuronic acid (GlcA), 4-*O*-methyl glucuronic acid (MeGlcA) and, in grasses, arabinose (Ara) side chains (Rennie and Scheller, 2014). Arabinose side chains can be further modified by esterification with ferulic acid or *p*-coumaric acid moieties. The amount and distribution of these decorations to the xylan backbone determine the properties of the polysaccharide *in muro*. Several engineering strategies targeting these side chains have been successfully employed to reduce biomass recalcitrance.

Approximately 40–60% of xylose residues comprising xylan are acetylated at the O-2 or O-3 position (Busse-Wicher et al., 2014), and reductions in acetylation correlate positively with conversion efficiency. Additionally, acetic acid released from xylan during pretreatment is a strong inhibitor of microbial fermentation downstream. To engineer a reduction in xylan acetylation, an acetyl xylan esterase from the lignocellulose-degrading fungus *Aspergillus niger* (*AnAXE1*) was overexpressed in Arabidopsis (**Figure 3**) (Pawar et al., 2016). After observing beneficial effects, the same gene was used to generate transgenic overexpressors in hybrid aspen (*Populus tremula L.* × *tremuloides Michx*) (Pawar et al., 2017). The transgenic trees developed normally with approximately 10% less xylan acetylation. This reduction in acetylation had the secondary effect of reducing the average xylan DP. Since no changes in xylan biosynthesis were observed, the authors propose that xylan with reduced acetylation may be more susceptible to cleavage by endogenous cell wall hydrolases. These modifications to xylan led to a 26% increase in saccharification efficiency compared to controls without pretreatment. The positive effect of *AnAXE1*-overexpression was confirmed after acid pretreatment of the biomass. Acid pretreatment removed xylan, and the difference in saccharification efficiency between transgenic and control lines was greatly reduced. The effects of xylan acetylation on recalcitrance, however, are not yet fully understood. The overexpression of a gene encoding a DUF231-containing protein in poplar (*PdDUF231A*) had beneficial effects on biomass composition, despite increasing xylan acetylation by 8% (Yang et al., 2017b). It is likely that the effect on xylan acetylation is an indirect one, though, since the PdDUF231A is phylogenetically more closely related to PMR5, a pectin acetyltransferase (Chiniquy et al., 2019), than to the known xylan acetyltransferases. *PdDUF231A*-overexpressing lines exhibited an 8–21% increase in total cellulose content and 6–8% reduction in lignin, while cell wall content of other cell wall sugars was unchanged. Metabolite analysis of the transgenic lines additionally suggested the increase in carbon flux to cellulose biosynthesis came at the expense of lignin biosynthesis. While the precise role of *PdDUF231A* is unknown, the combination of beneficial biomass changes it produced makes it an interesting subject of further research. Despite the successful studies on reducing xylan acetylation in plants, it is clear that xylan acetylation is important for plant growth and development (Xiong et al., 2013). Therefore, we should expect that the successful engineering of reduced xylan acetylation will depend on careful choice of promoters and target cell types.

The xylans of grasses are extensively arabinosylated, allowing the formation of many cross-linkages with other components in the cell wall and directly affecting the characteristics of cellulose microfibrils (Li et al., 2015). Reducing the number of xylan arabinose side chains has great potential to improve the biomass quality of grasses, perhaps the most important group of dedicated biomass crops. Two native arabinofuranosidase genes from rice (*OsARAF1* and *OsARAF3*), were used to engineer rice

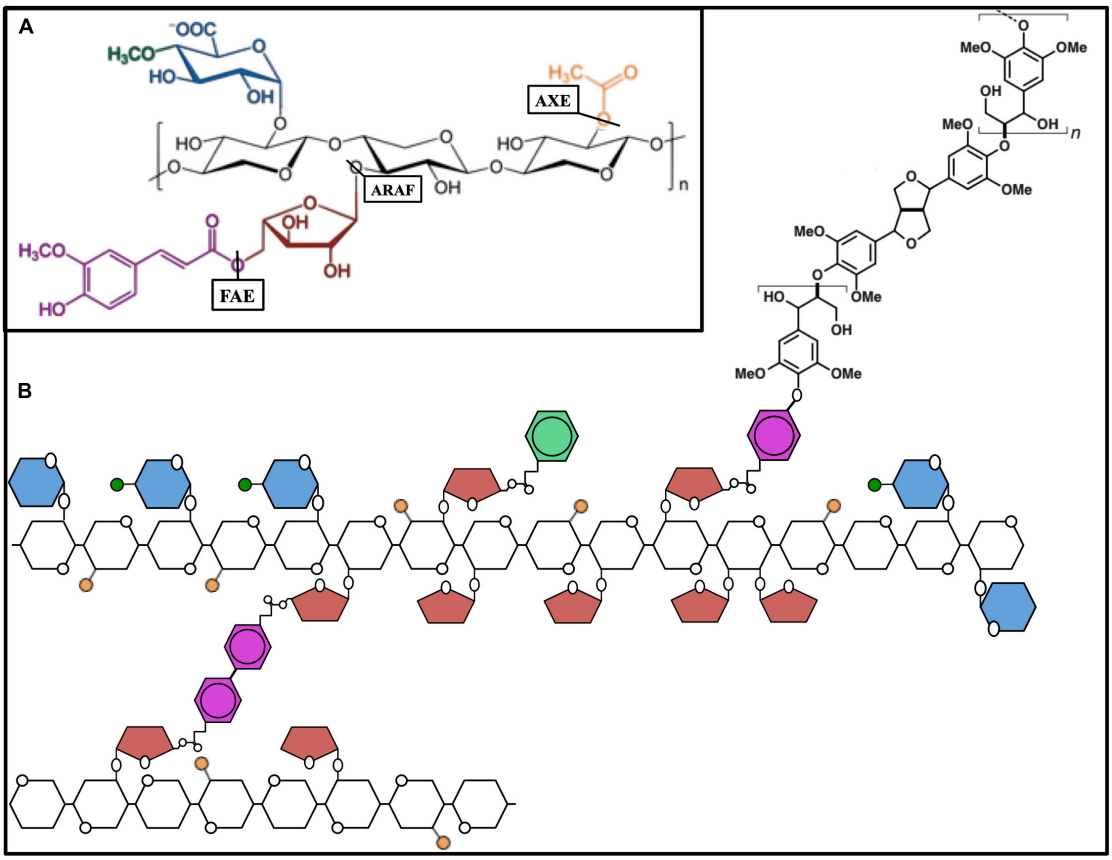

FIGURE 3 | The hemicellulose xylan and interactions. **(A)** Schematic molecular structure of xylan module with β-(1,4)-linked xylose residues (black) of the xylan backbone that are substituted with acetyl (orange), arabinose (red), and glucuronic acid (blue) residues. Arabinose is partially esterified with ferulic acid (magenta) and glucuronic acid is often 4-*O*-methylated (green). Acetyl xylan esterase (AXE), arabinofuranosidase (ARAF), and ferulic acid esterase (FAE) indicating the bonds they hydrolyze. **(B)** Schematic representation of the xylan chain, xylan-xylan diferulate cross-linking, and ferulic acid-mediated lignin polymerization.

lines accumulating high levels of the protein in the cell wall (**Figure 3**; Sumiyoshi et al., 2013). Transgenic lines had 19–25% less arabinose in matrix cell-wall polysaccharides and a 28–34% increase in cellulose content. It is likely both factors combined led to a 46–70% increase in enzymatic sugar release compared to wild type. Many xylan arabinose side chains are esterified with ferulic acid, which can form extensive diferulate covalent bonds with other glucuronoarabinoxylans. Ferulic acid can also act as a nucleation site for the polymerization of lignin (Grabber et al., 2004; Terrett and Dupree, 2019). These properties make ferulic acid a significant contributor to the recalcitrance of grass biomass. Diferulate cross-linking and the free-radical nucleation of lignin happen spontaneously and non-enzymatically. Thus, enzymes severing the connection between ferulic acid and arabinose are particularly useful for biomass engineering. Fortunately, we can again draw from ingenuity of lignocellulose-degrading fungi. *In planta* expression of a variety of ferulic acid esterases (FAEs) has demonstrated beneficial effects to reduce recalcitrance and increase biomass digestibility, but in many cases there are trade-offs to be considered. Expression of *Aspergillus nidulans* FAE in Arabidopsis and *Brachypodium distachyon* enhanced saccharification but increased susceptibility to fungal pathogens (Pogorelko et al., 2011; Reem et al., 2016). In alfalfa (*Medicago sativa*), expression of *A. niger* FAE did reduce arabinoxylan feruloylation, but unexpectedly led to increased lignification and decreased digestibility (Badhan et al., 2014). The authors noted, however, that saccharification efficiency did increase compared to wild type after chemical delignification, possibly due to a reduction of inter-arabinoxylan cross-linking. The beneficial effect of *in planta* expression of FAEs has been particularly well studied in tall fescue (*Festuca arundinacea)*, an important animal forage crop (de Buanafina et al., 2008, 2010). As fescue is typically studied for its nutritional value to livestock, an *in vitro* model using rumen microorganisms was used to assess cell wall digestibility and *AnFAE*-overexpressors demonstrated a 10–14% increase in digestibility. More recent work with fescue suspension culture cells showed that significantly less time was required for rumen microbes to digest the cell wall from cells overexpressing *AnFAE* than for control cells (Morris et al., 2017). Despite the use of very different analytical techniques, it is reasonable to believe these changes to cell wall composition could translate in bioenergy crops to improvements in enzymatic saccharification. Importantly, this increase in digestion rate could not be replicated by exogenous application of recombinant

AnFAE to control cell wall material. Much like the previously described transgenics expressing endoglucanases, the beneficial effects are most significant when the enzyme is active during cell wall development. The observed negative side effects of FAE expression such as increased susceptibility to pathogens could perhaps be mitigated by restricting expression of the transgene in the tissues most affected by the stress, like epidermal cells. Other labs have found success in rice and switchgrass using an alternative approach to reduce ferulate crosslinking in the cell wall by out-competing ferulic acid for arabinosyl side chains on xylan (Bartley et al., 2013; Li et al., 2018). The BAHD-type acyltransferase OsAT10 from rice is a putative *p*-coumaroyl transferase that increases *p*-coumaric acid esterification of arabinose moieties of xylan. Although it is also a phenylpropanoid-derived hydroxycinnamic acid, *p*-coumaric acid does not undergo oxidative coupling to monolignols in the polymerization of lignin (Marcia, 2009). In *OsAT10*-overexpressing lines of rice and switchgrass, the ratio of *p*-coumaroyl to feruloyl esterification of arabinoxylan increased by 150% and 75%, respectively. In rice, *OsAT10* overexpression had the additional effect of increasing cell wall glucose by 8–19% compared to wild type. Transgenic lines of both rice and switchgrass developed normally and both demonstrated an increase of up to 40% in total sugar yield after enzymatic saccharification.

Other Polysaccharides

While most studies have focused on cellulose and xylan, some studies have targeted other polysaccharides, which are clearly important in the development of SCWs despite their lower abundance. Overexpression of *Aspergillus aculeatus* xyloglucanase gene (*AaXEG2*) in poplar (*Populus alba*) led to increased growth and cellulose deposition and up to 81% more glucose released after enzymatic hydrolysis (Park et al., 2004; Kaida et al., 2009). Pectin modification is likewise a useful strategy as shown in the knock-down studies mentioned above (Biswal et al., 2015, 2018; Li et al., 2019b). The modification of pectin by overexpression of a pectate lyase gene from *Populus trichocarpa* (*PtPL1*) in hybrid aspen (*Populus tremula* × *tremuloides*) resulted in improved saccharification (Biswal et al., 2014). The overall composition of the biomass did not differ significantly from wild-type plants, indicating that the enzyme might act by loosening interactions between matrix wall components and increasing accessibility to cell wall-degrading enzymes. The homogalacturonan domain of the pectin backbone and/or the extent of its methyl-esterification have been targets for *in planta* modification. An *A. niger* polygalacturonase gene (*AnPGA2*) was used to engineer overexpressors in Arabidopsis and tobacco (Lionetti et al., 2010). While transgenic lines had a twofold increase in saccharification efficiency, they were also severely stunted in growth with 50–84% reduction in total biomass. This detrimental effect has been described previously, wherein it was demonstrated that despite the use of a strong promoter, the only plants that survived to be studied accumulated AnPGA2 at very low levels (Capodicasa et al., 2004). A polygalacturonase-inhibiting protein (PGIP) produced by the common bean plant (*Phaseolus vulgaris*) has been identified and well described to eliminate PGA2 activity (Leckie et al., 1999). If the growth defects of PGA2-transgenics were determined to be tissue- or developmental stage-dependent, it would be interesting to include PGIP with a promoter specific to that tissue or developmental stage. Then, the dramatic improvements to saccharification could possibly be engineered into bioenergy crops without sacrificing biomass yield. One study has explored controlled pectin modification by generating transgenic Arabidopsis lines expressing a pectate lyase from *Pectobacterium carotovorum* (*PcPL1*) and the *A. niger* polygalacturonase (*AnPGA2*) (Tomassetti et al., 2015). The β-estradiol inducible promoter (pEST) and the senescence promoter (pSAG12) were used to drive expression of *PcPL1* and *AnPGA2*, respectively. Tissue collected from *pEST:AnPGA2* transgenic plants post-induction yielded almost twice as much glucose as wild-type or non-induced plants. However, some lines expressed *AnPGA2* even in the absence of the inducer and suffered diminished growth. In contrast, all *pSAG12:AnPGA2* transgenic lines grew normally. Expression of the gene was undetectable in plants until they reached 7 weeks old, at the beginning of senescence, and mirrored the expression of the native *SAG12* gene. Stem and leaf biomass from 4 to 7-week-old *pSAG12:PGA2* plants and evaluated for saccharification efficiency. The amount of glucose released by cellulase digestion was comparable between 4-week-old *pSAG12:PGA2* and wild type plants. In contrast, significantly more glucose was released from *pSAG12:PGA2* plants compared to wild type after 7 weeks of growth. The increase was almost twofold from leaf tissue, and about 50% from stems.

Homogalacturonan is extensively methylesterified as it is synthesized and secreted. Once it reaches the cell wall, pectin methylesterases remove many of these methyl groups, revealing carboxylic acid moieties that subsequently form linkages with various components of the cell wall. Free carboxyl groups of de-methylesterified homogalacturonan can additionally form rigid, calcium-mediated "egg box" structures with adjacent stretches of homogalacturonan (Atmodjo et al., 2013). Therefore, maintaining the methylesterification of homogalacturonan could be a viable approach to reducing cell wall recalcitrance. In the previously referenced study by Lionetti et al., pectin methylesterase inhibitor genes from Arabidopsis (*AtPMEI-2*) and kiwi (*Actinidia chinensis*, *AcPMEI*) were overexpressed in Arabidopsis and wheat (*Triticum durum*), respectively (Lionetti et al., 2010). Saccharification efficiency of *AtPMEI-2* overexpressing Arabidopsis stem increased by 50%, while wheat plants overexpressing *AcPMEI* had 2.5-fold greater saccharification efficiency than controls. In contrast to polygalacturonase-overexpressing plants, the total biomass of Arabidopsis plants overexpressing *AtPMEI-2* was greater than wild type by 68%. This increase was not observed in *AcPMEI* overexpressing wheat plants, suggesting differing impacts of homogalacturonan methylesterification between monocots and dicots. The role of minor matrix polysaccharides in the development and structure of SCWs is poorly understood, but the successful results suggest that this would be a fertile ground for further research.

CONCLUSION AND FUTURE PERSPECTIVES

Plant biotechnological research can make a major contribution to meeting perhaps humanities' most common and critical goals. That is, mitigating the effects of climate change by reducing our net production of carbon dioxide and developing an alternative to the fundamentally unsustainable fossil resources on which we have become dependent. While traditional breeding techniques will be important in the development of dedicated bioenergy crops, genetic engineering allows for significant modifications to biomass to be made much more quickly. As our fundamental knowledge of plant cell biology and the regulation of cell wall development continues to grow, so too will our ability to approach the perfect balance of maximizing yield and quality while minimizing recalcitrance and deleterious side effects. Here we have described a number of dominant genetic engineering strategies to improve plant biomass for bioenergy conversion. Though they may have only been demonstrated in one or two species, they have potential for broad usability in the wide variety of dedicated bioenergy species currently being researched. In several of the studies we have reviewed here, the results observed in different species were somewhat contradictory. For example, overexpression of CesAs or Susy in different species resulted in both increased or decreased cellulose and lignin depending on the particular study. This suggests that direct translation of an engineering strategy from a model plant or from one bioenergy crop to another is not entirely straightforward. The reasons for these contradictory results are not clear. They could be due to differences in the physiology and different cell wall types of the various species. However, we find it more likely that differences in the promoters used in the various studies, including their spatial and temporal expression profiles can explain many of the observed differences. Studies that resolve the results obtained with a larger set of promoters in a single species are required and would help to better predict outcomes in future studies. Future studies also need to determine the impact of environmental factors on the outcomes. If indeed the transgene expression profiles are very important, then we may also expect that the effect on plant development, biomass composition and recalcitrance could be strongly influenced by growth conditions. For the practical deployment of approaches such as those reviewed here, it is very important to determine if observations in the laboratory and greenhouse translate to the field condition, and to understand the mechanisms responsible where differences are observed.

AUTHOR CONTRIBUTIONS

AB and HS developed the idea of the work and wrote the manuscript. AB illustrated **Figures 1**, **3**.

FUNDING

This work conducted by the Joint BioEnergy Institute and supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research under contract number DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the US Department of Energy.

ACKNOWLEDGMENTS

Special thanks to Julian Tan for the design and illustration of **Figure 2**.

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

Atmodjo, M. A., Hao, Z., and Mohnen, D. (2013). Evolving views of pectin biosynthesis. *Annu. Rev. Plant Biol.* 64, 747–779. doi: 10.1146/annurev-arplant-042811-105534

Aznar, A., Chalvin, C., Shih, P. M., Maimann, M., Ebert, B., Birdseye, D. S., et al. (2018). Gene stacking of multiple traits for high yield of fermentable sugars in plant biomass. *Biotechnol. Biofuels* 11:2. doi: 10.1186/s13068-017-1007-6

Badhan, A., Jin, L., Wang, Y., Han, S., Kowalczys, K., Brown, D. C., et al. (2014). Expression of a fungal ferulic acid esterase in alfalfa modifies cell wall digestibility. *Biotechnol. Biofuels* 7:39. doi: 10.1186/1754-6834-7-39

Barnes, W. J., and Anderson, C. T. (2018). Release, recycle, rebuild: cell-wall remodeling, autodegradation, and sugar salvage for new wall biosynthesis during plant development. *Mol. Plant* 11, 31–46. doi: 10.1016/j.molp.2017.08.011

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

Bhatia, R., Gallagher, J. A., Gomez, L. D., and Bosch, M. (2017). Genetic engineering of grass cell wall polysaccharides for biorefining. *Plant Biotechnol. J.* 15, 1071–1092. doi: 10.1111/pbi.12764

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

Biswal, A. K., Hao, Z., Pattathil, S., Yang, X., Winkeler, 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

Biswal, A. K., Soeno, K., Gandla, M. L., Immerzeel, P., Pattathil, S., Lucenius, J., et al. (2014). Aspen pectate lyase PtxtPL1-27 mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. *Biotechnol. Biofuels* 7:11. doi: 10.1186/1754-6834-7-11

Borland, A. M., Griffiths, H., Hartwell, J., and Smith, J. A. C. (2009). Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands. *J. Exp. Bot.* 60, 2879–2896. doi: 10.1093/jxb/erp118

Brandon, A. G., Birdseye, D. S., and Scheller, H. V. (2019). A dominant negative approach to reduce xylan in plants. *Plant Biotechnol. J.* doi: 10.1111/pbi.13198 [Epub ahead of print].

Brown, D. M., Goubet, F., Wong, V. W., Goodacre, R., Stephens, E., Dupree, P., et al. (2007). Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J.* 52, 1154–1168. doi: 10.1111/j.1365-313X.2007.03307.x

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

Brown, D. M., Zhang, Z., Stephens, E., Dupree, P., and Turner, S. R. (2009). Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in *Arabidopsis*. *Plant J.* 57, 732–746. doi: 10.1111/j.1365-313X.2008.03729.x

Brunecky, R., Selig, M. J., Vinzant, T. B., Himmel, M. E., Lee, D., Blaylock, M. J., et al. (2011). In planta expression of A. cellulolyticus Cel5A endocellulase reduces cell wall recalcitrance in tobacco and maize. *Biotechnol. Biofuels* 4:1. doi: 10.1186/1754-6834-4-1

Burton, R. A., Collins, H. M., Kibble, N. A. J., Smith, J. A., Shirley, N. J., Jobling, S. A., et al. (2011). Over-expression of specific HvCslF cellulose synthase-like genes in transgenic barley increases the levels of cell wall (1,3;1,4)-(-d-glucans and alters their fine structure. *Plant Biotechnol. J.* 9, 117–135. doi: 10.1111/j.1467-7652.2010.00532.x

Busse-Wicher, M., Gomes, T. C. F., 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/tpj.12575

Cai, X., Zhang, X., and Wang, D. (2011). Land availability for biofuel production. *Environ. Sci. Technol.* 45, 334–339. doi: 10.1021/es103338e

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 37, D233–D238. doi: 10.1093/nar/gkn663

Capodicasa, C., Vairo, D., Zabotina, O., McCartney, L., Caprari, C., Mattei, B., et al. (2004). Targeted modification of homogalacturonan by transgenic expression of a fungal polygalacturonase alters plant growth. *Plant Physiol.* 135, 1294–1304. doi: 10.1104/pp.104.042788

Chen, P., and Peng, L. (2013). "CHAPTER 6. the diversity of lignocellulosic biomass resources and their evaluation for use as biofuels and chemicals," in *Biological Conversion of Biomass for Fuels and Chemicals: Explorations from Natural Utilization Systems Energy and Environment Series*, eds J. Sun, S.-Y. Ding, and J. D. Peterson, (Cambridge: Royal Society of Chemistry), 83–113. doi: 10.1039/9781849734738-00083

Chen, X., Vega-Sánchez, M. E., Verhertbruggen, Y., Chiniquy, D., Canlas, P. E., Fagerström, A., et al. (2013). Inactivation of osirx10 leads to decreased xylan content in rice culm cell walls and improved biomass saccharification. *Mol. Plant* 6, 570–573. doi: 10.1093/mp/sss135

Chiniquy, D., Underwood, W., Corwin, J., Ryan, A., Szemenyei, H., Lim, C. C., et al. (2019). PMR5, an acetylation protein at the intersection of pectin biosynthesis and defense against fungal pathogens. *Plant J.* 100, 1022–1035. doi: 10.1111/tpj.14497

Christensen, U., and Scheller, H. V. (2012). Regulation of (1,3;1,4)-(-d-glucan synthesis in developing endosperm of barley lys mutants. *J. Cereal Sci.* 55, 69–76. doi: 10.1016/j.jcs.2011.10.005

Coleman, H. D., Yan, J., and Mansfield, S. D. (2009). Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13118–13123. doi: 10.1073/pnas.0900188106

Cosgrove, D. J. (2000). Loosening of plant cell walls by expansins. *Nature* 407, 321–326. doi: 10.1038/35030000

Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev.* 6, 850–861. doi: 10.1038/nrm1746

Damm, T., Commandeur, U., Fischer, R., Usadel, B., and Klose, H. (2016). Improving the utilization of lignocellulosic biomass by polysaccharide modification. *Process Biochem.* 51, 288–296. doi: 10.1016/j.procbio.2015.12.003

de Buanafina, M. M. O., Langdon, T., Hauck, B., Dalton, S., and Morris, P. (2008). Expression of a fungal ferulic acid esterase increases cell wall digestibility of tall fescue (*Festuca arundinacea*). *Plant Biotechnol. J.* 6, 264–280. doi: 10.1111/j.1467-7652.2007.00317.x

de Buanafina, M. M. O., Langdon, T., Hauck, B., Dalton, S., Timms-Taravella, E., and Morris, P. (2010). Targeting expression of a fungal ferulic acid esterase to the apoplast, endoplasmic reticulum or golgi can disrupt feruloylation of the growing cell wall and increase the biodegradability of tall fescue (*Festuca arundinacea*). *Plant Biotechnol. J.* 8, 316–331. doi: 10.1111/j.1467-7652.2009.00485.x

Donev, E., Gandla, M. L., Jönsson, L. J., and Mellerowicz, E. J. (2018). Engineering non-cellulosic polysaccharides of wood for the biorefinery. *Front. Plant Sci.* 9:1537. doi: 10.3389/fpls.2018.01537

Donohoe, B. S., Wei, H., Mittal, A., Shollenberger, T., Lunin, V. V., Himmel, M. E., et al. (2017). Towards an understanding of enhanced biomass digestibility by in planta expression of a family 5 glycoside hydrolase. *Sci. Rep.* 7:4389. doi: 10.1038/s41598-017-04502-1

Fan, C., Feng, S., Huang, J., Wang, Y., Wu, L., Li, X., et al. (2017). AtCesA8-driven OsSUS3 expression leads to largely enhanced biomass saccharification and lodging resistance by distinctively altering lignocellulose features in rice. *Biotechnol. Biofuels* 10:221. doi: 10.1186/s13068-017-0911-0

Fan, C., Wang, G., Wu, L., Liu, P., Huang, J., Jin, X., et al. (2019). Distinct cellulose and callose accumulation for enhanced bioethanol production and biotic stress resistance in OsSUS3 transgenic rice. *Carbohydr. Polym.* 232:115448. doi: 10.1016/j.carbpol.2019.115448

Fujii, S., Hayashi, T., and Mizuno, K. (2010). Sucrose synthase is an integral component of the cellulose synthesis machinery. *Plant Cell Physiol.* 51, 294–301. doi: 10.1093/pcp/pcp190

Gírio, F. M., Fonseca, C., Carvalheiro, F., Duarte, L. C., Marques, S., and Bogel-Łukasik, R. (2010). Hemicelluloses for fuel ethanol: a review. *Bioresour. Technol.* 101, 4775–4800. doi: 10.1016/j.biortech.2010.01.088

Gondolf, V. M., Stoppel, R., Ebert, B., Rautengarten, C., Liwanag, A. J., Loqué, D., et al. (2014). A gene stacking approach leads to engineered plants with highly increased galactan levels in *Arabidopsis*. *BMC Plant Biol.* 14:344. doi: 10.1186/s12870-014-0344-x

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

Grisolia, M. J., Peralta, D. A., Valdez, H. A., Barchiesi, J., Gomez-Casati, D. F., and Busi, M. V. (2017). The targeting of starch binding domains from starch synthase III to the cell wall alters cell wall composition and properties. *Plant Mol. Biol.* 93, 121–135. doi: 10.1007/s11103-016-0551-y

Guidi, W. E., Pitre, F., and Labrecque, M. (2013). "Short-rotation coppice of willows for the production of biomass in Eastern Canada," in *Biomass Now - Sustainable Growth and Use*, ed. M. D. Matovic, (London: InTech), doi: 10.5772/51111

Hall, M., Bansal, P., Lee, J. H., Realff, M. J., and Bommarius, A. S. (2010). Cellulose crystallinity–a key predictor of the enzymatic hydrolysis rate. *FEBS J.* 277, 1571–1582. doi: 10.1111/j.1742-4658.2010.07585.x

Hansen, S. F., Harholt, J., Oikawa, A., and Scheller, H. V. (2012). Plant glycosyltransferases beyond cazy: a perspective on DUF families. *Front. Plant Sci.* 3:59. doi: 10.3389/fpls.2012.00059

Hayashi, T., Yoshida, K., Park, Y. W., Konishi, T., and Baba, K. (2005). Cellulose metabolism in plants. *Int. Rev. Cytol.* 247, 1–34. doi: 10.1016/S0074-7696(05)47001-1

Heaton, E. A., Dohlemon, F. G., and Long, S. P. (2008). Meeting US biofuel goals with less land: the potential of *Miscanthus*. *Glob. Change Biol.* 14, 2000–2014. doi: 10.1111/j.1365-2486.2008.01662.x

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

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

Huang, J., Xia, T., Li, G., Li, X., Li, Y., Wang, Y., et al. (2019). Overproduction of native endo-(-1,4-glucanases leads to largely enhanced biomass saccharification and bioethanol production by specific modification of cellulose features in transgenic rice. *Biotechnol. Biofuels* 12:11. doi: 10.1186/s13068-018-1351-1

Jensen, J. K., Johnson, N. R., and Wilkerson, C. G. (2014). *Arabidopsis thaliana* IRX10 and two related proteins from psyllium and *Physcomitrella* patens are xylan xylosyltransferases. *Plant J.* 80, 207–215. doi: 10.1111/tpj.12641

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

Junior, H. J. E., Melo, R. X., de Sartori, M. M. P., Guerra, S. P. S., and Ballarin, A. W. (2016). Sustainable use of eucalypt biomass grown on short rotation coppice for bioenergy. *Biomass Bioenerg.* 90, 15–21. doi: 10.1016/j.biombioe.2016.03.037

Kaida, R., Kaku, T., Baba, K., Oyadomari, M., Watanabe, T., Nishida, K., et al. (2009). Loosening xyloglucan accelerates the enzymatic degradation of cellulose in wood. *Mol. Plant* 2, 904–909. doi: 10.1093/mp/ssp060

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-(-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

Keadtidumrongkul, P., Suttangkakul, A., Pinmanee, P., Pattana, K., Kittiwongwattana, C., Apisitwanich, S., et al. (2017). Growth modulation effects of CBM2a under the control of AtEXP4 and CaMV35S promoters in *Arabidopsis thaliana, Nicotiana tabacum* and *Eucalyptus camaldulensis*. *Transgen. Res.* 26, 447–463. doi: 10.1007/s11248-017-0015-4

Kimura, S., and Itoh, T. (2007). "Biogenesis and function of cellulose in the tunicates," in *Cellulose: Molecular and Structural Biology*, eds R. M. Brown, and I. M. Saxena, (Dordrecht: Springer Netherlands), 217–236. doi: 10.1007/978-1-4020-5380-1_13

Klose, H., Günl, M., Usadel, B., Fischer, R., and Commandeur, U. (2015). Cell wall modification in tobacco by differential targeting of recombinant endoglucanase from *Trichoderma reesei*. *BMC Plant Biol.* 15:54. doi: 10.1186/s12870-015-0443-3

Knoch, E., Dilokpimol, A., Tryfona, T., Poulsen, C. P., Xiong, G., Harholt, J., et al. (2013). A (-glucuronosyltransferase from *Arabidopsis thaliana* involved in biosynthesis of type II arabinogalactan has a role in cell elongation during seedling growth. *Plant J.* 76, 1016–1029. doi: 10.1111/tpj.12353

Lao, J., Oikawa, A., Bromley, J. R., McInerney, P., Suttangkakul, A., Smith-Moritz, A. M., et al. (2014). The plant glycosyltransferase clone collection for functional genomics. *Plant J.* 79, 517–529. doi: 10.1111/tpj.12577

Leckie, F., Mattei, B., Capodicasa, C., Hemmings, A., Nuss, L., Aracri, B., et al. (1999). The specificity of polygalacturonase-inhibiting protein (PGIP): a single amino acid substitution in the solvent-exposed beta-strand/beta-turn region of the leucine-rich repeats (LRRs) confers a new recognition capability. *EMBO J.* 18, 2352–2363. doi: 10.1093/emboj/18.9.2352

Lee, C., Zhong, R., Richardson, E. A., Himmelsbach, D. S., McPhail, B. T., and Ye, Z.-H. (2007). The PARVUS gene is expressed in cells undergoing secondary wall thickening and is essential for glucuronoxylan biosynthesis. *Plant Cell Physiol.* 48, 1659–1672. doi: 10.1093/pcp/pcm155

Lee, C., Zhong, R., and Ye, Z.-H. (2012). *Arabidopsis* family GT43 members are xylan xylosyltransferases required for the elongation of the xylan backbone. *Plant Cell Physiol.* 53, 135–143. doi: 10.1093/pcp/pcr158

Li, F., Zhang, M., Guo, K., Hu, Z., Zhang, R., Feng, Y., et al. (2015). High-level hemicellulosic arabinose predominately affects lignocellulose crystallinity for genetically enhancing both plant lodging resistance and biomass enzymatic digestibility in rice mutants. *Plant Biotechnol. J.* 13, 514–525. doi: 10.1111/pbi.12276

Li, G., Jones, K. C., Eudes, A., Pidatala, V. R., Sun, J., Xu, F., et al. (2018). Overexpression of a rice BAHD acyltransferase gene in switchgrass (*Panicum virgatum* L.) enhances saccharification. *BMC Biotechnol.* 18:54. doi: 10.1186/s12896-018-0464-8

Li, M., Wang, S., Liu, Y., Zhang, Y., Ren, M., Liu, L., et al. (2019a). Overexpression of PsnSuSy1, 2 genes enhances secondary cell wall thickening, vegetative growth, and mechanical strength in transgenic tobacco. *Plant Mol. Biol.* 100, 215–230. doi: 10.1007/s11103-019-00850-w

Li, M., Yoo, C. G., Pu, Y., Biswal, A. K., Tolbert, A. K., Mohnen, D., et al. (2019b). Downregulation of pectin biosynthesis gene GAUT4 leads to reduced ferulate and lignin-carbohydrate cross-linking in switchgrass. *Commun. Biol.* 2:22. doi: 10.1038/s42003-018-0265-6

Liebminger, E., Grass, J., Altmann, F., Mach, L., and Strasser, R. (2013). Characterizing the link between glycosylation state and enzymatic activity of the endo-(1,4-glucanase KORRIGAN1 from *Arabidopsis thaliana*. *J. Biol. Chem.* 288, 22270–22280. doi: 10.1074/jbc.M113.475558

Lionetti, V., Francocci, F., Ferrari, S., Volpi, C., Bellincampi, D., Galletti, R., et al. (2010). Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. *Proc. Natl. Acad. Sci. U.S.A.* 107, 616–621. doi: 10.1073/pnas.0907549107

Liu, L., Shang-Guan, K., Zhang, B., Liu, X., Yan, M., Zhang, L., et al. (2013). Brittle Culm1, a COBRA-like protein, functions in cellulose assembly through binding cellulose microfibrils. *PLoS Genet.* 9:e1003704. doi: 10.1371/journal.pgen.1003704

Loqué, 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

Marcia, M. D. (2009). Feruloylation in grasses: current and future perspectives. *Mol. Plant* 2, 861–872.

Matthysse, A. G., Deschet, K., Williams, M., Marry, M., White, A. R., and Smith, W. C. (2004). A functional cellulose synthase from ascidian epidermis. *Proc. Natl. Acad. Sci. U.S.A.* 101, 986–991. doi: 10.1073/pnas.0303623101

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.* 9:1114. doi: 10.3389/fpls.2018.01114

McCormick, A. J., Watt, D. A., and Cramer, M. D. (2009). Supply and demand: sink regulation of sugar accumulation in sugarcane. *J. Exp. Bot.* 60, 357–364. doi: 10.1093/jxb/ern310

McFarlane, H. E., Döring, 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

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

Mir, B. A., Mewalal, R., Mizrachi, E., Myburg, A. A., and Cowan, D. A. (2014). Recombinant hyperthermophilic enzyme expression in plants: a novel approach for lignocellulose digestion. *Trends Biotechnol.* 32, 281–289. doi: 10.1016/j.tibtech.2014.03.003

Morris, P., Dalton, S., Langdon, T., Hauck, B., and de Buanafina, M. M. O. (2017). Expression of a fungal ferulic acid esterase in suspension cultures of tall fescue (*Festuca arundinacea*) decreases cell wall feruloylation and increases rates of cell wall digestion. *Plant Cell Tissue Organ Cult.* 129, 181–193. doi: 10.1007/s11240-017-1168-9

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

Ndimande, S. (2013). *Increasing Cellulosic Biomass in Sugarcane*. Ph.d. Thesis, Stellenbosch University, Stellenbosch.

Niu, E., Fang, S., Shang, X., and Guo, W. (2018). Ectopic expression of GhCOBL9A, a cotton glycosyl-phosphatidyl inositol-anchored protein encoding gene, promotes cell elongation, thickening and increased plant biomass in transgenic *Arabidopsis*. *Mol. Genet. Genomics* 293, 1191–1204. doi: 10.1007/s00438-018-1452-3

Obembe, O. O., Jacobsen, E., Timmers, J., Gilbert, H., Blake, A. W., Knox, J. P., et al. (2007). Promiscuous, non-catalytic, tandem carbohydrate-binding modules modulate the cell-wall structure and development of transgenic tobacco (*Nicotiana tabacum*) plants. *J. Plant Res.* 120, 605–617. doi: 10.1007/s10265-007-0099-7

Park, S.-H., Ong, R. G., and Sticklen, M. (2016). Strategies for the production of cell wall-deconstructing enzymes in lignocellulosic biomass and their utilization for biofuel production. *Plant Biotechnol. J.* 14, 1329–1344. doi: 10.1111/pbi.12505

Park, Y. W., Baba, K., Furuta, Y., Iida, I., Sameshima, K., Arai, M., et al. (2004). Enhancement of growth and cellulose accumulation by overexpression of xyloglucanase in poplar. *FEBS Lett.* 564, 183–187. doi: 10.1016/S0014-5793(04)00346-1

Pawar, P. M.-A., Derba-Maceluch, M., Chong, S.-L., Gandla, M. L., Bashar, S. S., Sparrman, T., et al. (2017). In muro deacetylation of xylan affects lignin properties and improves saccharification of aspen wood. *Biotechnol. Biofuels* 10:98. doi: 10.1186/s13068-017-0782-4

Pawar, P. M.-A., Derba-Maceluch, M., Chong, S.-L., Gómez, L. D., Miedes, E., Banasiak, A., et al. (2016). Expression of fungal acetyl xylan esterase in *Arabidopsis thaliana* improves saccharification of stem lignocellulose. *Plant Biotechnol. J.* 14, 387–397. doi: 10.1111/pbi.12393

Peña, M. J., Zhong, R., Zhou, G.-K., Richardson, E. A., O'Neill, M. A., Darvill, A. G., et al. (2007). *Arabidopsis* irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19, 549–563. doi: 10.1105/tpc.106.049320

Petersen, P. D., Lau, J., Ebert, B., Yang, F., Verhertbruggen, Y., Kim, J. S., et al. (2012). Engineering of plants with improved properties as biofuels feedstocks by vessel-specific complementation of xylan biosynthesis mutants. *Biotechnol. Biofuels* 5:84. doi: 10.1186/1754-6834-5-84

Pogorelko, G., Fursova, O., Lin, M., Pyle, E., Jass, J., and Zabotina, O. A. (2011). Post-synthetic modification of plant cell walls by expression of microbial hydrolases in the apoplast. *Plant Mol. Biol.* 77, 433–445. doi: 10.1007/s11103-011-9822-9

Poovaiah, C. R., Mazarei, M., Decker, S. R., Turner, G. B., Sykes, R. W., Davis, M. F., et al. (2015). Transgenic switchgrass (*Panicum virgatum* L.) biomass is increased by overexpression of switchgrass sucrose synthase (PvSUS1). *Biotechnol. J.* 10, 552–563. doi: 10.1002/biot.201400499

Reem, N. T., Pogorelko, G., Lionetti, V., Chambers, L., Held, M. A., Bellincampi, D., et al. (2016). Decreased polysaccharide feruloylation compromises plant cell wall integrity and increases susceptibility to necrotrophic fungal pathogens. *Front. Plant Sci.* 7:630. doi: 10.3389/fpls.2016.00630

Ren, Y., Hansen, S. F., Ebert, B., Lau, J., and Scheller, H. V. (2014). Site-directed mutagenesis of IRX9, IRX9L and IRX14 proteins involved in xylan biosynthesis: glycosyltransferase activity is not required for IRX9 function in *Arabidopsis*. *PLoS One* 9:e105014. doi: 10.1371/journal.pone.0105014

Rennie, E. A., and Scheller, H. V. (2014). Xylan biosynthesis. *Curr. Opin. Biotechnol.* 26, 100–107. doi: 10.1016/j.copbio.2013.11.013

Roudier, F., Fernandez, A. G., Fujita, M., Himmelspach, R., Borner, G. H. H., Schindelman, G., et al. (2005). COBRA, an *Arabidopsis* extracellular glycosyl-phosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. *Plant Cell* 17, 1749–1763. doi: 10.1105/tpc.105.031732

Safra-Dassa, L., Shani, Z., Danin, A., Roiz, L., Shoseyov, O., and Wolf, S. (2006). Growth modulation of transgenic potato plants by heterologous expression of bacterial carbohydrate-binding module. *Mol. Breed.* 17, 355–364. doi: 10.1007/s11032-006-9007-4

Sanderson, M. A., and Adler, P. R. (2008). Perennial forages as second generation bioenergy crops. *Int. J. Mol. Sci.* 9, 768–788. doi: 10.3390/ijms9050768

Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315

Somerville, C., Youngs, H., Taylor, C., Davis, S. C., and Long, S. P. (2010). Feedstocks for lignocellulosic biofuels. *Science* 329, 790–792. doi: 10.1126/science.1189268

Stein, O., and Granot, D. (2019). An overview of sucrose synthases in plants. *Front. Plant Sci.* 10:95. doi: 10.3389/fpls.2019.00095

Sumiyoshi, M., Nakamura, A., Nakamura, H., Hakata, M., Ichikawa, H., Hirochika, H., et al. (2013). Increase in cellulose accumulation and improvement of saccharification by overexpression of arabinofuranosidase in rice. *PLoS One* 8:e78269. doi: 10.1371/journal.pone.0078269

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

Terrett, O. M., and Dupree, P. (2019). Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Curr. Opin. Biotechnol.* 56, 97–104. doi: 10.1016/j.copbio.2018.10.010

Tomassetti, S., Pontiggia, D., Verrascina, I., Reca, I. B., Francocci, F., Salvi, G., et al. (2015). Controlled expression of pectic enzymes in *Arabidopsis thaliana* enhances biomass conversion without adverse effects on growth. *Phytochemistry* 112, 221–230. doi: 10.1016/j.phytochem.2014.08.026

United Nations Environment Programme, (2007). *Global Environment Outlook: Environment for Development*, 4th Edn. New York, NY: UNEP.

Urbanowicz, B. R., Peña, 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

Vega-Sánchez, M. E., Loqué, D., Lao, J., Catena, M., Verhertbruggen, Y., Herter, T., et al. (2015). Engineering temporal accumulation of a low recalcitrance polysaccharide leads to increased C6 sugar content in plant cell walls. *Plant Biotechnol. J.* 13, 903–914. doi: 10.1111/pbi.12326

Wang, G., Gao, Y., Wang, J., Yang, L., Song, R., Li, X., et al. (2011). Overexpression of two cambium-abundant Chinese fir (*Cunninghamia lanceolata*) (-expansin genes ClEXPA1 and ClEXPA2 affect growth and development in transgenic tobacco and increase the amount of cellulose in stem cell walls. *Plant Biotechnol. J.* 9, 486–502. doi: 10.1111/j.1467-7652.2010.00569.x

Wang, J., Nayak, S., Koch, K., and Ming, R. (2013). Carbon partitioning in sugarcane (*Saccharum* species). *Front. Plant Sci.* 4:201. doi: 10.3389/fpls.2013.00201

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

Wei, Z., Qu, Z., Zhang, L., Zhao, S., Bi, Z., Ji, X., et al. (2015). Overexpression of poplar xylem sucrose synthase in tobacco leads to a thickened cell wall and increased height. *PLoS One* 10:e0120669. doi: 10.1371/journal.pone.0120669

Willis, J. D., Smith, J. A., Mazarei, M., Zhang, J.-Y., Turner, G. B., Decker, S. R., et al. (2016). Downregulation of a UDP-Arabinomutase gene in switchgrass (*Panicum virgatum* L.) results in increased cell wall lignin while reducing arabinose-glycans. *Front. Plant Sci.* 7:1580. doi: 10.3389/fpls.2016.01580

Wu, A.-M., Hörnblad, E., Voxeur, A., Gerber, L., Rihouey, C., Lerouge, P., et al. (2010). Analysis of the *Arabidopsis* IRX9/IRX9-L and IRX14/IRX14-L pairs of glycosyltransferase genes reveals critical contributions to biosynthesis of the hemicellulose glucuronoxylan. *Plant Physiol.* 153, 542–554. doi: 10.1104/pp.110.154971

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

Yang, Y., Yoo, C. G., Guo, H.-B., Rottmann, W., Winkeler, K. A., Collins, C. M., et al. (2017a). Overexpression of a domain of unknown function 266-containing protein results in high cellulose content, reduced recalcitrance, and enhanced plant growth in the bioenergy crop populus. *Biotechnol. Biofuels* 10:74. doi: 10.1186/s13068-017-0760-x

Yang, Y., Yoo, C. G., Winkeler, K. A., Collins, C. M., Hinchee, M. A. W., Jawdy, S. S., et al. (2017b). Overexpression of a domain of unknown function 231-containing protein increases O-xylan acetylation and cellulose biosynthesis in populus. *Biotechnol. Biofuels* 10:311. doi: 10.1186/s13068-017-0998-3

Zhou, Y., Li, S., Qian, Q., Zeng, D., Zhang, M., Guo, L., et al. (2009). BC10, a DUF266-containing and Golgi-located type II membrane protein, is required for cell-wall biosynthesis in rice (*Oryza sativa* L.). *Plant J.* 57, 446–462. doi: 10.1111/j.1365-313X.2008.03703.x

Conflict of Interest: AB and HS are inventors on a patent application related to the work described here on xylan reduction.

frontiers in Plant Science

REVIEW
published: 02 April 2020
doi: 10.3389/fpls.2020.00299

White Mustard (*Sinapis alba* L.) Oil in Biodiesel Production: A Review

Petar M. Mitrović[1], Olivera S. Stamenković[2], Ivana Banković-Ilić[2], Ivica G. Djalović[1], Zvonko B. Nježić[3], Muhammad Farooq[4,5,6], Kadambot H. M. Siddique[6]* and Vlada B. Veljković[2,7]*

[1] Institute of Field and Vegetable Crops, Novi Sad, Serbia, [2] Faculty of Technology, University of Niš, Leskovac, Serbia, [3] Institute of Food Technology, University of Novi Sad, Novi Sad, Serbia, [4] Department of Crop Sciences, College of Agricultural and Marine Sciences, Sultan Qaboos University, Al-Khoud, Oman, [5] Department of Agronomy, University of Agriculture, Faisalabad, Faisalabad, Pakistan, [6] UWA School of Agriculture and Environment, The UWA Institute of Agriculture, The University of Western Australia, Perth, WA, Australia, [7] Serbian Academy of Sciences and Arts, Belgrade, Serbia

White mustard (*Sinapis alba* L.) seed oil is used for cooking, food preservation, body and hair revitalization, biodiesel production, and as a diesel fuel additive and alternative biofuel. This review focuses on biodiesel production from white mustard seed oil as a feedstock. The review starts by outlining the botany and cultivation of white mustard plants, seed harvest, drying and storage, and seed oil composition and properties. This is followed by white mustard seed pretreatments (shelling, preheating, and grinding) and processing techniques for oil recovery (pressing, solvent extraction, and steam distillation) from whole seeds, ground seed or kernels, and press cake. Novel technologies, such as aqueous, enzyme-assisted aqueous, supercritical CO_2, and ultrasound-assisted solvent extraction, are also discussed. The main part of the review considers biodiesel production from white mustard seed oil, including fuel properties and performance. The economic, environmental, social, and human health risk/toxicological impacts of white mustard-based biodiesel production and use are also discussed.

Keywords: biodiesel, white mustard seed, oil recovery, transesterification, *Sinapis alba* L

OPEN ACCESS

Edited by:
Mengzhu Lu,
Chinese Academy of Forestry, China

Reviewed by:
Abdurrahman Saydut,
Dicle University, Turkey
Steven Lim,
Tunku Abdul Rahman University, Malaysia

****Correspondence:***
Ivica G. Djalović
maizescience@yahoo.com;
ivica.djalovic@ifvcns.ns.ac.rs
Kadambot H. M. Siddique
kadambot.siddique@uwa.edu.au

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *15 October 2019*
Accepted: *28 February 2020*
Published: *02 April 2020*

Citation:
Mitrović PM, Stamenković OS, Banković-Ilić I, Djalović IG, Nježić ZB, Farooq M, Siddique KHM and Veljković VB (2020) White Mustard (Sinapis alba L.) Oil in Biodiesel Production: A Review. Front. Plant Sci. 11:299. doi: 10.3389/fpls.2020.00299

INTRODUCTION

Sinapis alba L. (white or yellow mustard, also known as *Brassica hirta*) is an annual plant of the family Brassicaceae that originates from the Mediterranean region (Katepa-Mupondwa et al., 2005). It is found worldwide as a cultivated plant species as well as a weed. It is a winter–spring plant that can be grown in short cycles, commonly in rotation with other cereal crops, with the possibility of second-crop cultures (Falasca and Ulberich, 2011). In Europe, white mustard is the most used mustard species (Monsalve et al., 1993) and in North America, it is the only species in commercial production for the food processing and condiment industries (Katepa-Mupondwa et al., 2005).

White mustard has many cropping applications, including edible oilseeds (Raney et al., 1995), fast-growing salads (Rahman et al., 2018), condiments, fodder, and green manure (Krstić et al., 2010). The plant can extract toxic heavy metals from soil (Kos et al., 2003). Young seedling leaves, which are rich in vitamin A, C, and E, are edible as fresh and tasty salad leaves and have a medicinal value to purify blood (Rahman et al., 2018). White mustard seed has significant agronomic value due to its high protein and oil contents and low starch content (Balke and Diosady, 2000). Its well-balanced amino acid profile makes the seed an attractive source of food-grade proteins. It is

widely used as a binding agent and protein extender in meat processing and for hot dog mustard, mayonnaise, and salad dressings. The seeds have strong disinfectant properties and can be used as a food preservative (Rahman et al., 2018). Its essential oil can be used to preserve foods due to its potent antimicrobial activity (Peng et al., 2014). Industrially, white mustard seed oil is used as a lubricant and for lighting (Falasca and Ulberich, 2011). Moreover, the seed is used in traditional medicine for its anti-tumor, antiviral, and analgesic activities (Peng et al., 2013); it also has expectorant, stimulant, and antimicrobial activities that are useful for digestive and respiratory ailments (Sujatha et al., 2013).

Recently, white mustard seed oil has garnered interest for its use as a feedstock for biodiesel production (Ahmad et al., 2008; Issariyakul et al., 2011; Issariyakul and Dalai, 2012; Ciubota-Rosie et al., 2013; Sáez-Bastante et al., 2016). The oil itself can be used as an alternative fuel (Nieschlag and Wolff, 1971; Alam and Rahman, 2013). In addition, oil meal—a byproduct of the biodiesel industry from white mustard seeds—can be used for animal feed (Thacker and Petri, 2009) or further extracted to produce additional oil, thus improving economic benefits.

This review focuses on biodiesel production from white mustard seed oil as a feedstock and starts by discussing the botany, cultivation, and use of white mustard plants, seed harvest, drying and storage, and seed oil composition, properties, and uses. This is followed by the pretreatment of white mustard seeds (shelling, preheating, grinding) and processing techniques for oil recovery (pressing, solvent extraction, steam distillation) from whole seeds, ground seeds or kernels, and oil meals (press cakes). Novel technologies, such as aqueous, enzyme-assisted aqueous, supercritical CO_2, and ultrasound-assisted solvent extraction, are also covered. The main part of the review considers biodiesel production from white mustard seed oil, including its fuel properties and performance. The economic, environmental, social, and human health risk/toxicological impacts of white mustard-based biodiesel production and use are also discussed.

WHITE MUSTARD

Mustard Production

White mustard is an annual, broad-leaved, yellow-flowered, cool-season plant from the Brassicaceae family, that grows up to 100 cm high, with a relatively short growing season of about 85–95 days (**Figure 1**). The flowers, which bloom from May to June, are yellow with four petals. Mustard tolerates drought and heat, so it is well suited to production in drier soil zones. It is typically grown in rotation with cereal crops for its young leaves, seeds, or green manure. The use of white mustard in crop rotations is desirable due to its effect on residue conditions in the field, soil moisture conditions, and disease, weed, and insect problems. Ideally, white mustard is grown after a cereal crop. Mustard is commonly grown on summer fallow or stubble in dry and moist areas, respectively. Varietal selection involves various factors, including expected price, yield potential, and agronomic characteristics.

FIGURE 1 | White mustard (*Sinapis alba* L.): **(a)** experimental field, **(b)** leaves, **(c)** flowers, and **(d)** seeds.

Young mustard leaves have a sharp flavor and are used in salads. The seeds are yellow to yellowish-brown and reveal an odor only when mixed with liquid. The heat and aroma of mustard come from sinalbine (a glycoside) and the essential oil, respectively. The advantages of cultivating white mustard as green manure include the long-term supply of soil organic matter, stable soil structure, and increased soil fertility, capacity for soil water storage, humus content, and soil microorganism activity.

White Mustard Seed Production

White mustard is globally cultivated on 60,000–80,000 ha annually, producing up to 685,000 t of seed (**Figure 2A**). Nepal and Canada are the world's top mustard seed producers, with 159,710 t and 121,600 t, respectively, in 2017 (**Figure 2B**); these numbers represent around 27.6% and 21.0%, respectively, of world supply. Therefore, the global average production yield of white mustard seeds is 770–930 kg/ha. In Serbia, the average yield of mustard seeds is 1,500–1,800 kg/ha, with yields often exceeding 2,000 kg/ha. In comparison, the average yields of corn and rapeseed from 2015–2018 were 4,500 kg/ha and 2,800 kg/ha, respectively.

The amount of nitrogen fertilizer applied to the soil has a greater effect on grain yield and harvest index in white mustard than plant density (Sáez-Bastante et al., 2016). White mustard produces less oil per hectare than rapeseed or soybean. In addition, nitrogen fertilizer dose has been positively correlated with total oil extracted. Both plant density and nitrogen fertilizer dose influence fatty acid composition.

Knowledge of thermal and physical properties is essential for identifying appropriate processing equipment and optimizing transport and storage conditions. Additionally, specific heat capacity, thermal conductivity, and thermal diffusivity are important for determining the sensory quality of food products (Ikegwu and Ezeh, 2012; Mahapatra et al., 2013) and heat transfer characteristics (Sirisomboon and Posom, 2012; Jibril et al., 2016). Furthermore, the physical properties (bulk density, true density, porosity, surface area, length, and width) of food products and

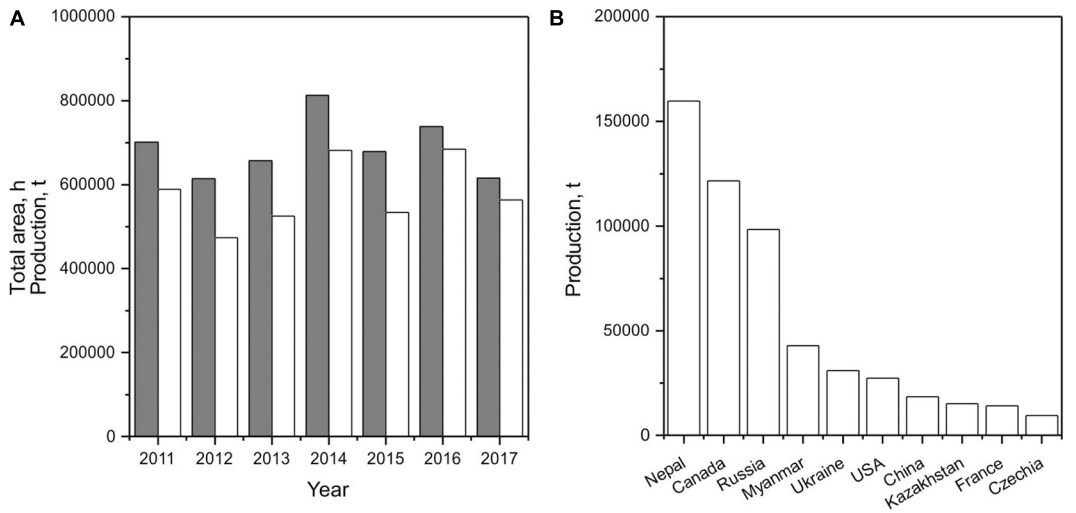

FIGURE 2 | Mustard seed production in the world **(A)** and the world's top producing countries **(B)** (gray, total area in ha; white, production in tons; source: http://www.fao.org/faostat/en/).

mechanical behavior under compression are needed to design processing equipment and identify the optimal conditions for harvesting, handling, sorting, storing, and processing (heating, drying, and cooling) seeds (Tavakoli H. et al., 2009; Tavakoli M. et al., 2009; Sangamithra et al., 2016). The thermal properties of white mustard seeds are cultivar-dependent, but the mechanical properties are not (Ropelewska et al., 2018). The minimum force required to break the white mustard seed coat and the average deformation are 17.5 N and 0.21 mm, respectively (Szczyglak and Zuk, 2012). Excessive breaking or deforming force reduces the quality of the processed seeds and increases shelling costs.

WHITE MUSTARD OIL PRODUCTION

The overall process of white mustard oil production consists of seed harvesting, pre-cleaning, drying, storage and pretreatment, and oil recovery, refinement, and packaging. White mustard seed processing is schematically presented in **Figure 3**.

Harvesting, Drying, and Storage of White Mustard Seed

White mustard plants should be harvested when the leaves start to yellow, and the pods start to turn brown (McKenzi and Carcamo, 2010). The pods must not be left on the plant for too long as they shatter when fully ripe, resulting in the loss of seed. Since white mustard plants are relatively resistant to pod shattering, they can be swathed or straight combined (McKenzi and Carcamo, 2010). For swathing, at least 75% of the seeds should be yellow (Sask Mustard, 2019). Harvesting can be undertaken manually using sickles or with a combine harvester. The seeds are either removed from the pods by hand, or the flower heads are placed in paper bags for a couple of weeks prior to seed maturity. A gentle shake of the bags releases most of the seeds. Modern combines have eliminated the need for hand-cutting plants. The use of a swath roller minimizes seed losses from wind damage by compacting the swath. When threshing with a combine, the lowest cylinder speed should be used to reduce pod cracking. Threshing should be carried out when the seed moisture content is less than 9.5% (McKenzi and Carcamo, 2010).

After seed harvest, caution is needed to preserve the oil quality by avoiding fat breakdown. Seeds are first pre-cleaned from external impurities, such as dust, plant leaves, stones, and ferrous particles (ABC Machinery, 2019). The removal of impurities allows for a high-quality product and prolonged life. For safe storage, seeds are commonly dried to remove water and to ensure high-quality oil recovery. Seeds are either dried by the sun or using hot-air convective drying—air and seed temperatures should not exceed 65°C and 45°C, respectively (Sask Mustard, 2019). Seeds are dried to less than 9% moisture content (McKenzi and Carcamo, 2010).

Dried seeds can be stored below 18°C for an extended period with appropriate aeration and precautions against rodent and insect infestation (Sask Mustard, 2019). Adequate ventilation or aeration of seeds during storage will maintain a low moisture content and reduce the risk of microbial development and general seed deterioration. Some companies crush and grind the seeds with roller mills before solvent extraction, which is then passed through sieves to separate the shells and bran.

Pretreatment of White Mustard Seeds

In traditional processing, white mustard seeds are crushed and ground to a flour. In industrial oil production, the pretreatment involves shelling (decortication) and preheating (ABC Machinery, 2019). The oil-bearing portion of the seed (kernels) is separated from the shell by hand or a shelling machine, which gently cracks the seeds. The kernels are preheated by roasting or cooking to liquefy the oil in the oily organs and facilitate its release during recovery. Preheating

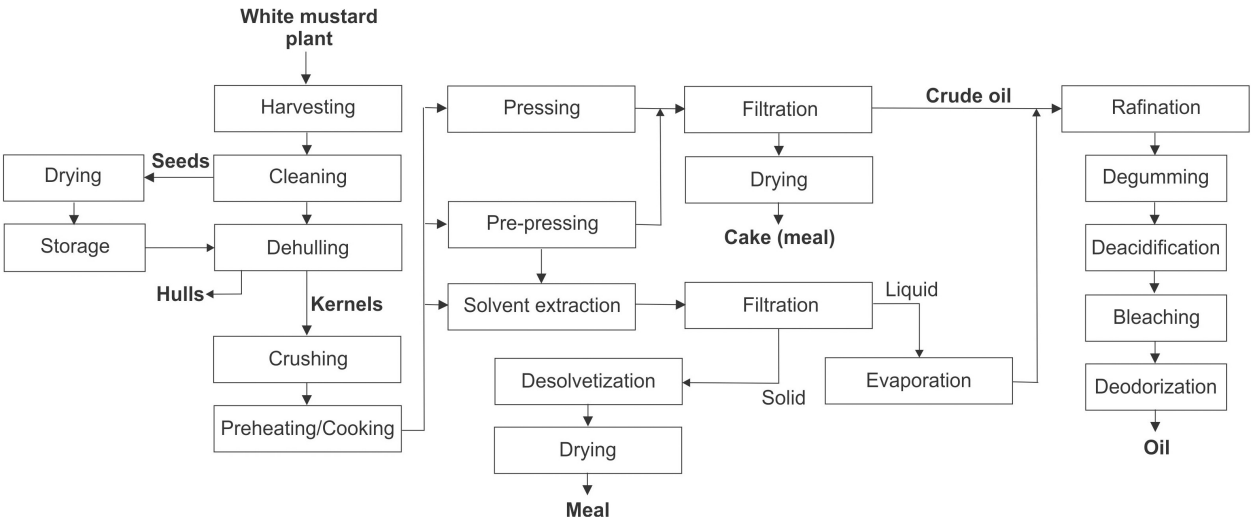

FIGURE 3 | Schematic presentation of white mustard seed processing.

increases the oil yield and protein availability in the seed cake. Machines offer easier and faster production of white mustard oil than manual operations. Before oil extraction, the kernels are usually ground to increase surface area and maximize oil yield.

Oil Recovery From White Mustard Seed

White mustard oil is recovered from the seed either by pressing (expelling), solvent extraction preceding grinding, enzymatic extraction, aqueous extraction, or a combination of pressing and solvent extraction (prepress–solvent extraction), while steam distillation is rarely used. An overview of the methods used for the recovery of the white mustard oil is in **Table 1**. While whole seeds are used for mechanical oil extraction, white mustard seeds are usually dried and ground before solvent extraction or steam distillation. After recovery, the solvent is commonly removed from the oil by vacuum evaporation. The pressed oil is then filtered and dried by heating or vacuum evaporation. Sometimes, the white mustard seed oil is subjected to acid degumming, neutralization, and solid separation. The white mustard oil machine has made oil extraction easier, thus making the oil more affordable.

Oil Pressing

Oil pressing extracts oil from whole seeds by physical (mechanical) pressing. Mechanical oil recovery from whole white mustard seeds involves cold pressing (Ciubota-Rosie et al., 2013; Stamenković et al., 2018), hot pressing (Nie et al., 2016), or expelling (Ahmad et al., 2013; Sultana et al., 2014). Oil pressing has many advantages, including simplicity, ease of operation, flexibility, and fewer processing operations and machines, and produces good quality oil and cake. However, it is less efficient than the Soxhlet extraction (Stamenković et al., 2018). The cake from oil pressing contains up to 8–9% oil (Ciubota-Rosie et al., 2013). A commercial white mustard oil pressing line usually includes a sheller, cleaning sieve, cooker, oil expeller, and filter (ABC Machinery, 2019).

Solvent Oil Extraction

For solvent extraction, white mustard seeds are dried, ground, and then subjected to extraction. Various solvents and extraction techniques are used, including the Soxhlet extraction with petroleum ether (Ali and McKay, 1982; Yaniv et al., 1994) or *n*-hexane (Seal et al., 2008; Ciubota-Rosie et al., 2013; Singh et al., 2014; Kozlowska et al., 2016; Sáez-Bastante et al., 2016; Stamenković et al., 2018), Smalley-Butt extraction (Seal et al., 2008), traditional maceration with *n*-hexane (Stamenković et al., 2018), shaking extraction using chloroform/methanol (Kozlowska et al., 2016), continuous one-step maceration with *n*-hexane (Ciubota-Rosie et al., 2009), ultrasonic extraction with diethyl ether, ethyl acetate, and petroleum ether (Peng et al., 2013), supercritical CO_2 extraction (Barthet and Daun, 2002; Seal et al., 2008), aqueous extraction (Ataya Pulido, 2010; Jung and Diosady, 2012; Tabtabaei and Diosady, 2012, 2013; Tabtabaei et al., 2014), and enzymatic aqueous extraction (Tabtabaei and Diosady, 2013). After extraction, the solvent is usually removed from the oil by evaporation under reduced pressure. Besides the Soxhlet apparatus, the Butt tube extractor (Seal et al., 2008) and the FOSFA method (Barthet and Daun, 2002) have been used to measure oil content in white mustard seeds.

Soxhlet extraction

The Soxhlet extraction is a traditional method for oil extraction that has several disadvantages, including the use of costly, non-selective, hazardous, and flammable high-purity organic solvents, toxic emissions during extraction, and a laborious and time-consuming procedure (Gayas and Kaur, 2017). However, this method provides the highest oil yields from ground white mustard seeds due to the high solvent:seed ratio (usually 10:1 mL/g), long processing time (usually 6 h or longer), and high extraction temperature (boiling point). Using *n*-hexane (53°C),

TABLE 1 | Summary of white mustard seed oil extraction methods and results.

Method	Material	Pretreatment of seed		Extraction	Refinement		References
		Drying	Milling	Technique/solvent/yield/time	Solvent or solid residue removal	Drying	
Pressing	Seed	No	No	Press with 8 mm nozzles/13.28%	Vacuum filtration	Vacuum evaporation	Stamenković et al., 2018
	Seed	Oven (<70°C)	Grinder cutting (0.13–0.50 mm)	Pressing (5–8 kg/h)/oil content of cake reduced to 8–9%	Acid degumming and chemical deacidification, followed by centrifugation	Vacuum evaporation	Ciubota-Rosie et al., 2013
	Seed	–	No	Press extractor	Filtration	–	Mejia-Garibay et al., 2015
	Seed	–	No	Electric oil expeller	Electric filter apparatus	Heating at 120°C using a hot plate	Ahmad et al., 2013
	Seed	–	No	Electric oil expeller	Suction filtration	Heating above 100°C for 1 h	Sultana et al., 2014
	Seed	–	No	Pressing while applying heat (drive speed setting: 5–6)	Fiberglass filter disk under vacuum	–	Nie et al., 2016
Solvent extraction, batch	Seed (moisture contents: 3.78%)	No	Electric milling (1 min) before extraction (mean particle diameter: 0.44 mm)	Soxhlet apparatus/*n*-hexane (seed:solvent 1:10 g/mL)/20.64%/6 h	Vacuum filtration	Vacuum evaporation	Stamenković et al., 2018
			Manual crushing and electric milling (1 min) of press cake (mean particle diameter: 0.47 mm)	Soxhlet apparatus/*n*-hexane (seed:solvent 1:10 g/mL)/8.58%/6 h			
	Seed	Oven (<70°C)	Grinder cutting (0.13–0.50 mm)	Soxhlet apparatus/*n*-hexane (55°C)/39.2%/6 h or 43.5/18 h	Acid degumming and chemical deacidification, followed by centrifugation	Vacuum evaporation	Ciubota-Rosie et al., 2013
	Seed	–	Pestle and mortar	Soxhlet apparatus/*n*-hexane/31.6%/24 h	–	Rotary evaporation under vacuum	Singh et al., 2014
	Seed	–	Coffee mill	Soxhlet apparatus/*n*-hexane (70°C; 10:1 v/w)/25.30%/8 h	Anhydrous magnesium sulfate placed over a filter paper	Rotary evaporator at 40°C	Kozlowska et al., 2016
				Extraction (Folch method, room temperature, shaking)/ chloroform/methanol (2:1 v/v; 10:1 v/w)/29.80%/2 h	Whatman No. 1 paper filter into a separatory funnel with 1 M KCl solution; after gentle shaking, mixture left overnight for separation into two phases	Rotary evaporation under vacuum at 40°C	
	Seed	–	Pestle and mortar	Soxhlet apparatus/*n*-hexane	–	Rotary evaporation	Sáez-Bastante et al., 2016
	Dry, clean seeds	–	Crushing	Soxhlet apparatus/petroleum ether (60–80°C)/32.5%	–	Vacuum evaporation (<40°C)	Ali and McKay, 1982
	Seed	Dried overnight (50°C)	–	Soxhlet apparatus/petroleum ether (60°C, seed:solvent 5 g:100 mL)/21.1 ± 0.4%/16 h	–	Evaporation	Yaniv et al., 1994

(Continued)

TABLE 1 | Continued

Method	Material	Pretreatment of seed		Extraction	Refinement		References
		Drying	Milling	Technique/solvent/yield/time	Solvent or solid residue removal	Drying	
	Seed	–	Coffee mill	Smalley-Butt apparatus/*n*-hexane/∼30%/4 h+2 h	–	Vacuum evaporation	Seal et al., 2008
	Seed	–	Grinder	Magnetically stirred beaker/*n*-hexane/overnight	Vacuum filtration	Vacuum evaporation	Nie et al., 2016
Solvent extraction, continuous	Seed	–	Grinding (0.2–0.3 mm)	Continuous stirred extractor (10:1 *n*-hexane:seed mass, 29°C, 30 h)/78% oil recovery	–	–	Ciubota-Rosie et al., 2009
Aqueous extraction	Dehulled flour	–	–	4:1 water:flour (g/g), 3 min blending, pH 11.00 ± 0.05, room temperature, 30 min extraction, centrifugation (∼9000 *g* for 20 min), re-extraction/washing stage of solid residue under same conditions/75% oil recovery	–	–	Balke, 2006
	Partially dehulled flour	–	–	Procedure of Balke (2006) (39% oil recovery in emulsion) and oil extraction from emulsion using isopropyl alcohol: Single extraction (25:1–31:1 isopropyl alcohol:oil mass)/90–94% oil recovery	–	–	Ataya Pulido, 2010
				Three-stage extraction (3:1 water:flour)/94% oil recovery			
				Four-stage extraction (2:1 water:flour)/96.3% oil recovery			
	Pre-ground dehulled flour (<100 mesh)	–	–	Procedure of Balke (2006) Four-stage PSE using fresh isopropyl alcohol at each stage (2:1 solvent:oil)/92.3% oil recovery			Jung and Diosady, 2012
				Four-stage PSER reusing extracted water-rich phase, containing isopropyl alcohol (2:1 solvent:oil)/84.0% oil recovery			
	Dehulled flour	–	–	Procedure of Balke (2006) Fully alkaline two-stage aqueous extraction (64.6% oil recovery in emulsion) and oil extraction from emulsion using cyclic ethers.	–	–	Tabtabaei and Diosady, 2012
				Tetrahydrofuran (4:1)/97% oil recovery and 5% water			
				1,4-dioxane (9:1)/95% oil recovery and 99% water			
	Dehulled flour	–	–	Procedure of Balke (2006) (57.5% oil recovery in emulsion) and oil extraction from emulsion using organic solvents with partial solubilities for oil	–	–	Tabtabaei et al., 2013
				30:1 dimethylformamide:oil mass/38% oil recovery			
	Dehulled flour	–	–	Fully alkaline two-stage aqueous extraction (flour slurry, 4:1 water:flour mass) + three-stage emulsion extraction with:	–	–	Tabtabaei et al., 2014
				0.75:1 tetrahydrofuran /100% oil recovery and 3.5% water			
				0.5:1 1-4-dioxane/85.8% oil recovery and 76% water			
	Dehulled flour (<100 mesh)	–	–	Fully alkaline two-stage aqueous extraction (flour slurry, 4:1 water:flour mass) + emulsion extraction with tetrahydrofuran (0.5:1 and 0.75:1, producing miscella I and II, respectively) (Tabtabaei et al., 2014) + adsorption (zeolite 4A)	–	–	Tabtabaei et al., 2015

(Continued)

TABLE 1 | Continued

Method	Material	Pretreatment of seed		Extraction	Refinement		References
		Drying	Milling	Technique/solvent/ yield/time	Solvent or solid residue removal	Drying	
				Single-stage aqueous extraction with tetrahydrofuran (4:1, producing miscella III) (Tabtabaei and Diosady, 2012) + adsorption (zeolite 4A)			
				Batch adsorption (10:1.5 miscella:zeolite, shaking rate 125 cycles/min), water removal/time: miscella I 72.4%/40 min, miscella II 98.8%/4 h, and miscella III 98.7%/4 h			
				Fixed-bed adsorption (2.5 cm column dimeter, 30 cm length), water removal/breakthrough time/bed capacity: miscella I (flow rate 1.6 mL/min) 100/44.6 h/0.222 g/g; miscella II (flow rate 2.0 mL/min) 98.6%/14.2 h/0.244 g/g			
Enzymatic aqueous extraction	Dehulled flour	–	–	Sequential two-stage aqueous extraction (flour slurry, 4:1 water:flour mass): (1) pH 4.8–5.0, 25°C, 30 min; (2) pH 11, 30 min, room temperature/30.0% oil, 58.7% water, 8.3% protein, and 5.2% phospholipids	–	–	Tabtabaei and Diosady, 2013
				Sequential two-stage enzymatic aqueous extraction: (1) pH 4.8, 40°C, 3% carbohydrase of flour mass (Viscozyme L, Pectinex Ultra SP-L, Celluclast each 1%), 3 h; (2) pH 11, room temperature, 3 h/35.3% oil, 52.2% water, 9.1% protein, and 5.9% phospholipids			
				Sequential two-stage aqueous extraction+emulsion extraction: (3) 3:1 water:emulsion mass, pH 11, 25°C, 30 min/80.0% oil, 20.6% water, 0.6% protein, and 0.6% phospholipids			
				Sequential two-stage aqueous extraction+emulsion extraction: (4) 3:1 water:emulsion mass, pH 11, 25°C, 3 h/80.1% oil, 20.8% water, 0.7% protein, and 0.8% phospholipids			
				Fully alkaline two-stage aqueous extraction (Tabtabaei and Diosady, 2012)/55.5% oil, 39.8% water, 3.2% protein, and 3.3% phospholipids			
Ultrasound-assisted extraction	Seed	Dry	Grinding (18 mesh)	Flask in ultrasonic batch (room temperature)/diethyl ether (10 g seed:40 mL solvent)/8.96%/30 min	Filtration under reduced pressure	Evaporation	Peng et al., 2013
Supercritical CO_2 extraction	Seed	–	Mill equipped with 1.0-mm sieve	Single extraction (2 mL CO_2/min, 51.7 MPa, 100°C, 60 min)/21.59 ± 0.29%	–	–	Barthet and Daun, 2002
				Multiple extraction (5 × 20 min)/28.63 ± 0.57%			
				Multiple extraction with modifier (2 × 30 min + 30 min with 15% ethanol)/28.60 ± 0.49%			
	Seed	–	Mill with diatomaceous earth	Single extraction (41.4 MPa, 80°C, 90 min)/∼30%	–	Vacuum evaporation	Seal et al., 2008
				Two-step extraction with modifier (60 min with 15% ethanol)/∼30%			

(Continued)

TABLE 1 | Continued

Method	Material	Pretreatment of seed		Extraction	Refinement		References
		Drying	Milling	Technique/solvent/yield/time	Solvent or solid residue removal	Drying	
Pressing + solvent extraction	Seed	Oven (<70°C)	Grinder cutting (0.13–0.50 mm)	Pressing (5–8 kg/h) + solvent extraction (*n*–hexane)/> 40% (92% of total oil)/2 h	Acid degumming and chemical deacidification, followed by centrifugation	Vacuum evaporation	Ciubota-Rosie et al., 2013
	Seed	Seed moisture content: 3.78%	Electric milling (1 min) of seeds (mean particle diameter: 0.44 mm)	Seed pressing + Soxhlet extraction (*n*-hexane; 1:10 g/mL seed:solvent)/19.90%/6 h	Vacuum filtration	Vacuum evaporation	Stamenković et al., 2018
			Manual crushing and electric milling (1 min) of press cake (mean particle diameter: 0.47 mm)	Seed pressing + maceration (*n*-hexane; 70°C; 1:6.5 g/mL seed:solvent)/20.48%/5 min			
				1:8.5 g/mL seed:solvent/20.50%/5 min			
Steam distillation	Seed	Dry	Grinding (18 mesh)	Steam distillation/water vapor/6.48%	–	–	Peng et al., 2013

Ciubota-Rosie et al. (2013) reported the highest oil yields of 39.2% and 43.5% after 6 h and 18 h of Soxhlet extraction, respectively, while much lower oil yields (21–32%) have been reported by other researchers (Seal et al., 2008; Singh et al., 2014; Stamenković et al., 2018). Petroleum ether extractions resulted in oil yields of 21.2 ± 0.4% from the 11 best lines of white mustard (Yaniv et al., 1994). Soxhlet extraction from a press cake yielded 8.58% oil, which was 41.6% of the oil content from the processed seed (Stamenković et al., 2018). Differences in oil yields of white mustard seeds using Soxhlet extraction are attributed to variations in oil content in cultivars of different geographic origin, particle sizes of processed ground seeds, solubilities of solvents used, and extraction temperatures.

Aqueous Extraction

Aqueous extraction uses water as an extracting solvent to recover oil from oilseeds. It is emerging as an alternative to hexane extraction due to fewer related health, safety, and environmental problems (Rosenthal et al., 1996). It is also beneficial for the simultaneous recovery of oil and high-quality proteins for industrial applications (Tabtabaei and Diosady, 2012). However, low oil yields and the stable emulsion formation have prevented the commercialization of aqueous extraction because an additional demulsification step is needed to recover the oil fully.

Many researchers have used aqueous oil extraction from white mustard seed and flour. Using a two-step process, Balke (2006) achieved the highest oil and protein yields of 85% and 95%, respectively, from fully dehulled white mustard flour. Ataya Pulido (2010) recovered only 39% oil from partially dehulled flour in the form of an oil-in-water emulsion. The low oil yield was ascribed to mucilage, a polysaccharide present in the tested flour, with good emulsifying properties, which prevented oil separation from the solids. Balke and Diosady (2000) developed a rapid aqueous extraction process for mucilage removal from whole white mustard seeds prior to grinding and oil separation. The seed coat, including mucilage, can also be removed readily by mechanical dehulling.

Oil-in-water emulsions are successfully destabilized with a freeze-thaw treatment, while a heat treatment followed by centrifugation and pH adjustment to the isoelectric point of white mustard protein could not break the formed emulsions (Ataya Pulido, 2010). Organic solvents could be used to fully or partially dissolve both oil and water to recover free oil from the emulsion, such as isopropyl alcohol (Ataya Pulido, 2010; Jung and Diosady, 2012) and dimethylformamide with partial solubilities for oil (Tabtabaei et al., 2013), and tetrahydrofuran and 1,4-dioxane with complete solubilities for oil (Tabtabaei and Diosady, 2012).

The cost-effective technologies for recovering miscella with high oil and low water contents from the emulsion are desirable. Tabtabaei et al. (2014) developed multi-stage extractions of the emulsion using lower mass ratios of tetrahydrofuran- or 1,4-dioxane:oil to produce miscellas with low water content. Using 0.5:1 and 0.75:1 tetrahydrofuran:oil mass ratios, 97% of the oil was recovered in the oil-rich miscella that contained only 1% and 1.5% water, respectively. Having low-free fatty acid and phosphorus contents, the produced miscella might be suitable

feedstock for biodiesel production by direct base-catalyzed transesterification. Tetrahydrofuran—added to the oil-in-water emulsions—produces miscellas containing about 1–2% water, thus preventing the direct conversion of the miscella to biodiesel. These miscellas are successfully dehydrated by adsorption on zeolite 4A at room temperature using batch or continuous fixed-bed systems to the water content lower than 0.3% specified for biodiesel feedstock (Tabtabaei et al., 2014).

Based on the weaknesses of the aqueous extraction process mentioned above, an improved emulsion destabilization process that concurrently extracts oil and water in separate phases with enhanced solvent usage efficiency is needed. The reported results have shown that organic solvents with complete solubility for oil (*n*-hexane, petroleum ether, diethyl ether, and ethyl acetate) are more efficient for oil recovery from white mustard emulsion than those having partial solubility, providing high oil recovery at a lower solvent:oil mass ratio.

Enzymatic Aqueous Extraction

The stable oil-in-water emulsions produced by aqueous extraction can be destabilized using different enzymes (alone or in combination) that hydrolyze certain emulsifiers (Jeevan Kumar et al., 2017). This process, known as enzymatic aqueous extraction, is hampered by the high cost of enzyme production and downstream processing, long incubation time, and additional demulsification step.

Tabtabaei and Diosady (2013) applied a three-step process to recover the oil from dehulled white mustard flour using enzymatic aqueous extraction. In the first step, the flour slurry (4:1 water:flour weight ratio) was extracted in the presence of 3% carbohydrases (Viscozyme L, Pectinex Ultra SP-L, Celluclast) at pH 4.8 and 40°C for 3 h. The second step involved alkaline extraction (pH 11) of the solid residue at room temperature for 30 min. The third step was the extraction of oil from the collected emulsion using a 3:1 water:emulsion mass ratio at pH 11 and 25°C for 30 min. A two-stage alkaline aqueous extraction at pH 11 and sequential two-stage aqueous extraction at pH 4.8 and 11 were performed without enzymes using the same procedure. This alkaline treatment produced unstable emulsions and increased oil dispersion into the skim (to about 26%). The protease Protex 6 L treatment (2.5%) recovered >91% oil in the emulsions while the phospholipase treatment had no effect on free oil or protein recovery by isoelectric precipitation. However, the enzymatic aqueous extraction of dehulled white mustard flour does not offer adequate improvements in protein recovery to justify the additional effort and cost.

Novel Solvent Extraction Methods

Novel solvent extraction methods, such as ultrasound-assisted and supercritical CO_2 extraction, have been rarely used for white mustard seed oil recovery, despite having numerous advantages over traditional methods related to time and energy consumption, safety hazards, low-quality oil and meal, environmental risks, and toxicological consequences (Reverchon and De Marco, 2006; Koubaa et al., 2016). Ultrasonication fragments or disrupts the seed particles immersed in the extraction vessel, thus accelerating diffusion, enhancing overall mass transfer, and reducing processing time and temperature (Koubaa et al., 2016). Peng et al. (2013) reported the greater efficiency of ultrasound-assisted oil extraction from white mustard seeds by diethyl ether (8.96% oil yield) than by ethyl acetate (7.63%) or petroleum ether (7.54%). The known benefits of liquid CO_2 are its non-toxic and non-explosive nature, availability, ease of removal, and preservation of oil quality. Seal et al. (2008) used neat CO_2 and a mix of CO_2 and 15% ethanol to extract oil from white mustard seed using a modified two-step process. Both extraction fluids yielded about 30% oil, but the ethanol mix reduced the processing time. Barthet and Daun (2002) improved the efficiency of oil recovery from ground white mustard seeds using multiple extractions. Five consecutive extractions (5 × 20 min) without a modifier or a combination of double extractions with neat CO_2 (2 × 30 min) followed by a third extraction with 15% modifier (30 min) produced higher oil yields (about 28.6%) than a single extraction (21.59%). However, the supercritical CO_2 method extracted 25% less oil than the standard FOSFA method, suggesting seed matrix effects on the oil extraction.

Other Oil Extraction Methods

In addition to the above-mentioned major methods, the oil can be also recovered from white mustard seed by steam distillation, a two-step process combining oil pressing and solvent extraction and continuous single-stage solvent extraction. Steam distillation yielded 6.48% oil from white mustard seed, which was less than that from ultrasound-assisted extraction (8.96%) (Peng et al., 2013). However, steam distillation has high equipment costs, is time-consuming, and controlling the temperature is difficult (Gayas and Kaur, 2017). In the two-step process, seeds are first pressed to remove most of the oil and then the residual oil is extracted from the press cake using a solvent. Stamenković et al. (2018) combined pressing with either Soxhlet extraction or maceration using *n*-hexane to extract total oil yields of 19.6% and 20.5%, respectively. Ciubota-Rosie et al. (2013) reported that pressing followed by solvent extraction using *n*-hexane (for 2 h) yielded >41% oil and produced a press cake with low oil content (<2%). The continuous single-stage extraction recovers 78% oil from ground white mustard seeds using *n*-hexane (Ciubota-Rosie et al., 2009). This extraction was modeled using the generalized reduced gradient method to determine the optimum conditions for maximum efficacy.

Comparison of Various Oil Recovery Methods

Excluding comparisons with standard methods, the various oil recovery methods have rarely been compared. For instance, Peng et al. (2013) reported higher efficiency of oil extraction using ultrasound-assisted extraction with diethyl ether than with steam distillation. Stamenković et al. (2018) compared cold pressing, Soxhlet extraction, and combined pressing and solvent extraction. Oil yields obtained from a Soxhlet extraction using *n*-hexane (25.30 ± 1.24%) for 8 h was lower than that from the Folch method (29.80 ± 2.95%) using chloroform/methanol (2:1, v/v) with 2 h shaking at room temperature (Kozlowska et al., 2016). This result was attributed to the extraction of polar

TABLE 2 | Oil yields obtained from various oil sources by different extraction methods.[a]

Extraction method	Source	Oil yield[b] (g/100 g)	%
Soxhlet	Seed	20.64 ± 0.18	100.0
Soxhlet	Press cake	8.58 ± 0.06	41.6
Cold pressing	Seed	13,28 ± 0.11	64.3
Cold pressing/Soxhlet	Seed/press cake	19.90 ± 0.04	96.4
Cold pressing/maceration (70°C, 6.5:1, 5 min)	Seed/press cake	20.48	99.2
Cold pressing/maceration (70°C, 8.5:1, 5 min)	Seed/press cake	20.50	99.3

[a]*Adapted from Stamenković et al. (2018);* [b]*Mean value ± standard deviation.*

materials (phospholipids), apart from neutral triacylglycerols. As shown in **Table 2**, seed cold pressing followed by press cake maceration under optimal extraction conditions recovered > 99% oil, which was close to the reference Soxhlet method and much higher than that of seed cold pressing alone (41.6%).

Table 3 summarizes the results from a study by the University of Toronto on the destabilization of emulsions from dehulled white mustard flour using several methods. At lower solvent:oil mass ratios, the single extraction using tetrahydrofuran or 1,4-dioxane recovered more oil than dimethylformamide or isopropyl alcohol. Similar results were obtained with multiple-stage extractions using much lower solvent:oil mass ratios with both types of solvents. The use of recycled isopropyl alcohol recovered less oil than the other solvents, but the water content in the oil-rich phase decreased substantially due to improved oil and water separation. The isopropyl alcohol usage efficacy, as represented by oil extracted per isopropyl alcohol used, increased by a factor of 2.4 when the recycled solvent was used, which would reduce processing costs (Tabtabaei and Diosady, 2013).

Mechanisms, Optimization, Kinetics, and Thermodynamics of Ground Press Cake Maceration

Only Stamenković et al. (2018) have studied the mechanisms, optimization, kinetics, and thermodynamics of oil extraction from ground press cake, remaining after pressing whole white mustard seeds by maceration using *n*-hexane.

Mechanism of Ground Press Cake Maceration

Typically, for oily plant material, maceration of ground press cake at a constant temperature increases oil yields rapidly within the first minute before decelerating (up to about the 5th min) to reach a plateau (next 10 min) (Stamenković et al., 2018). Maceration reached practical saturation within 5 min. The speedy first-extraction step (oil washing) grinds and washes out the oil from the external surfaces of seed particles. In the second step (oil diffusion), the oil diffuses from the interior of the particles and dissolves in the solvent.

Optimization of Ground Press Cake Maceration

A response surface 3D plot was used to visualize the effects of the process factors and their interactions on the oil yield obtained within 5 min (Stamenković et al., 2018). Generally, increasing both the extraction temperature and solvent-to-seed cake ratio increased oil yield. The maximum oil yield was achieved at an extraction temperature close to 70°C and a solvent:seed cake ratio of between 6.5:1 and 8.5:1 mL/g. Taking the lowest solvent amount as the criterion of choice, the selected optimal extraction conditions were 6.5:1 mL/g, 70°C, and 5 min. The best-predicted oil yield of 7.29 g/100 g matched the actual oil yield (7.20 ± 0.13 g/100 g), being 84% of the oil yield obtained by Soxhlet extraction.

Kinetics of Ground Press Cake Maceration

Based on the observed extraction mechanism, the kinetics of oil maceration is described by the simplified phenomenological model (Stamenković et al., 2018):

$$q = q_\infty[1 - (1 - f) \cdot e^{-k_d \cdot t}] \quad (1)$$

where q is oil yield at time t, q_∞ is maximum oil yield at saturation, f is the fraction of oil extracted by washing (washable oil) and k_d is diffusion rate constant. This model supposes instantaneous oil washing, followed by oil diffusion. At $t = 0, f = q/q_\infty$.

The saturation oil yield, q_∞, increases with an increasing extraction temperature and solvent:seed cake ratio due to increased oil solubility at higher temperatures and an increased amount of solvent that dissolves a larger amount of the oil, respectively. The washable oil fraction, f, increases with increasing solvent:seed ratio at a constant temperature due to the positive effect of the increased amount of solvent on washing. The diffusion rate constant, k_d, increases with an increasing solvent:seed cake ratio and extraction temperature, which was attributed to the reduced viscosity of the liquid phase. Also, mass transfer was facilitated at higher solvent:seed cake ratios by the increased concentration driving force. The modified Arrhenius equation was used to correlate the diffusion rate constant with solvent:seed cake ratio and temperature. The activation energy value (5.99 kJ/mol) was close to that for the hempseed oil maceration by *n*-hexane (5.75 kJ/mol) (Kostić et al., 2014).

Thermodynamics of Ground Press Cake Maceration

Maceration thermodynamics involves analysis of enthalpy, entropy, and Gibbs free energy changes, as well as the temperature extraction coefficient. This analysis is based on oil content determined by the Soxhlet extraction, oil yield during press cake maceration, and oil content in the exhausted press cake at saturation, which are used to calculate the distribution coefficient (K_D). According to the Van't Hoff equation, the dependence of $\ln K_D$ on the reciprocal absolute temperature at different solvent:seed ratios is linear, allowing the calculation of enthalpy, entropy, and Gibbs free energy changes. The enthalpy and entropy changes for the maceration of ground press cake using *n*-hexane were positive, ranging from 5.2–12.5 kJ/mol and 29–47 J/(mol K), respectively (Stamenković et al., 2018). This implies that ground press cake maceration is endothermic and

TABLE 3 | Summary of emulsion destabilization results.

Extraction process[a]	Demulsification[b]	Solvent[c]	Solvent:oil mass ratio	Recovery by final oil extract (%)		Final oil extract composition (%)			References
				Oil recovery	Water recovery	Oil	Water	Solvent	
Two-stage alkaline	None	None	–	–	–	57.5 ± 4.0	38.4 ± 3.9		Tabtabaei and Diosady, 2012
	Single extraction	Tetrahydrofuran	4:1	97.2 ± 0.9	<2	23	5	72	
		1,4-dioxane	9:1	95 ± 3	99	9	7	84	
		Dimethylformamide	30:1	38 ± 3	–	1	3	96	Tabtabaei et al., 2013
		Isopropyl alcohol	31:1	94.0	6.1	2.8	2.1	95.0	Ataya Pulido, 2010
	Three-stage extraction	Tetrahydrofuran	0.75:1	100	3.5	55.9 ± 1.5	1.5 ± 0.1	42.6 ± 1.6	Tabtabaei et al., 2014
		1,4-dioxane	0.5:1	85.9	0	86.5 ± 0.5	0.2 ± 0.0	13.3 ± 0.5	
	Four-stage extraction	Isopropyl alcohol	2:1	97	100.0	10.0	7.5	82.5	Ataya Pulido, 2010
		Isopropyl alcohol	0.2:1	92.3 ± 2.3	4.6 ± 0.6	21.1 ± 1.4	0.6 ± 0.1	78.3 ± 1.4	Jung and Diosady, 2012
		Isopropyl alcohol (recycled)	0.2:1	84.0 ± 0.9	1.0 ± 0.1	92.9 ± 0.8	0.6 ± 0.0	6.5 ± 0.8	
	None	None	–	<5 (pH 2–11)[d] 94.7 ± 4.2[e] 4.0 ± 0.2[f]	–	55.5 ± 4.6	39.8 ± 3.9	–	Tabtabaei and Diosady, 2013
Sequential two-stage	None	None	–	30 (pH 4–6)[d] 92.7 ± 2.6[e] 22.5 ± 0.7[f]	–	30.0 ± 2.5	58.7 ± 2.2	–	
	Single extraction	Water	3:1	96.5 (pH 4.5)[d] 97.8 ± 0.9[e] 99.4 ± 0.6[f]	–	80.0 ± 3.4	20.6 ± 3.3	–	
Sequential two-stage enzymatic	None	None	–	51 (pH 3)[d] 91.3 ± 2.4[e] 41.2 ± 1.7[f]	–	35.3 ± 2.1	52.2 ± 1.0	–	
	Single extraction	Water	3:1	91.1 (pH 4.5)[d] 94.6 ± 0.1[e] 94.3 ± 0.1[f]	–	80.1 ± 2.2	20.8 ± 3.8	–	

[a]Two-stage alkaline extraction: Water:flour ratio: 4:1 g/g; blending time: 3 min; pH 11.00 ± 0.05; room temperature; extraction time: 30 min; centrifugation (~9000 g for 20 min); re-extraction/washing stage of solid residue under same conditions; Sequential acid/alkaline two-stage: Slurry (water:flour ratio: 4:1 g/g) first extracted at native pH (4.8–5.0) and 25°C for 30 min followed by second-stage alkaline extraction (pH 11) for another 30 min at room temperature; Sequential two-stage – enzymatic: Slurry (water:flour ratio: 4:1 g/g) first extracted at pH 4.8 and 40°C in presence of 3% carbohydrases of the flour mass for 3 h, followed by second-stage alkaline extraction (pH 11) at room temperature for 30 min. [b]None – no emulsion destabilization; Single extraction: single-stage emulsion destabilization by adding a solvent; multiple extractions: multiple-stage emulsion destabilization by adding a solvent. [c]None – no solvent was added. [d]After pH adjustment. [e]After protease treatment (Protex 6 L, 2.5%, pH 9.0, 60°C, 3 h). [f]After phospholipase treatment (G-ZYME G999, 2.5%, pH 7.5, 40°C, 3 h).

TABLE 4 | Comparison of thermodynamic quantities for oil extraction from different oily materials.[a]

Plant material	*T* (°C)	Δ*S*° (J/mol K)	Δ*H*° (kJ/mol)	Δ*G*° (kJ/mol)	References
White mustard seed cake	20–70	29-47	5.2-12.5	−4.8 to −1.4	Stamenković et al., 2018
Olive cake	20–50	12.9	59.3	–6.3 to –4.5	Meziane and Kadi, 2008
Soybean flakes	50–100	48.2–95.4	137–296	–10 to –4	Rodrigues et al., 2010
Sunflower seeds	30–60	11.2	36.75–39.60	–1.1 to –0.8	Topallar and Gecgel, 2000
Cotton seeds	15–45	43.2–85.8	190.9–331.3	–21.0 to –10.4	Saxena et al., 2011
Hemp seeds	20–70	6.17–10.54	33.09–44.19	–5.17 to –2.41	Kostić et al., 2014

[a]*Adapted from Stamenković et al. (2018).*

irreversible. Values of the enthalpy and entropy changes for oil maceration from ground white mustard cake are similar to those for oil extraction from olive cake, hemp seeds, and sunflower seeds but much lower than those for cotton seeds and soybean flakes (**Table 4**).

The negative Gibbs free energy change (from –4.8 kJ/mol to –1.4 kJ/mol) means that the process is feasible and spontaneous (Stamenković et al., 2018). The spontaneity of ground press cake maceration is favored by an increasing solvent:seed cake ratios and maceration temperatures.

The temperature extraction coefficient defines the increase in oil yield for every 10°C increase in extraction temperature. For white mustard seed cake, its values were 1.040, 1.021, and 1.011 for maceration at solvent:seed cake ratios of 3:1, 6.5:1, and 10:1 mL/g, respectively (Stamenković et al., 2018). These values are similar to those reported for oil extraction from olive cake (1.02–1.14) (Meziane and Kadi, 2008) and hemp seeds (1.012–1.027) (Kostić et al., 2014).

Oil Refinement

The final step in oil extraction is the refining process, which is comprised of several operations, including degumming, alkali treatment, bleaching, and deodorization (Vaisali et al., 2015). Degumming removes phosphatides and mucilaginous gum while the alkali-refining treatment eliminates free fatty acids, color bodies, and metallic pro-oxidants. Bleaching removes pigments and residual soaps and improves the taste of the oil. Deodorization is carried out through high-vacuum steam distillation to remove unwanted odor and taste from the degummed and/or neutralized oil.

Fatty Acid Profile and Physicochemical Properties of White Mustard Oil

The oil contents, fatty acid profiles, and physicochemical properties of white mustard oil obtained from seeds and press cake by various extraction techniques are in **Table 5**. Generally, all oils contain the same fatty acids, thus indicating no influence of the extraction method on their composition. For seeds from Serbia, the content of total saturated fatty acids (SFA) is very low (2.0–4.1%) and increases by thermal treatment during the Soxhlet extraction (Stamenković et al., 2018). Among the SFAs, palmitic acid is the most abundant (about 0.7–3.4%), with an exceptionally high content of palmitic acid reported for a Canadian oil (23.7%). The monounsaturated fatty acids (MUFA) include oleic (C18:1), eicosenoic (C20:1), erucic (C22:1), and nervonic (C24:1) acids, with erucic acid the most abundant (32.8–60.3%). The high content of erucic acid is a unique property of white mustards originating from various regions, particularly Europe. Oleic acid is the second most abundant fatty acid with a content of 9.1–43.4%. The highest oleic acid content is characteristic for white mustard oils from North and South America. The primary polyunsaturated fatty acids (PUFA) of white mustard oils are linoleic (C18:2) and linolenic (C18:3) acids. The ratio of oleic to linoleic fatty acids (stability index) in white mustard oil depends on the extraction technique and decreases with thermal treatment as the content of linoleic acid increases. For the same reason, the linoleic acid/linoleic acid ratio increases after thermal treatment.

Generally, the extraction technique has little effect on the physicochemical properties of white mustard oil, except for the acid value, which is higher for oil obtained through solvent extraction (Stamenković et al., 2018). This is attributed to the higher temperature of solvent extraction affecting oil acidity caused by the hydrolysis of acylglycerols (Adeeko and Ajibola, 1990). However, the acid and iodine values of oils depend on the oily feedstock (seed or press cake), which are higher for the oil from press cake (Stamenković et al., 2018) due to the pressing, milling, and solvent extraction process that increases free fatty acid formation (Adeeko and Ajibola, 1990). The saponification value depends on neither extraction technique nor feedstock (Stamenković et al., 2018).

BIODIESEL PRODUCTION FROM WHITE MUSTARD SEED OIL

White mustard seed oil is a promising oily feedstock for biodiesel production (Ciubota-Rosie et al., 2013). In many countries, it is considered unsuitable for human consumption (Wendlinger et al., 2014). While white mustard seed can be used as a spice, its widespread use in the food industry is restricted by its strong, hot taste and high erucic acid content. Therefore, its use as an alternative feedstock for biodiesel production will not compete with its use as human food. Indeed, the transesterification of erucic acid provides alkyl esters with great lubricant properties for better engine operation (Issariyakul et al., 2011). Furthermore, white mustard can grow spontaneously on abandoned land or under cultivation, typically in rotation with cereal crops (Falasca and Ulberich, 2011; Rahman et al., 2018). It can also grow on different soil types, is resistant to many diseases and insect pests, and can endure extreme weather conditions without substantial

TABLE 5 | Variability of oil content, fatty acid profile and physicochemical properties in white mustard seed and press cake of different origin.[a]

Origin	Serbia				Romania		Spain	Great Britain	Israel	India	Canada			Mexico
Extraction technique	**Soxhlet/ *n*–hexane**	**Cold pressing**	**Soxhlet/ *n*–hexane[b]**	**Maceration/ *n*–hexane[b]**	**Soxhlet/ *n*–hexane**	**Cold pressing**	**Soxhlet/ *n*–hexane**	**Soxhlet/ petroleum ether**	**Soxhlet/ petroleum ether**	**Soxhlet/ *n*–hexane**	**Soxhlet/ *n*–hexane**	**Commercial**	**Cold pressing**	**Cold pressing**
Yield, %	20.6	13.2	8.58	7.20	43.5	~ 35	25	–	19.5	31.6	35.1	–	–	22.3
C14:0	–	–			–	–	–	–	–	–	0.3	0.05	–	–
C16:0	0.86	0.73	1.09	1.72	1.6	1.5	2.82	3.1	3.0	3.36	23.7	2.80	2.81	2.10
C16:1	–	–	0.11	0.14	–	–	–	–	–		–	0.16	–	0.09
C16:2	–	–			–	–	–	–	–		–	0.06	–	–
C18:0	0..35	0.3	0.38	0.61	0.7	0.9	–	0.7		1.12	1.6	1.09	1.06	0.80
C18:1	11.63	13.95	12.60	14.86	12.4	12.0	17.61	9.1	15.8	22.12	43.4	26.08	24.89	19.62
C18:2	6.03	5.98	7.46	8.86	12.0	12.3	7.82	11.7	9	10.78	30.1	11.64	9.21	8.43
C18:3	7.00	7.37	8.03	8.76	8.7	8.9	10.99	12.5	8.6	12.52	0.2	8.61	10.8	21.64
C20:0	0.33	0.28	0.35	0.52	0.7	0.6	–	0.7			0.6	0.70	–	0.41
C20:1	7.00	7.41	7.32	9.59	6.7	6.6	5	10.8	5.8	11.91	–	10.44	10.63	*nd*
C20:2	0.22	0.17	0.24	0.29	0.3	0.3	–	0.7	–		–	–	–	0.25
C22:0	0.57	0.39	0.56	0.78	0.7	0.6	–	Tr	–		–	0.57	–	–
C22:1	60.29	59.98	56.21	49.81	55.0	55.0	55.76	46.5	50.8	38.16	–	32.81	34.94	40.80
C22:2	0.40	0.32	0.42	1.26	0.5	0.4	–	0.4	–		–	–	–	–
C24:0	0.50	0.30	0.69	0.52	0.1	0.1	–	Tr	–		–	–	–	nd
C24:1	4.79	2.95	4.59	2.31	0.6	0.6	–	3.6	–		–	–	–	1.25
SFA, %	2.61	2.00	3.05	4.14	3.80	3.70	2.82	4.50	3.00	4.48	26.20	5.21	3.87	3.31
MUFA,	83.7	84.3	80.82	76.70	74.7	74.2	78.4	70.0	72.4	72.2	43.4	69.5	70.5	61.8
PUFA, %	13.7	13.8	16.14	19.17	21.5	21.9	18.8	25.3	17.6	23.3	30.3	20.3	20.0	30.3
ALC	20.9	20.8	19.5	20.8	20.4	20.4	20.3	20.2	18.8	19.7	17.5	18.6	18.5	18.8
TUD, %	118.0	119.3	121	119	126.4	126.9	127.0	133.1	116.2	131.3	104.2	118.4	121.3	144.0
OLR	1.93	2.33	1.69	2.13	1.03	0.98	2.25	0.78	1.76	2.05	1.44	2.24	2.70	2.33
LLR	0.86	0.81	0.93	0.84	1.38	1.38	0.71	0.94	1.05	0.86	–	1.35	0.85	0.39
AV	2.73	1.95	4.09	4.43	1.58	1.23	–	1.23	–	–	2.19-	0.85	–	–
SV	180.72	180.65	178.29	179.65	–	–	–	174	–	184.7	175	–	–	–
IV	101.78	100.58	107.49	108.21	102.3	102.3	–	105.4	–	112	106.2-	–	–	–
CV	–	–			–	–	–	–	–	50.6	–	–	–	–
Reference	Stamenković et al., 2018				Ciubota-Rosie et al., 2013		Sáez-Bastante et al., 2016	Ali and McKay, 1982	Yaniv et al., 1994	Singh et al., 2014	Sengupta and Bhattacharyya, 1996	Issariyakul et al., 2011	Nie et al., 2016	Mejia-Garibay et al., 2015

[a]*Adapted from Stamenković et al. (2018). SFA, saturated fatty acids; MUFA, monosaturated fatty acids; PUFA, polyunsaturated fatty acids; ALC, average length chain; TUD, total unsaturation degree; OLR, oleic/linoleic ratio; LLR, linoleic /linoleic (ω–6/ω–3) ratio; AV, acid value; SV, saponification value; IV, iodine value and CV; cetane value.* [b]*Press cake.*

harm (Sask Mustard, 2019). Considering the maximum grain production, the oil content and its conversion into methyl esters, Sáez-Bastante et al. (2016) estimated the white mustard oil-based biodiesel production of about 480-486 L ha^{-1}. This biodiesel output is similar to the biodiesel production from soybean oil (400-500 L ha^{-1}) but much lower than the outputs from sunflower (1,000 L ha^{-1}), rapeseed (1,200 L ha^{-1}), or palm (5,000 L ha^{-1}) oil (Worldwatch, 2006). Recently, Jaime et al. (2018) have shown that white mustard could replace rapeseed for biodiesel production in the Mediterranean basin and other Western European countries where their cultivation is expected to be compromised by climate change. Another advantage of white mustard, compared to the other traditional feedstocks for biodiesel production, is the possibility îf its growth on abandoned land with marginal cultivation requirements.

The literature reveals that white mustard oil is used as a feedstock for biodiesel production through a transesterification reaction in the presence of homogeneous or heterogeneous base catalysts while no acid catalyst has been applied. The homogeneous base-catalyzed transesterification is suitable for biodiesel production from white mustard oil due to its low acid value. Indeed, biodiesel has been mainly produced from white mustard oil using homogeneous base catalysts, while a heterogeneous catalyst was applied only in a study (**Table 6**). It can be speculated that the researchers have preferred to use homogeneous base catalysts because they provide fast reactions and a high ester yield under mild reaction conditions, compared to both homogeneous acid and solid base catalysts. It may be expected that solid base catalysts will get more attention in the future because of their well-known advantages over homogeneous catalysts (for instance, easy recovery from the reaction mixture and reusability).

A summary of biodiesel production from white mustard oil is in **Table 6**. In these reactions, triacylglycerols (TAG) from white mustard oil are converted into fatty acid alkyl esters, most frequently fatty acid methyl esters (FAME). Generally, as is the case for other oily feedstocks (Živković et al., 2017), the reaction efficiency and ester yield are influenced by many factors, including the type of alcohol, initial alcohol:oil molar ratio, type and amount of catalyst, reaction temperature, mixing intensity, and reaction time.

Base-Catalyzed Transesterification of White Mustard Seed Oil

To date, biodiesel production from white mustard oil has mainly used homogeneous base catalysts (**Table 6**). Alkali hydroxides (KOH and NaOH) are more frequently used than alkali methoxides. NaOH is catalytically more active than KOH, providing higher ester yields (92% vs. 84%) under the same reaction conditions (Sultana et al., 2014). The use of catalysts, ranging from 0.1 to 1.8% (based on oil mass), under different reaction conditions resulted in various methyl ester yields, making it difficult to select the optimal catalyst concentration. Generally, low catalyst amounts will not complete the reaction, while excess catalyst amounts favor soap formation, both of which result in lower ester yields (Yesilyurt et al., 2019). Catalyst amounts of about 1% (based on oil mass) are used most often. Methanol is the main alcohol used, but others include ethanol, propanol, and 1-butanol (Issariyakul et al., 2011). The methanol:oil molar ratio ranges from 2:1 to 14:1, with 6:1 the most frequently applied ratio. At lower methanol amounts, the reaction reaches equilibrium at lower FAME contents, whereas higher amounts result in faster reactions and higher final ester contents. This is attributed to a shift in the reaction equilibrium toward TAG conversion and the reduced density and viscosity of the reaction mixture (Kostić et al., 2018). However, excess alcohol can cause difficulties in glycerol separation from the esters phase, lowering FAME yield (Yesilyurt et al., 2019). Glycerol formation has a small effect on ester yield, as the two-step reaction with glycerol removal in between only increased ester yield by 2% (Issariyakul et al., 2011). The transesterification reaction is carried out at different temperatures (22–75°C) but most frequently at close to alcohol boiling point. The FAME content in the esters phase is dependent on the quality of the white mustard oil (Ciubota-Rosie et al., 2013) and the reaction conditions. Generally, FAME purity and yields are higher with commercial (refined) oils (Oshodi et al., 2014) than crude oils (Ahmad et al., 2008, 2013; Sultana et al., 2014). Ciubota-Rosie et al. (2013) reported that achieving FAME contents above 80% was difficult with crude oil due to the presence of phosphorus compounds, gums, and free fatty acids that emulsify or cause sediment, making ester synthesis, separation, and purification difficult.

Quicklime is the only heterogeneous catalyst that has been used for methanolysis of white mustard oil (Kostić et al., 2018). At optimal reaction conditions, the quicklime-catalyzed reaction is slower than the KOH-catalyzed reaction due to greater mass-transfer limitations in the three-phase reaction system, which controls the overall process rate. A heterogeneous reaction carried out with a methanol:oil molar ratio of 12:1 and 10% quicklime (based on oil mass) for 50 min provided almost the same TAG conversion as the KOH-catalyzed reaction under milder reaction conditions for a shorter time (methanol:oil molar ratio 6:1, 1% KOH, and 20 min). However, heterogeneous reactions are a cheaper, more straightforward, and more environmentally friendly process (Kostić et al., 2018).

Modeling and Optimization of Biodiesel Production From White Mustard Oil

Useful tools for improving biodiesel production from any oily feedstock include statistical modeling and optimization. Ester yield is influenced by the reaction conditions, namely initial alcohol:oil molar ratio, catalyst type and loading, reaction temperature, and time. Therefore, it is important to know the impact of these process factors on ester yield and optimization. Response surface methodology (RSM), combined with specific experimental designs, has been widely used to optimize biodiesel production from various oily feedstocks, but rarely for improving white mustard oil-based biodiesel production (Kostić et al., 2018; Yesilyurt et al., 2019).

Yesilyurt et al. (2019) applied a central composite design combined with the RSM to analyze and optimize methanolysis of white mustard oil catalyzed by NaOH in an experimental domain

TABLE 6 | Review of the alcoholysis reaction of white mustard oil.

Type of alcohol	Alcohol:oil molar ratio	Catalyst/amount (% of oil)	Temperature (°C)	Type, volume of reactor/ agitation speed (rpm)	Yield (purity) (%/time)	References
Methanol	2:1–10:1	NaOH/0.2–1.0	50–70	Three-neck flask, 250 mL/magnetic/800 rpm	96.5%/62.12 min[a]	Yesilyurt et al., 2019
Methanol	6:1	KOH/1	40–60	Stirred reactor, –/600 rpm	–	Issariyakul and Dalai, 2012
Methanol	6:1	CH_3OK/0.2:1[b]	60	Stirred reactor, 500 mL/–	(>98)/1.5 h; (99)/4 h	Ciubota-Rosie et al., 2013
Methanol	6:1	NaOH/0.8	70	Stirred reactor, 2 L/–	(82)/2 h	Ahmad et al., 2013
Methanol	2:1–10:1	NaOH/0.1–0.9	50–75	Stirred reactor, 500 mL/600 rpm	92/75 min[c]	Sultana et al., 2014
	6:1	KOH/0.5	65		84/75 min[c]	
Methanol[c]	14:1	NaOH/1.2	Room	Stirred reactor, 200 mL/magnetic	(99.3)/10 min[d]	Tabtabaei et al., 2015
Methanol	12% oil relative to KOH/ methanol solution	KOH/0.5[e]	22	–	–	Nie et al., 2016
Methanol	25:6 mL/mL	KOH/1.8	65	Erlenmeyer flask/magnetic/ 300 rpm	96.56/2 h	Oshodi et al., 2014
Methanol	6:1	KOH/0.3	60	–	–/2 h	Sarala et al., 2012
Methanol Ethanol, propanol, 1-butanol	6:1	KOH/1 CH_3ONa/0.5 and 1	60	Stirred reactor, –/600 rpm	(66)/1.5 h[f]	Issariyakul et al., 2011
					(66)/1.5 h[f,g]	
Methanol	–	NaOH (150 mL, 1 M)	55	Glass container, jerked/–	–/5 min	Alam and Rahman, 2013
Methanol	6:1–8:1	NaOH/0.7	60	–	85%/8:1	Ahmad et al., 2008
Methanol	6:1	KOH/1	60	Three-neck flask, 250 mL/magnetic/400 and 900 rpm	(98.7%)/20 min	Kostić et al., 2018
	6:1–12:1	Quicklime/2–10			(98.5%)/50 min[h]	

[a]Optimal conditions (methanol:oil molar ratio 7.41:1, NaOH 0.63% of oil, 61.84°C); [b]Catalyst:oil molar ratio; [c]Optimal conditions (methanol:oil molar ratio 6:1, NaOH 0.5% of oil, 65°C); [d]In the presence of THF (THF:methanol 1:1 mL/mL); [e]wt% to methanol mass; [f]Methanol, two-step process with glycerol removing between steps; [g]1% CH_3ONa; [h]Optimal conditions (methanol:oil molar ratio 12:1, CaO 10% of oil) and after ester purification

(methanol:oil molar ratio 2:1–10:1, NaOH loading 0.2–1.0% of oil mass, reaction temperature 50–70°C, and reaction time 30–90 min). According to the developed second-order polynomial model, the methanol:oil molar ratio, temperature, and time had a significant, positive influence on ester yield, while NaOH loading was statistically insignificant. All the quadratic terms and the interaction of methanol:oil molar ratio with reaction time had significant, adverse effects on ester yield, while the interactions of catalyst loading with reaction temperature and reaction time were significant and positive. The optimal reaction conditions for the highest ester yield were methanol:oil molar ratio 7.4:1, catalyst concentration 0.63 wt.%, reaction temperature 61.84°C, and reaction time 62.62 min. The predicted FAME yield of 96.7% agreed with the experimental value of 96.5%. The quicklime-catalyzed methanolysis of white mustard oil was statistically analyzed and optimized for the methanol:oil molar ratio (6:1–12:1), catalyst amount (2–10%, of oil mass), and reaction time (30–50 min) using to 3^3 full factorial design with replication combined with RSM (Kostić et al., 2018). The experimental data were modeled by a quadratic model, the adequacy and reliability of which was proven by statistical criteria. All three individual factors, as well as the interactions of catalyst amount with methanol:oil molar ratio and reaction time and the quadratic terms for catalyst amount, had a significant influence on ester yield. All three factors also had a positive effect on the FAME content. Based on the reduced quadratic model, complete TAG conversion could be obtained at catalyst amounts >9.8% and methanol:oil molar ratios ranging from 6.1:1 to 11.6:1 for 50 min.

Kinetic Modeling of Biodiesel Production From White Mustard Seed Oil

Although reaction kinetics is fundamental for process design and development, the kinetics of white mustard oil transesterification has been rarely studied (Issariyakul and Dalai, 2012; Kostić et al., 2018). To develop a kinetic model for white mustard oil methanolysis catalyzed by KOH, Issariyakul and Dalai (2012) proposed a mechanism involving three consecutive, reversible reactions following the second-order reaction rate law. To avoid mass transfer limitations in the initial reaction period, the reaction mixture was vigorously stirred (600 rpm). The reaction rate constants of the forward reactions were at least one order of magnitude higher than the reverse rate constants.

Increasing the temperature increased the forward reaction constants, while the effect of temperature on the reverse reaction rate constant was more complex. Based on the values of the rate constants, it was concluded that the rate-determining step was the conversion of TAG to diacylglycerides with an activation energy of 26.8 kJ/mol. The lower activation energy of the white mustard oil methanolysis reaction, compared to palm oil methanolysis (30.2 kJ/mol), was attributed to its lower content of saturated fatty acids. Agreement between experimental and predicted data confirmed the reliability and accuracy of the developed kinetic model.

Kostić et al. (2018) used two different reaction mechanisms for white mustard oil methanolysis over quicklime: (a) the first-order reaction with respect to TAG in the heterogeneous and pseudo-homogeneous regimes and (b) the changing mechanism combined with the TAG mass transfer limitation. The first model confirmed that TAG mass-transfer controlled the reaction in the initial, heterogeneous reaction regime, while in the later pseudo-homogeneous regime, the chemical reaction is the rate-determining step. The volumetric TAG mass-transfer coefficient was dependent on the initial catalyst and methanol concentrations, while the apparent reaction rate was constant. The second model, which combined the changing mechanism and first-order reaction rate law with respect to TAG and FAME, was also verified for the whole reaction period. Its parameters—reaction rate constant and "pure" TAG affinity for the active catalyst sites—were lower than those determined for sunflower oil methanolysis over quicklime, which was attributed to differences in the composition of the oily feedstocks. The mean relative percentage deviations of the TAG conversion degree for the more straightforward and more complex models were 3.0% and 16.1%, respectively, indicating the validity of both kinetic models. Therefore, while the more straightforward model is not applicable in the middle reaction period, it can be successfully used for simulation of white mustard oil methanolysis over quicklime.

FUEL PROPERTIES OF WHITE MUSTARD-BASED BIODIESELS

The most important physicochemical and fuel properties of white mustard-based biodiesels reported in the literature, and standard biodiesel properties according to EN14214, are in **Table 7**. Generally, most of the properties fulfill the standard biodiesel quality, except for purified biodiesels obtained from NaOH- and CH_3OK-catalyzed methanolysis of white mustard oil. The flashpoint, sulfur content, and acid value of white mustard-based biodiesels obtained using NaOH were not in accordance with EN14214 but satisfied the ASTM standard (Ahmad et al., 2013; Sultana et al., 2014). Since there are no reports on exhaust gas emissions for white mustard oil-based biodiesel, this important issue is not considered here.

The higher biodiesel viscosity than the EN14214 standard limit, which could damage the injection system due to poorer atomization of the fuel spray, was attributed to the high molecular weight and large chemical structure of pure biodiesel B100 (Ahmad et al., 2013; Sultana et al., 2014) and the presence of long FAMEs (mainly methyl erucate), which comprised more than 60% of the mustard oil fatty acid profile (Ciubota-Rosie et al., 2013). Due to the lower oxidation stability (2 h) of white mustard oil-based biodiesel than the minimum required by the EN standard (6 h), antioxidants should be added (Ciubota-Rosie et al., 2013). The amount of group II metals (Ca + Mg) in the purified biodiesel from the quicklime-catalyzed white mustard oil transesterification was above the EN 14214 standard limit (Kostić et al., 2018). Hence, the biodiesel purification process should be improved to reduce calcium and magnesium contents further.

Sáez-Bastante et al. (2016) predicted important biodiesel properties using mathematical models based on the chemical properties of hydrocarbon chains. The prediction values for cetane number, density, and cold filter plugging point agreed well with the European standard limits, but kinematic viscosity did not. An increase in the unsaturation degree (i.e., concentrations of linoleic and linolenic acid) improves some biodiesel properties, such as kinematic viscosity and cold filter plugging point, and reduces others, such as cetane number, calorific value, and oxidation stability (Sáez-Bastante et al., 2016). Therefore, a compromise is needed to use white mustard oil as a feasible feedstock for biodiesel production.

The use of different esters (methyl, ethyl, propyl, and butyl) derived from white mustard oil by homogeneous base-catalyzed transesterification was tested for diesel additives (Issariyakul et al., 2011). Most of the properties of the distilled methyl, ethyl, and propyl esters satisfied the European (EN14214) and United States (ASTM D 6751) specifications, but butyl esters had higher acid values than the proposed limits. All of these esters showed potential as a lubricant additive for diesel fuel, particularly methyl esters. Moreover, the diesel/biodiesel blend had a higher lubricant potential than commercial diesel.

OTHER PRODUCTS AND USES OF WHITE MUSTARD

The white mustard plant—aerial parts, seeds, oil and oil components, and essential oil—has a variety of applications in agriculture, food, and other industries, including medicine, culinary, and phytoremediation. The most important uses for white mustard are intercropping (Farooq et al., 2011; Paulsen, 2011; Rahman et al., 2018), biofumigation (Viuda-Martos et al., 2007; Arriaga-Madrid et al., 2017; Berlanas et al., 2018), phytoremediation (Kos et al., 2003; Jankowski et al., 2014; Popoviciu et al., 2017; Bulak et al., 2018), oilseed crop with high-quality properties (Raney et al., 1995), as a protein and amino acid source (Bell et al., 2000; Sarker et al., 2015), and as a condiment crop (Katepa-Mupondwa et al., 2005). White mustard oil also has many industrial applications, such as the production of bio-polyols for synthesis of rigid polyurethane-polyisocyanurate foams (Paciorek-Sadowska et al., 2018), edible biopolymer films for food packaging (Hendrix et al., 2012), and particle and interior boards, including furniture (Dukarska et al., 2011). Also, non-edible white mustard seed oil is used as a lubricant and for lighting (Falasca and Ulberich, 2011).

TABLE 7 | Properties of purified biodiesel obtained by base-catalyzed transesterification of white mustard oils.

	Catalyst (conc.)/Alcohol									
	CaO (10%[a])	KOH (1%[a])	KOH (1%[a])				NaOH (0.5%[a])	NaOH (0.72%[a])	CH_3OK (0.2:1[b])	EN14214 limit
Property	**Methanol**	**Methanol**	**Methanol**	**Ethanol**	**Propanol**	**Butanol**	**Methanol**	**Methanol**	**Methanol**	**(min/max)**
FAME content (%)	98.9	98.7	99.8	99.7	99.7	98.0		82	>98	96.5 min
Density at 15°C (kg/m^3)	881.1	880.1	900	900	900	900	834.3	899	878	860/900
Viscosity at 40°C (mm^2/s)	4.15	4.13	4.2	4.5	5.0	5.5	5.45	6.72	5.67	3.50/5.00
Flash point (°C)							90	110	178	101 min
Sulfur content (mg/kg)			0	14.7	9.5	10.4	43.2	130	0.21	10 max
Cetane number									60	51 min
Water content (mg/kg)	235	217	231	62	187	345			223	500 max
Oxidation stability at 110°C (h)									2	6.0 min
Acid value (mg KOH/g)	0.44	0.47	0.4	0.5	0.6	4.0	0.8		0	0.50 max
Iodine value (g I_2/100 g)	102.9	104.7							102.3	120 max
Methanol content (%)									0	0.20 max
Monoglyceride content (%)	0.5	0.4							0.15	0.80 max
Diglyceride content (%)	0.1	0.1							0.05	0.20 max
Triglyceride content (%)	0.2	0.2	0	0	0	0			0.02	0.20 max
Free glycerol (%)									0.0008	0.02 max
Total glycerol (%)			0.121						0.05	0.25 max
Group I metals (Na+K) (mg/kg)		3.7								5.0 max
Group II metals (Ca+Mg) (mg/kg)	15.5									5.0 max
Phosphorus content (mg/kg)			9	4	10	8			0.87	4.0 max
Cold filter plugging point (°C)									-5	-5 max
Cloud point (°C)							-10	3	5	Not specified
Pour point (°C)							-13	-6		Not specified
Reference	Kostić et al., 2018		Issariyakul et al., 2011				Sultana et al., 2014	Ahmad et al., 2013	Ciubota-Rosie et al., 2013	

[a]*Catalyst:oil weight ratio,* [b]*Catalyst:oil molar ratio.*

The essential oil isolated from white mustard seeds has the potential for food preservation (Graumann et al., 2008). Being a potent natural antioxidant, white mustard seeds can be used to treat many diseases involving free radicals (Thangi et al., 2016). White mustard has many medicinal uses, such as an emetic and diuretic, as well as for treating inflammatory conditions (arthritis and rheumatism), cardiovascular disease, cancer, and diabetes (Khan and Abourashed, 2010); however, there are limited clinical trials to support its use for any indication (Anonymous, 2019). Some parts of the plant can be used as forage, lignocellulosic raw material (Dukarska et al., 2011), green manure (Krstić et al., 2010), or biomass fuel (Maj et al., 2019). Maj et al. (2019) assessed the biomass energy traits of various crop species, including white mustard, intended as forecrops with low gross and net calorific values (lower than agro-biomass or forest biomass). White mustard had the highest heat of combustion (15.55 MJ/kg), indicating its potential as an additional bioenergy source.

ECONOMIC, ENVIRONMENTAL, AND SOCIAL CONSIDERATIONS OF WHITE MUSTARD CULTIVATION, PROCESSING, AND UTILIZATION

Besides technical issues, economic, environmental, social, human health risk/toxicological, and policy aspects are the main issues when assessing the cost-effectiveness and sustainability of white mustard cultivation, processing, and utilization. A few studies have investigated the socio-economic implications of white mustard cultivation and processing (Withers et al., 2000; Sharma, 2015; Vach et al., 2016), but only Tabtabaei et al. (2015) has reported on white mustard oil-based biodiesel production costs.

According to Withers et al. (2000), white mustard is a viable alternative crop for crop rotations for at least three reasons: (a) better utilization of existing equipment on farms, (b) increased diversification, and (c) contribution to weed and disease control with fewer chemicals than other crops. Sharma (2015) analyzed the socio-economic characteristics of mustard growers, cost and return of both mustard cultivation and oil production, and profitability of different patterns of marketing, and suggested how to increase the economic viability of mustard cultivation in the Morena district, India. Mustard cultivation and processing is profitable on every scale, with opportunities to increase yield and profit through the efficient management and adoption of improved seed varieties, fertilizers, and plant-protective chemicals. Vach et al. (2016) investigated the economic efficiency of three crops (white mustard, winter wheat, and spring barley) cultivated using conventional, conservation with minimum tillage, and no-tillage methods. No-tillage produced the highest average white mustard seed yield but seed yields in other tillage methods were not significantly lower. Among the tested crops, white mustard had the lowest profitability. Regarding tillage methods, cost-effective and easy-to-manage systems with lower tillage intensity level should be prioritized.

A major environmental benefit of oil recovery from white mustard seeds is related to the reduction in waste generation. An oil press unit produces no pollutants and thus has no direct relation to environmental benefits. However, any improvements in the efficiency of energy uptake by these units will have indirect environmental benefits, such as reduced electricity consumption. The only waste from the oil press unit is the solid press cake (meal), which can be used as an animal feed, for compost, or as a solid biofuel.

Tabtabaei et al. (2015) analyzed the preliminary costs of an integrated process for producing food-grade protein products, high-purity methyl esters, and fiber-rich solid residue from dehulled white mustard seed flour. This process involved two-stage aqueous processing of white mustard flour at pH 11 followed by membrane separation technologies to produce protein products, namely soluble and precipitated protein isolates, from the protein-rich skim fraction (Tabtabaei and Diosady, 2012, 2013). This analysis compared the cost of white mustard seeds with the values of the protein isolates as primary products and the biodiesel and fiber-rich residue as byproducts (Tabtabaei et al., 2015). While the ultimate process will be selected on its total implementation costs, the minimum requirement for process viability is that the cost of the products recovers the cost of the raw materials.

In the absence of studies directly related to white mustard seed oil-based biodiesel production, many reports on the production and use of biodiesel from other feedstocks can be used to estimate socio-economic, environmental, and toxicological implications (Živković et al., 2017). As biodiesel from other oily feedstocks, white mustard oil-based biodiesel is expected to have several positive impacts on sustainable development, including improvements in energy security, stimulation of economic development, and contribution to environmental protection. The basic requirements for success regarding these impacts include defining policy, objectives, tasks, operating manuals, responsible workers, and deadlines for each step in the manufacturing process.

Economic, environmental, and social implications of white mustard seed oil-based biodiesel production are expected to be the same or similar to those of other oilseed crops (Živković et al., 2017). First, white mustard oil, as a renewable source, will contribute to the substitution of non-renewable diesel fuel. White mustard plants can mitigate climate change by consuming CO_2—a dominant greenhouse gas—during photosynthesis, reducing the negative impacts on air, water, land, and biodiversity, and promoting rural economic development. Second, the use of non-edible white mustard oil as a feedstock for biodiesel production does not contribute to the food versus fuel controversy. Finally, biodiesel production from white mustard oil will be closely connected to agricultural production and may contribute to energy security.

Characteristic pollution parameters during the production of white mustard oil should be similar to those of other crops (Živković et al., 2017). The increased use of mineral fertilizers causes ecological damage, reduces the quality of water for human consumption, and pollutes waterways. In addition, the energy embedded in chemicals (fertilizers, agrochemicals, and methanol) must be included in the life cycle assessment of white mustard-based biodiesel. Non-renewable energy—consumed

during white mustard cultivation and processing, oil extraction, biodiesel production and purification, transportation of raw materials, inputs and distribution—must be considered when evaluating the overall influence of biodiesel on the environment. There are no published life cycle analyses of white mustard oil-based biodiesel production and use.

POTENTIAL AND FUTURE OUTLOOK OF WHITE MUSTARD OIL FOR BIODIESEL

White mustard plant parts, especially seeds, have potential for economically valuable applications and research on their practical application as a bioenergy and bioproduct source. Innovative research will help to improve biodiesel production from white mustard oil and contribute to developing novel processes to produce other biofuels and value-added products. Further investigations are needed to reflect the kinetics of the transesterification of white mustard oil and optimize this reaction for the type and concentration of alcohol and catalyst, type of reactor, and reaction temperature. It is especially important to test the low-cost, active, and stable solid catalysts obtained from natural or waste sources using continuous operation. Novel biodiesel production processes, including unconventional heating methods (ultrasonication and microwave irradiation), continuous reactors with improved mass transfer characteristics, and integration of reaction and separation phases in a single stage, should be developed to advance the economy of the overall biodiesel production process. Possible uses of other parts of the white mustard plant to produce other types of biofuels by liquefaction, gasification, and pyrolysis should be evaluated. These methods face significant challenges for the commercial utilization of white mustard biomass for biofuel production. The generation of high-value products from white mustard should be considered. New emerging technologies that synergistically combine various conversion processes and provide multiple products, called biorefineries, are expected to address the technical and economic obstacles of existing biomass conversion processes. Biorefineries need to develop or improve constituent processes to optimize the integrated system, and provide heat and power supplies for, at least, energy self-sufficiency.

Research is needed to optimize suitable biomass properties without compromising the ability of white mustard plants to grow in diverse environments. Natural genetic variation could be used to improve the bioenergy properties of white mustard plants. In addition, agricultural management (including fertilizer type, time of harvesting, and biomass storage) is critical, and will impact biomass properties. Finally, intensive cultivation of white mustard is needed for making it attractive and economically favorable for biodiesel production on a commercial basis. This is especially important for the Mediterranean basin and Central Europe, as they are expected to become unsuitable for rapeseed in the near future, as shown by the models developed by Jaime et al. (2018). According to these models, the increases in aridity and average annual temperature will expand the climatically appropriate areas for the white mustard cultivation in the Mediterranean basin while favorable areas for the rapeseed cultivation will reduce remarkably in Western Europe. Because of its good potential as a biofuel crop and with potential for genetic improvements, white mustard could replace rapeseed crops for future biodiesel production in the above areas.

CONCLUSION

This review considers all stages of biodiesel production from white mustard seed oil, from seed harvest, drying, storage, and pretreatment via oil recovery to transesterification reaction. White mustard seed oil is a promising feedstock for biodiesel production for several reasons: (a) plants can be cultivated on different soil types, usually in rotation with cereal crops, resists many diseases and insect pests, and endures extreme weather conditions without substantial harm, (b) the oil is considered unsuitable for human consumption in many countries due to its high erucic acid content, (c) the biodiesel has excellent lubricant properties for better engine operation, and (d) biodiesel production can be integrated with protein and oil recoveries into an economically justified process.

AUTHOR CONTRIBUTIONS

PM, OS, IB-I, ID, ZN, MF, KS, and VV wrote the manuscript.

FUNDING

This work has been funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia, Serbia (Project III 45001, Projects 31025, and Project 31073) and partly by the Serbian Academy of Sciences and Arts (Project F-78). This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

ABC Machinery (2019). *Mustard Oil Mill Plant*. Anyang: ABC Machinery.

Adeeko, K. A., and Ajibola, O. O. (1990). Processing factors affecting yield and quality of mechanically expressed groundnut oil. *J. Agric. Eng. Res.* 45, 3–43.

Ahmad, M., Ajub Khan, M., Zafar, M., Hasan, A., Ahmad, Z., Akhter, G., et al. (2008). Base catalyzed transesterification of *Brassica alba* oil and its association with mineralogy to environment friendly biodiesel. *Asian J. Chem.* 20, 6402–6410.

Ahmad, M., Zafar, M., Rashid, S., Sultana, S., Sadia, H., and Ajab Khan, M. (2013). Production of methyl ester (biodiesel) from four plant species of *Brassicaceae*: optimization of the transesterification process. *Int. J. Green Energy* 10, 362–369. doi: 10.1080/15435075.2012.655352

Alam, M. M., and Rahman, K. A. (2013). Biodiesel from mustard oil: a sustainable engine fuel substitute for Bangladesh. *Int. J. Renew. Energ. Dev.* 2, 141–149.

Ali, A., and McKay, J. E. (1982). The chemical and physical characteristics and fatty acid composition of seed oils extracted from *Cruciferous* species cultivated in Pakistan. *Food Chem.* 8, 225–231. doi: 10.1016/0308-8146(82)90044-9

Anonymous (2019). *Mustard*. Available at: https://www.drugs.com/npp/mustard.html (accessed January 31, 2020).

Arriaga-Madrid, D. A., Guevara González, R., Feregrino Pérez, A., Contreras Medina, L. M., and Cortez, P. A. (2017). "Production of glucosinolates in different organs of white mustard plant (*Sinapis alba* L.) as a result of the application of hydrogen peroxide," in *Proceedings of the 13th International Engineering Congress, CONIIN, Universidad Autonoma de Queretaro Santiago de Queretaro*, Mexico.

Ataya Pulido, V. M. (2010). *The Production of a Potential Feedstock for Biodiesel using Water and Isopropyl Alcohol to Extract Yellow Mustard Oil*. M.A.Sc. thesis, University of Toronto, Toronto.

Balke, D. T. (2006). *The Production of Higher Value Food Ingredients from White Mustard Seed Via Aqueous Extraction*. Ph.D. Thesis. University of Toronto, Toronto.

Balke, D. T., and Diosady, L. L. (2000). Rapid aqueous extraction of mucilage from whole white mustard seed. *Food Res. Int.* 33, 347–356. doi: 10.1016/s0963-9969(00)00055-7

Barthet, V. J., and Daun, J. K. (2002). An evaluation of supercritical fluid extraction as an analytical tool to determine fat in canola, flax, solin, and mustard. *J. Am. Oil Chem. Soc.* 79, 245–251. doi: 10.1007/s11746-002-0468-8

Bell, J. M., Rakow, G., and Downey, R. K. (2000). Comparisons of amino acid and protein levels in oil-extracted seeds of *Brassica* and *Sinapis* species, with observations on environmental effects. *Can. J. Anim. Sci.* 80, 169–174. doi: 10.4141/a97-117

Berlanas, C., Andrés-Sodupe, M., López-Manzanares, B., Maldonado-González, M. M., and Gramaje, D. (2018). Effect of white mustard cover crop residue, soil chemical fumigation and Trichoderma spp. Root treatment on black-foot disease control in grapevine. *Pest. Manag. Sci.* 74, 2864–2873. doi: 10.1002/ps.5078

Bulak, P., Lata, L., Plak, A., Wiącek, D., Strobel, W., Walkiewicz, A., et al. (2018). Electromagnetic field pretreatment of *Sinapis alba* seeds improved cadmium phytoextraction. *Int. J. Phytoremediation* 20, 338–342. doi: 10.1080/15226514.2017.1381943

Ciubota-Rosie, C., Diaconescu, R., Volf, I., and Macoveanu, M. (2009). Modeling the extraction process of oil from seeds of white mustard (*Sinapis alba*). *Environ. Eng. Manage. J.* 8, 1429–1432. doi: 10.30638/eemj.2009.208

Ciubota-Rosie, C., Macoveanu, M., Fernández, C. M., Ramos, M. J., Pérez, A., and Moreno, A. (2013). *Sinapis alba* seed as a prospective biodiesel source. *Biomass Bioenergy* 51, 83–90. doi: 10.1016/j.biombioe.2013.01.008

Dukarska, D., Lecka, J., and Szafoni, K. (2011). Straw of white mustard (*Sinapis alba*) as an alternative raw material in the production of particle boards resinated with UF resin. *Acta Sci. Pol.* 10, 19–28.

Falasca, L. S., and Ulberich, A. (2011). "Argentina's semiarid lands aptitude to cultivate non-traditional species for biodiesel production," in *Biodiesel: Blends, Properties and Applications*, eds J. M. Marchetti, and Z. Fang (New York, NY: Nova Science Publishers, Inc), 123–150.

Farooq, M., Jabran, K., Cheema, Z. A., Wahid, A., and Siddique, K. H. M. (2011). The role of allelopathy in agricultural pest management. *Pest Manag. Sci.* 67, 494–506. doi: 10.1002/ps.2091

Gayas, B., and Kaur, G. (2017). Novel oil extraction methods in food industry: a review. *J. Oilseed Brassica* 8, 1–11. doi: 10.1080/10408398.2013.818933

Graumann, G., Richard, H., and Holley, A. (2008). Inhibition of *Escherichia coli* O157:H7 in ripening dry fermented sausage by ground yellow mustard. *J. Food Protect.* 71, 486–493. doi: 10.4315/0362-028x-71.3.486

Hendrix, K. M., Morra, M. J., Lee, H. B., and Min, S. C. (2012). Defatted mustard seed meal-based biopolymer film development. *Food Hydrocoll.* 26, 118–125. doi: 10.1016/j.foodhyd.2011.04.013

Ikegwu, O. J., and Ezeh, C. Q. (2012). Thermal properties of *Kerstingiella geocarpa* seeds as influenced by moisture content. *Niger. Food J.* 30, 100–105. doi: 10.1016/S0189-7241(15)30042-4

Issariyakul, T., Dalai, A. K., and Desai, P. (2011). Evaluating esters derived from mustard oil (*Sinapis alba*) as potential diesel additives. *J. Am. Oil Chem. Soc.* 88, 391–402. doi: 10.1007/s11746-010-1679-6

Issariyakul, T., and Dalai, A. K. (2012). Comparative kinetics of transesterification for biodiesel production from palm oil and mustard oil. *Can. J. Chem. Eng.* 90, 342–350. doi: 10.1002/cjce.20679

Jaime, R., Alcántara, J. M., Manzaneda, A. J., and Rey, P. J. (2018). Climate change decreases suitable areas for rapeseed cultivation in Europe but provides new opportunities for white mustard as an alternative oilseed for biofuel production. *PLoS One* 13:e0207124. doi: 10.1371/journal.pone.0207124

Jankowski, K. J., Budzynski, W. S., and Kijewski, L. (2014). Concentrations of copper, zinc and manganese in the roots, straw and oil cake of white mustard (*Sinapis alba* L.) and Indian mustard (*Brassica juncea* (L.) Czern. et Coss.) depending on sulphur fertilization. *Plant Soil Environ.* 60, 364–371. doi: 10.17221/225/2014-pse

Jeevan Kumar, S. P., Rajendra Prasad, S., Rintu, B., Agarwal, D. K., Kulkarni, K. S., and Ramesh, K. V. (2017). Green solvents and technologies for oil extraction from oilseeds. *Chem. Cent. J.* 11:9. doi: 10.1186/s13065-017-0238-8

Jibril, A. N., Yadav, K. C., Binni, M. I., and Kabir, M. H. (2016). Study on effect of moisture content on thermal properties of bambara groundnut (*Vigna subterranea* L. Verdc.) seed. *Int. Res. J. Eng. Technol.* 3, 773–782.

Jung, Y. T., and Diosady, L. L. (2012). Application of a ternary phase diagram to the phase separation of oil-in-water emulsions using isopropyl alcohol. *J. Am. Oil Chem. Soc.* 89, 2127–2134. doi: 10.1007/s11746-012-2127-6

Katepa-Mupondwa, F., Gugel, R. K., and Raney, J. P. (2005). Genetic diversity for agronomic, morphological and seed quality traits in *Sinapis alba* L. *Can. J. Plant Sci.* 86, 1015–1025. doi: 10.4141/p05-185

Khan, I. A., and Abourashed, E. A. (2010). *Leung's Encyclopedia of Common Natural Ingredients used in Food, Drugs, and Cosmetics*, 3rd Edn. Hoboken, NJ: John Wiley & Sons, Inc, 406–409.

Kos, B., Grèman, H., and Leštan, D. (2003). Phytoextraction of lead, zinc and cadmium from soil by selected plants. *Plant Soil Environ.* 49, 548–553. doi: 10.17221/4192-pse

Kostić, M., Djalović, I., Stamenković, O., Mitrović, P., Adamović, D., Kulina, M., et al. (2018). Kinetic modeling and optimization of biodiesel production from white mustard (*Sinapis alba* L.) seed oil by quicklime-catalyzed transesterification. *Fuel* 223, 125–139. doi: 10.1016/j.fuel.2018.03.023

Kostić, M., Joković, N., Stamenković, O., Rajković, K., Milić, P., and Veljković, V. (2014). The kinetics and thermodynamics of hempseed oil extraction by n-hexane. *Ind. Crop. Prod.* 52, 679–686. doi: 10.1016/j.indcrop.2013.11.045

Koubaa, M., Mhemdi, H., Barba, F. J., Roohinejad, S., Greiner, R., and Vorobiev, E. (2016). Oilseed treatment by ultrasounds and microwaves to improve oil yield and quality: an overview. *Food Res. Int.* 85, 59–66. doi: 10.1016/j.foodres.2016.04.007

Kozlowska, M., Gruczyńska, E., Ścibisz, I., and Rudzińska, M. (2016). Fatty acids and sterols composition, and antioxidant activity of oils extracted from plant seeds. *Food Chem.* 213, 450–456. doi: 10.1016/j.foodchem.2016.06.102

Krstić, D., Ćupina, B., Antanasović, S., Erić, P., Đabilovski, R., Manojlović, M., et al. (2010). Potential of white mustard (*Sinapis alba* L. subsp. alba) as a green manure crop. *Crucif. Newslett.* 29, 12–13.

Mahapatra, A. K., Melton, S. L., and Isang, E. M. (2013). Effect of moisture content on thermal properties of cowpea flours. *Agric. Eng. Int. CIGR J.* 15, 251–255.

Maj, G., Krzaczek, P., Stamirowska-Krzaczek, E., and Lipinska, H. (2019). Assessment of energy and physicochemical biomass properties of selected forecrop plant species. *Renew. Energy* 143, 520–529. doi: 10.1016/j.renene.2019.04.166

McKenzi, R. H., and Carcamo, H. (2010). *Mustard Production for Alberta*. Available at: http://www1.agric.gov.ab.ca/\protect\T1\textdollardepartment/deptdocs.nsf/all/agdex12947/\protect\T1\textdollarfile/143_20-1.pdf (accessed September 23, 2019).

Mejia-Garibay, B., Guerrero-Beltrán, J. Á, Palou, E., and López-Malo, A. (2015). Physical and antioxidant characteristics of black (*Brassica nigra*) and yellow mustard (*Brassica alba*) seeds and their products. *Arch. Latinoam. Nutr.* 65, 128–135.

Meziane, S., and Kadi, H. (2008). Kinetics and thermodynamics of oil extraction from olive cake. *J. Am. Oil Chem. Soc.* 85, 391–396. doi: 10.1007/s11746-008-1205-2

Monsalve, R. I, Gonzalez De La Peña, M. A., Menéndez-Arias, L., Lopez-Otin, C., Villalba, M., and Rodriguez, R. (1993). Characterization of a new oriental-mustard (*Brassica juncea*) allergen, Bra j IE: detection of an allergenic epitope. *Biochem. J.* 293, 625–632. doi: 10.1042/bj2930625

Nie, J., Wang, S., Emami, S., Falk, K., Shen, J., and Reaney, M. J. T. (2016). Unusually low pour point of fatty acid methyl esters with low saturated fatty

acid content. *Eur. J. Lipid Sci. Technol.* 118, 1486–1494. doi: 10.1002/ejlt.201400500

Nieschlag, H. J., and Wolff, I. A. (1971). Industrial uses of high erucic oils. *J. Am. Oil Chem. Soc.* 48, 723–727. doi: 10.1007/bf02638529

Oshodi, O. A., Chukwuneke, C. E., and Linus, O. (2014). The viscometric analysis of biodiesel from mustard and coconut oil. *Eur. Chem. Bull.* 3, 946–948.

Paciorek-Sadowska, J., Borowicz, M., Czupryński, B., Tomaszewska, E., and Liszkowska, J. (2018). New bio–polyol based on white mustard seed oil for rigid PUR–PIR foams. *Pol. J. Chem. Tech.* 20, 24–31. doi: 10.2478/pjct-2018-0019

Paulsen, H. M. (2011). Improving green-house gas balances of organic farms by the use of straight vegetable oil from mixed cropping as farm own fuel and its competition to food production. *Landbauforschung Volkenrode* 3, 209–216.

Peng, C., Zhang, T., Zhao, G., and Wang, S. (2013). Analysis on fat–soluble components of *Sinapis semina* from different habitats by GC–MS. *J. Pharm. Anal.* 3, 402–407. doi: 10.1016/j.jpha.2013.04.007

Peng, C., Zhao, S. Q., Zhang, J., Huang, G. Y., Chen, L. Y., and Zhao, F. Y. (2014). Chemical composition, antimicrobial property and microencapsulation of Mustard (*Sinapis alba*) seed essential oil by complex coacervation. *Food Chem.* 165, 560–568. doi: 10.1016/j.foodchem.2014.05.126

Popoviciu, D. R., Pirjol, T. N., and Miclau, L. S. (2017). Phytotoxic effect and bioaccumulation of chromium in white mustard (*Sinapis alba* L.) seedlings. *Rev. Chim. (Bucharest)* 68, 40–42. doi: 10.37358/RC.17.1.5384

Rahman, M., Amina Khatun, A., Liu, L., and Barkla, B. J. (2018). Brassicaceae mustards: traditional and agronomic uses in Australia and New Zealand. *Molecules* 23:231. doi: 10.3390/molecules23010231

Raney, J. P., Rakow, G., and Olson, T. (1995). "Development of low erucic, low glucosinolate *Sinapis alba*," in *Proceedings of the 9th International Rapeseed Congress*, Cambridge, 416–418.

Reverchon, E., and De Marco, I. (2006). Review: supercritical fluid extraction and fractionation of natural matter. *J. Supercrit. Fluids* 38, 146–166. doi: 10.1016/j.supflu.2006.03.020

Rodrigues, C. E. C., Aracava, K. K., and Abreu, F. N. (2010). Thermodynamic and statistical analysis of soybean oil extraction process using renewable solvent. *Food Sci. Technol. Int.* 45, 2407–2414. doi: 10.1111/j.1365-2621.2010.02417.x

Ropelewska, E., Jankowski, K. J., Zapotoczny, P., and Bogucka, B. (2018). Thermophysical and chemical properties of seeds of traditional and double low cultivars of white mustard. *Zemdirbyste Agric.* 105, 257–264.

Rosenthal, A., Pyle, D. L., and Niranjan, K. (1996). Aqueous and enzymatic processes for edible oil extraction. *Enzyme Microb. Technol.* 19, 402–420. doi: 10.1016/s0141-0229(96)80004-f

Sáez-Bastante, J., Fernández-García, P., Saavedra, M., López-Bellido, L., Dorado, M. P., and Pinzi, S. (2016). Evaluation of *Sinapis alba* as feedstock for biodiesel production in Mediterranean climate. *Fuel* 184, 656–664. doi: 10.1016/j.fuel.2016.07.022

Sangamithra, A., Swamy, G. J., Prema, S. R., Nandini, K., Kannan, K., Sasikala, S., et al. (2016). Moisture dependent physical properties of maize kernels. *Int. Food Res. J.* 23, 109–115. doi: 10.1021/acs.jafc.5b05698

Sarala, R., Rajendran, M., and Devadasan, S. R. (2012). Emission characteristics of mustard oil methyl ester (MOME) – diesel fuel blends on a C.I Engine. *Int. J. Green Chem. Bioprocess.* 2, 6–10.

Sarker, A. K., Saha, D., Begum, H., Zaman, A., and Rahman, M. M. (2015). Comparison of cake compositions, pepsin digestibility and amino acids concentration of proteins isolated from black mustard and yellow mustard cakes. *AMB Express* 5:22. doi: 10.1186/s13568-015-0110-y

Sask Mustard (2019). *Mustard Production Manual Saskatchewan Mustard Development Commission, Government of Saskatchewan*. Available at: https://saskmustard.com/production-manual/. (accessed September 23, 2019).

Saxena, D. K., Sharma, S. K., and Sambi, S. S. (2011). Kinetics and thermodynamics of cottonseed oil extraction. *Grasas Aceites* 62, 198–205. doi: 10.3989/gya.090210

Seal, C. E., Kranner, I., and Pritchard, D. H. W. (2008). Quantification of seed oil from species with varying oil content using supercritical fluid extraction. *Phytochem. Anal.* 19, 493–498. doi: 10.1002/pca.1072

Sengupta, R., and Bhattacharyya, D. K. (1996). Enzymatic extraction of mustard seed and rice bran. *J. Am. Oil Chem. Soc.* 73, 687–692. doi: 10.1007/bf02517941

Sharma, R. (2015). *Economics of Production, Marketing and Processing of Mustard in Morena District (Madhya Pradesh)*. M.Sc. Thesis, The Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalaya, Gwalior.

Singh, B. K., Bala, M., and Rai, P. K. (2014). Fatty acid composition and seed meal characteristics of *Brassica* and allied genera. *Nat. Acad. Sci. Lett.* 37, 219–226. doi: 10.1007/s40009-014-0231-x

Sirisomboon, P., and Posom, J. (2012). Thermal properties of *Jatropha curcas* L. kernels. *Biosyst. Eng.* 113, 402–409. doi: 10.1016/j.biosystemseng.2012.09.013

Stamenković, S., Djalović, I., Kostić, M., Mitrović, M., and Veljković, V. (2018). Optimization and kinetic modeling of oil extraction from white mustard (*Sinapis alba* L.) seeds. *Ind. Crop Prod.* 121, 132–141. doi: 10.1016/j.indcrop.2018.05.001

Sujatha, R., Mariajancyrani, J., and Chandramohan, G. (2013). Preliminary phytochemical investigation and antimicrobial activity of *Sinapis alba*. *Sch. J. App. Med. Sci.* 1, 138–141.

Sultana, S., Khalid, A., Ahmad, M., Zuhairi, A. A., Teong, L. K., and Zafar, M. (2014). The production, optimization, and characterization of biodiesel from a novel source: *Sinapis alba* L. *Int. J. Green Energy* 11, 280–291. doi: 10.1080/15435075.2013.772520

Szczyglak, P., and Zuk, Z. (2012). Method of determining the minimum breaking strength of the mustard seed coat. *Techn. Sci.* 15, 243–249.

Tabtabaei, S., Boocock, D. G. B., and Diosady, L. L. (2014). Biodiesel feedstock from emulsions produced by aqueous processing of yellow mustard. *J. Am. Oil Chem. Soc.* 91, 1269–1282. doi: 10.1007/s11746-014-2448-8

Tabtabaei, S., Boocock, D. G. B., and Diosady, L. L. (2015). Biodiesel production from mustard emulsion by a combined destabilization/adsorption process. *J. Am. Oil Chem. Soc.* 92, 1205–1217. doi: 10.1007/s11746-015-2677-5

Tabtabaei, S., and Diosady, L. L. (2012). The isolation of yellow mustard oil using water and cyclic ethers. *J. Am. Oil Chem. Soc.* 89, 935–945. doi: 10.1007/s11746-011-1971-0

Tabtabaei, S., and Diosady, L. L. (2013). Aqueous and enzymatic extraction processes for the production of food–grade proteins and industrial oil from dehulled yellow mustard flour. *Food Res. Int.* 52, 547–556. doi: 10.1016/j.foodres.2013.03.005

Tabtabaei, S., Diosady, L. L., and Ataya Pulido, V. M. (2013). Destabilization of yellow mustard emulsion using organic solvents. *J. Am. Oil Chem. Soc.* 90, 707–716. doi: 10.1007/s11746-013-2202-7

Tavakoli, H., Rajabipour, A., and Mohtasebi, S. S. (2009). Moisture-dependent some engineering properties of soybean grains. *CIGR J.* 11, 1–14.

Tavakoli, M., Tavakoli, H., Rajabipour, A., Ahmadi, H., and Gharib-Zahedi, S. M. T. (2009). Moisture–dependent physical properties of barley grains. *Int. J. Agr. Biol. Eng.* 2, 84–91.

Thacker, P. A., and Petri, D. (2009). The effects of canola or mustard biodiesel press cake on nutrient digestibility and performance of broiler chickens. *Asian Aust. J. Anim. Sci.* 22, 531–539.

Thangi, J., Shashitha, K. N., Ashwini, H. A., and Shlini, P. (2016). A correlation study of antioxidant potential's from *Synapis alba*. *Indo Am. J. Pharm. Sci.* 3, 234–239.

Topallar, H., and Gecgel, U. (2000). Kinetics and thermodynamics of oil extraction from sunflower seeds in the presence of aqueous acidic hexane solutions. *Turk. J. Chem.* 24, 247–253.

Vach, M., Strašil, Z., and Javùrek, M. (2016). Economic Efficiency of selected crops cultivated under different technology of soil tillage. *Sci. Agric. Bohem.* 4, 40–46. doi: 10.1515/sab-2016-0007

Vaisali, C., Charanyaa, S., Belur, P. D., and Regupathi, I. (2015). Refining of edible oils: a critical appraisal of current and potential technologies. *Int. J. Food Sci. Tech.* 50, 13–23. doi: 10.1111/ijfs.12657

Viuda-Martos, M., Ruiz-Navajas, Y., Fernández-López, J., and Pérez-Alvare, J. A. (2007). Antifungal activities of thyme clove and oregano essential oils. *J. Food Saf.* 27, 91–101.

Wendlinger, C., Hammann, S., and Vetter, W. (2014). Various concentrations of erucic acid in mustard oil and mustard. *Food Chem.* 153, 393–397. doi: 10.1016/j.foodchem.2013.12.073

Withers, R., Berglund, S., Esser, A., Brown, J., and Smathers, B. (2000). *Economics of Yellow Mustard in the Pacific Northwest*. Moscow: University of Idaho Cooperative Extension.

Worldwatch (2006). *Biofuels for Transportation – Global Potential and Implications for Sustainable Energy in the 21st Century, Report Prepared for the German Federal Ministry for Food, Agriculture and Consumer Protection*. Washington, DC: Worldwatch Institute.

Yaniv, Z., Schafferman, D., Elber, Y., Ben-Moshe, E., and Zur, M. (1994). Evaluation of *Sinapis alba*, native to Israel, as a rich source of erucic acid in seed oil. *Ind. Crop Prod.* 2, 137–142. doi: 10.1016/0926-6690(94)90095-7

Yesilyurt, M. K., Arslan, M., and Eryilmaz, T. (2019). Application of response surface methodology for the optimization of biodiesel production from yellow mustard (*Sinapis alba* L.) seed oil. *Int. J. Green Energy* 16, 60–71. doi: 10.1080/15435075.2018.1532431

Živković, S., Veljković, M., Banković-Ilić, I., Krstić, I., Konstantinović, S., Ilić, S., et al. (2017). Technological, technical, economic, environmental, social, human health risk, toxicological and policy considerations of biodiesel production and use. *Renew. Sust. Energ. Rev.* 79, 222–247. doi: 10.1016/j.rser.2017.05.048

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 April 2020
doi: 10.3389/fpls.2020.00380

Hybrid Aspen Expressing a Carbohydrate Esterase Family 5 Acetyl Xylan Esterase Under Control of a Wood-Specific Promoter Shows Improved Saccharification

Zhao Wang[1†], Prashant Mohan-Anupama Pawar[2†‡], Marta Derba-Maceluch[2], Mattias Hedenström[1], Sun-Li Chong[3‡], Maija Tenkanen[3], Leif J. Jönsson[1] and Ewa J. Mellerowicz[2]*

[1] Department of Chemistry, KBC Chemical-Biological Centre, Umeå University, Umeå, Sweden, [2] Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, Umeå, Sweden, [3] Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland

Fast-growing broad-leaf tree species can serve as feedstocks for production of bio-based chemicals and fuels through biochemical conversion of wood to monosaccharides. This conversion is hampered by the xylan acetylation pattern. To reduce xylan acetylation in the wood, the *Hypocrea jecorina* acetyl xylan esterase (*Hj*AXE) from carbohydrate esterase (CE) family 5 was expressed in hybrid aspen under the control of the wood-specific *Pt*GT43B promoter and targeted to the secretory pathway. The enzyme was predicted to deacetylate polymeric xylan in the vicinity of cellulose due to the presence of a cellulose-binding module. Cell-wall-bound protein fractions from developing wood of transgenic plants were capable of releasing acetyl from finely ground wood powder, indicative of active AXE present in cell walls of these plants, whereas no such activity was detected in wild-type plants. The transgenic lines grew in height and diameter as well as wild-type trees, whereas their internodes were slightly shorter, indicating higher leaf production. The average acetyl content in the wood of these lines was reduced by 13%, mainly due to reductions in di-acetylated xylose units, and in C-2 and C-3 mono-acetylated xylose units. Analysis of soluble cell wall polysaccharides revealed a 4% reduction in the fraction of xylose units and an 18% increase in the fraction of glucose units, whereas the contents of cellulose and lignin were not affected. Enzymatic saccharification of wood from transgenic plants resulted in 27% higher glucose yield than for wild-type plants. Brunauer–Emmett–Teller (BET) analysis and Simons' staining pointed toward larger surface area and improved cellulose accessibility for wood from transgenic plants compared to wood from wild-type plants, which could be achieved by *Hj*AXE deacetylating xylan bound to cellulose. The results show that CE5 family can serve as a source of enzymes for *in planta* reduction of recalcitrance to saccharification.

Keywords: acetyl xylan esterase, hybrid aspen, *Populus*, xylan, acetyl, enzymatic saccharification

OPEN ACCESS

Edited by:
Mengzhu Lu,
State Key Laboratory of Tree Genetics and Breeding, Research Institute of Tropical Forestry, Chinese Academy of Forestry, China

Reviewed by:
Alexander Andrew Myburg,
University of Pretoria, South Africa
Heather D. Coleman,
Syracuse University, United States

****Correspondence:***
Leif J. Jönsson
leif.jonsson@umu.se

† These authors have contributed equally to this work

‡ Present address:
Prashant Mohan-Anupama Pawar, Regional Centre for Biotechnology, NCR Biotech Science Cluster, Faridabad, India
Sun-Li Chong, State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, China

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *17 December 2019*
Accepted: *17 March 2020*
Published: *08 April 2020*

Citation:
Wang Z, Pawar PM-A, Derba-Maceluch M, Hedenström M, Chong S-L, Tenkanen M, Jönsson LJ and Mellerowicz EJ (2020) Hybrid Aspen Expressing a Carbohydrate Esterase Family 5 Acetyl Xylan Esterase Under Control of a Wood-Specific Promoter Shows Improved Saccharification. Front. Plant Sci. 11:380. doi: 10.3389/fpls.2020.00380

INTRODUCTION

Bioconversion of woody biomass has potential to provide advanced biofuels and bio-based materials. However, wood is relatively resistant to chemical and biological conversion processes, which necessitates costly processing and reduces recovery of desired products (Himmel et al., 2007; Zhao et al., 2012; McCann and Carpita, 2015). The main constituents of this biomass, i.e., lignin, hemicelluloses, and cellulose, and their interactions in the cell wall, all contribute to the complex structure of wood and its recalcitrance (Ragauskas, 2013).

The importance of acetylation for the recalcitrance of hemicelluloses received attention relatively long time ago (Biely et al., 1985; Grohmann et al., 1989). In woody biomass from hardwoods, most acetyl groups are on Xyl*p* (xylopyranosyl) units of xylan. The fraction of acetyl groups in hardwoods varies from around 3.5 to 4.4% (w/w) on dry-weight basis (Pawar et al., 2013), where 40–70% of the Xyl*p* units can be acetylated at the C-2 and/or C-3 positions of Xyl*p* units (Teleman et al., 2000, 2002; Kabel et al., 2003; Goncalves et al., 2007). Distribution of acetylated Xyl*p* units along the xylan chain of woody dicots is regulated by the activity of acetyl transferase ESK1/TBL29 (Grantham et al., 2017). Typically, every second unit is mono- or di-acetylated, and this pattern enables the xylan backbone to interact with the hydrophilic face of cellulose microfibrils in two-fold screw conformation (Busse-Wicher et al., 2014; Chong et al., 2014; Grantham et al., 2017).

Acetylation on xylan chains may contribute to biomass recalcitrance by changing surface hydrophobicity and thereby inhibiting productive binding of hydrolytic enzymes, and by causing steric hindrance of enzymes targeting cellulose and xylan (Pan et al., 2006; Busse-Wicher et al., 2014). Acetyl groups are hydrolyzed by pretreatment of lignocellulosic biomass yielding acetic acid, which is, however, a quantitatively significant inhibitor of microbial fermentation processes (Jönsson and Martin, 2016). Reduced acetylation of plant cell walls would therefore increase the accessibility of polysaccharides and reduce the inhibition of fermenting microbes.

Strong reduction of acetylation, especially in xylan, obtained by knocking out components of a xylan acetylation machinery typically causes dwarfism, reduced mechanical strength of the stem, collapsed vessels, and stunted plant growth (Lee et al., 2011b; Manabe et al., 2013; Yuan et al., 2013, 2016). Plants with severely reduced acetylation may therefore not necessarily exhibit increased sugar yield after enzymatic saccharification (Lee et al., 2011b; Xiong et al., 2013; Yuan et al., 2013). However, moderate decrease of xylan acetylation in hybrid aspen was not only well-supported by plants but also lead to better saccharification (Pawar et al., 2017b). On the other hand, excess xylan acetylation in rice, while providing some beneficial effects on saccharification, disrupted the structure of the secondary cell wall and lead to growth defects (Zhang et al., 2017).

Microbial enzymes with acetyl xylan esterase (AXE) activity could be used *in planta* to reduce xylan acetylation. These enzymes are grouped in at least eight Carbohydrate Esterase (CE) families that differ with regard to protein structure and other properties (Biely, 2012; Pawar et al., 2013). Previous studies have shown that introduction of an *Aspergillus niger* AXE1 (*An*AXE1) from CE1 in *Arabidopsis* or in hybrid aspen and targeting the enzyme to the cell wall for post-synthetic xylan deacetylation significantly improved the cellulose digestibility without changing the growth properties of these plants (Pawar et al., 2016, 2017a). Post-synthetic xylan deacetylation was considered as a more promising strategy than synthetic xylan deacetylation in the Golgi, since the latter could induce excess glucuronidation (Donev et al., 2018) caused by the promiscuous activity of glucuronyl transferases GUX1 and GUX2 (Grantham et al., 2017). These results encourage further trials with microbial enzymes capable of deacetylation of xylan in cell walls.

Here we are testing the AXE from the filamentous fungus *Hypocrea jecorina* (formerly *Trichoderma reesei*), *Hj*AXE, from family CE5. Compared to CE1 AXEs, which have broad specificity to different poly- and oligosaccharides, the AXEs from CE5 are thought to be more specific to polymeric xylan (Biely et al., 2011). Subtle differences were observed between CE1 and CE5 AXEs *in vitro* when deacetylating different acetylated xylo-oligosaccharides (Koutaniemi et al., 2013). Moreover, unlike CE1 AXEs, CE5 *Hj*AXE has a C-terminal cellulose binding domain (Margolles-Clark et al., 1996). These features expectedly would affect the performance of the members of these families when expressed *in planta*. A pairwise alignment of *An*AXE1 and *Hj*AXE (data not shown) indicated that the amino-acid sequence identity was < 20%, which further accentuates the difference between CE1 and CE5 enzymes.

We found that transgenic hybrid aspen expressing *Hj*AXE has normal growth in the greenhouse whereas its xylan is deacetylated by approximately 13% compared to the wild type (WT). The wood of such plants had improved bioprocessing properties along with increased cellulose accessibility. These results support the suitability of CE5 AXEs for post-synthetic xylan deacetylation.

MATERIALS AND METHODS

Plant Material

Transgenic hybrid aspen (*Populus tremula* L. × *tremuloides* Michx.) lines were generated as described previously (Ratke et al., 2015). The lines harbored the codon-optimized cDNA of *Hypocrea jecorina AXE* (*HjAXE*), GenBank accession CAA93247 (Margolles-Clark et al., 1996) cloned behind the wood-specific promoter in *pK-GT43B-GW7* (Ratke et al., 2015), and were denoted as WP:CE5. Twenty independent transgenic lines were screened *in vitro* for expression of *Hj*AXE, and the 11 best lines were further screened in the greenhouse. The three most highly expressing lines were finally selected and grown in the greenhouse along with WT control for 8 weeks. The growth conditions were as follows: light photoperiods 18 h [using HQI-TS 400W/DH metal halogen lamps (Osram, Munich, Germany) to supplement daylight when necessary], 20/15°C (day/night) temperatures, and 60–70% relative humidity. The plants were watered daily, fertilized once per week with Rika-S (Weibulls Horto, Hammenhög, Sweden) and shifted weekly to avoid any position effects. Stem height was periodically measured, and the

average internode length for internodes 19–35, and the stem diameter for internodes 20 and 40 were determined at the time of harvest.

Transcript Level

Total RNA was extracted from developing xylem tissue of hybrid aspen by using the Cetyl Trimethyl Ammonium Bromide (CTAB) extraction method (Chang et al., 1993). The cDNA was synthesized from 1 μg of RNA using a cDNA biosynthesis Bio-Rad kit (Bio-Rad Laboratories AB, Sundbyberg, Sweden). Diluted cDNA (20–30 times) was used for transcript analysis of transgenic plants. The expression was normalized to ubiquitin (Potri.005G198700) and tubulin (Potri.001G464400), and presented relative to the levels in the lowest-expressing line as previously described (Pawar et al., 2017a). The primers of reference and target genes are provided in **Supplementary Table S1**.

Acetyl Esterase Activity Assay

Soluble and wall-bound fractions of proteins were isolated from the developing wood of transgenic and WT hybrid aspen using the method described by Biswal et al. (2014) and tested for acetyl esterase activity using naturally acetylated aspen wood components as esterase substrates (Margolles-Clark et al., 1996). Aspen wood powder (particle size < 50 μm) was suspended in 50 mM sodium citrate buffer (pH 5.0) at 20 g/L and 2 μL of this suspension was incubated with 10 μg of total protein extract in a total volume of 400 μL (in the same buffer) for 24 h at 45°C. After reaction, the mixtures were denatured for 5 min at 100°C, centrifuged briefly and 10 μL of the supernatants were analyzed for the content of acetic acid by using a K-ACET kit (Megazyme, Bray, Ireland). Reaction mixtures containing denatured (10 min at 100°C) instead of fresh protein were used as negative controls. Results are presented as μmol of acetic acid produced from wood powder by 1 mg of protein in 1 h at 45°C. Three trees per each transgenic line and per WT were analyzed.

Cell Wall Compositional Analysis

Wood from internodes 19–35 was freeze-dried, and then ground to a rough wood powder (particle size < 0.5 mm), which was then ball-milled to a fine wood powder as previously described (Gandla et al., 2015). The fine wood powder was analyzed by using Fourier transform infra-red (FTIR) spectroscopy and pyrolysis gas chromatography combined with mass spectrometry (Py-GC/MS) as previously described (Gandla et al., 2015; Pawar et al., 2017a). The data were analyzed by using SIMCA-P (Umetrics AB, Umeå, Sweden).

The acetyl content was determined according to Gille et al. (2011) by saponification of the fine wood powder. The released acetic acid was analyzed by using HPAEC (high-performance anion-exchange chromatography) as previously described (Wang et al., 2018).

Alcohol insoluble residue (AIR) of the fine wood powder was prepared (Pawar et al., 2017a). The AIR was used to determine Klason lignin, acid-soluble lignin (ASL), Updegraff cellulose, and trimethylsilyl (TMS) monosaccharides content of non-cellulosic polysaccharides as described by Gandla et al. (2015).

Xylan Structure Analyses

NMR (Nuclear Magnetic Resonance) Spectroscopy Analysis

Acetylated xylan polymer was prepared from AIR by delignification and DMSO extraction (Pawar et al., 2017a). 2D 1H-13C HSQC was used to analyze the xylan polymer and spectra were acquired from a Bruker Avance III HD 850 MHz spectrometer as described by Pawar et al. (2017a).

OLIMP (OLIgosaccharide Mass Profiling) Analysis

AIR residue was heat-treated at 60°C for 1 h to deactivate acetyl xylan esterase and digested by pure GH10 endo-1,4-β-D-xylanase from *Aspergillus aculeatus* (*Aa*GH10) [kind gift from Novozymes A/S (Bagsværd, Denmark)]. The released xylo-oligosaccharides were desalted and separated into neutral and acidic fractions using a Graphitized Carbon SPE column (Thermo Scientific) (Chong et al., 2014). The mass spectra were acquired with atmospheric pressure matrix-assisted laser desorption/ionization-ion trap mass spectrometry (AP-MALDI-ITMS) as described by Chong et al. (2011).

Pretreatment and Saccharification

The rough wood powder was sieved and the fraction with a particle size of 0.1–0.5 mm was used for pretreatment and saccharification. Reaction mixtures containing 50 mg wood (dry weight) and 1% (w/w) sulfuric acid were pretreated at 165°C for 10 min using a single-mode microwave system (Initiator Exp, Biotage, Uppsala, Sweden). Saccharification of pretreated and non-pretreated wood samples was performed by enzymatic digestion of 50 mg wood (dry weight) (or, for pretreated material, the solid residue remaining after the pretreatment of 50 mg wood) using a 1:1 (v/v) mixture of Celluclast 1.5L and Novozyme 188. The load of enzyme protein corresponded to 1 mg per 50 mg wood. The total mass of the reaction mixture was 1000 mg and the medium consisted of sodium citrate buffer (0.5 M, pH 5.2). Reaction mixtures were incubated for 72 h in 2-mL Sarstedt safe-seal micro-centrifuge tubes in an orbital shaker set at 170 rpm and 45°C. Aliquots withdrawn after 2 h were analyzed by using a glucometer (Accu-Chek Aviva, Roche Diagnostics, Risch-Rotkreuz, Switzerland), and data were used to calculate the glucose production rate (GPR). The yields of arabinose, galactose, glucose, mannose, and xylose in pretreatment liquids and in enzymatic hydrolyzates after 72 h incubation were determined using HPAEC, as previously described (Wang et al., 2018).

Brunauer–Emmett–Teller (BET) Analysis

The surface area of non-pretreated and acid-pretreated sieved rough wood powder (0.1–0.5 mm) was analyzed with a single-point BET procedure using a TriStar 3000 analyzer (Micromeritics, Atlanta, GA, United States). A SmartPrep Degasser (Micromeritics) was utilized prior to the analysis with TriStar 3000 to remove potential adsorbed contaminants. The BET method is based on Langmuir theory and adsorption of nitrogen gas.

Simons' Staining

Simons' staining estimates the accessibility of cellulosic materials to enzymes based on solute exclusion (Yu et al., 1995; Arantes and Saddler, 2011). A modified Simons' staining assay (Chandra and Saddler, 2012) was used to analyze cellulose accessibility. The analysis was performed using the non-pretreated and acid-pretreated sieved rough wood powder (0.1–0.5 mm). Direct Blue (DB, Pontamine Fast Sky Blue 6BX) and Direct Orange (DO, Pontamine Fast Sky Orange 6RN) dyes were obtained from Pylam Products (Garden City, NY, United States).

Statistical Analysis

JMP® Pro program[1] with analysis of variance (ANOVA) was used for data analysis. *Post hoc* Dunnett-test was used to compare individual transgenic lines with the WT and contrast test was used to compare all transgenic lines with the WT.

RESULTS

Growth of CE5-Expressing Hybrid Aspen

Hybrid aspen lines 11, 14B, and 14C, expressing *Hj*AXE (Margolles-Clark et al., 1996) harboring the plant signal peptide of aspen cellulase *Ptxt*CEL9B3 under control of wood-specific promoter (Ratke et al., 2015), denoted WP:CE5, were grown in the greenhouse for 2 months. The three selected lines had the highest transgene expression from 20 obtained lines as described in Section "Materials and Methods." Transgene expression in developing wood tissues was approximately two times lower in line 11 than in lines 14B and 14C (**Figure 1A**). Acetyl esterase activity was studied in wall-bound and soluble protein fractions using naturally acetylated aspen wood as a substrate (Margolles-Clark et al., 1996). This method was considered more specific for xylan acetyl esterases acting on polymeric xylan as compared to methods using synthetic esters as substrates. The activity was detected only in wall-bound protein fraction extracted from developing wood of WP:CE5 lines (**Figure 1B**). It was varying within the range of 0.56–0.73 μmol mg^{-1} protein h^{-1}, consistent with the transgene expression levels. No activity was detected in wall-bound protein extracts of WT plants. No or negligible activity was recorded in the soluble protein fraction in transgenic lines and no activity was found in this fraction in the WT (**Figure 1C**). These data indicated that the expressed protein was active as acetyl esterase and associated with cell walls, as expected.

Plant morphology, height growth, and diameter growth of transgenic lines were similar to WT, whereas their internodes were shorter (**Figures 1D–F**). Since both transgenic and WT plants produced one leaf per internode, this indicates that the transgenic plants produced more leaves. Overall, the results show that the CE5 enzyme is expressed and active in the cell walls of selected transgenic lines without negatively affecting their growth or development.

[1]www.jmp.com

Lignin and Carbohydrates Contents

To investigate potential effects of expression of CE5 AXE on the wood cell wall chemical composition, the contents of lignin and Updegraff cellulose, the yields of trimethylsilyl (TMS) monosaccharides, and the composition of wood pyrolyzates were analyzed (**Table 1**). The contents of Klason and acid-soluble lignin, and Updegraff cellulose did not show consistent differences among the transgenic lines compared to WT (**Table 1**). Analysis using Py-GC/MS (**Table 1**) showed a small increase in H lignin units (which are a minor lignin component) in transgenic lines. For Py-GC/MS analysis, this was the only statistically significant difference for all transgenic lines. Thus, the wet chemistry and PyGC/MS results showed in agreement no distinct changes in crystalline cellulose or in the main lignin components.

In contrast, the monosaccharide composition did show some changes in transgenic lines compared to WT (**Table 1**). The xylose (Xyl) content of the transgenic lines was ∼4% lower than in the WT, whereas the glucose (Glc) and rhamnose (Rha) contents were ∼18 and ∼8% higher, respectively. These results indicate changes in the composition of the matrix polysaccharides, a decrease in xylan content, and an increase in glucan in transgenic lines.

To further reveal potential subtle chemical changes in cell walls of WP:CE5 plants, diffuse reflectance Fourier-transform infrared (FTIR) spectroscopy was applied to the ground wood samples. The spectral data analyzed by orthogonal projections to latent structures-differential analysis (OPLS-DA) showed that there was a clear separation between transgenic hybrid aspen lines and the WT (**Figure 2A**). The bands at 1240, 1370, and 1740 cm^{-1}, which originate from C-O stretching, CH_2 bending and C = O stretching vibrations (Gorzsas et al., 2011), respectively, all present in acetyl ester groups, strongly contributed to this separation and these signals were less abundant in the transgenic plants than in the WT (**Figures 2B,C**). This result was similar to that obtained with the previously studied transgenic plants overexpressing *AnAXE1* (Pawar et al., 2016, 2017a). On the other hand, the band at 1600 cm^{-1}, reflecting aromatic C = C vibrations found abundantly in lignin, and the region around 1640 cm^{-1}, reflecting water, were more abundant in WT plants, which was in contrast to results observed with *AnAXE1*-expressing *Arabidopsis* and aspen (Pawar et al., 2016, 2017a). The latter results suggest some changes in lignin structure and cell wall hydration that are specific to CE5 overexpressors.

Xylan Acetylation

Quantitation of acetic acid released through saponification of wood showed that the transgenic plants had reduced acetyl content by 10 to 16%, compared to the WT (**Figure 3A**).

The acetylation was further investigated using 2D HSQC NMR spectroscopic analysis of DMSO-extracted xylan (**Supplementary Data Sheets S2, S3**), which revealed the presence of acetylated and non-acetylated Xyl*p* residues (**Figure 3B**). The signals shown in green were used to obtain the relative content of acetylated Xyl*p* units whereas the signals

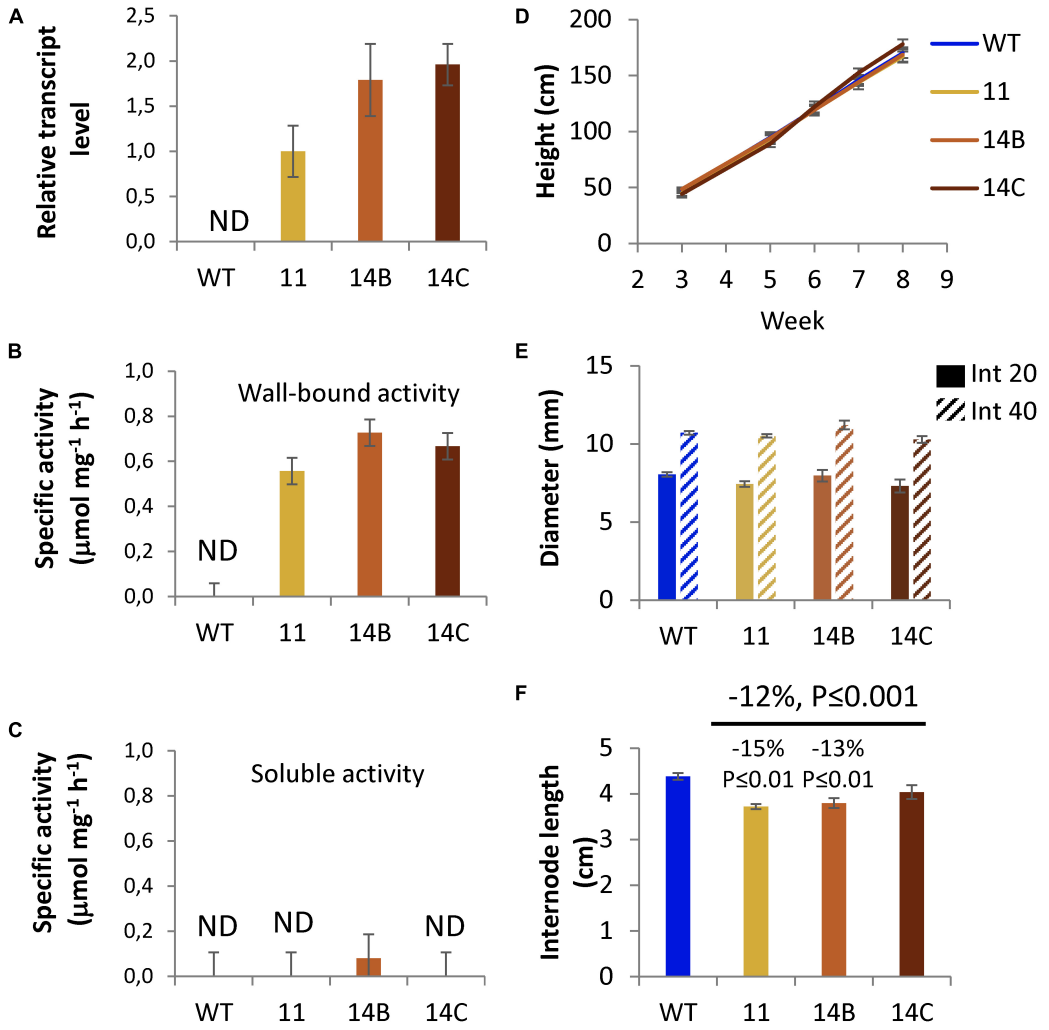

FIGURE 1 | Growth and transgene expression in selected WP:CE5 lines (11, 14B, and 14C) used in this study compared to wild type (WT). **(A)** Transgene transcript levels in developing wood determined by RT-qPCR; the expression relative to two reference genes was normalized to the levels observed in line 11. AXE specific activity in wall-bound **(B)** and soluble **(C)** protein fractions isolated from developing wood, determined by release of acetic acid from acetylated aspen wood powder used as substrate. Data in **(A–C)** are means ± *SE*, *N* = 3, biological replicates. **(D)** Height growth of transgenic lines during 8 weeks in the greenhouse. **(E)** Diameter of internodes 20 and 40 at week 8. **(F)** Average internode length at week 8. Data in **(D–F)** are means ± *SE*, *N* = 6 trees per transgenic lines and 23 trees per WT. *P*-values correspond to post-ANOVA Dunnett test (for individual lines) or contrast analysis comparing all transgenic lines to WT.

shown in blue represent different non-acetylated Xyl*p*. For the WT, 49% of the total Xyl*p* units were acetylated. That included 22% monoacetylation at position C-2 (X2), 16% monoacetylation at position C-3 (X3), 6% di-acetylation at positions C-2 and C-3 (X23), and 5% acetylation at C-3 and meGlcA (X3G2) (**Figure 3C**). Transgenic lines exhibited reductions in signals from all acetylated Xyl*p*, by 16% for X23, 14% for X2, 5% for X3G2, and 11% for X3. In total, the content of acetylated Xyl*p* units was reduced by 12% in transgenic lines compared to the WT (**Figure 3C**). Moreover, there was a 44% increase in the content of non-acetylated Xyl*p* units preceeding either non-acetylated or C-2 acetylated units (X-X/X2). The results suggest that the CE5 enzyme acted on Xyl*p* positions 2 and 3, and could also deacetylate 2,3-double-acetylated Xyl*p* as well as position 3 in glucuronosylated Xyl*p* units of aspen wood xylan.

Moreover, to investigate the changes in the pattern of glucuronoxylan acetylation in the transgenic plants, heat-treated AIR samples from lines 14A and 14B and from the WT were treated with the *Aa*GH10 endo-1,4-β-xylanase, and the released xylo-oligosaccharides (XOS) were analyzed by using oligosaccharide mass profiling (by AP-MALDI-ITMS; Chong et al., 2011). The treatment released acidic XOS with a degree of polymerization (DP) of three to seven and with up to seven acetyl groups (**Figure 3D**). The transgenic lines showed a prominent change in the distribution of acidic XOS compared to the WT. Whereas the most abundant acidic XOS in the transgenic lines had a DP of three and one acetyl group, the most abundant acidic XOS in the WT had a DP of four and two acetyl groups, although XOS with a DP of five with four acetyl groups and XOS with a DP of six with four or five acetyl groups were also common. In

TABLE 1 | Lignin and carbohydrates contents of wood of transgenic and wild-type (WT) hybrid aspen analyzed with three methods.[a]

Line	Lignin and cellulose contents[b]		
	Klason lignin	**Acid-soluble lignin**	**Cellulose**
11	18.8 ± 1.1	6.0 ± 0.3	39.6 ± 2.8
14B	16.7 ± 1.1**	6.5 ± 0.4	43.7 ± 4.1
14C	18.1 ± 1.0	6.6 ± 0.6	44.3 ± 2.7
Trans	17.7 ± 1.4	6.4 ± 0.5	42.9 ± 3.8
WT	18.2 ± 0.9	6.1 ± 0.8	44.2 ± 4.1

Line	Non-cellulosic polysaccharide composition (mol%)[c]								
	Ara	**Rha**	**Xyl**	**Man**	**MeGlcA**	**Gal**	**GalA**	**Glc**	**GlcA**
11	2.3 ± 0.2	2.7 ± 0.1***	59.3 ± 1.7***	4.8 ± 0.1**	4.8 ± 0.4	3.5 ± 0.4**	5.6 ± 0.3	13.9 ± 0.4***	2.8 ± 0.2
14B	2.2 ± 0.1	2.7 ± 0.1*	60.1 ± 1.6***	4.3 ± 0.2	4.8 ± 0.4	2.6 ± 0.3	5.3 ± 0.2	15.4 ± 1.5***	2.7 ± 0.2
14C	2.2 ± 0.1	2.6 ± 0.1	61.7 ± 1.4	4.6 ± 0.3	5.1 ± 0.4	2.9 ± 0.4	5.3 ± 0.3	12.7 ± 1.1	2.7 ± 0.3
Trans	2.2 ± 0.1	2.7 ± 0.1***	60.4 ± 1.8***	4.5 ± 0.3	4.9 ± 0.4	2.9 ± 0.5	5.4 ± 0.3	14.1 ± 1.5***	2.7 ± 0.2
WT	2.2 ± 0.1	2.5 ± 0.1	63.2 ± 1.6	4.4 ± 0.2	4.9 ± 0.5	2.7 ± 0.7	5.2 ± 0.3	12.0 ± 0.9	2.6 ± 0.2

Line	Py-GC/MS analysis of milled wood[d]						
	C	**G**	**S**	**H**	**L**	**S/G ratio**	**C/L ratio**
11	74.3 ± 1.3	8.2 ± 0.5	12.1 ± 0.5	2.0 ± 0.1**	22.5 ± 1.1	1.47 ± 0.04**	3.31 ± 0.22
14B	75.6 ± 2.0	7.6 ± 0.7	11.8 ± 1.1	1.7 ± 0.2	21.3 ± 1.9	1.56 ± 0.05	3.58 ± 0.45
14C	74.4 ± 1.4	8.0 ± 0.6	12.5 ± 0.7	1.9 ± 0.2	22.6 ± 1.4	1.55 ± 0.04	3.30 ± 0.27
Trans	74.8 ± 1.7	7.9 ± 0.7	12.1 ± 0.9	1.9 ± 0.2**	22.1 ± 1.6	1.53 ± 0.06	3.41 ± 0.35
WT	74.5 ± 1.6	8.0 ± 0.7	12.5 ± 0.8	1.8 ± 0.2	22.5 ± 1.5	1.56 ± 0.08	3.33 ± 0.31

*[a] The analyses included two or three samples from each of the transgenic lines and eight samples from WT. Each sample consisted of a pool of two trees, and each sample was analyzed as technical triplicates. Statistical significance is based on ANOVA with post hoc Dunnett test for each transgenic line and on contrast for all transgenic lines (Trans) together: ***P ≤ 0.001; **P ≤ 0.01; *P ≤ 0.05. [b] Lignin and Updegraff cellulose content of transgenic and wild-type hybrid aspen. Data given as mass fractions in% (g/100 g dry weight). [c] Monosaccharide composition of non-cellulosic polymers determined by TMS in alcohol-insoluble wood residue of transgenic and WT hybrid aspen. Data given as mol% (mol/100 mol). [d] Py-GC/MS analysis of milled wood of transgenic and wild-type hybrid aspen. Data for C (carbohydrate), G (guaiacyl lignin units), S (syringyl lignin units), H (p-hydroxyphenyl lignin units), and L (total lignin, i.e., the sum of G, S, H, and generic phenolic constituents as defined by Gerber et al., 2012) are given as fractions in% of the total signals from the Py-GC/MS analysis. The ratio of syringyl and guaiacyl lignin units and the ratio of carbohydrate- and lignin-derived constituents are also included.*

comparison with the WT, the distribution of acidic XOS products with a DP of four and five exhibited a shift toward less acetylated products for the transgenic plants. Furthermore, only the WT yielded noticeable quantities of acidic XOS with a DP of six or seven. This indicates that the glucuronoxylan of the transgenic plants was more accessible to the xylanase treatment and that the main products were less acetylated. The neutral XOS had a DP of two to five and up to five acetyl groups in both genotypes, but the transgenic lines had more DP 2 products without acetyl groups compared to WT (**Figure 3D**). The results indicate that the CE5 enzyme caused reduced acetylation of xylan for both neutral and meGlcA-substituted domains in glucuronoxylan.

Saccharification

In the screening of the original 11 transgenic lines, the three lines (11, 14B, and 14C) showing the highest transcript levels for the transgene were also the only lines that gave improved glucose yield in enzymatic saccharification of non-pretreated wood, suggesting that a threshold of transgene expression was needed to achieve measurable saccharification gains. Analytical enzymatic saccharification of non-pretreated wood (**Figure 4A**) resulted in a glucose yield of 137 mg/g for the WT (34.1% glucan recovery based on Updegraff cellulose). For the transgenic lines (average for 11, 14B, and 14C) there was a 27% improvement in glucose yield compared to the WT (glucose yield: 174 mg/g; glucan recovery: 44.6%). Although the transgenic lines contained similar fractions of mannose and galactose as the WT and although they had 5% lower content of xylose, enzymatic saccharification of non-pretreated transgenic lines resulted in higher yields of xylose, mannose, and galactose compared to the WT (11–21% improvement, **Figure 4A**). Glucose production rate analysis after 2 h of enzymatic hydrolysis also pointed toward higher rates of glucose release for transgenic plants (**Supplementary Figure S1**). The results show that expression of CE5 AXE improved the digestibility of non-pretreated transgenic aspen with regard to both cellulose and hemicellulose.

With acid pretreatment, the transgenic aspen gave 8% higher glucose yield in the pretreatment liquid than the WT (data not shown), and also 3% higher glucose yield in the enzymatic saccharification (**Figure 4B**). Most of the xylan was hydrolyzed in the acid pretreatment (data not shown), which resulted in higher xylan recoveries for the transgenic aspen (59.1%) than for the WT (55.6%) even though the total xylose yield was not affected (**Figure 4C**). When the sugars released in both the

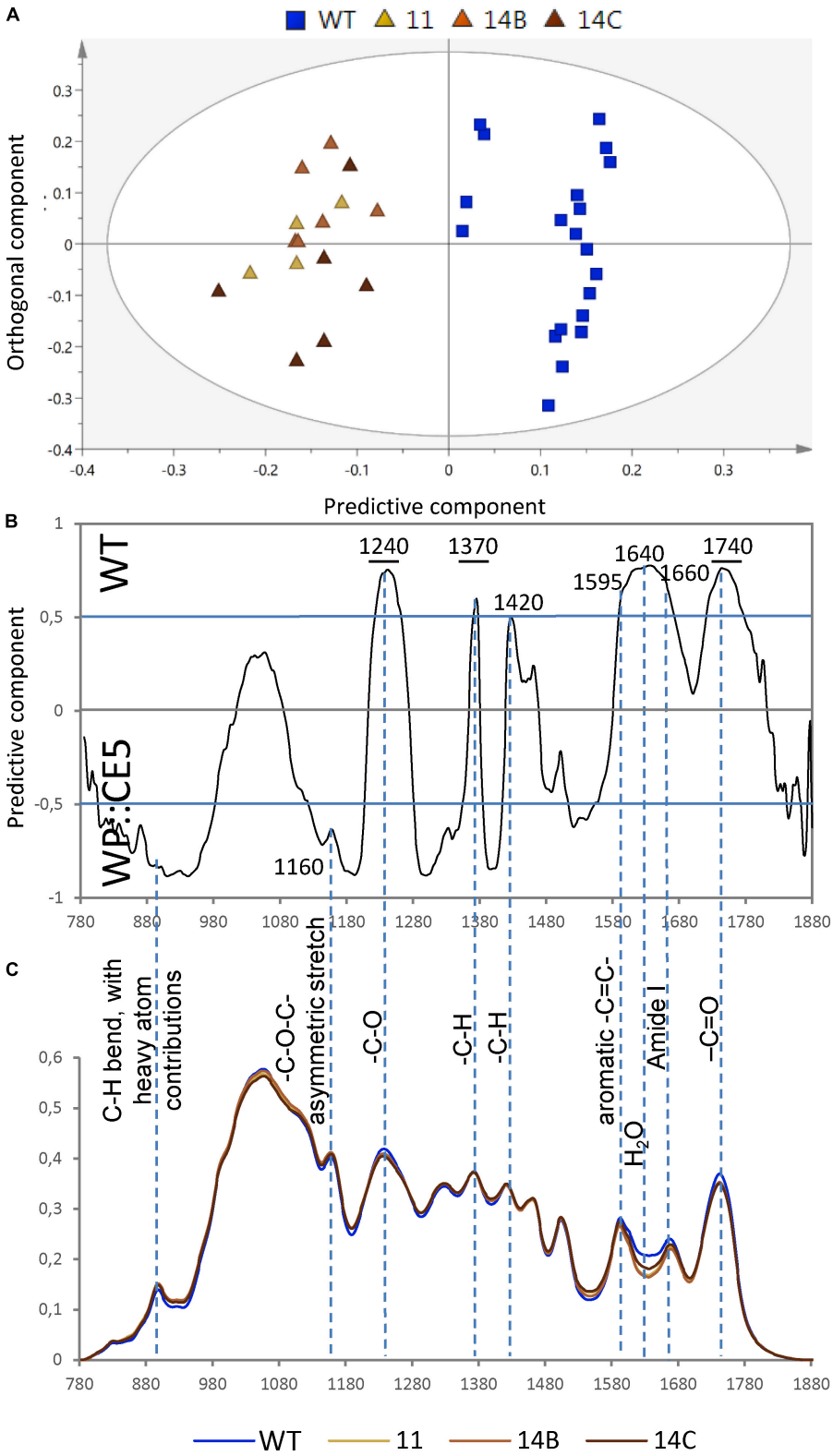

FIGURE 2 | Diffuse reflectance Fourier-transform infrared (FTIR) spectra of wood of WP:CE5 transgenic (lines 11, 14B, 14C) and wild type (WT). Dotted lines show bands that are significantly different (≥50% correlation) in the transgenic lines compared to WT, according to OPLS-DA (orthogonal projections to latent structures – discriminant analysis) models using 1 + 1 (predictive + orthogonal) components. Model components are: R2X (cum) = 0.788, R2Y (cum) = 0.891, Q2 (cum) = 0.878. Score plot **(A)**, loadings **(B)** and the corresponding average spectra **(C)**.

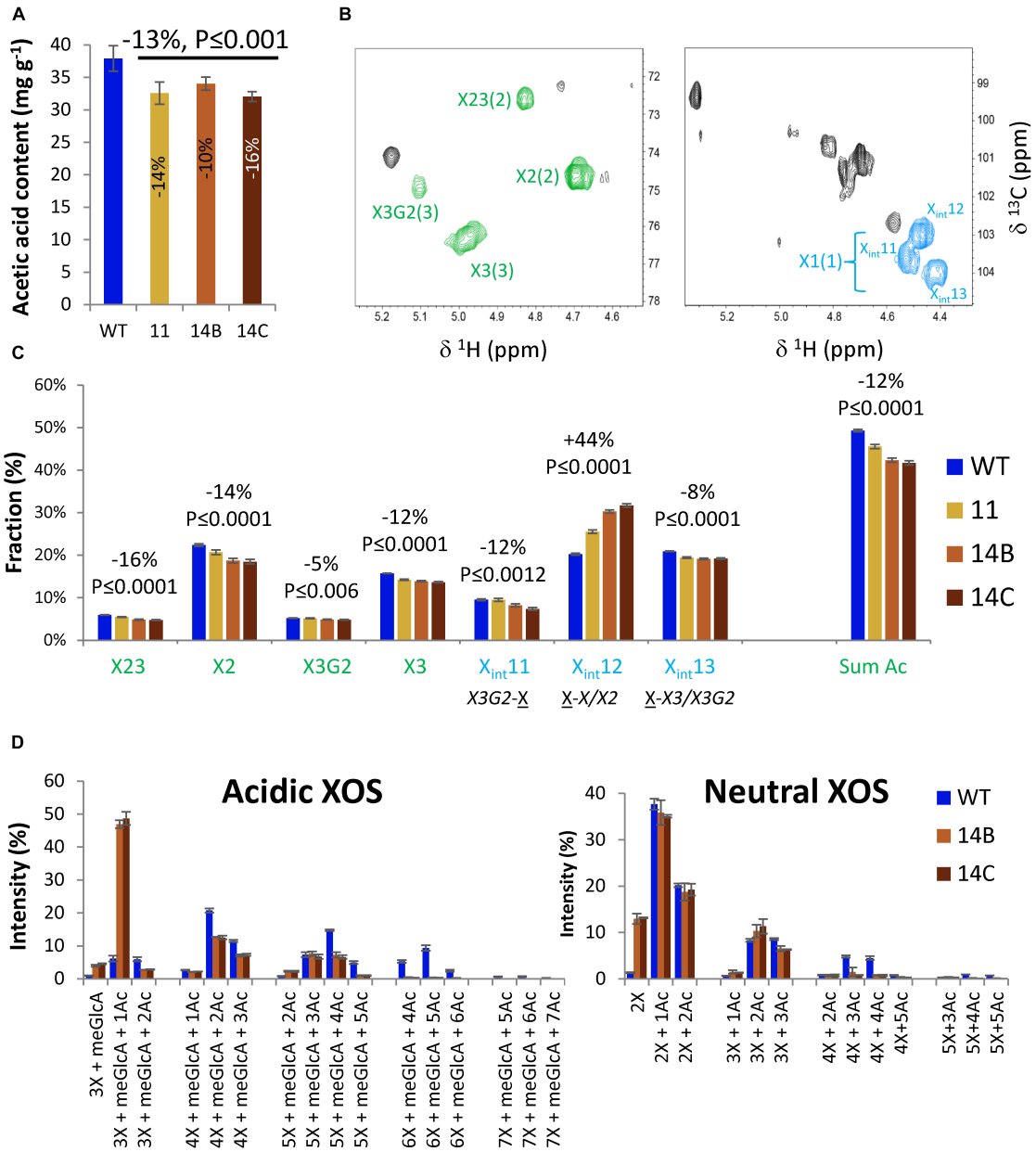

FIGURE 3 | Effects of WP:CE5 expression (lines 11, 14B, and 14C) on xylan acetylation in aspen wood. **(A)** Cell wall acetyl content determined by acetic acid release. **(B)** 2D qHSQC NMR spectra of extracted xylan showing signals from acetylated (green) and non-acetylated (blue) Xyl*p* units in xylan, which were used for quantification in **(C)**. X23-diacetylated Xyl*p*; X2-Xyl*p* monoacetylated at position 2; X3G2-Xyl*p* acetylated at position 3 and glucuronosylated at position 2; numbers in parenthesis correspond to carbon number in Xyl*p;* Xint – different internal Xyl*p* signals as assigned by Grantham et al. (2017). **(D)** OLIMP analysis of acidic and neutral xylo-oligosaccharides (XOS) released by endoxylanase *Aa*GH10. Data are means ± *SE*, *N* = 2 or more biological × 3 technical replicates for **(A)**, 2 or more biological × 2 technical replicates for **(C)**, and 2 biological replicates for **(D)**. *P*-values in **(A,C)** correspond to post-ANOVA contrast analysis comparing all transgenic lines to WT.

pretreatment liquid and the enzymatic hydrolyzate are combined, the WT reached 65.7% glucan recovery and 58.6% xylan recovery, whereas the transgenic aspen reached 70.2% glucan recovery and 61.7% xylan recovery. The results show that the transgenic aspen had improved glucan recovery and yields, and indicate that the xylan in the transgenic aspen was easier to degrade during the pretreatment.

Wood Nanostructure

Two methods were employed to compare the properties of the cell wall nanostructure of the transgenic plants and the WT, viz. Brunauer–Emmett–Teller (BET) analysis (Brunauer et al., 1938) and Simons' staining (Chandra and Saddler, 2012). While BET analysis was carried out using air-dried wood powder or pretreated wood powder, Simons' staining was carried out using

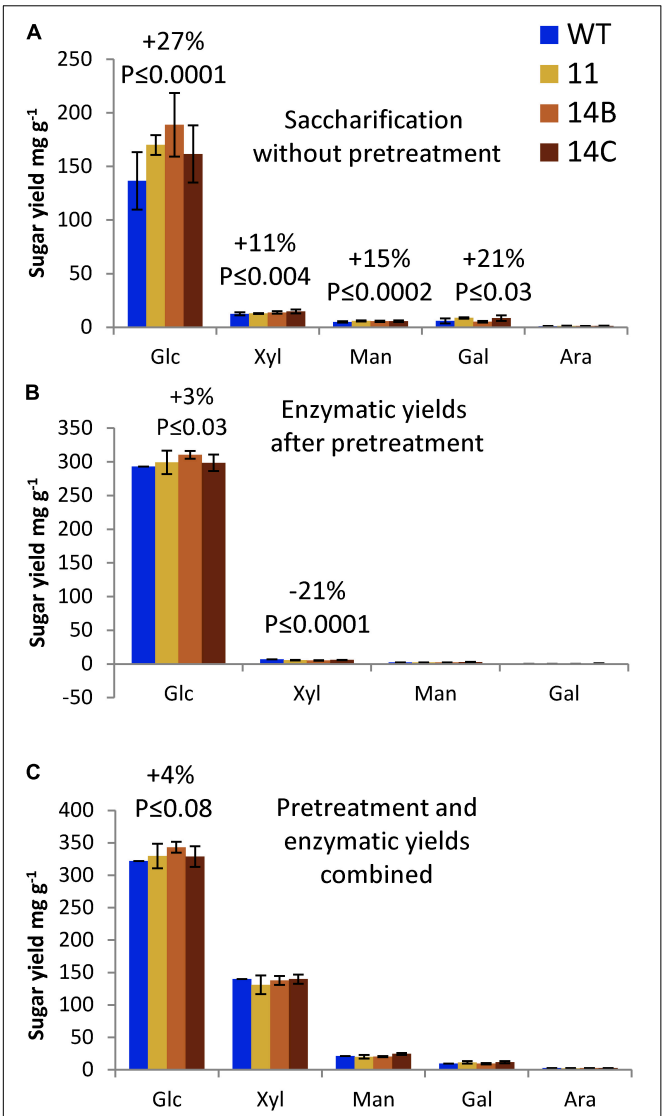

FIGURE 4 | Effects of WP:CE5 expression on saccharification. **(A)** Sugar yields of enzymatic hydrolysis without pretreatment. **(B)** Sugar yields of enzymatic hydrolysis after acid pretreatment. **(C)** Combined sugar yields of pretreatment liquid and enzymatic hydrolyzate. Data are means ± *SE*, *N* = 3 technical replicates of a pooled material from four trees. *P*-values correspond to ANOVA contrast analysis comparing all transgenic lines to WT.

wood powder or pretreated wood powder suspended in water to simulate the conditions during a saccharification reaction.

BET analysis, which measures physical adsorption of nitrogen gas to the solid phase, provided information about the specific surface area (**Table 2**). On an average, both non-pretreated and pretreated wood of the transgenic plants exhibited larger surface area than the WT. The average surface area was 11% larger for the transgenic plants than for the WT for both materials. The increase of the surface area after pretreatment agrees with the fact that the susceptibility to enzymatic saccharification also increased after pretreatment.

TABLE 2 | BET surface area (m^2/g) of wood of transgenic and wild-type (WT) hybrid aspen with and without acid pretreatment.[a,b]

Wood samples	Non-pretreated wood	Acid-pretreated wood
11	2.06 ± 0.04***	3.22 ± 0.09
14B	1.91 ± 0.13*	3.29 ± 0.26*
14C	1.75 ± 0.07	3.33 ± 0.08**
Trans	1.91 ± 0.15***	3.28 ± 0.15**
WT	1.72 ± 0.07	2.96 ± 0.02

[a]*Analyses were based on samples consisting of pooled wood from four separate trees. Each sample was analyzed as technical triplicates.* [b]*Statistical significance is based on ANOVA with post hoc Dunnett test for each transgenic line and on contrast for all transgenic lines (Trans) together: ***$P \leq 0.01$; **$P \leq 0.05$; *$P \leq 0.1$.*

The Simons' stain assay (**Figure 5**) is based on two dyes: blue (DB), which has a molecular mass of 993 Da, and orange (DO), which is filtered to obtain its high-molecular-mass (>100 kDa) fraction (Yu et al., 1995). Using these dyes, the accessible surface area and the porosity of the samples can be evaluated. For non-pretreated wood, transgenic plants exhibited 10% higher adsorption of DO than the WT, and 8% higher adsorption of total dyes (**Figure 5A**). However, this difference between the transgenic plants and the WT disappeared after pretreatment (**Figure 5B**). Comparing the adsorption ratio DO/DB, the transgenic plants showed 7% higher values than the WT without pretreatment. All samples showed a decreased DO/DB ratio after pretreatment, as the pretreatment resulted in increased adsorption of DB.

Data from both BET and Simons' staining agree that transgenic plants had an advantage over the WT with regard to structural properties, such as larger surface area and better accessibility. In summary, the activity of the CE5 AXE enlarged the surface area of the transgenic aspen and improved its accessible surface area, which was especially evident for non-pretreated wood.

DISCUSSION

Biomass recalcitrance to enzymatic saccharification is one of the main problems that need to be solved in order to implement technologies delivering green chemicals and fuels from renewable plant biomass. Since reducing xylan acetylation was suggested as one of the most promising strategies for reducing lignocellulose recalcitrance and increasing bioethanol yields (Donev et al., 2018), we studied the effectiveness of *in planta* expression of an AXE from family CE5 to reduce biomass acetylation. Here we present analyses of transgenic aspen expressing *Hj*AXE driven by a wood-specific promoter, and targeted to cell walls using a plant signal peptide. *Hj*AXE represents a distinct AXE belonging to a different CE family than the previously used *An*AXE1 from CE1 (Pawar et al., 2016, 2017a). Although both these enzymes were active on polymeric xylan, there were subtle differences in their activities *in vitro* (Koutaniemi et al., 2013), and substantial differences with respect to their amino-acid sequences, the three-dimensional structure of their catalytic sites, and their domain structures (Margolles-Clark et al., 1996; Hakulinen et al., 2000).

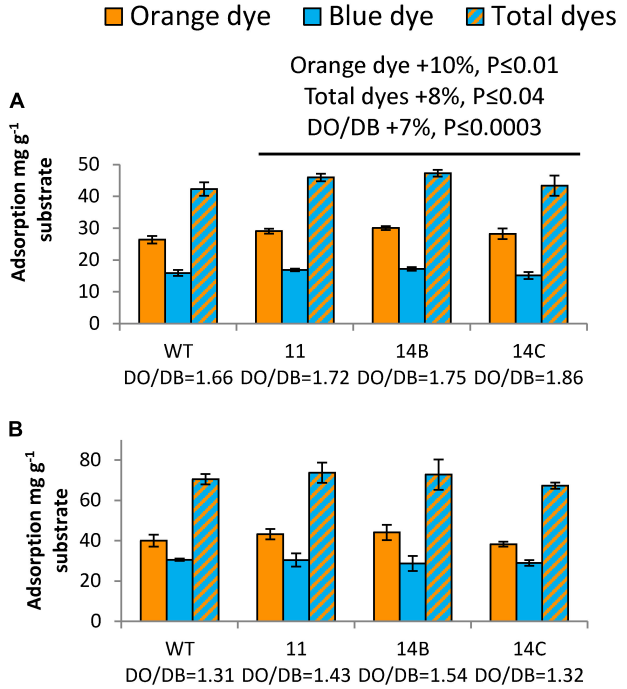

FIGURE 5 | Effects of WP:CE5 expression on wood nanostructure. Simons' staining of non-pretreated **(A)** and acid-pretreated aspen wood **(B)**. Data are means ± *SE*, *N* = 3 technical replicates. *P*-values correspond to ANOVA contrast analysis comparing all transgenic lines to WT.

CE5 AXE Can Efficiently Deacetylate Cell Wall Xylan *in vivo*

The transgenic plants expressing the CE5 member *Hj*AXE exhibited decreased xylan acetylation in their wood by 10–16% compared to the WT (**Figure 3A**), which is comparable with the reductions observed in previously studied plants expressing the CE1 enzyme *An*AXE1 (Pawar et al., 2016, 2017a) or plants with suppressed RWA gene expression (Pawar et al., 2017b). While these reductions were relatively modest, they had significant impact on xylan digestibility by endoxylanases. Overexpression of *Hj*AXE prominently affected the distribution of acidic xylo-oligosaccharides released by endoxylanase *Aa*GH10 (**Figure 3D**). As in *An*AXE1-expressing plants and RWA-suppressed plants, the 2D qHSQC NMR signals from *Hj*AXE-expressing plants indicated that Xyl*p* units were deacetylated at both C-2 and C-3 positions of Xyl*p* (**Figures 3B,C**). This is consistent with previous reports on deacetylation by *Hj*AXE *in vitro* where activity at positions 2 and 3 was detected, and a preference for monoacetylated Xyl*p* (Biely et al., 1997; Hakulinen et al., 2000; Biely et al., 2011; Neumüller et al., 2015). It was suggested that the deacetylation of glucuronosylated Xyl*p* would be difficult based on the structure of the catalytic site (Hakulinen et al., 2000). In agreement, no such activity was observed *in vitro* when acetylated oligosaccharides were used as substrates (Neumüller et al., 2015). However, a significant decrease in X3G2 units accompanied by significant decrease in internal Xylp preceeded by X3G2 (X3G2-X) in WP:CE5 (**Figure 3C**), which was not observed in any of the previously studied transgenic plants, suggested a possibility of deacetylation at C-3 position in the glucuronosylated Xyl*p* units by *Hj*AXE. This conclusion is in line with the observed increase in completely deacetylated 3X + MeGlc xylo-oligosaccharides observed in WP:CE5 lines (**Figure 3D**), which would be expected assuming that *Hj*AXE could deacetylate glucuronosylated Xyl*p*. While this is a novel observation, these results might reflect endogenous GH10 activity (Derba-Maceluch et al., 2015), which creates a free C-4 position at the non-reducing end enabling the migration of an acetyl group to this position (Mastihubova and Biely, 2004). However, we cannot exclude that *Hj*AXE could exhibit additional specificities when acting on native xylan associated with cellulose microfibrils, especially that their cellulose-binding domain could precondition them on cellulose-bound substrates.

HjAXE Does Not Impair Plant Growth

Transgenic plants expressing *Hj*AXE exhibited good height and diameter growth in the greenhouse during a 2-month cultivation period. The only difference from WT was a shorter internode length, which indicates an increased leaf production. It would need to be further studied if such an increase could result in more biomass production by the transgenic lines, and if the field performance of these lines is satisfactory.

HjAXE Targeted to Secondary Cell Walls Improves Woody Biomass Saccharification

Xylan of transgenic plants was shown to be more accessible to hydrolysis by *Aa*GH10 xylanase (**Figure 3D**), confirming previous observations of synergy between AXEs and xylanases in xylan hydrolysis (Biely et al., 1986). Interestingly, the WP:CE5 lines had reduced Xyl content in matrix polysaccharides (**Table 1**), indicative of increased hydrolysis of the deacetylated xylan *in muro* by cell-wall-residing native GH10 enzymes, xylanases and/or transglycosylases (Derba-Maceluch et al., 2015). Similar reductions were observed in other transgenic lines with reduced xylan acetylation (Pawar et al., 2017a,b). If cell-wall-residing xylan was indeed partially hydrolyzed by plant GH10 enzymes, then increased cell-wall porosity, and increased accessibility to cellulose leading to increased glucose yields would be expected. Our analyses demonstrate both.

Increased surface area and porosity of lignocellulose from transgenic plants with reduced acetylation was for the first time demonstrated in this study using BET analysis and Simons' staining (**Table 2** and **Figure 5**). We speculate that increased surface area and porosity is a direct consequence of removal of cell-wall-residing xylan by endogenous GH10 enzymes in the cell walls of transgenic plants. Interestingly, the decrease in xylan biosynthesis that reduces xylan content is also known to improve saccharification but only when coupled with reduction in xylan chain length (Lee et al., 2011a; Ratke et al., 2018).

A 27% increase in glucose yield and 11% increase in xylose yield in enzymatic saccharification of non-pretreated wood, as well as a 3% increase in glucose yield after acid pretreatment and enzymatic saccharification were observed in

WP:CE5 lines (**Figure 4**). Similarly increased sugar yield of enzymatic saccharification without pretreatment was reported for transgenic aspen and *Arabidopsis* with reduced xylan acetylation by either supressing native RWA genes or by overexpressing CE1 AXE (Pawar et al., 2016, 2017a,b). Similarly, the positive influence was smaller after acid pretreatment (Pawar et al., 2017a,b). The observed reductions in recalcitrance could be a direct consequence of increased porosity and accessibility due to xylan deacetylation by CE5 AXE acting in close proximity to cellulose.

Moreover, altered xylan acetylation is thought to highly affect cell wall architecture since the acetylation pattern was shown to mediate xylan binding to cellulose microfibrils (Grantham et al., 2017) as well as xylan covalent linkages to lignin (Giummarella and Lawoko, 2016). This could be the basis of the apparently paradoxical observation that increased acetylation in mutant rice (Zhang et al., 2017) and in transgenic poplar (Yang et al., 2017) resulted in improved saccharification. Applying tools to probe cell wall porosity, and cellulose-xylan as well as lignin-xylan interactions in these acetylation-altered plants could possibly give more definite explanations regarding their recalcitrance behavior.

CONCLUSION

Expression *in planta* of *Hj*AXE from family CE5 leads to reduced xylan acetylation and approximately 30% increased glucose yields in enzymatic saccharification of wood without pretreatment, as well as 3% improved glucose yields even when using industrially relevant pretreatment conditions. Plants expressing *Hj*AXE show good growth, and similar improvement of saccharification and reduction in xylan content as the previously studied plants expressing *An*AXE1 from family CE1, although the cell wall chemotypes and de-acetylation patters show subtle differences between these two types of transgenics. Increased cell wall nanoporosity likely plays a key role in reducing the recalcitrance by *Hj*AXE expression. Structural analyses based on BET and Simons' staining emerge as useful tools for understanding differences in recalcitrance of engineered transgenic wood.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

PP and MD-M prepared plant material, analyzed transcript levels, and determined acetyl esterase activity. ZW and PP performed compositional analysis. MH performed NMR analyses. S-LC performed OLIMP analyses. ZW performed pretreatment and enzymatic saccharification, BET analysis, and Simons' staining. EM and ZW performed statistical analyses. EM, LJ, and MT conceived the study and supervised the experiments. ZW, LJ, and EM wrote the manuscript with contributions from all authors.

FUNDING

This work was supported by the Swedish Energy Agency, Bio4Energy, Formas, SSF program ValueTree, and the Kempe Foundations.

ACKNOWLEDGMENTS

We are grateful to Vimal K. Balasubramanian for help with grinding of wood samples, RNA extraction, and cDNA preparation, and to Anais Meynard for helping with acetic acid analysis during initial screening of transgenic lines. We are also grateful for the services offered by technical platforms at the KBC Chemical-Biological Centre in Umeå including the NMR Core Facility supported by the Knut and Alice Wallenberg Foundation program NMR for Life (www.nmrforlife.se), ViSP (Vibrational Spectroscopy Core Facility) (András Gorzsás), and the Biopolymer Analytical Facility (Junko Takahashi Schmidt).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00380/full#supplementary-material

DATA SHET S1 | Table S1 and Figure S1.

DATA SHEET S2 | NMR Data Part 1.

DATA SHEET S3 | NMR Data Part 2.

REFERENCES

Arantes, V., and Saddler, J. N. (2011). Cellulose accessibility limits the effectiveness of minimum cellulase loading on the efficient hydrolysis of pretreated lignocellulosic substrates. *Biotechnol. Biofuels* 4:3. doi: 10.1186/1754-6834-4-3

Biely, P. (2012). Microbial carbohydrate esterases deacetylating plant polysaccharides. *Biotechnol. Adv.* 30, 1575–1588. doi: 10.1016/j.biotechadv.2012.04.010

Biely, P., Côté, G., Kremnicky, L., Greene, R., and Tenkanen, M. (1997). Action of acetylxylan esterase from Trichoderma reesei on acetylated methyl glycosides. *FEBS Lett.* 420, 121–124. doi: 10.1016/s0014-5793(97)01500-7

Biely, P., MacKenzie, C. R., Puls, J., and Schneider, H. (1986). Cooperativity of esterases and xylanases in the enzymatic degradation of acetyl xylan. *Bio Technol.* 4, 731–733. doi: 10.1038/nbt0886-731

Biely, P., Mastihubová, M., Tenkanen, M., Eyzaguirre, J., Li, X. L., and Vršanská, M. (2011). Action of xylan deacetylating enzymes on monoacetyl derivatives of 4-nitrophenyl glycosides of beta-D-xylopyranose and alpha-l-arabinofuranose. *J. Biotechnol.* 151, 137–142. doi: 10.1016/j.jbiotec.2010.10.074

Biely, P., Puls, J., and Schneider, H. (1985). Acetyl xylan esterases in fungal cellulolytic systems. *FEBS Lett.* 186, 80–84. doi: 10.1016/0014-5793(85)81343-0

Biswal, A. K., Soeno, K., Gandla, M. L., Immerzeel, P., Pattathil, S., Lucenius, J., et al. (2014). Aspen pectate lyase Ptxt PL1-27 mobilizes matrix polysaccharides

from woody tissues and improves saccharification yield. *Biotechnol. Biofuels* 7:11. doi: 10.1186/1754-6834-7-11

Brunauer, S., Emmett, P. H., and Edward, T. (1938). Adsorption of gases in multimolecular layers. *J. Am. Chem. Soc.* 60, 309–319. doi: 10.1016/j.jhazmat.2010.01.120

Busse-Wicher, M., Gomes, T. C. F., 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/tpj.12575

Chandra, R., and Saddler, J. N. (2012). Use of the Simons' staining technique to assess cellulose accessibility in pretreated substrates. *Ind. Biotechnol.* 8, 230–237. doi: 10.1089/ind.2012.0016

Chang, S. J., Puryear, J., and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* 11, 113–116. doi: 10.1007/bf02670468

Chong, S. L., Nissila, T., Ketola, R. A., Koutaniemi, S., Derba-Maceluch, M., Mellerowicz, E. J., et al. (2011). Feasibility of using atmospheric pressure matrix-assisted laser desorption/ionization with ion trap mass spectrometry in the analysis of acetylated xylooligosaccharides derived from hardwoods and *Arabidopsis thaliana*. *Anal. Bioanal. Chem.* 401, 2995–3009. doi: 10.1007/s00216-011-5370-z

Chong, S. L., Virkki, L., Maaheimo, H., Juvonen, M., Derba-Maceluch, M., Koutaniemi, S., et al. (2014). O-Acetylation of glucuronoxylan in *Arabidopsis thaliana* wild type and its change in xylan biosynthesis mutants. *Glycobiology* 24, 494–506. doi: 10.1093/glycob/cwu017

Derba-Maceluch, M., Awano, T., Takahashi, J., Lucenius, J., Ratke, C., Kontro, I., et al. (2015). Suppression of xylan endotransglycosylase PtxtXyn10A affects cellulose microfibril angle in secondary wall in aspen wood. *New Phytol.* 205, 666–681. doi: 10.1111/nph.13099

Donev, E., Gandla, M. L., Jönsson, L. J., and Mellerowicz, E. J. (2018). Engineering non-cellulosic polysaccharides of wood for the biorefinery. *Front. Plant Sci.* 9:1537. doi: 10.3389/fpls.2018.01537

Gandla, M. L., Derba-Maceluch, M., Liu, X.-K., Gerber, L., Master, E. R., Mellerowicz, E. J., et al. (2015). Expression of a fungal glucuronoyl esterase in *Populus*: effects on wood properties and saccharification efficiency. *Phytochemistry* 112, 210–220. doi: 10.1016/j.phytochem.2014.06.002

Gerber, L., Eliasson, M., Trygg, J., Moritz, T., and Sundberg, B. (2012). Multivariate curve resolution provides a high-throughput data processing pipeline for pyrolysis-gas chromatography/mass spectrometry. *J. Anal. Appl. Pyrol.* 95, 95–100. doi: 10.1016/j.jaap.2012.01.011

Gille, S., Cheng, K., Skinner, M. E., Liepman, A. H., Wilkerson, C. G., and Pauly, M. (2011). 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

Giummarella, N., and Lawoko, M. (2016). Structural basis for the formation and regulation of lignin-xylan bonds in birch. *ACS Sustain. Chem. Eng.* 4, 5319–5326. doi: 10.1021/acssuschemeng.6b00911

Goncalves, V. M. F., Evtuguin, D. V., and Domingues, M. R. M. (2007). Structural characterization of the acetylated heteroxylan from the natural hybrid *Paulownia elongata/Paulownia fortune*. *Carbohyd. Res.* 343, 256–266. doi: 10.1016/j.carres.2007.11.002

Gorzsas, A., Stenlund, H., Persson, P., Trygg, J., and Sundberg, B. (2011). Cell-specific chemotyping and multivariate imaging by combined FT-IR microspectroscopy and orthogonal projections to latent structures (OPLS) analysis reveals the chemical landscape of secondary xylem. *Plant J.* 66, 903–914. doi: 10.1111/j.1365-313X.2011.04542.x

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

Grohmann, K., Mitchell, D. J., Himmel, M. E., Dale, B. E., and Schroeder, H. A. (1989). The role of ester groups in resistance of plant cell wall polysaccharides to enzymatic hydrolysis. *Appl. Biochem. Biotechnol.* 2, 45–61. doi: 10.1007/bf02936472

Hakulinen, N., Tenkanen, M., and Rouvinen, J. (2000). Three-dimensional structure of the catalytic core of acetylxylan esterase from *Trichoderma reesei*: insights into the deacetylation mechanism. *J. Struct. Biol.* 132, 180–190. doi: 10.1006/jsbi.2000.4318

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

Jönsson, L. J., and Martin, C. (2016). Pretreatment of lignocellulose: formation of inhibitory by-products and strategies for minimizing their effects. *Bioresour. Technol.* 199, 103–112. doi: 10.1016/j.biortech.2015.10.009

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. *Carbohyd. Res.* 338, 69–77. doi: 10.1016/s0008-6215(02)00351-8

Koutaniemi, S., van Gool, M. P., Juvonen, M., Jokela, J., Hinz, S. W., Schols, H. A., et al. (2013). Distinct roles of carbohydrate esterase family CE16 acetyl esterases and polymer-acting acetyl xylan esterases in xylan deacetylation. *J. Biotechnol* 168, 684–692. doi: 10.1016/j.jbiotec.2013.10.009

Lee, C., Teng, Q., Huang, W., Zhong, R., and Ye, Z. H. (2011a). Molecular dissection of xylan biosynthesis during wood formation in poplar. *Mol. Plant.* 4, 730–747. doi: 10.1093/mp/ssr035

Lee, C. H., Teng, Q., Zhong, R. Q., and Ye, Z. H. (2011b). 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

Manabe, Y., Verhertbruggen, Y., Gille, S., Harholt, J., Chong, S. L., Pawar, P. M. A., 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

Margolles-Clark, E., Tenkanen, M., Söderlund, H., and Penttilä, M. (1996). Acetyl xylan esterase from *Trichoderma reesei* contains an active-site serine residue and a cellulose-binding domain. *Eur. J. Biochem.* 237, 553–560. doi: 10.1111/j.1432-1033.1996.0553p.x

Mastihubova, M., and Biely, P. (2004). Lipase-catalysed preparation of acetates of 4-nitrophenyl-D-xylopyranoside and their use in kinetic studies of acetyl migration. *Carbohyd. Res.* 339, 1353–1360. doi: 10.1016/j.carres.2004.02.016

McCann, M. C., and Carpita, N. C. (2015). Biomass recalcitrance: a multi-scale, multi-factor, and conversion-specific property. *J. Exp. Bot.* 66, 4109–4118. doi: 10.1093/jxb/erv267

Neumüller, K. G., Carvalho de Souza, A., van Rijn, J. H. J., Streekstra, H., Gruppen, H., and Schols, H. A. (2015). Positional preferences of acetyl esterases from different CE families towards acetylated 4-O-methyl glucuronic acid-substituted xylo-oligosaccharides. *Biotechnol. Biofuels* 8, 7. doi: 10.1186/s13068-014-0187-6

Pan, X. J., Gilkes, N., and Saddler, J. N. (2006). Effect of acetyl groups on enzymatic hydrolysis of cellulosic substrates. *Holzforschung* 60, 398–401. doi: 10.1515/hf.2006.062

Pawar, P. M., Derba-Maceluch, M., Chong, S. L., Gandla, M. L., Bashar, S. S., Sparrman, T., et al. (2017a). In muro deacetylation of xylan increases lignin extractability and improves saccharification of aspen wood. *Biotechnol. Biofuels* 10:98. doi: 10.1186/s13068-017-0782-4

Pawar, P. M., Ratke, C., Balasubramanian, V. K., Chong, S. L., Gandla, M. L., Adriasola, M., et al. (2017b). Downregulation of RWA genes in hybrid aspen affects xylan acetylation and wood saccharification. *New Phytol.* 214, 1491–1505. doi: 10.1111/nph.14489

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.-A., Derba-Maceluch, M., Chong, S.-L., Gómez, L. D., Miedes, E., Banasiak, A., et al. (2016). Expression of fungal acetyl xylan esterase in *Arabidopsis thaliana* improves saccharification of stem lignocellulose. *Plant Biotechnol. J.* 14, 387–397. doi: 10.1111/pbi.12393

Ragauskas, A. J. (2013). *Materials for Biofuels*. Singapore: World Scientific.

Ratke, C., Pawar, P. M., Balasubramanian, V. K., Naumann, M., Duncranz, M. L., Derba-Maceluch, M., et al. (2015). Populus GT43 family members group into distinct sets required for primary and secondary wall xylan biosynthesis and include useful promoters for wood modification. *Plant Biotechnol. J.* 13, 26–37. doi: 10.1111/pbi.12232

Ratke, C., Terebieniec, B. K., Winestrand, S., Derba-Maceluch, M., Grahn, T., Schiffthaler, B., et al. (2018). Downregulating aspen xylan biosynthetic

GT43 genes in developing wood stimulates growth via reprograming of the transcriptome. *New Phytol.* 219, 230–245. doi: 10.1111/nph.15160

Teleman, A., Lundqvist, J., Tjerneld, F., Stålbrand, H., and Dahlman, O. (2000). Characterization of acetylated 4-O-methylglucuronoxylan isolated from aspen employing H-1 and C-13 NMR spectroscopy. *Carbohyd. Res.* 329, 807–815. doi: 10.1016/s0008-6215(00)00249-4

Teleman, A., Tenkanen, M., Jacobs, A., and Dahlman, O. (2002). Characterization of O-acetyl-(4-O-methylglucurono) xylan isolated from birch and beech. *Carbohyd. Res.* 337, 373–377. doi: 10.1016/s0008-6215(01)00327-5

Wang, Z., Winestrand, S., Gillgren, T., and Jönsson, L. J. (2018). Chemical and structural factors influencing enzymatic saccharification of wood from aspen, birch and spruce. *Biomass Bioenerg.* 109, 125–134. doi: 10.1016/j.biombioe.2017.12.020

Xiong, G. Y., 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

Yang, Y., Yoo, C. G., Winkeler, K. A., Collins, C. M., Hinchee, M. A. W., Jawdy, S. S., et al. (2017). Overexpression of a domain of unknown function 231-containing protein increases O-xylan acetylation and cellulose biosynthesis in *Populus*. *Biotechnol. Biofuels* 10:311. doi: 10.1186/s13068-017-0998-3

Yu, X. C., Minor, J. L., and Atalla, R. H. (1995). Mechanism of action of Simons' stain. *TAPPI J.* 78, 175–180.

Yuan, Y. X., Teng, Q., Zhong, R., Haghighat, M., Richardson, E. A., and Ye, Z.-H. (2016). 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. X., Teng, Q., Zhong, R. Q., 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

Zhao, X., Zhang, L., and Liu, D. (2012). Biomass recalcitrance. Part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuel. Bioprod. Bior.* 6, 465–482. doi: 10.1002/bbb.1331

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 24 April 2020
doi: 10.3389/fpls.2020.00479

Direct Measurement of Plant Cellulose Microfibril and Bundles in Native Cell Walls

Bo Song[1†], Shuai Zhao[1,2,3†], Wei Shen[1,2], Cynthia Collings[1,2] and Shi-You Ding[1,2]*

[1] Department of Plant Biology, Michigan State University, East Lansing, MI, United States, [2] Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI, United States, [3] State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, College of Life Science and Technology, Guangxi University, Nanning, China

Plants use rigid cellulose together with non-cellulosic matrix polymers to build cell walls. Cellulose microfibrils comprise linear β(1,4)-glucan chains packed through inter- and intra-chain hydrogen-bonding networks and van der Waals forces. Due to its small size, the number of glucan chains and their arrangement in a microfibril remains elusive. Here we used atomic force microscopy (AFM) to directly image primary cell walls (PCWs) and secondary cell walls (SCWs) from fresh tissues of maize (*Zea mays*) under near-native conditions. By analyzing cellulose structure in different types of cell walls, we were able to measure the individual microfibrils in elongated PCWs at the sub-nanometer scale. The dimension of the microfibril was measured at 3.68 ± 0.13 nm in width and 2.25 ± 0.10 nm in height. By superimposing multiple AFM height profiles of these microfibrils, the overlay area representing the cross-section was estimated at 5.6 ± 0.4 nm^2, which fitted well to an 18-chain model packed as six sheets with 234432 conformation. Interestingly we found in PCW, all these individual microfibrils could be traced back to a bundle in larger imaging area, suggesting cellulose are synthesized as large bundles in PCWs, and then split during cell expansion or elongation. In SCWs where cell growth has ceased we observed nearly-parallel twined or individual microfibrils that appeared to be embedded separately in the matrix polymers without the splitting effect, indicating different mechanisms of cellulose biosynthesis in PCW and SCW. The sub-nanometer structure of the microfibril presented here was measured exclusively from elongated PCWs, further study is required to verify if it represents the inherent structure synthesized by the cellulose synthase complex in PCWs and SCWs.

Keywords: cellulose microfibril, atomic force microscopy, direct imaging, primary cell wall, secondary cell wall, cellulose synthesis

OPEN ACCESS

Edited by:
Jin Zhang,
Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Jenny C. Mortimer, Lawrence Berkeley National Laboratory, United States
Li Tan, University of Georgia, United States

****Correspondence:***
Shi-You Ding
sding@msu.edu

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *18 December 2019*
Accepted: *31 March 2020*
Published: *24 April 2020*

Citation:
Song B, Zhao S, Shen W, Collings C and Ding S-Y (2020) Direct Measurement of Plant Cellulose Microfibril and Bundles in Native Cell Walls. Front. Plant Sci. 11:479. doi: 10.3389/fpls.2020.00479

INTRODUCTION

Plant growth and development relies on the regulation of cell wall biogenesis. As the main skeletal component, cellulose forms interwoven microfibril networks to constitute the multilayer (lamellae) architecture observed for plant cell walls (Somerville et al., 2004). During cell growth and development, the biosynthesis and dynamic arrangement of the cellulose microfibrils play a key role in maintaining the mechanical properties and physiological functions of the cell walls (Cosgrove, 2016; Zhang et al., 2017, 2019). Cellulose has relatively simple chemistry that comprises a number of

linear homopolymeric chains of β(1,4)-D-glucosyl residues packed through intra- and inter-chain hydrogen bonding networks and van der Waals forces to form *para*-crystalline microfibrils. The native structures of cellulose have been determined based on non-plant sources of large (20–40 nm) cellulose crystals (Nishiyama et al., 2002, 2003). In plant however, the cellulose microfibril has a small cross-sectional dimension (2–3 nm), in which the number of chains and how they pack into a microfibril is unknown.

Traditional high-resolution imaging techniques, such as electron microscopy (McCann et al., 1990; Xu et al., 2006) and field emission SEM (Carpita et al., 2001a,b; Zheng et al., 2017), have been extensively used to measure the cellulose microfibrils in plants, resulting in diameters in a range of ˜3–50 nm depending on cell wall types. This wide range of size distribution probably represents microfibril bundles that either exist in native cell walls or are formed during sample preparation (Carpita et al., 2001a,b; Ding and Himmel, 2006; Ding et al., 2012; Zhang et al., 2013, 2016, 2017; Zheng et al., 2017). Despite the daunting challenges of directly measuring the cellulose microfibril, analytic methods, such as nuclear magnetic resonance (NMR) and diffraction-based techniques, have been widely used to characterize the physicochemical properties of plant cellulose. Early works by Chanzy et al. (1978, 1979) proposed a 15–25-chain cellulose fibril based primarily on diffraction data. Newman et al. (2013) used various analytic approaches, such as solid-state NMR, small-angle X-ray scattering, synchrotron wide-angle X-ray scattering, and computer simulation techniques and proposed 18-chain models with mixed cross-sectional shapes and possible microfibril twinning in the cell wall. Controversially, a model containing at least 24-chains has also been proposed by other researchers based on analysis of similar techniques (Wang and Hong, 2016) when the microfibril twinning effect is not considered.

Native cellulose in plant cell walls often appear to be bundles with variable sizes and closely associated with hemicelluloses. Cellulose structure could also be continuously modified during cell expansion, elongation and cell wall thickening and lignification (Busse-Wicher et al., 2014). Analysis of cellulose often requires chemical treatment and/or dehydration processes, which further alter cellulose structure (O'Neill et al., 2017), it is therefore difficult to interpret the diffraction data and calculate the fundamental structure of a microfibril based on ensemble average measurement.

Discovery of cellulose synthase (*CESA*) genes in different plant species and biochemical studies of cellulose synthase complexes (CSC) (Kimura et al., 1999) have provided new insights into prediction of microfibril structure. Plant cellulose is synthesized in plasma membrane by multiprotein CSCs. Observations using freeze fracture electron microscopy (FF-TEM) (Mueller et al., 1976; Giddings et al., 1980; Mueller and Brown, 1980; Nixon et al., 2016) and immuno-EM (Kimura et al., 1999) have suggested that the CSC appears to be a six-lobed rosette containing multiple CESAs. It has been therefore postulated that the number of chains in a microfibril should be 6-fold, assuming the CSC comprises only active CESAs and each CESA synthesizes one cellulose chain, thus 18- (Jarvis, 2013; Newman et al., 2013; Nixon et al., 2016), 24- (Fernandes et al., 2011; Thomas et al., 2013; Wang and Hong, 2016) or 36-mer (Scheible et al., 2001; Doblin et al., 2002) models of CSCs and corresponding microfibril models containing 18-, 24-, and 36-chains, respectively have been proposed. Computational simulations (Oehme et al., 2015) and density functional theory calculations (Kubicki et al., 2018) have suggested that 18-chain is more favorable than 24- or 36-chain models (Haigler and Roberts, 2019).

Atomic force microscopy (AFM) uses a sharp tip to probe the surface features by raster scanning, which offers a non-destructive approach to characterize biological materials from cellular to molecular scales (Dufrene et al., 2017). The resolution of early AFM works was comparable to electron microscopy, and showed additional fine details of microfibril arrangement when imaging plant cell walls (Kirby et al., 1996). However the quality of an AFM image relies substantially upon the sharpness of the tip, firmness and flatness of the sample, and imaging environments. AFM technique has recently been greatly improved by the development of ultra-sharp probes (1–2 nm) combined with new operation modes for imaging in liquid (Pyne et al., 2014; Shiotari and Sugimoto, 2017; Wang et al., 2017), resulting in typically an order of magnitude enhancement of spatial resolution compared to that obtained in ambient conditions (Fukuma et al., 2007; Israelachvili, 2011; Voitchovsky, 2013; Miller et al., 2016). While imaging plant tissue, the quality of AFM images could be affected by the 3-dimensional structure of the cell wall at the cellular (micrometer) scale and the complex architecture at the molecular (nanometer) scale. Previous studies have shown that the sizes of the microfibril appear larger under dehydrated condition than that observed in water (Pesacreta et al., 1997; Thimm et al., 2000). Our early studies (Ding and Himmel, 2006; Ding et al., 2012, 2014) have revealed 3–5 nm microfibril, the uncertainty of measured dimension is likely due to over-estimation of the actual size under dehydrated condition. Recently, imaging cell walls in aqueous buffer have estimated the width of the cellulose microfibril at ˜3.5 nm (Zhang et al., 2013, 2016, 2017). However, the quality of these published images has yet to be sufficient to resolve the dimension and the cross-sectional shape of the microfibril.

In this study, we took the advantages of AFM imaging in aqueous condition and a combined effort of sample preparation, pre-selection of tips and systematic adjustment of imaging parameters to optimize the quality of images. The primary cell walls (PCWs) and secondary cell walls (SCWs) from fresh tissues of maize (*Zea mays*) (**Figure 1**) were directly imaged and analyzed at the sub-nanometer scale.

RESULTS

Optimization of AFM Image Quality

To minimize the potential alteration of cell wall structure, we used a double-edged razor blade to hand-cut fresh tissues longitudinally to yield ˜5–10-μm slices containing a single layer of cells, and the sample was washed by water, mounted on a glass slide pre-coated by poly-lysine and then imaged in water (see section "Materials and Methods"). With the aid of in-line optical microscope, we positioned the AFM tip onto specific cell

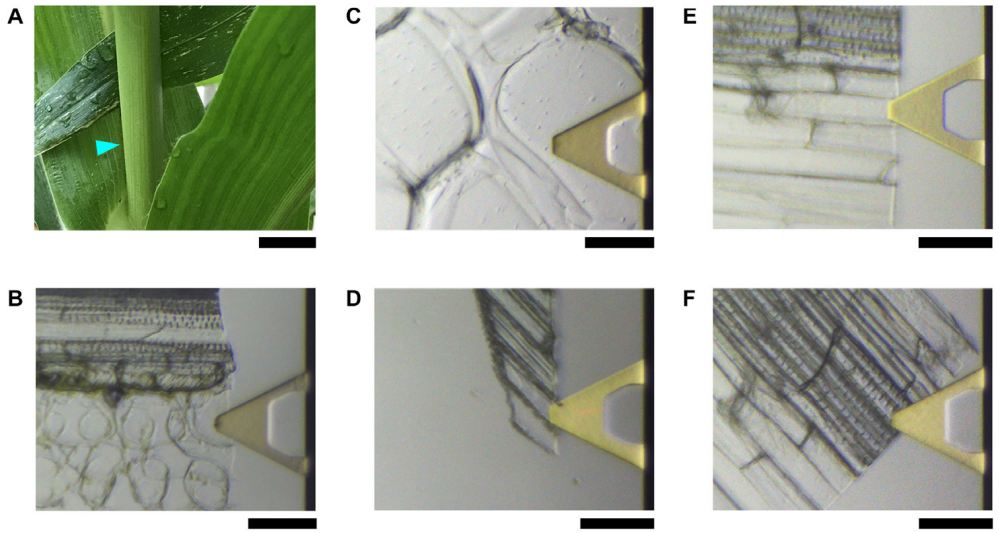

FIGURE 1 | Maize plant and tissues used for AFM imaging. **(A)** Four-week vegetative growing maize plant. The cyan triangle indicates the approximate location of the third internode where the tissues containing a single layer of cells in **(B–F)** are obtained. Light microscopy of gold-plated (yellow) triangles in **(B–F)** are the AFM cantilever, in which the tip is located underneath; photos are taken in real-time during imaging in water showing the actual areas where AFM images are taken. **(B)** Mature Vascular bundle sheath cells that have thin primary cell wall (PCW). **(C)** Thickened and expanded PCW in stem pith parenchyma. **(D)** Thickened and elongated PCW in the rind fiber under the epidermis. **(E)** Secondary cell walls (SCWs) of fibers in vascular bundle. **(F)** SCWs in xylem vessel cells. Scale bar = 2 cm **(A)**, 50 μm **(B–F)**.

wall types (**Figure 1**) based on their morphological structures, so that different PCWs from expanded or elongated parenchyma cells, and SCWs from vessel and sclerenchyma fiber cells could be imaged reproducibly using different AFM tips and imaging parameters. We scanned initially in large sizes, i.e., 5 – 10 μm to localize areas of interest and "zoom-in" progressively from 1 μm to 100 nm with 512–1024 scan lines to allow observation of different scales from overall microfibril arrangement to sub-nanometer features of individual microfibrils. For each type of cell walls, we optimized the imaging process by altering scan sizes, scan rates, and applied forces, and repeatedly imaged the same area or the same type of walls but from independent sample preparations to compare the consistency of measured features. Using this strategy we could assess the tip quality with optimized parameter settings by means of identifying artifacts generated from the sample itself, environmental noise and mechanical drift.

In many cases, we observed particle shape features that appeared ultra-soft, presumably debris of cytosolic materials in large scan areas (**Figure 2**), which we intentionally excluded when imaging in small areas to reduce possible tip contamination and thus focused on tuning the imaging parameters optimized to observe fine details of the cellulose microfibrils.

The cell wall samples used in this study were only washed by water; it is assumed that cellulose microfibrils and matrix polymers, such as hemicelluloses and pectins are co-localized in native cell walls (Simmons et al., 2016). Previous studies have shown that the matrix polymers are not normally detectable by an AFM tip in liquid, due to the mobile nature of these polymers with the force applied by the probe (Zhang et al., 2016). Since the accuracy of AFM measurement is largely determined by the sufficient force required to gain imaging contrast without mechanical deformation of the sample (Leung et al., 2012), in this study we found the constant force less than 500 pN was critical to obtain high quality image showing sub-nanometer characteristics of the microfibril. At force ranging from 200 to 500 pN, we were able to obtain relatively sharp images of the microfibrils and minimized the effect of less-defined amorphous structures between the microfibrils (**Figures 3**, **4**).

Different Types of Cell Walls

Five types of cells were imaged extensively from fresh tissues of the stem of a vegetative growing maize plant (**Figure 1A**), including, (1) vascular bundle sheath cells (**Figure 1B**). Maize is a C4 plant, the bundle sheath cells contain chloroplast where the major step of photosynthesis, Calvin-Benson-Bassham cycle occurs to fix carbon into sugars. Therefore these expanded cells contain thin PCWs. (2) Pith parenchyma (**Figure 1C**). These cells normally expanded and have thickened PCWs. (3) Rind fiber cells immediately under the epidermis (**Figure 1D**). These cells are elongated and have thickened PCWs. (4) Sclerenchyma fibers adjacent to vascular bundle sheath (**Figure 1E**). These cells are elongated and have thickened lignified SCWs. (5) Vessels (**Figure 1F**). These cells are expanded and elongated and have thickened lignified SCWs.

In all observed PCWs cellulose appeared to be bundles with variable widths between 5 to 30 nm, which was consistent with previous studies (McCann et al., 1990; Carpita et al., 2001b; Ding and Himmel, 2006; Ding et al., 2012; Zhang et al., 2016). These large bundles split into small bundles and individual microfibrils (**Figure 3**). Some amorphous substances appeared to be bridging between microfibrils, which was similar with the observation by EM techniques (McCann et al., 1990). The

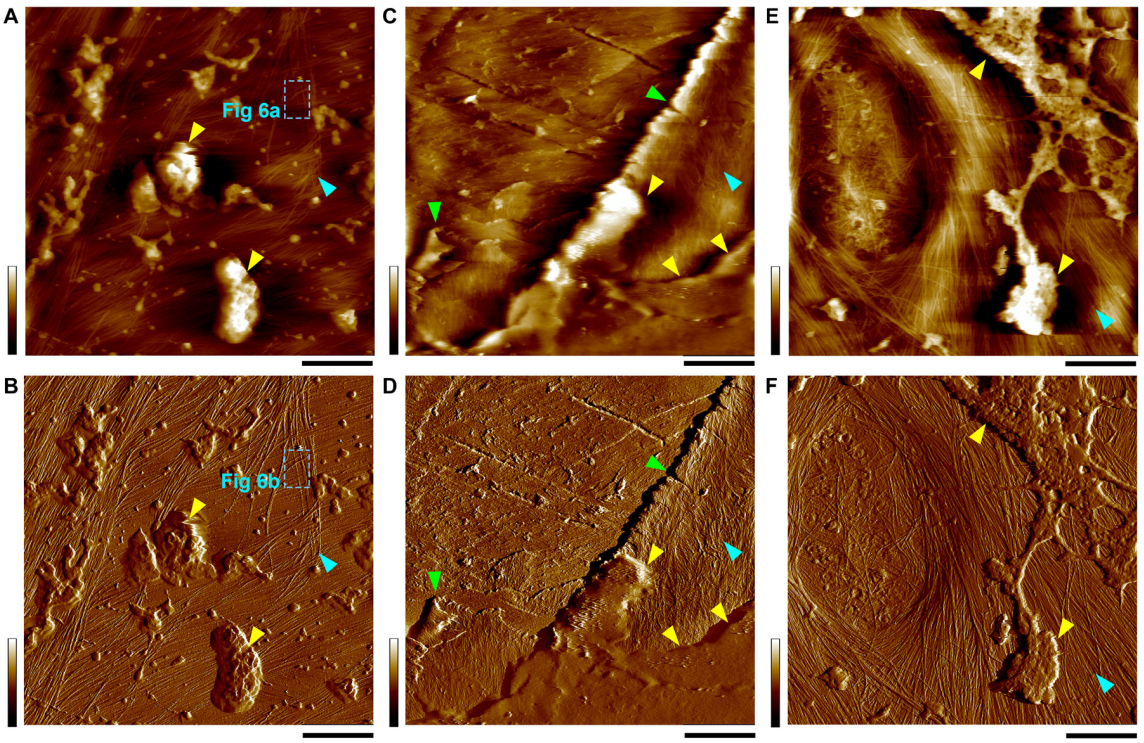

FIGURE 2 | Height **(A,C,E)** and PeakForce error **(B,D,F)** images taken simultaneously in thickened PCWs. Yellow triangles indicate the amorphous substances, likely cytoplasmic debris that appears to be clearly distinguishable on wall surface, specifically in the PeakForce error images. Cyan triangles indicate the inner surface of the cell wall where cellulose microfibril networks can be observed. The boxes with cyan dashed lines **(A,B)** indicate the location of images in **Figures 6A,B**. Green triangles **(C,D)** Indicate the broken layers of the cell wall lamellae probably due to sample cutting during preparation. Occasionally, a primary pit field can also be observed **(E,F)**. Scale bar = 400 nm **(A,B,E,F)** and 2 μm **(C,D)**. Color bar = 100 nm **(A,E)**, 2 nN **(B)**, 200 nm **(C)**, 4 nN **(D)**, and 700 pN **(F)**.

SCWs were imaged from sclerenchyma fibers (**Figures 1E**, **4A,B**) and xylem vessels (**Figures 1F**, **4C–F**), in which predominately individual or twinned microfibrils were observed. The microfibrils in SCWs appeared to be near-parallel and independently embedded in the matrix polymers without further splitting effect (**Figure 4**). We found that such fibril splitting effect could be a signature feature to distinguish PCW and SCW, which is likely an indication of cell growth. We speculate that these large cellulose bundles are synthesized by multiple CSCs, which split as cell volume increases during cell expansion or elongation. All individual microfibrils observed in this study indeed could be traced back to a bundle of larger scales (**Figure 3**). The SCWs are deposited after cell growth has ceased, cellulose microfibrils are synthesized as individuals without further splitting.

The vessel wall is featured with large pits and pit cavities (**Figure 4C**). The pit membrane is formed before SCW deposition and is considered to be PCW, where only large bundles were observed (**Figure 4F**).

In the SCWs of the vessel, the microfibrils were observed both from the side (perpendicular) wall of the pit cavity (**Figure 4D**) and the surface (**Figure 4E**). The microfibrils appeared to be wider from the side view (**Figure 4D**) than those from the top view (**Figure 4E**), indicating that the cross-section of a microfibril is asymmetrical and the narrow side is vertically arranged on the wall surface (**Figures 4B,D**). Microfibril ends are also observed in the surface of SCWs (**Figure 4E**).

Measurement of Individual Microfibrils

In survey of all AFM observations in 100–200 nm scan sizes from different types of cell walls, we found it was extremely difficult to measure the size and the cross-sectional shape of individual microfibrils. The measurement uncertainties include: (1) Bundling. In the PCWs, the overall structure of these bundles appeared to be ribbon-like and the size varied substantially in different walls. Although it was possible to estimate the numbers of microfibrils in a bundle based on its sequential splitting, the measured data was insufficient to calculate the accurate size of the microfibril due to unknown confirmation of these microfibrils in the bundle. (2) Dangling. During cell wall synthesis, cellulose microfibrils are deposited by layers and form complicated 3-dimentional networks. While the AFM tip detects the dangling microfibril with applied force, non-linear tip dilation artifacts may be generated thus increase the baseline noise. (3) Matrix polymers. The cell walls we imaged in this study were simply washed by water to minimize potential alteration of the native structure of the microfibril, however, the downside was the substantial amount of matrix polymers associated on the surface of the microfibril, especially in the SCWs (**Figure 4**), which could significantly

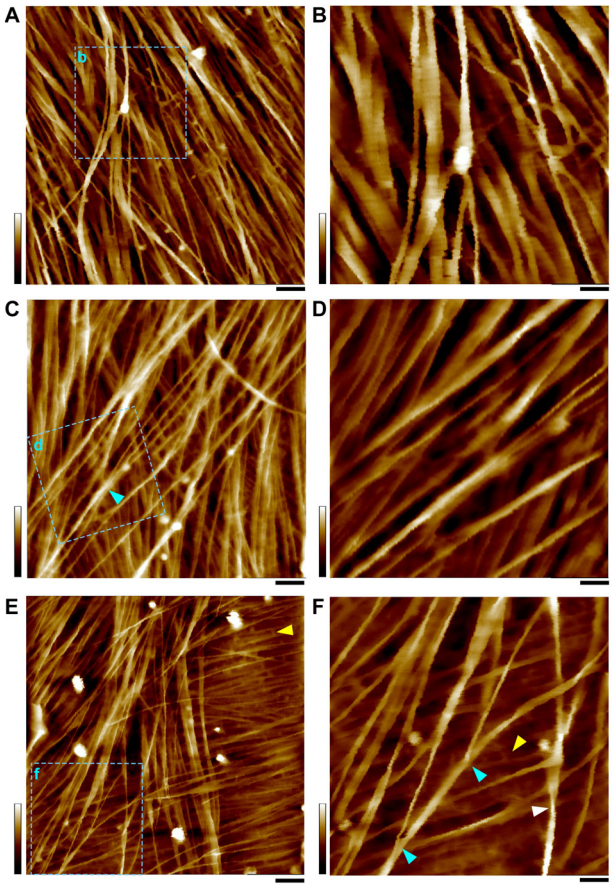

FIGURE 3 | AFM height images of PCWs. **(A,B)** Thin and expanded PCW from bundle sheath cell (**Figure 1B**). **(C,D)** Thickened and expanded PCW from stem pith parenchyma (**Figure 1C**). **(E,F)** Thickened and elongated PCW from rind fiber (**Figure 1D**). The scan areas of **(B,D,F)** are indicated as boxes with cyan dashed-lines in **(A,C,E)**, respectively. Cell walls are prepared from the stem of living maize plant (**Figure 1A**) and imaged in water. A mixture of variable sizes of microfibril bundles are observed that may split into smaller bundles or individual microfibrils (cyan triangles in **C,F**). Amorphous substance (McCann et al., 1990) appeared to be matrix wall materials between microfibrils are indicated by yellow triangles **(E,F)**. Twisting microfibrils (white triangle in **F**) are found on wall surface. Scale bars = 50 nm **(A,C,E)** and 20 nm **(B,D,F)**. Color bars = 20 nm **(A,C,E)** and 15 nm **(B,D,F)**.

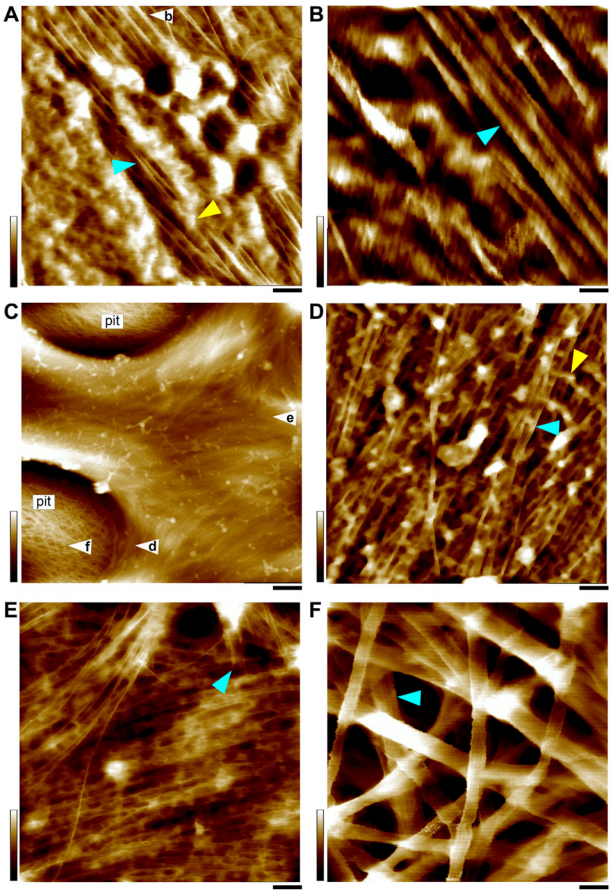

FIGURE 4 | AFM height images of SCWs. **(A,B)** SCW of the fiber in the vascular bundle (**Figure 1E**). **(C–F)** Xylem vessel wall (**Figure 1F**). The SCWs contain predominately individual or twinned microfibrils indicated by cyan triangles in **(A,B,D,E)**, except the pit membrane area **(C,F)** that is the PCW containing large bundles. Apparent amorphous matrix components are indicated as yellow triangles in **(A,D)**. **(B)** "Zoom-in" image of the area in (white triangle in **A**), showing twinned microfibrils (cyan triangle). **(C)** The SCW between two large pits. **(D–F)** Are imaged from the areas indicated in (white triangles in **C**), respectively. **(D)** The microfibrils are imaged from the vertical wall of the pit cavity showing side view of twinned microfibrils. **(E)** The microfibrils in the surface are mostly parallel-arranged individual microfibrils and the termini are indicated by cyan triangle. **(F)** Large bundles (cyan triangle) in the pit membrane. Cell walls are prepared from the stem of a living maize plant (**Figure 1A**). Scale bars = 50 nm **(A,D,E,F)**, 20 nm **(B)**, and 500 nm **(C)**. Color bars = 15 nm **(A,D,E)**, 5 nm **(B)**, 20 nm **(F)**, 250 nm **(C)**.

increase the uncertainty of measuring the microfibril at the sub-nanometer scale. To address these issues, we developed several strategies to optimize measurement accuracy at the sub-nanometer scale.

We used pre-selected tips that were approximately 1 nm in radius and repeatedly imaged the same sample at least by three new tips to ensure image reproducibility in 100–200 nm scan scales. By examining the same microfibril imaged by different tips we found that the first couple of images appeared ultra-sharp when a new tip was used, and quickly became blurred in details after a few scans even though large features appeared to be the same, suggesting that in the case of imaging a small area it was possible that the geometry of the tip apex was critical rather than the overall tip size that could be easily worn out or contaminated (Santos et al., 2011). We further analyzed only the images taken by new tips.

By exploring the high quality AFM images, we found only in the case of a microfibril that run across the top of another microfibril (**Figure 5A**), thus provided a relative firm base locally to allow highly stable data acquisition, which could be found in the surface of elongated PCWs. We developed a simple three-line method (**Figure 5**) to measure the height and width of the microfibril based on raw image data.

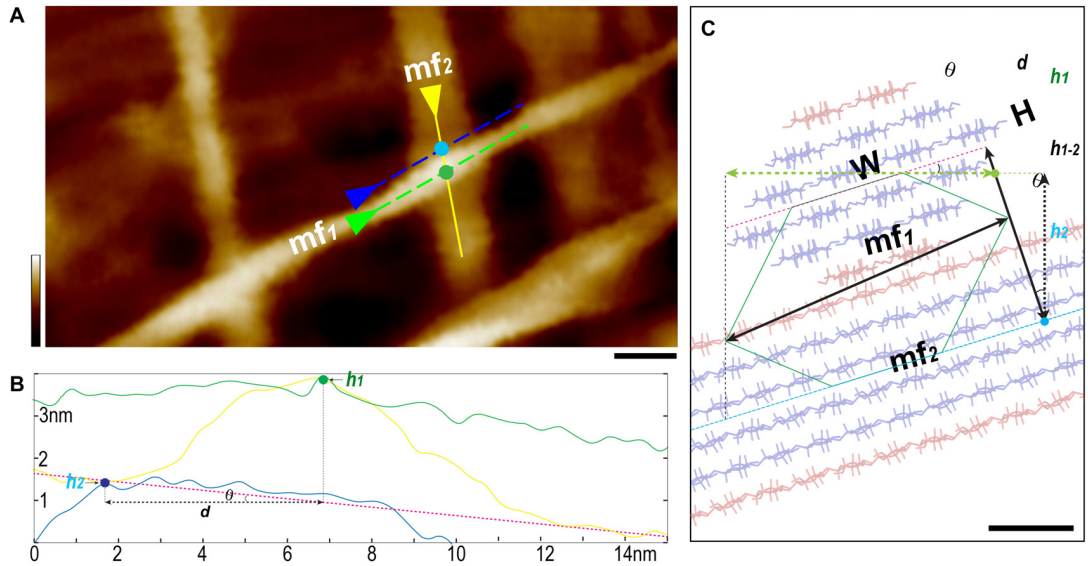

FIGURE 5 | Three-line method to measure individual microfibril. **(A)** A typical AFM height image of thickened and elongated PCW scanned in 100 × 50 nm^2 area with 1024 scan lines. Two well-defined microfibrils indicated as "mf*1* on the top of mf*2*" fashion are selected for measurement. **(B)** Using the section tool of AFM analysis software, three lines are drawn as shown in yellow, green and blue in **(A)** to generate height profiles. The *Rmax* value of mf*1* can be read out from the line profile (green line, **B**). **(C)** The 18-chain microfibril model (red-blue sticks, W, width; H, height) is used to show the relative arrangement of two microfibrils mf*1* (end view) and mf*2* (side view), red sticks present the hydrophobic surface chains. Based on this geometry, the height difference (h_{1-2}) between mf*1* (measured height point h_1, green dots in **A,B**) and mf*2* (measured height point h_2, light blue dots in **A,B**) is calculated as $h_1 - h_2$, and the tilting angle θ of the mf*2* is determined by applying the first-order fitting (pink dash line in **B,C**). The distance (*d*) is between the two height-measuring lines (green and blue dash lines in **A**). The height of mf*1* is calculated as $H = [h_{1-2}+d^*\tan(\theta)]^* \cos(\theta)$. Scale bar = 10 nm **(A)** and 1 nm **(C)**. Color bar = 10 nm **(A)**.

The first line was drawn on the top of the target microfibril (**Figure 5A**, **mf*1***) along its long axis. The *Rmax* value (maximum vertical distance between the highest and lowest data points after the planefit) was read from the line profile of height image, which was used to determine if there were matrix polymers directly associated with the microfibril in the measured area. Considering the theoretical size of a sugar, such as a glucose molecule is 0.7–1 nm, the relatively mobile matrix polysaccharides could contribute at least 1 nm to the *Rmax*. If we choose the area that has *Rmax* smaller than 0.5 nm, it is plausible to assume the microfibril is clean cellulose. Our results suggest that in many cases the microfibrils are not fully covered by matrix polymers, especially these on the PCW surface (**Figures 5**, **6**). Further measurement of the microfibril was only carried out in the area where apparently no matrix polymer was directly associated with the microfibril (*Rmax* < 0.5 nm).

The microfibril in the interwoven networks was often tilted with respect to the cell wall surface, which required a local plane correction. We drew a second line along the long axis of the bottom microfibril (**Figure 5A**, **mf*2***), and the third line that was parallel to but immediately adjacent to **mf*1***, so that the height (**H**) of **mf*1*** can be calculated based on the tilting angle (θ) and measured height values, respectively (**Figures 5B,C**).

It is known when imaging using tapping mode, such as PeakForce tapping used in this study, while the tip scans the edge of a feature (i.e., the microfibril), a momentary spike in the error signal appears before the controller can adjust the tip height. Therefore in a PeakForce error image that was taken simultaneously with the height image, the point of the spike could be used to estimate the edge of the feature (Andersson, 2006). Using this method, we drew a single line perpendicular across the microfibril in both height (**Figure 6A**) and PeakForce error (**Figure 6B**) images that were acquired simultaneously, and overlaid these two line profiles to determine the microfibril edges (**Figure 6C**, points ***a*** and ***b***), the measured width value (***w***) could be read as horizontal distance between the point ***a*** and ***b***, and the actual width (**W**) could then be calculated based on the three-line method (**Figure 5**).

Previously we have found that the microfibril occasionally appeared twisted or laid down in different conformations on the wall (**Figure 6**), so that both height and width values of the same microfibril could be measured accurately based on height profiles (Ding et al., 2012). In this study, we further measured the microfibrils laid down on the wall surface in different conformations, i.e., horizontal (**Figure 6C**) vs. vertical (**Figure 6D**) in the same image, and the different line profiles suggest that the microfibril has an asymmetrical cross-sectional shape, which agreed with the observation in SCWs (**Figures 4B,D**).

Considering the AFM tip scanning on the surface of the cell walls only detects at maximum half of the microfibril surface – when the microfibril lies down in a nearly horizontal conformation (**Figure 6C**). We aligned and overlaid the edges (**Figure 6C**, points ***a*** and ***b***) of multiple height profiles, so that the cross-sectional area could be estimated (**Figure 6E**).

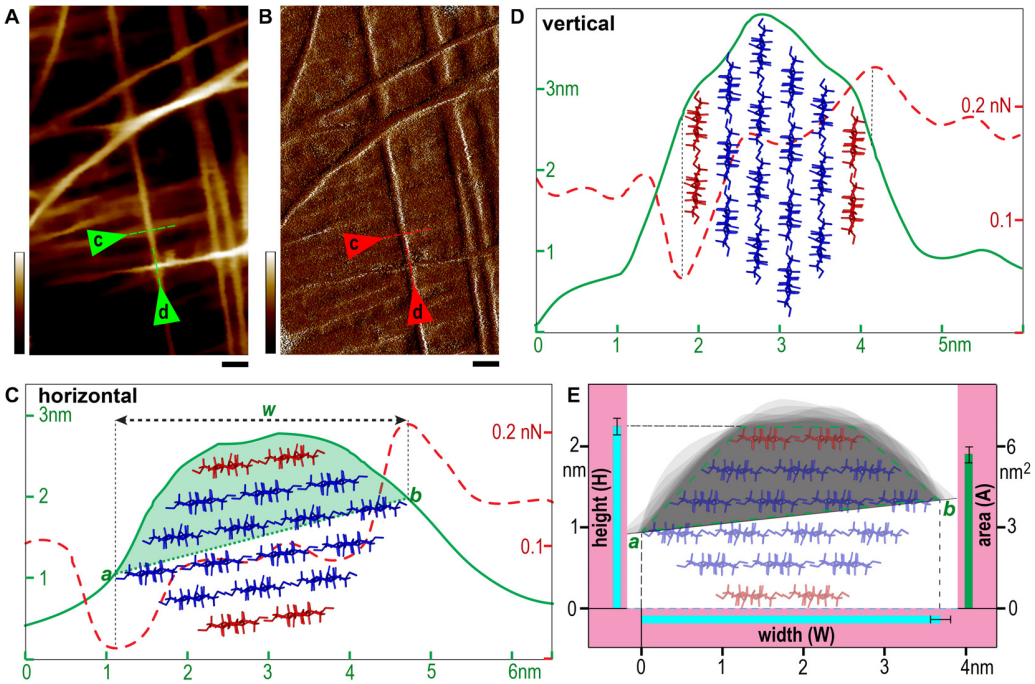

FIGURE 6 | Microfibril measurement. **(A,B)** Typical height **(A)** and PeakForce error **(B)** images taken in 100 nm × 200 nm scan area. **(C,D)** Line profiles of two representative microfibrils indicated by triangles and dashed lines in **(A,B)** laid down on wall surface in nearly horizontal **(C)** or vertical **(D)** conformations. Profiles of heights are presented as solid green lines and PeakForce errors are dashed red lines. The edges of the microfibril (points *a* and *b* in the line profiles in **C**) are determined based on the spike PeakForce error signals (Andersson, 2006), and *w* indicates the horizontal distance. The half cross-sectional area of the microfibril (light green fill in **C**) is estimated between the line *a*-*b* and the height profile. **(E)** The heights (**H**), widths (W) and areas (**A**) are calculated using the three-line method described in **Figure 5** presented as mean values with standard deviation (black bar) based on measured data from different microfibrils. The gray shaded area in (**E**, light green) indicates the overlay of multiple line profiles of individual microfibrils by aligning the points *a* and *b* in **(C)**. The data were measured based on a total of 40 raw images from different areas obtained from 30 sample preparations (only fresh sample and a new tip is used for each experiment). The proposed 18-chain microfibril model is showed as red (two chains in the hydrophobic surface) and blue sticks, which is used to show an empirical fit of the model to the experimental data **(C–E)**, and the half cross-sectional area is illustrated as a quadrilateral with green dashed lines in **(E)**. Scale bar = 20 nm **(A,B)**. Color bar = 10 nm **(A)** and 150 pN **(B)**.

Using these strategies, individual microfibrils were calculated with width (**W**), height (**H**) and cross-sectional area at 3.68 ± 0.13 nm (n = 33), 2.25 ± 0.10 nm (n = 63), and 5.6 ± 0.4 nm^2 (n = 15), respectively. We then built microfibril models based on recent proposed models containing 18 or 24 chains (Fernandes et al., 2011; Jarvis, 2013; Newman et al., 2013; Thomas et al., 2013; Zhang et al., 2013; Nixon et al., 2016; Wang and Hong, 2016), assuming that plant microfibrils exhibit the same native cellulose Iβ structure (Nishiyama et al., 2002, 2003), and the chains have relatively regular arrangement. Theoretical heights and widths were estimated based on the conformations laid down on a surface (**Figure 7**). We found the 18-chain model arranged in 6-layer as 234432 (**Figure 7G**) fitted favorably into the data presented in this study (**Figure 6E**), compared with other 18-chain models, such as 34443 (Kubicki et al., 2018), 12333321 and 24-chain models.

DISCUSSION

Plant cellulose has been analyzed for decades by analytic methods and high resolution imaging approaches (Murdock, 1930; Ioyelovich, 1991; Wada et al., 2004; Agarwal et al., 2010; Barnette et al., 2011; Fernandes et al., 2011; Chunilall et al., 2013; Jarvis, 2013; Newman et al., 2013; Thomas et al., 2013; Wang and Hong, 2016). We and many other groups (Carpita et al., 2001a; Ding and Himmel, 2006; Ding et al., 2012; Zhang et al., 2013, 2016, 2017; Zheng et al., 2017) have demonstrated that the microfibrils often appear to be bundles especially in PCWs. Considering the facts that a mixture of bundles in variable sizes and individual microfibrils co-exists in different layers of any given cell wall, and the amount of matrix components may also affect the crystalline features of cellulose (Martinez-Sanz et al., 2017), the diffraction data measured from ensemble averaging of these mixed cellulose structures may not represent the fundamental structure of the microfibril. Indeed, previous studies (Fernandes et al., 2011; Jarvis, 2013; Thomas et al., 2013; Wang and Hong, 2016) have suggested 24-chain models by assuming that each microfibril is independent and has unique surface chains, however, they have also noted that an 18-chain model could fit into these data if the microfibrils coalesce even partly in their length without distinguishable surface chains.

We also noticed that even though the AFM operation conditions were extensively optimized in this study, most of

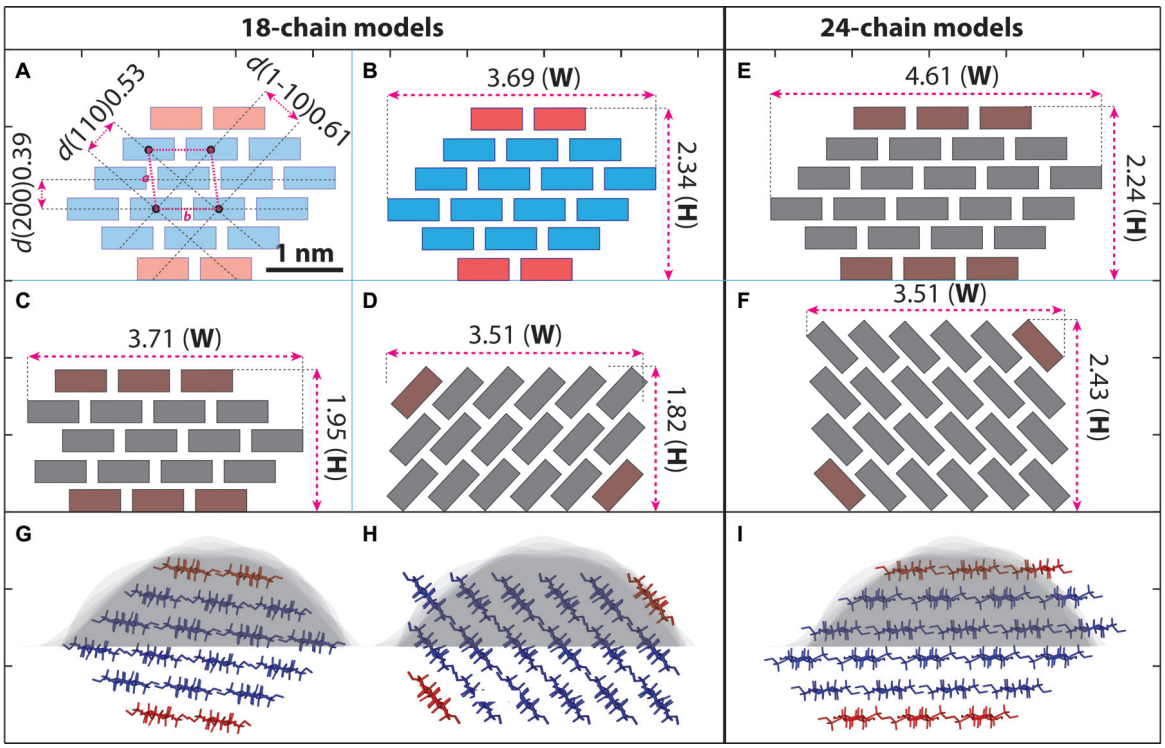

FIGURE 7 | Hypothetical models of cellulose microfibril and empirical fitting to measured data. **(A–F)** Cross-section shapes and calculated height (H) and width (W) with 18-chain and 24-chain. Cellulose chains are presented as simplified boxes. **(A)** The 18-chain model is built based on the cellulose Iβ structure (Nishiyama et al., 2002). The two chains in the hydrophobic surface are shown in red. **(B–D)** Calculated size of the 18-chains 234432 model **(B)**, 34443 model **(C),** and 12333321 model **(D)**. **(E,F)** 24-chain models in 345543 **(E)** and 123444321 **(F)** conformations. **(G–I)** The 234432 model **(B)** fits the best to our AFM measurement (gray background, see also **Figure 6E**) compared with 12333321 **(D)** and 345543 **(E)** models. Scale bar = 1 nm.

images were still not suitable for determining the size and cross-section shape of the microfibril, because of the complicated three-dimensional architecture and the associated matrix components in native cell walls. We examined more than a thousand images taken from PCWs and SCWs, but we were only able to find a dozen of them that could be used to measure the microfibril at the sub-nanometer accuracy. Despite our data appeared to be preferable to the 18-chain 234432 model, which was measured exclusively from elongated PCWs, further study must be carried out to verify whether this microfibril structure is inherently synthesized by the CSC, or it is a resulting structure from splitting of large bundles.

In the SCWs, the microfibrils appeared to be twin or individuals with similar width, but the height value could not be accurately determined due to each microfibrils were independently embedded in the matrix materials (**Figure 4**). Although biochemical studies have shown that cellulose synthesis in the PCW and the SCW are carried out by different sets of CESAs (Taylor et al., 2003, 2004), the overall conserved sequences of these CESA proteins seem to suggest the same microfibril is produced in different walls. In this study the measurement of the microfibril has been performed in PCWs, the question remains unanswered if the same structure of microfibril is synthesized in SCWs as in PCWs. The large bundles and splitting effect of cellulose microfibrils observed in the PCWs are clearly different from the near-parallel twined or individuals in the SCWs, which suggest there are different mechanisms of CSC assembly during cellulose synthesis in different walls. Future studies on the structure and their molecular interactions between CESA proteins are required to better understand the fundamental mechanism of cellulose biosynthesis in plants.

MATERIALS AND METHODS

General Chemicals

All chemicals and reagents, unless specifically noted, were purchased from Sigma–Aldrich (St. Louis, MO, United States).

Preparation of Plant Cell Wall Material

For growth of maize plant, kernels of sweet corn seeds (Burpee Garden Products Company, #65681 Early Sunglow Hybrid) were directly sown in a pot filled with a mixture of humus soil, vermiculite, and perlite. The pot was placed in a growth chamber set to 30°C, 80% humidity, and a 12-h light/dark photoperiod. Maize plants were grown for 4 weeks prior to collecting tissues for further experiments.

For AFM imaging, we used the third internode of maize plants. All plant tissues were prepared by hand-cutting longitudinally

using a double-blade razor. These sections were put in a petri dish with 20 ml ddH_2O, washed at least three times by exchanging fresh water in the petri dish. These sections were then transferred onto a glass slide (EMS, Hatfield, PA, United States) or fresh-cut mica (Ted Pella, Redding, CA, United States) that was pre-coated with poly-lysine and checked by bright field light microscopy to select samples with a relatively uniform thickness of approximately 5–10 μm containing a single layer of cells (**Figure 1**). Extra water was then carefully removed by filter paper, and 100 μl fresh water was added immediately. The sample was allowed to settle down in water for at least 30 min in room temperature before AFM imaging. To ensure reproducibility, all samples were prepared from fresh plant tissue and discarded after imaging. In order to image all different types of cell walls from living plant, we grew at least 10 plants every week continuously in last the 3 years, at least a hundred plants have been used to generate the AFM images reported in this paper.

AFM Operation

We used two AFM systems: The Dimension FastScan and the MultiMode 8-HR with NanoScope V controller (Bruker Nano, Santa Barbara, CA, United States). Both microscopes were installed with a vibration and acoustic isolation system. These two systems were used to compare imaging of the same cell wall sample as a quality control for fine features; we found when imaging in the scale of 100–200 nm scan scales using the same controller and imaging mode, the image quality was not distinguishable. Since the FastScan system has an open stage that allows for navigating the scanner in a large range, the entire cell wall surface can be imaged sequentially. Therefore, the majority of the image data presented in this study was carried out using the FastScan system.

The Fastscan AFM system allowed us to scan a field of up to 25 μm and the entire cell wall could be imaged by moving the sample stage (10–100 μm depending on wall type).

A standard 35-μm scanner was used with the ScanAsyst^{TM} imaging mode and probes SCANASYST-FLUID+ (Bruker, Camarillo, CA, United States) for imaging under fluid. The AFM control software (Nanoscope V9.3) was used in all imaging experiments. The PeakForce was manually controlled in values between 50 pN to 3 nN depending on surface features and the gain was automatically adjusted. Before AFM imaging, the scanner was carefully calibrated using calibration standards (Bruker, Camarillo, CA, United States) for x/y and z direction respectively. The system was warmed up for at least 2 h before imaging to minimize the creep phenomenon of the AFM scanner. During imaging, the x-y closed loop was always on to avoid image distortion caused by the hysteresis effect. A built-in optics system with a digital camera (5MP) was used to aid the positioning of the AFM tip to a desirable location and types of cell walls (**Figure 1**). Plant cell wall samples were kept in water during AFM imaging, images were taken at 512 × 512 and 1024 × 1024 lines with a scan rate of 0.5–3 Hz. At least five images with different scan sizes of 0.1, 0.2, 0.5, 1, and 5 μm, sometimes 10–20 μm if needed, were taken in the same scan area and on the same piece of cell wall, and at least five different areas were measured, both height and PeakForce error images were recorded simultaneously.

The cantilever of AFM probe (SCANASYST-FLUID+) was 70 μm long, 10 μm wide, and 600 nm thick. The spring constant is 0.7 N/m, and their resonant frequency in an aqueous solution is 150 KHz. The average tip radius was 2 nm. We preselected the AFM tips using the built-in function of "Tip Qualification" from the software Nanoscope Analysis v1.8 (Bruker Nano, Santa Barbara, CA, United States) to check the tip radius, and only the probes with tip radius of less than 2 nm were selected for the imaging experiment.

AFM provides a 3-D profile by raster scanning and recording the small interaction forces between a sharp tip and the sample surface. An AFM image therefore represents combined information of the tip geometry and the actual surface features of the sample (Santos et al., 2011). In this study, image deconvolution is particularly difficult due to the 3-D complexity of the relatively stiff cellulose networks and the surrounding matrix polymers that are highly mobile. In addition, the intrinsic resolution limit of AFM in imaging small features (smaller than the tip radius) can cause an overestimation of the width due to the tip broadening effect, and an underestimation of the height (height loss) due to sample deformation or intrinsic signal spread-out by the interaction of tip-surface-sample geometry (Santos et al., 2011).

A recently developed rapid force-distance (FD) curve-based imaging mode, PeakForce Tapping^{TM}, is applied for minimizing the height loss, which allows precise control of probe-to-sample interaction and provides the lowest available imaging forces to achieve the highest resolution imaging (Pyne et al., 2014). Compared with regular AFM imaging technique, such as non-contact mode, the PeakForce tapping mode minimizes the height loss effect caused by tip-sample-surface interaction (Santos et al., 2011) by maintaining a constant contact force at the pN level which is suitable for the measurement of biological samples owing to its exceptional low imaging forces.

We adjusted the setpoint of contact force to be the smallest value as long as the images were reproducible to minimize the sample compression or deformation during imaging. The contact force may affect the measurement accuracy in both vertical and lateral directions, but minimal contact force with sharp image is likely to get the AFM measurement close to its actual value (Pyne et al., 2014). In our experiment, we found that by applying the PeakForce setpoint between 50 to 230 pN, most of cell walls could be imaged in high resolution, except scanning large areas, such as 5–10 μm, in which the force could be increased to a range of 200-3 nN.

Measurement of the Height and Width of the Microfibrils

The software Nanoscope Analysis v1.8 (Bruker Nano, Santa Barbara, CA, United States) was used for AFM image processing and analysis. The height and PeakForce error images were analyzed, which were flattened at 3rd order and filtered with the lowpass filter (filter size less than 3 pixels) for images presented in all figures. The data scale was also manually adjusted

according to the color bars presented in each image. For height and width measurement, we used only raw images in 100–200 nm scan areas with 1024 scan lines, no off-line flatten or filter were applied.

DATA AVAILABILITY STATEMENT

All data are available from the corresponding author upon reasonable request.

AUTHOR CONTRIBUTIONS

S-YD conceptualized the project, conducted AFM, analyzed the data, and wrote the manuscript. BS, SZ, WS, and CC conducted AFM and data analysis. All authors revised the manuscript.

FUNDING

This work was supported by U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under Award Number DE-SC0019072 and by the Great Lakes Bioenergy Research Center (DE-SC0018409). SZ also supported by the Training Program for 1000 Young and Middle-aged Backbone Teachers of Guangxi Higher Education Institution in 2019.

ACKNOWLEDGMENTS

The authors thank Drs. Edward Bayer, Kenneth Keegstra, Yair Shachar-Hill, and Curtis Wilkerson for critical review of the manuscript.

REFERENCES

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: 10.1021/jf304465k

Andersson, S. B. (2006). "An algorithm for boundary tracking in AFM," in *Proceedings of the American Control Conference 2006*, (Minneapolis, MN: IEEE), 508–513.

Barnette, A. L., Bradley, L. C., Veres, B. D., Schreiner, E. P., Park, Y. B., Park, J., et al. (2011). Selective detection of crystalline cellulose in plant cell walls with sum-frequency-generation (SFG) vibration spectroscopy. *Biomacromolecules* 12, 2434–2439. doi: 10.1021/bm200518n

Busse-Wicher, M., Gomes, T. C. F., 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 Journal* 79, 492–506. doi: 10.1111/tpj.12575

Carpita, N. C., Defernez, M., Findlay, K., Wells, B., Shoue, D. A., Catchpole, G., et al. (2001a). Cell wall architecture of the elongating maize coleoptile. *Plant Physiol.* 127, 551–565. doi: 10.1104/pp.010146

Carpita, N. C., Defernez, M., Findlay, K., Wells, B., Shoue, D. A., Catchpole, G., et al. (2001b). Cell wall architecture of the elongating maize coleoptile. *Plant Physiol.* 127, 551–565.

Chanzy, H., Imada, K., Mollard, A., Vuong, R., and Barnoud, F. (1979). Crystallographic aspects of sub-elementary cellulose fibrils occurring in the wall of rose cells cultured in vitro. *Protoplasma* 100, 317–322.

Chanzy, H., Imada, K., and Vuong, R. (1978). Electron diffraction from the primary wall of cotton fibers. *Protoplasma* 94, 299–306. doi: 10.1007/bf01276778

Chunilall, V., Bush, T., and Larsson, P. T. (2013). "Supra-molecular structure and chemical reactivity of cellulose I studied using CP / MAS 13 C-NMR," in *Cellulose - Fundamental Aspects*, (New York, NY: Intech).

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

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., 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

Ding, S. Y., Zhao, S., and Zeng, Y. N. (2014). Size, shape, and arrangement of native cellulose fibrils in maize cell walls. *Cellulose* 21, 863–871. doi: 10.1007/s10570-013-0147-5

Doblin, M. S., Kurek, I., Jacob-Wilk, D., and Delmer, D. P. (2002). Cellulose biosynthesis in plants: from genes to rosettes. *Plant Cell Physiol.* 43, 1407–1420. doi: 10.1093/pcp/pcf164

Dufrene, Y. F., Ando, T., Garcia, R., Alsteens, D., Martinez-Martin, D., Engel, A., et al. (2017). Imaging modes of atomic force microscopy for application in molecular and cell biology. *Nat. Nanotechnol.* 12, 295–307. doi: 10.1038/nnano.2017.45

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

Fukuma, T., Higgins, M. J., and Jarvis, S. P. (2007). Direct imaging of individual intrinsic hydration layers on lipid bilayers at Angstrom resolution. *Biophys. J.* 92, 3603–3609. doi: 10.1529/biophysj.106.100651

Giddings, T. H. Jr., Brower, D. L., and Staehelin, L. A. (1980). Visualization of particle complexes in the plasma membrane of Micrasterias denticulata associated with the formation of cellulose fibrils in primary and secondary cell walls. *J. Cell Biol.* 84, 327–339. doi: 10.1083/jcb.84.2.327

Haigler, C. H., and Roberts, A. W. (2019). Structure/function relationships in the rosette cellulose synthesis complex illuminated by an evolutionary perspective. *Cellulose* 26, 227–247. doi: 10.1007/s10570-018-2157-9

Ioyelovich, M. Y. (1991). Supermolecular structure of native and isolated cellulose. *Vysokomolekulyarnye Soedineniya Seriya A* 33, 1786–1792.

Israelachvili, J. N. (2011). *Intermolecular and Surface Forces*, 3rd Edn. Cambridge, MA: Academic Press, 1–674.

Jarvis, M. C. (2013). Cellulose biosynthesis: counting the chains. *Plant Physiol.* 163, 1485–1486. doi: 10.1104/pp.113.231092

Kimura, S., Laosinchai, W., Itoh, T., Cui, X., Linder, C. R., and Brown, R. M. (1999). Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant vigna angularis. *Plant Cell* 11, 2075–2086.

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

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

Leung, C., Bestembayeva, A., Thorogate, R., Stinson, J., Pyne, A., Marcovich, C., et al. (2012). Atomic force microscopy with nanoscale cantilevers resolves different structural conformations of the DNA double helix. *Nano Letters* 12, 3846–3850. doi: 10.1021/nl301857p

Martinez-Sanz, M., Pettolino, F., Flanagan, B., Gidley, M. J., and Gilbert, E. P. (2017). Structure of cellulose microfibrils in mature cotton fibres. *Carbohydr Polym* 175, 450–463. doi: 10.1016/j.carbpol.2017.07.090

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.

Miller, E. J., Trewby, W., Payam, A. F., Piantanida, L., Cafolla, C., and Voitchovsky, K. (2016). Sub-nanometer resolution imaging with amplitude-modulation atomic force microscopy in liquid. *J. Vis. Exp.* 118, 54924.

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, 315–326. doi: 10.1083/jcb.84.2.315

Mueller, S. C., Brown, R. M., and Scott, T. K. (1976). Cellulosic microfibrils - nascent stages of synthesis in a higher plant-cell. *Science* 194, 949–951. doi: 10.1126/science.194.4268.949

Murdock, C. C. (1930). The form of the x-ray diffraction bands for regular crystals of colloidal size. *Physical Review* 35, 8–23. doi: 10.1103/physrev.35.8

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

Nishiyama, Y., Langan, P., and Chanzy, H. (2002). Crystal structure and hydrogen-bonding system in cellulose 1 beta from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* 124, 9074–9082. doi: 10.1021/ja025 7319

Nishiyama, Y., Sugiyama, J., Chanzy, H., and Langan, P. (2003). Crystal structure and hydrogen bonding system in cellulose I(alpha) 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

Oehme, D. P., Downton, M. T., Doblin, M. S., Wagner, J., Gidley, M. J., and Bacic, A. (2015). Unique aspects of the structure and dynamics of elementary Ibeta cellulose microfibrils revealed by computational simulations. *Plant Physiol.* 168, 3–17. doi: 10.1104/pp.114.254664

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

Pesacreta, T. C., Carlson, L. C., and Triplett, B. A. (1997). Atomic force microscopy of cotton fiber cell wall surfaces in air and water: Quantitative and qualitative aspects. *Planta* 202, 435–442. doi: 10.1007/s004250050147

Pyne, A., Thompson, R., Leung, C., Roy, D., and Hoogenboom, B. W. (2014). Single-Molecule Reconstruction of Oligonucleotide Secondary Structure by Atomic Force Microscopy. *Small* 10, 3257–3261. doi: 10.1002/smll.2014 00265

Santos, S., Barcons, V., Christenson, H. K., Font, J., and Thomson, N. H. (2011). The intrinsic resolution limit in the atomic force microscope: implications for heights of nano-scale features. *PLoS ONE* 6, e23821. doi: 10.1371/journal.pone. 0023821

Scheible, W. R., Eshed, R., Richmond, T., Delmer, D., and Somerville, C. (2001). Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in Arabidopsis Ixr1 mutants. *Proc. Natl. Acad. Sci. USA* 98, 10079–10084. doi: 10.1073/pnas.191361598

Shiotari, A., and Sugimoto, Y. (2017). Ultrahigh-resolution imaging of water networks by atomic force microscopy. *Nat. Commun.* 8, 14313. doi: 10.1038/ncomms14313

Simmons, T. J., Mortimer, J. C., Bernardinelli, O. D., Poppler, 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.

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

Taylor, N. G., Gardiner, J. C., Whiteman, R., and Turner, S. R. (2004). Cellulose synthesis in the Arabidopsis secondary cell wall. *Cellulose* 11, 329–338. doi: 10.1023/b:cell.0000046405.11326.a8

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

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., Sturcova, 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

Voitchovsky, K. (2013). Anharmonicity, solvation forces, and resolution in atomic force microscopy at the solid-liquid interface. *Phys Rev E Stat Nonlin Soft Matter Phys* 88, 022407.

Wada, M., Chanzy, H., Nishiyama, Y., and Langan, P. (2004). Cellulose III I crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. *Macromolecules* 37, 8548–8555. doi: 10.1021/ma0485585

Wang, L., Wang, H., Wagner, M., Yan, Y., Jakob, D. S., and Xu, X. G. (2017). Nanoscale simultaneous chemical and mechanical imaging via peak force infrared microscopy. *Sci Adv* 3, e1700255. doi: 10.1126/sciadv.1700255

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

Xu, P., Donaldson, L. A., Gergely, Z. R., and Staehelin, L. A. (2006). Dual-axis electron tomography: a new approach for investigating the spatial organization of wood cellulose microfibrils. *Wood Science and Technology* 41, 101–116. doi: 10.1007/s00226-006-0088-3

Zhang, T., Mahgsoudy-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., Tang, H., Vavylonis, D., and Cosgrove, D. J. (2019). Disentangling loosening from softening: insights into primary cell wall structure. *Plant J.* 100, 1101–1117. doi: 10.1111/tpj.14519

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

Zheng, Y., Cosgrove, D. J., and Ning, G. (2017). High-Resolution Field Emission Scanning Electron Microscopy (FESEM) Imaging of Cellulose Microfibril Organization in Plant Primary Cell Walls. *Microsc Microanal* 23, 1048–1054. doi: 10.1017/s143192761701251x

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.

frontiers in Plant Science

REVIEW
published: 25 May 2020
doi: 10.3389/fpls.2020.00652

Transcriptional and Post-transcriptional Regulation of Lignin Biosynthesis Pathway Genes in *Populus*

Jin Zhang[1,2], Gerald A. Tuskan[1,2], Timothy J. Tschaplinski[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

Lignin is a heterogeneous polymer of aromatic subunits derived from phenylalanine. It is polymerized in intimate proximity to the polysaccharide components in plant cell walls and provides additional rigidity and compressive strength for plants. Understanding the regulatory mechanisms of lignin biosynthesis is important for genetic modification of the plant cell wall for agricultural and industrial applications. Over the past 10 years the transcriptional regulatory model of lignin biosynthesis has been established in plants. However, the role of post-transcriptional regulation is still largely unknown. Increasing evidence suggests that lignin biosynthesis pathway genes are also regulated by alternative splicing, microRNA, and long non-coding RNA. In this review, we briefly summarize recent progress on the transcriptional regulation, then we focus on reviewing progress on the post-transcriptional regulation of lignin biosynthesis pathway genes in the woody model plant *Populus*.

Keywords: lignin biosynthesis, plant cell wall, transcriptional regulation, post-transcriptional regulation, transcription factor

OPEN ACCESS

Edited by:
Chandrashekhar Pralhad Joshi, Michigan Technological University, United States

Reviewed by:
Jae-Heung Ko, Kyung Hee University, South Korea
Hairong Wei, Michigan Technological University, United States

****Correspondence:***
Jin Zhang
zhangj1@ornl.gov
Jin-Gui Chen
chenj@ornl.gov

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *12 February 2020*
Accepted: *28 April 2020*
Published: *25 May 2020*

Citation:
Zhang J, Tuskan GA, Tschaplinski TJ, Muchero W and Chen J-G (2020) Transcriptional and Post-transcriptional Regulation of Lignin Biosynthesis Pathway Genes in Populus. Front. Plant Sci. 11:652. doi: 10.3389/fpls.2020.00652

INTRODUCTION

Lignin is one of the most abundant biopolymers, accounting for ∼30% of the organic carbon in the biosphere. As a principal component of secondary cell walls, lignin provides plants with structural integrity and a response mechanism to environmental stimuli, e.g., pathogen attack. In addition, lignin supports transport of water and solutes through the vascular system. The lignin structure varies between plant species, between cell types within a single plant, and between different parts of the wall of a single cell. The lignin polymer is primarily comprised of three major monomers: *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) monolignols that are synthesized via the phenylpropanoid pathway (Raes et al., 2003). From *Arabidopsis* genome-wide analysis and mutant/transformation studies, at least 14 structural genes have been characterized and shown to be involved in the monolignol biosynthesis pathway (Goujon et al., 2003a).

Although the regulatory mechanism of lignin biosynthesis has been studied in several plant species (Zhong et al., 2006; Zhong and Ye, 2011; Xie et al., 2018b; Zhang et al., 2018a), many aspects of its regulation remain unresolved. Identification of *cis*-acting elements in monolignol biosynthetic genes provides an understanding of the transcriptional regulation of lignin

biosynthesis. Promoter analysis and electrophoretic mobility shift assay have revealed that the SNBE (Zhong et al., 2010a) and AC elements (Zhong and Ye, 2011) (corresponding to the NAC and MYB transcription factor-binding motif, respectively) are necessary for coordinated monolignol pathway gene activation. However, a comprehensive understanding of the transcriptional and post-transcriptional regulation of lignin biosynthesis in woody species is still lacking. In this review, we summarize the current understanding of the regulation of lignin biosynthesis pathway genes at the transcriptional level, then focus on the emerging area of post-transcriptional regulation.

TRANSCRIPTIONAL REGULATION OF LIGNIN BIOSYNTHESIS PATHWAY GENES

Structural Genes of Monolignol Biosynthesis

The monolignol biosynthesis pathway has been well studied in several model plant species, such as the model herbaceous species *Arabidopsis* and the model woody species *Populus*. Monolignols are synthesized from phenylalanine via the phenylpropanoid pathway, which includes a series of enzymes controlling alternate linear steps, ultimately providing precursors for numerous secondary metabolites (Fraser and Chapple, 2011). Wang et al. (2018) demonstrated the importance of phenylpropanoid biosynthetic enzymes for lignin biosynthesis in *Populus* using 221 independent transgenic lines derived from 21 lignin biosynthetic genes. These enzymes belong to an assembly of genes and gene families, including phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate:CoA ligase (4CL), *p*-coumaroyl-shikimate/quinate 3-hydroxylase (C3H), hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT), caffeoyl-CoA *O*-methyltransferase (CCoAOMT), 5-hydroxyconiferyl aldehyde *O*-methyltransferase (AldOMT), coniferyl aldehyde/ferulate 5-hydroxylase (CAld5H/F5H), cinnamoyl-CoA reductase (CCR), cinnamyl alcohol dehydrogenase (CAD), caffeoyl shikimate esterase (CSE), and caffeic acid *O*-methyltransferase (COMT) (**Figure 1**). PAL, C4H and 4CL play important roles to provide precursors for various downstream metabolites (**Figure 1**). Down-regulation of PAL, C4H or 4CL can significantly decrease lignin content in both *Arabidopsis* and *Populus* (Rohde et al., 2004; Chen and Dixon, 2007; Vanholme et al., 2008; Wang et al., 2018). Recently, a C3H enzyme is identified as a bifunctional peroxidase that oxidizes both ascorbate and 4-coumarate in the model plants *Brachypodium distachyon* and *Arabidopsis* by directly catalyzing the 3-hydroxylation of 4-coumarate to caffeate in lignin biosynthesis pathway (Barros et al., 2019).

Populus is a promising feedstock for biofuels and other value-added products due to its fast growth and high efficiency

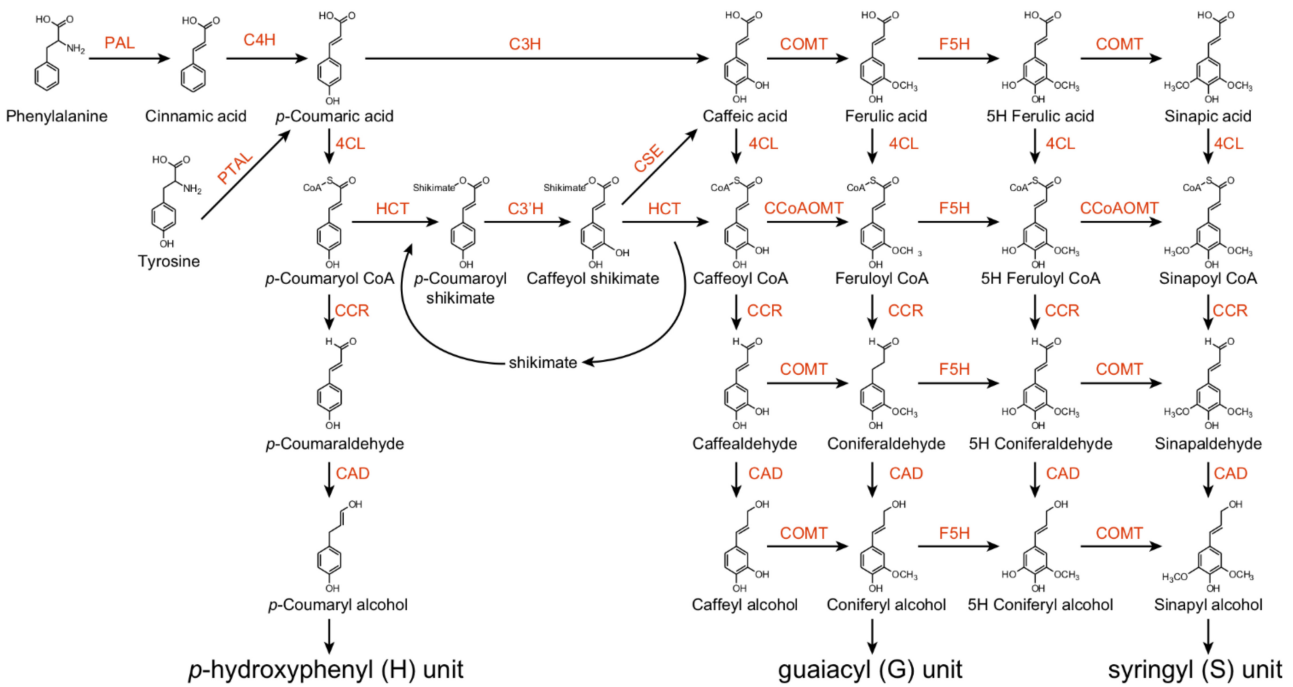

FIGURE 1 | The monolignol biosynthetic pathway in *Populus*. PAL, L-phenylalanine ammonia-lyase; PTAL, bifunctional L-phenylalanine/L-tyrosine ammonia-lyase; C4H, cinnamate 4-hydroxylase; C3H, *p*-coumarate 3-hydroxylase; COMT, caffeate/5-hydroxyferulate 3-O-methyltransferase; F5H, ferulate 5-hydroxylase/coniferaldehyde 5-hydroxylase; 4CL, 4-hydroxycinnamate:CoA ligase; HCT, *p*-hydroxycinnamoyl CoA:shikimate/quinate hydroxycinnamoyltransferase; C3′H, *p*-coumaroyl shikimate/quinate 3′-hydroxylase; CSE, caffeoyl shikimate esterase; CCoAOMT, caffeoyl CoA 3-O-methyltransferase; CCR, cinnamoyl CoA reductase; CAD, cinnamyl alcohol dehydrogenase.

of biofuel conversion. In addition, abundant public genomics, and transcriptomics resources of *Populus* provide the basis for functional study. Here we focus on *Populus* to explore the transcriptional and post-transcriptional regulation of lignin biosynthetic genes. On the basis of findings reported in literature, we build a conceptual network of the enzymes that control monolignol biosynthesis in *Populus*. As shown in **Table 1**, the 21 enzymes reported by Wang et al. (2018), and three other enzymes [CSE1, CSE2 (Vanholme et al., 2013) and COMT2 (Marita et al., 2001)], play important roles in monolignol biosynthesis in *Populus* and *Arabidopsis*. We analyzed the expression profiles of the structural genes in monolignol biosynthesis pathway across various tissues and during wood formation in *Populus* based on the *Populus* Gene Expression Atlas database (different tissues of buds, male catkins, female catkins, leaf, root and stem of *P. trichocarpa*, 72 RNA-Seq libraries)[1] and AspWood database (micro meter-scale profile of *P. tremula* cambial growth and wood formation, 137 RNA-Seq libraries) (Sundell et al., 2017).

Broad expression evidence from key enzymes in the lignin biosynthetic pathway provides a hypothetical foundation for their functions in various tissues. For example, Kim et al. (2019) performed a series of wood-forming tissue-specific transcriptome analyses from a hybrid poplar and identified critical pathway genes for secondary wall biosynthesis in mature developing xylem. Wood formation is a process of plant secondary growth, which originates from the cambium meristem cells, eventually forming a tree's main stem or truck. Most of the genes involved in this process are highly expressed in the developing xylem. In contrast, *CAD2* and *AldOMT2* are highly expressed in maturing xylem and cambium, respectively (**Figure 2**). In a promoter-GUS histochemistry analysis, the GUS driven by promoter of *Eucalyptus gunnii CAD2* is expressed in all lignifying cells including vessel elements, xylem fibers and paratracheal parenchyma cells of the xylem tissues in the transgenic *Arabidopsis* floral stem and root (Baghdady et al., 2006). The expression pattern and function of *AldOMT2* homologs remains unclear.

TABLE 1 | Monolignol biosynthetic genes in *Populus*.

Gene ID	Gene family	Enzyme	Substrate
Potri.006G126800	*PAL*	PAL1	Phe
Potri.008G038200	*PAL*	PAL2	Phe
Potri.016G091100	*PAL*	PAL3	Phe
Potri.010G224100	*PAL*	PAL4	Phe
Potri.010G224200	*PAL*	PAL5	Phe
Potri.013G157900	*C4H*	C4H1	Cinnamic acid
Potri.019G130700	*C4H*	C4H2	Cinnamic acid
Potri.001G036900	*4CL*	4CL3	4-coumaric acid, caffeic acid, ferulic acid, 5-hydroxyferulic acid
Potri.003G188500	*4CL*	4CL5	Caffeic acid, 4-coumaric acid, ferulic acid, 5-hydroxyferulic acid, sinapic acid
Potri.006G033300	*C3H*	C3H3	4-coumaroyl shikimic acid, 4-coumaric acid
Potri.003G183900	*HCT*	HCT1	4-coumaroyl-CoA, 4-coumaroyl shikimic acid, caffeoyl-CoA, caffeoyl shikimic acid
Potri.001G042900	*HCT*	HCT6	4-coumaroyl-CoA, 4-coumaroyl shikimic acid, caffeoyl-CoA, caffeoyl shikimic acid
Potri.009G099800	*CCoAOMT*	CCoAOMT1	Caffeoyl-CoA
Potri.001G304800	*CCoAOMT*	CCoAOMT2	Caffeoyl-CoA
Potri.008G136600	*CCoAOMT*	CCoAOMT3	Caffeoyl-CoA
Potri.015G119600	*AldOMT*	AldOMT2	Caffealdehyde, 5-hydroxyconiferaldehyde, caffeyl alcohol, 5-hydroxyconiferyl alcohol, 5-hydroxyferulic acid, caffeic acid
Potri.005G117500	*CAld5H/F5H*	CAld5H1, F5H1	Coniferyl alcohol, coniferaldehyde, ferulic acid
Potri.007G016400	*CAld5H/F5H*	CAld5H2, F5H2	Coniferyl alcohol, coniferaldehyde, ferulic acid
Potri.003G181400	*CCR*	CCR2	Feruloyl-CoA, 4c-oumaroyl-CoA, caffeoyl-CoA
Potri.009G095800	*CAD*	CAD1	Coniferaldehyde, 4-coumaraldehyde, sinapaldehyde
Potri.016G078300	*CAD*	CAD2	Sinapaldehyde, coniferaldehyde
Potri.003G059200	*CSE*	CSE1	Caffeoyl shikimate
Potri.001G175000	*CSE*	CSE2	Caffeoyl shikimate
Potri.012G006400	*COMT*	COMT2	Caffeic acid, caffeoyl-CoA, caffeoyl aldehyde, caffeoyl alcohol

Transcription Factors Involved in the Lignin Biosynthesis Pathway

A hierarchical transcriptional regulatory network for lignin biosynthetic genes has been established over the past 10 years (Zhao et al., 2010; Zhong and Ye, 2011; Lin et al., 2017; Zhang et al., 2018a; Chen et al., 2019). This network involves members of several transcription factor (TF) families including MYBs and NACs. A recent study identified a novel TF (i.e., PtrEPSP-TF) encoding a homolog of 5-enolpyruvylshikimate 3-phosphate (EPSP) synthase in the shikimate pathway, which possesses a helix-turn-helix motif in the N terminus and can function as a transcriptional repressor to regulate gene expression in the phenylpropanoid pathway in *Populus* (Xie et al., 2018a). Correspondingly, the expression of lignin-related TFs is affected by several other genes. For example, overexpression of a serine hydroxymethyltransferase (*PtSHMT2*) decreases the lignin content in transgenic poplar (Zhang et al., 2019a). Overexpression of a prefoldin chaperonin β subunit gene *PdPFD2.2* increases lignin S/G ration in poplar (Zhang et al., 2019b). This suggests that the molecular regulation of lignin biosynthesis is not unidirectional and is more complex than that was previously reported.

[1]https://phytozome.jgi.doe.gov/phytomine/aspect.do?name=Expression

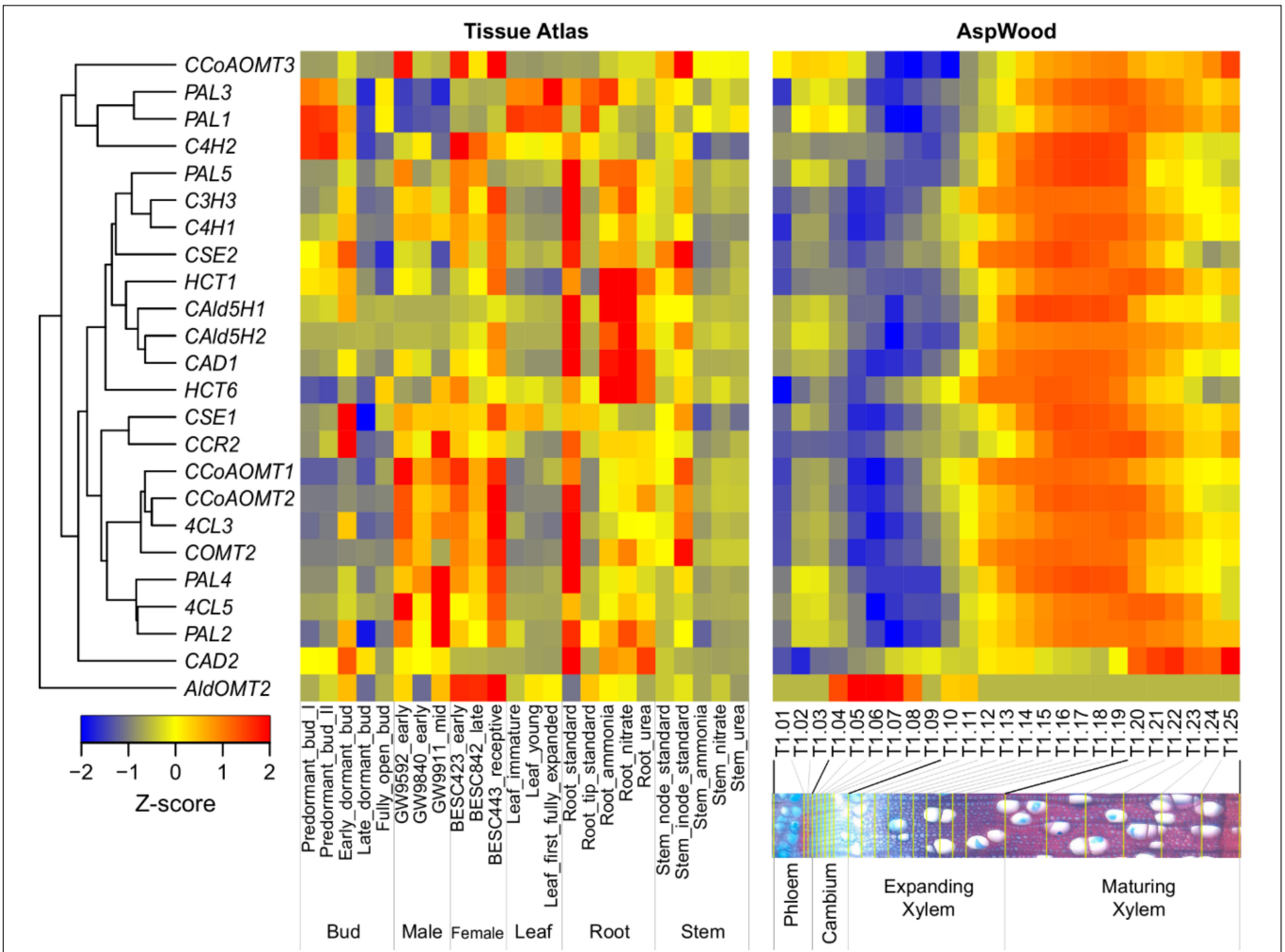

FIGURE 2 | Hierarchical clustering of the expression profiles of monolignol biosynthetic genes across various tissues (left) and during wood formation (right) in *Populus*. The expression data of different tissues and wood formation were obtained from *Populus* Gene Expression Atlas (https://phytozome.jgi.doe.gov/phytomine/aspect.do?name=Expression) and AspWood (http://aspwood.popgenie.org/aspwood-v3.0/), respectively. The tissue atlas dataset includes tissues collected from buds, male catkins, female catkins, leaf, root and stem. The AspWood dataset includes samples collected from phloem, cambium, expanding xylem and maturing xylem. Gene expression was normalized by Z-score. Red and blue represent high and low expression, respectively.

Recently, Gunasekara et al. (2018) developed a novel algorithm called triple-gene mutual interaction (TGMI) for identifying the pathway regulators using high-throughput gene expression data, which calculates the mutual interaction measure for each triple gene grouping (two pathway genes and one TF) and then examines its statistical significance using bootstrap. Implementing this algorithm, Gunasekara et al. (2018) analyzed pathway regulators of lignin biosynthesis using a compendium dataset that comprised 128 microarray samples from *Arabidopsis* stem tissues under short-day conditions. In this review, we also applied the TGMI algorithm to identify regulators of lignin biosynthesis in *Populus* based on the tissue-specific *Populus* Gene Expression Atlas and AspWood datasets (209 RNA-Seq samples in total). As anticipated, a series of known lignin biosynthesis-related TFs (87 TFs from 10 families), such as members in NAC and MYB families, were correlated with the lignin biosynthetic genes (**Figures 3**, **4**). In addition, we identified several novel TFs that were highly correlated with the monolignol biosynthetic genes, expanding our view of the transcriptional regulatory network affecting lignin biosynthesis. Individual classes of these TFs are presented in **Figures 3**, **4**.

Transcriptional Regulation of Lignin Biosynthetic Genes

PAL

To further understand the transcriptional regulation between TFs and lignin biosynthetic genes, we generated a heatmap to reveal the correlation between lignin biosynthetic genes and known lignin-related TFs (**Figure 4**). PAL genes showed strong correlation with MYB TFs. During secondary cell wall formation, *MYB46* and *MYB83* and their orthologs in several plant species, including *Arabidopsis*, *Populus*, and *Eucalyptus*, have been identified as the direct targets of SNDs (SECONDARY

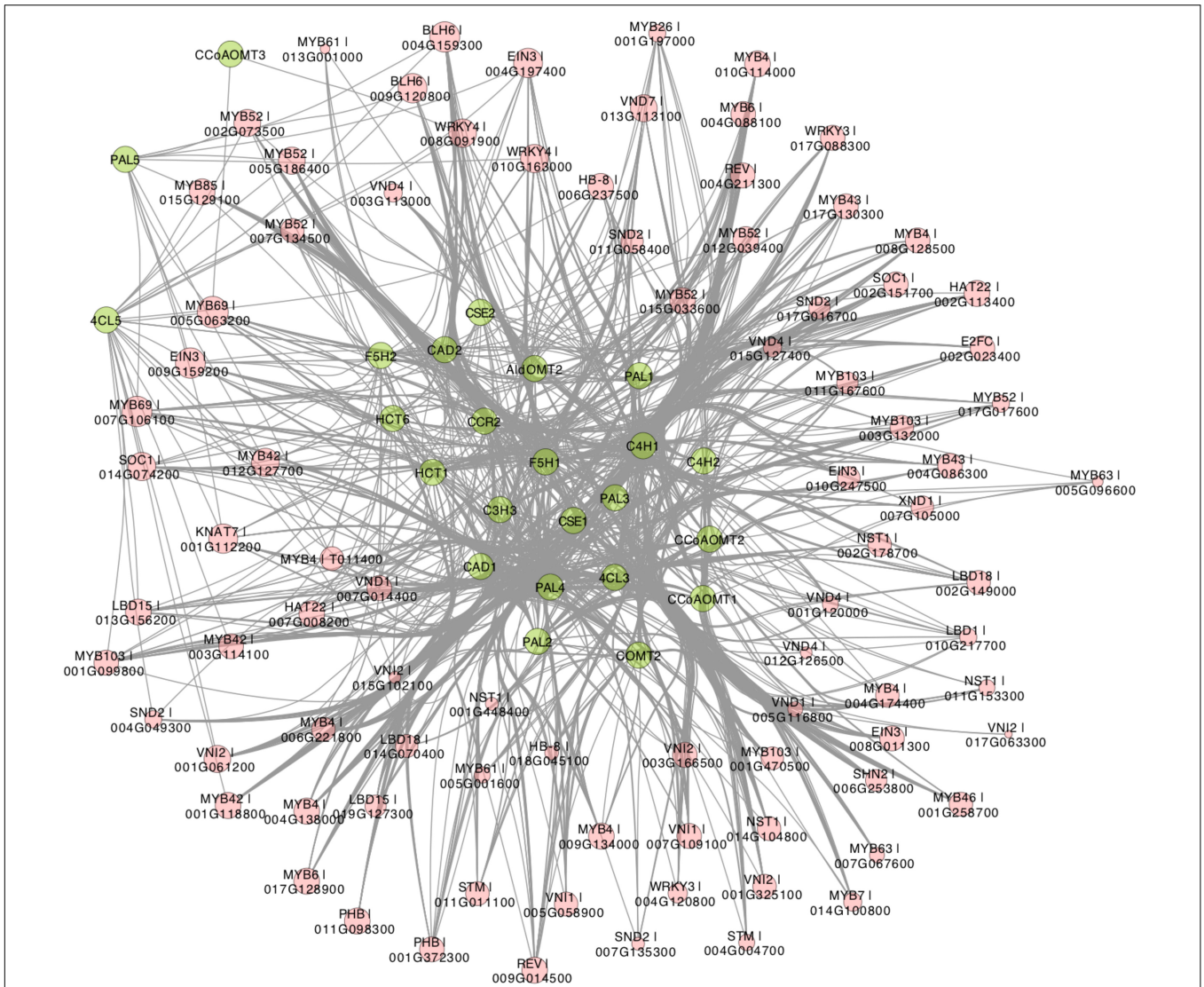

FIGURE 3 | Regulatory network generated by triple-gene mutual interaction (TGMI) algorithm for the *Populus* lignin biosynthesis pathway using the RNA-Seq data from *Populus* Gene Expression Atlas and AspWood datasets. Green nodes represent monolignol biosynthetic genes. Red nodes are transcription factors (TF) and node size represent frequency of TF. Edges represent regulatory relationships from TGMI algorithm.

WALL-ASSOCIATED NAC DOMAIN PROTEINS) and VNDs (VASCULAR-RELATED NAC DOMAINS) and function as the second-layer master switches (McCarthy et al., 2010; Zhong and Ye, 2011; Kim et al., 2013; Ko et al., 2014; Zhang et al., 2018a). Overexpression of *MYB46* and *MYB83* caused ectopic deposition of secondary cell walls through activation of the lignin, cellulose and xylan biosynthetic genes (Zhong and Ye, 2011). Electrophoretic mobility shift assay and chromatin immunoprecipitation analysis showed that MYB46 directly binds to the promoters of PAL (Kim et al., 2014). Similarly, their orthologs in *Populus*, *PtrMYB3* (Potri.001G267300) and *PtrMYB20* (Potri.009G061500), activate the biosynthesis pathways of lignin, cellulose and xylan in both *Arabidopsis* and *Populus*, including *PAL* genes (McCarthy et al., 2010). In *Populus tomentosa, PtoMYB216* (GenBank: JQ801749, ortholog of Potri.013G001000), a homolog of *Arabidopsis AtMYB61* and *AtMYB85*, was specifically expressed during secondary wall formation in wood. The expression of *PAL4* was induced in the transgenic plants overexpressing *PtoMYB216* (Tian et al., 2013). *PtoMYB156* (GenBank: KT990214, ortholog of Potri.009G134000) is a homolog of *AtMYB4*, which functions as phenylpropanoid/lignin biosynthesis repressor. Overexpression of *PtoMYB156* in poplar also resulted in downregulation of *PtoPAL1* (Yang et al., 2017). Four additional MYB TFs (MYB20, MYB42, MYB43 and MYB85) were recently reported as transcriptional regulators that directly activate lignin biosynthetic genes during secondary wall formation in *Arabidopsis*. Quadruple mutant *myb20/42/43/85* plants exhibited reduced transcript levels of *PAL* (Geng et al., 2020). From these results, MYB TFs appear to be regulated by a series

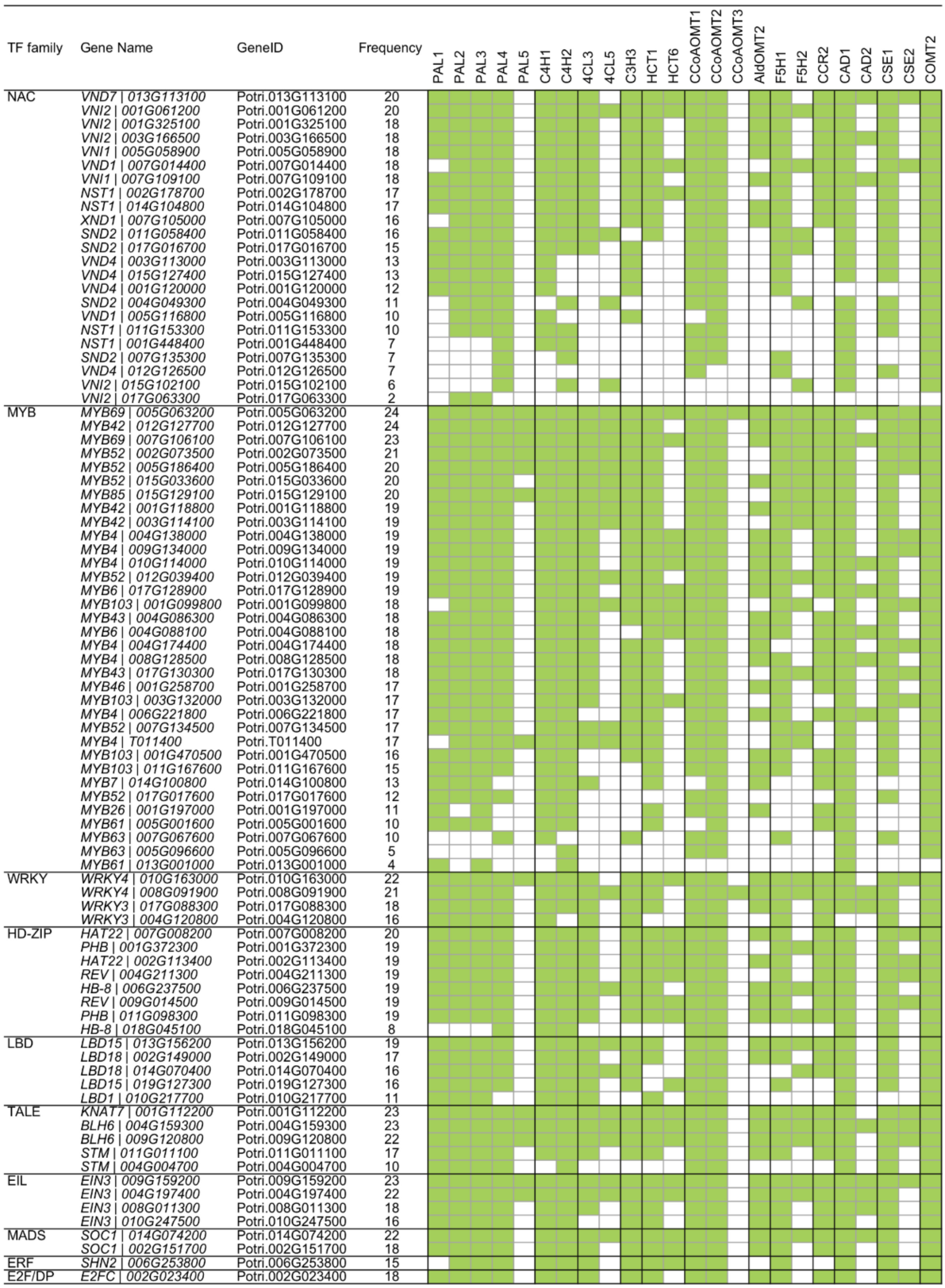

FIGURE 4 | Regulatory relationship of transcription factor (TF) and monolignol biosynthetic genes generated by triple-gene mutual interaction (TGMI) algorithm. Green blocks represent statistically significant interactions.

of master switches during secondary cell wall biosynthesis. The transcriptional regulation of *PAL* is likely regulated by a hierarchical or more complex pattern, in addition to the direct regulation by these MYB TFs.

C4H

As shown in **Figure 4**, *C4H1* was correlated with the TGMI-based expression of 32 *MYB* TFs. Recently, a transcriptional regulatory network (TRN) of wood formation based on a *P. trichocarpa* wood-forming cell system with quantitative transcriptomics and chromatin binding assays was constructed (Chen et al., 2019). In the TRN, *PtrC4H1* was regulated by *PtrWBLH2* (a wood Bel-like homeodomain protein), which is a direct target of *PtrMYB021* and *PtrMYB074*. Comparably, in *P. tomentosa*, *C4H2* is directly activated by *PtoMYB216* through AC elements (Tian et al., 2013). In addition, the expression of *C4H* was repressed by *MYB* transcriptional repressors. In *Arabidopsis*, *AtMYB4* downregulates the expression of *C4H* (Jin et al., 2000). Ectopic expression of *E. gunnii EgMYB1* in *Populus* repressed the expression of *PtaC4H2* in wood tissue (Legay et al., 2010). Moreover, *Arabidopsis WRKY12* is a transcriptional repressor that can directly bind to the promoter of *NST2*, a master regulator of lignin biosynthesis. Loss-of-function mutants of *WRKY12* in *Arabidopsis*, and its ortholog in *Medicago*, result in ectopic deposition of lignin, xylan, and cellulose in pith cells (Wang et al., 2010). Its homolog in *Populus*, *PtrWKRY19* (Potri.014G050000), is highly expressed in stems, especially in pith. Finally, *PtrWRKY19* can repress the expression of *PtoC4H2* through W-box elements (Yang et al., 2016).

4CL

4CL is the third step in the phenylpropanoid pathway and it is important for not only monolignol biosynthesis but also the generation of other secondary metabolites (Tsai et al., 2006). Based on the regulatory network, the two *4CL* genes (*4CL3* and *4CL5*) were correlated with multiple NAC and MYB TFs (**Figure 4**). In *Populus*, the expression of *4CL5* was upregulated in transgenic plants overexpressing *PtrMYB152* (GenBank: XM_002302907, ortholog of Potri.017G130300), a homolog of *AtMYB58/63/85* (Li et al., 2014). Similarly, *4CL5* could be activated by another MYB member *PtoMYB216* (Tian et al., 2013). The promoters of *4CL* genes include AC elements that provide binding sites for secondary cell-wall-related MYB genes. In several plant species, NAC TFs have been reported to regulate the expression of *4CL* genes. In support of these observations, *EjNAC1* had trans-activation activities on promoter of *Ej4CL1* (Xu et al., 2015) and the expression of *4CL* was repressed in *Medicago nst* mutant (Zhao et al., 2010). However, whether *4CL* genes are direct targets of NAC TFs in *Populus* remains unknown.

C3H

The regulatory network pattern in **Figure 4** reveals that *C3H* has a similar pattern to the *4CL* genes, indicating the transcriptional regulation of *C3H* might be similar with *4CL* genes. As expected, the expression of *C3H3* was also activated by *PtoMYB216* and *PtrMYB152* (Tian et al., 2013; Li et al., 2014). Still, studies of other species revealed that C3H could be regulated by other TF families. Switchgrass *PvMYB4* is a transcriptional repressor and binds to the AC elements. The expression of *C3H* was activated by overexpressing *PvMYB4* in transgenic tobacco and switchgrass (Shen et al., 2012). In *Medicago nst* mutant, the expression of *C3H* was repressed due to loss-of-function of *NST* (Zhao et al., 2010). In addition, the expression of *C3H* was induced by overexpressing *GbERF1-like*, a *Gossypium barbadense* ethylene response-related factor, in transgenic cotton and *Arabidopsis* (Guo et al., 2016). The AC elements provide the binding sites for the direct TF regulation.

HCT

HCT is involved in the production of methoxylated monolignols that are precursors to G- and S-unit lignin. HCT-downregulated plants are strikingly enriched in H lignin units, a minor component of lignin (Wagner et al., 2007). In *P. trichocarpa*, *HCT1* and *HCT6* display xylem-specific expression, which is regulated by *PtrWBLH2* and *PtrWBLH1*, respectively (Chen et al., 2019). A recent study using genome-wide association studies (GWAS) and expression quantitative trait loci (eQTL)/expression quantitative trait nucleotide (eQTN) studies identified a defense-related *HCT2* that was regulated by *WRKY* TFs (Zhang et al., 2018b), implying that other TF families might be also involved in the transcriptional regulation of *HCT* gene family under alternate developmental circumstances. Heterologous expressing *SbbHLH1*, a *Sorghum bicolor* basic helix-loop-helix gene, reduced the lignin content through repress the expression of *HCT* in transgenic *Arabidopsis* (Yan et al., 2013).

CCoAOMT

As shown in **Figure 4**, three *CCoAOMT* genes were highly positively correlated with seven TFs in NAC family. It has been reported that NAC TFs function as master regulators in the lignin biosynthesis pathway. The SECONDARY WALL NACs (SWNs) consists of two types NACs: SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN (SND)/NAC SECONDARY WALL THICKENING PROMOTING FACTOR (NST) and VASCULAR-RELATED NAC DOMAINS (VNDs) (Zhang et al., 2018a). In *Arabidopsis*, ectopic overexpression of *SND1* significantly induced the expression of CCoAOMT (Zhong et al., 2006). In *Populus*, six SND1 homologs, named *PtrWND1-6* (WOOD ASSOCIATED NAC DOMAIN), are highly expressed in the developing xylem. Overexpression of *PtrWND2B* and *PtrWND6B* in *Arabidopsis* causes ectopic deposition of secondary cell wall through activation of the lignin, cellulose and xylan biosynthetic genes (Zhong et al., 2010b). In *Populus*, the transcript of *CCoAOMT1* was induced by overexpressing WND3A (Yang et al., 2019). Zhou et al. (2014) demonstrated that the promoter of *CCoAOMT1* is directly activated by *Arabidopsis* VND1-5. Similar results were also found in *Arabidopsis* transgenic lines expressing PtrWND6B. A transactivation assay indicates *CCoAOMT* is direct target of PtrWND6B (Zhong et al., 2010b). In addition, MYB TFs were also involved in transcriptional regulation of *CCoAOMT*. As direct target of PtrWND2, PtrMYB3 and PtrMYB20

(homologous of *Arabidopsis* MYB46/83) were able to activate the promoters of *PtrCCoAOMT1* through *Arabidopsis* protoplast transactivation analysis (McCarthy et al., 2010).

CAld5H/F5H

F5H is a cytochrome P450 (CYP)-dependent monooxygenase, it is specifically required for S-unit lignin biosynthesis and diverts G-unit into the S-unit pathway (Humphreys et al., 1999). Using *P. trichocarpa* wood-forming cell system, three TFs (*PtrMYB090*, *PtrMYB161* and *PtrWBLH2*) were identified as upstream regulator of *F5H* genes in *Populus* (Chen et al., 2019). In *Medicago*, the expression of F5H is directly regulated by the secondary cell wall master switch NST1/SND1 (Zhao et al., 2010). In addition, *MYB103* is required for the expression of *F5H* and S-lignin biosynthesis in *Arabidopsis*. The S-lignin content, as well as transcript level of *F5H*, are strongly decreased in the *myb103* mutants, whereas the G-lignin content was concomitantly increased (Öhman et al., 2013).

CCR and CAD

CCR and CAD catalyze the final steps of monolignol biosynthesis (**Figure 1**). In many species, CCR and CAD exhibit similar expression patterns in vascular tissues. The expression of *PtrCAD1* was repressed by *PtrMYB174* in *Populus* (Chen et al., 2019). Other studies indicated that *CCR2* and *CAD* were activated by *PtoMYB216* and *PtrMYB152* (Tian et al., 2013; Li et al., 2014). Using promoter deletion analysis, Rahantamalala et al. (2010) identified an 80-bp region and a 50-bp region in the promoters of *E. gunnii EgCAD2* and *EgCCR* that contains MYB elements, respectively. In addition, heterologous expressing *Vitis vinifera VvWRKY2* activate the expression of *CCR* and *CAD* in transgenic tobacco (Guillaumie et al., 2010).

CSE

CSE is a recently identified novel enzymatic step in the lignin biosynthetic pathway (Vanholme et al., 2013). Similar to other MYB46/83 regulated genes, *CSE* has M46RE motifs in the promoter region, and its expression is induced by *MYB46* (Kim et al., 2014). In *Populus*, it is directly regulated by *PtrWBLH1*, a downstream regulator of *PtrMYB021* (homolog of *Arabidopsis MYB46*) (Chen et al., 2019). In addition, the regulatory network indicated that *CSE1* is negatively correlated with a *WRKY* TF in *Populus* (**Figure 4**), but whether *WRKY* directly regulates *CSE* needs to be confirmed.

COMT

COMT is critical for the S-unit lignin biosynthesis (Goujon et al., 2003b). In *Arabidopsis*, *COMT* is directly regulated by a lignin-specific MYB *AtMYB58* through binding to the AC elements (Zhou et al., 2009). A similar regulatory pattern is also observed in *Populus*. That is, *COMT2* is activated by *PtoMYB170*, *PtrMYB090* and *PtrMYB152*, but not *PtoMYB216* (Tian et al., 2013; Li et al., 2014; Xu et al., 2017; Chen et al., 2019). In addition, the promoter of *Arabidopsis COMT* could be bound by BP, a knotted1-like homeobox (KNOX) gene (Mele et al., 2003). The TGMI analysis indicated that *COMT2* is highly associated with TFs in HD-ZIP and LBD families, in addition to NAC and MYB TFs (**Figure 4**). However, experimental evidence will be required to verify this regulatory relationship.

POST-TRANSCRIPTIONAL REGULATION OF LIGNIN BIOSYNTHESIS PATHWAY GENES

Post-transcriptional regulation of lignin biosynthesis pathway genes plays important roles in molecular regulation at the RNA level, including controlling alternative splicing, RNA capping, poly-A tail addition, and mRNA stability (Sullivan and Green, 1993). To date, studies of the post-transcriptional regulation of lignin pathway have been focused on transcriptional regulatory genes. In this section, we summarize recent progress on the post-transcriptional regulation of regulatory genes in lignin pathway.

Alternative Splicing

Alternative splicing, as a post-transcriptional regulation mechanism, allows organisms to increase their proteomic diversity and regulate gene expression. It has been reported that alternative splicing of key regulators and enzymes play a critical role in the lignin biosynthesis pathway. A previous study analyzed the transcriptome of 20 *P. trichocarpa* individuals and found that ∼40% xylem genes are alternatively spliced, which include cell wall-related genes C2H2 TF and glycosyl transferases (Bao et al., 2013). Xu et al. (2014) compared the inter-species conservation of alternative splicing events in the developing xylem of *Populus* and *Eucalyptus* and found that ∼28% of alternative splicing genes were putative orthologs in these two species. Alternative splicing can also affect the expression of downstream genes. For example, retention of intron 2 of *Populus PtrWND1B/PtrSND1*, by alternative splicing, resulted in loss of DNA binding and transactivation activities (Li et al., 2012). This alternative splicing event appears to regulate secondary cell wall thickening and the expression of the lignin-related gene *4CL1*. Similar alternative splicing was also observed in its orthologs in *Eucalyptus*, but not in *Arabidopsis* (Zhao et al., 2014). In addition, other members in the VND- and SND-type NAC family are regulated by alternative splicing. For example, retained introns of *PtrSND1-A2* and *PtrVND6-C1* play reciprocal cross-regulation of the two families during wood formation (Lin et al., 2017).

microRNA

microRNAs (miRNAs) are a class of small non-coding RNAs with a 21-23 ribonucleotide RNA sequence that play central roles in gene expression regulation through directing mRNA cleavage or translational inhibition. Several miRNAs, such as miRNA397, miRNA408, miRNA857, and miRNA528, have been reported to target laccase (LAC) genes, encoding a class of blue copper oxidase proteins involved in lignin polymerization (Sunkar and Zhu, 2004; Lu et al., 2013). In *Populus*, the expression of 17 *PtrLACs* are down-regulated and lignin content is decreased by overexpression of *Ptr-miRNA397a* (Lu et al., 2013). *Arabidopsis LAC4* controls both lignin biosynthesis and seed yield, and its expression is controlled by *miRNA397* member *At-miRNA397b*. Overexpression of *At-miRNA397b* reduced lignin deposition through repression of the biosynthesis of both S- and G-lignin subunits (Wang et al., 2014). In addition, overexpression a wounding-responsive *miRNA828* can enhance lignin deposition

and H_2O_2 accumulation through repressed expression of *IbMYB* and *IbTLD* in sweet potato (Lin et al., 2012). *Acacia mangium* miRNA166 is differentially expressed between phloem and xylem, where it targets HD-ZIP III type TFs to regulate the expression of *C4H*, *CAD*, and *CCoAOMT* (Ong and Wickneswari, 2012). In maize, *Zm-miRNA528*, induced by excess nitrogen and repressed by nitrogen deficiency, targets *LAC3* and *LAC5* and regulates the biosynthesis of S-, G-, and H-subunits (Sun et al., 2018). Finally, in *Arabidopsis*, *miRNA858a* directly regulates the expression of several *MYBs* during flavonoid biosynthesis. Overexpression of *miRNA858a* results in ectopic deposition of lignin in transgenic plants (Sharma et al., 2016). Collectively, these results indicate that miRNAs play important regulatory roles during multiple levels of lignin biosynthesis.

Long Non-coding RNA

Long non-coding RNAs (lncRNAs) refer to transcripts that lack coding potential and are greater than 200 nucleotides (Kapranov et al., 2007). Chen et al. (2015) performed a genome-wide identification of lncRNA in tension wood, opposite wood and normal wood xylem of *P. tomentosa* and identified 16 genes targeted by lncRNAs that are involved in wood formation processes, including lignin biosynthesis (Chen et al., 2015). In a similar study, the interaction of NEEDED FOR RDR2-INDEPENDENT DNA METHYLATION (NERD) and its regulatory lncRNA NERDL, which is partially located within the promoter region of NERD, is involved in the wood formation processes in *Populus* (Shi et al., 2017). In cotton, Dt subgenome-specific lncRNAs are enriched in lignin catabolic processes. Wang et al. (2015) suggests that these lncRNAs may regulate lignin biosynthesis by regulating the expression of *LAC4* (Wang et al., 2015). Although these studies imply the potential roles of lncRNAs in lignin biosynthesis, the underlying regulatory mechanism remain unverified.

CONCLUDING REMARKS

In this review, we provide a comprehensive summary of the current knowledge of the transcriptional regulation of lignin biosynthetic genes and post-transcriptional regulation of regulatory genes in lignin biosynthesis in *Populus*. Lignin content has been reported as important factor in biomass recalcitrance for bioethanol conversion and production. Although many genes that play a regulatory role in the lignin biosynthesis pathway were captured in TGMI analysis, some previously reported lignin pathway regulators were missing, possibly due to limited data in our analysis. To overcome this issue and to capture other regulatory genes, multiple datasets, pooled from various tissues types during specific rapid developmental processes, should be investigated. In addition, GWAS and eQTL/eQTN analyses may provide further supportive lucidity in discovering novel regulators and regulatory mechanisms in lignin biosynthesis. Revealing the transcriptional and post-transcriptional regulatory mechanisms in lignin biosynthesis will help clarify the parameters of the lignin biosynthesis, ultimately improving the application of lignocellulose in biofuels and bioenergy. Understanding the increasingly complex lignin regulatory network will provide an important theoretical basis for basic plant biology and utilization of plant biomass.

AUTHORS' NOTE

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).

AUTHOR CONTRIBUTIONS

JZ collected and synthesized the data from literature and wrote the manuscript. GT, TT, 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

Baghdady, A., Blervacq, A. S., Jouanin, L., Grima-Pettenati, J., Sivadon, P., and Hawkins, S. (2006). *Eucalyptus gunnii* CCR and *CAD2* promoters are active in lignifying cells during primary and secondary xylem formation in *Arabidopsis thaliana*. Plant Physiol. Biochem. 44, 674–683. doi: 10.1016/j.plaphy.2006.10.027

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

Barros, J., Escamilla-Trevino, L., Song, L., Rao, X., Serrani-Yarce, J. C., Palacios, M. D., et al. (2019). 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase. *Nat. Commun.* 10, 1994.

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, H., Wang, J. P., Liu, H. Z., Li, H. Y., Lin, Y. C. J., Shi, R., et al. (2019). Hierarchical transcription factor and chromatin binding network for wood formation in black cottonwood (*Populus trichocarpa*). *Plant Cell* 31, 602–626. doi: 10.1105/tpc.18.00620

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

Fraser, C. M., and Chapple, C. (2011). The phenylpropanoid pathway in *Arabidopsis*. *Arabidopsis Book* 9:e0152. doi: 10.1199/tab.0152

Geng, P., Zhang, S., Liu, J., Zhao, C., Wu, J., Cao, Y., et al. (2020). MYB20, MYB42, MYB43 and MYB85 regulate phenylalanine and lignin biosynthesis during secondary cell wall formation. *Plant Physiol.* 182, 1272–1283. doi: 10.1104/pp.19.01070

Goujon, T., Sibout, R., Eudes, A., Mackay, J., and Jouanin, L. (2003a). Genes involved in the biosynthesis of lignin precursors in *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 41, 677–687. doi: 10.1016/s0981-9428(03)00095-0

Goujon, T., Sibout, R., Pollet, B., Maba, B., Nussaume, L., Bechtold, N., et al. (2003b). A new *Arabidopsis thaliana* mutant deficient in the expression of *O*-methyltransferase impacts lignins and sinapoyl esters. *Plant Mol. Biol.* 51, 973–989.

Guillaumie, S., Mzid, R., Mechin, 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

Gunasekara, C., Zhang, K., Deng, W., Brown, L., and Wei, H. (2018). TGMI: an efficient algorithm for identifying pathway regulators through evaluation of triple-gene mutual interaction. *Nucleic Acids Res.* 46:e67. doi: 10.1093/nar/gky210

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

Humphreys, J. M., Hemm, M. R., and Chapple, C. (1999). New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10045–10050. doi: 10.1073/pnas.96.18.10045

Jin, H. L., 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

Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316, 1484–1488. doi: 10.1126/science.1138341

Kim, M. H., Cho, J. S., Jeon, H. W., Sangsawang, K., Shim, D., Choi, Y. I., et al. (2019). Wood transcriptome profiling identifies critical pathway genes of secondary wall biosynthesis and novel regulators for vascular cambium development in populus. *Genes (Basel)* 10, E690.

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

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

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

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., Lu, W. X., Liu, R., Tian, Q. Y., Sun, Y. M., et al. (2014). A poplar R2R3-MYB transcription factor, PtrMYB152, is involved in regulation of lignin biosynthesis during secondary cell wall formation. *Plant Cell Tissue Organ Cult.* 119, 553–563. doi: 10.1007/s11240-014-0555-8

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

Lin, J. S., Lin, C. C., Lin, H. H., Chen, Y. C., and Jeng, S. T. (2012). MicroR828 regulates lignin and H2O2 accumulation in sweet potato on wounding. *New Phytol.* 196, 427–440. doi: 10.1111/j.1469-8137.2012.04277.x

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. U.S.A.* 114, E9722–E9729.

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. U.S.A.* 110, 10848–10853. doi: 10.1073/pnas.1308936110

Marita, J. M., Ralph, J., Lapierre, C., Jouanin, L., and Boerjan, W. (2001). NMR characterization of lignins from transgenic poplars with suppressed caffeic acid *O*-methyltransferase activity. *J. Chem. Soc. Perkin Trans.* 1, 2939–2945. doi: 10.1039/b107219f

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

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

Öhman, D., Demedts, B., Kumar, M., Gerber, L., Gorzsás, 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

Ong, S. S., and Wickneswari, R. (2012). Characterization of microRNAs expressed during secondary wall biosynthesis in *Acacia mangium*. *PLoS ONE* 7:e49662. doi: 10.1371/journal.pone.0049662

Raes, J., Rohde, A., Christensen, J. H., Van De Peer, Y., and Boerjan, W. (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiol.* 133, 1051–1071. doi: 10.1104/pp.103.026484

Rahantamalala, A., Rech, P., Martinez, Y., Chaubet-Gigot, N., Grima-Pettenati, J., and Pacquit, V. (2010). Coordinated transcriptional regulation of two key genes in the lignin branch pathway-CAD and CCR-is mediated through MYB-binding sites. *BMC Plant Biol.* 10:130. doi: 10.1186/1471-2229-10-130

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

Sharma, D., Tiwari, M., Pandey, A., Bhatia, C., Sharma, A., and Trivedi, P. K. (2016). MicroRNA858 is a potential regulator of phenylpropanoid pathway and plant development. *Plant Physiol.* 171, 944–959.

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, W., Quan, M., Du, Q., and Zhang, D. (2017). The interactions between the long non-coding RNA *NERDL* and its target gene affect wood formation in *Populus tomentosa*. *Front. Plant Sci.* 8:1035. doi: 10.3389/fpls.2017.01035

Sullivan, M. L., and Green, P. J. (1993). Post-transcriptional regulation of nuclear-encoded genes in higher plants: the roles of mRNA stability and translation. *Plant Mol. Biol.* 23, 1091–1104. doi: 10.1007/bf00042344

Sun, Q., Liu, X., Yang, J., Liu, W., Du, Q., Wang, H., et al. (2018). MicroRNA528 affects lodging resistance of maize by regulating lignin biosynthesis under nitrogen-luxury conditions. *Mol. Plant* 11, 806–814. doi: 10.1016/j.molp.2018.03.013

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., and Zhu, J.-K. (2004). Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* 16, 2001–2019. doi: 10.1105/tpc.104.022830

Tian, Q. Y., Wang, X. Q., Li, C. F., Lu, W. X., Yang, L., Jiang, Y. Z., 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

Tsai, C. J., Harding, S. A., Tschaplinski, T. J., Lindroth, R. L., and Yuan, Y. N. (2006). Genome-wide analysis of the structural genes regulating defense phenylpropanoid metabolism in *Populus*. *New Phytol.* 172, 47–62. doi: 10.1111/j.1469-8137.2006.01798.x

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.

Wagner, A., Ralph, J., Akiyama, T., Flint, H., Phillips, L., Torr, K., et al. (2007). Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyltransferase in *Pinus radiata*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11856–11861. doi: 10.1073/pnas.0701428104

Wang, C. Y., Zhang, S. C., Yu, Y., Luo, Y. C., Liu, Q., Ju, C. L., et al. (2014). MiR397b regulates both lignin content and seed number in *Arabidopsis* via modulating a laccase involved in lignin biosynthesis. *Plant Biotechnol. J.* 12, 1132–1142. doi: 10.1111/pbi.12222

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, J. P., Matthews, M. L., Williams, C. M., Shi, R., Yang, C. M., Tunlaya-Anukit, S., et al. (2018). Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat. Commun.* 9:1579.

Wang, M., Yuan, D., Tu, L., Gao, W., He, Y., Hu, H., et al. (2015). Long noncoding RNAs and their proposed functions in fibre development of cotton (*Gossypium* spp.). *New Phytol.* 207, 1181–1197. doi: 10.1111/nph.13429

Xie, M., Muchero, W., Bryan, A. C., Yee, K., Guo, H.-B., 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

Xu, C. Z., Fu, X. K., Liu, R., Guo, L., Ran, L. Y., Li, C. F., 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, P., Kong, Y., Song, D., Huang, C., Li, X., and Li, L. (2014). Conservation and functional influence of alternative splicing in wood formation of *Populus* and *Eucalyptus*. *BMC Genomics* 15:780. doi: 10.1186/1471-2164-15-780

Xu, Q., Wang, W. Q., Zeng, J. K., Zhang, J., Grierson, D., Li, X., et al. (2015). A NAC transcription factor, EjNAC1, affects lignification of loquat fruit by regulating lignin. *Postharvest Biol. Technol.* 102, 25–31. doi: 10.1016/j.postharvbio.2015.02.002

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. Bot.* 64, 3021–3032. doi: 10.1093/jxb/ert150

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.

Yang, L., Zhao, X., Yang, F., Fan, D., Jiang, Y. Z., and Luo, K. M. (2016). PtrWRKY19, a novel WRKY transcription factor, contributes to the regulation of pith secondary wall formation in *Populus trichocarpa*. *Sci. Rep.* 6:18643.

Yang, Y., Yoo, C. G., Rottmann, W., Winkeler, K. A., Collins, C. M., Gunter, L. E., et al. (2019). PdWND3A, a wood-associated NAC domain-containing protein, affects lignin biosynthesis and composition in *Populus*. *BMC Plant Biol.* 19:486. doi: 10.1186/s12870-019-2111-5

Zhang, J., Li, M., Bryan, A. C., Yoo, C. G., Rottmann, W., Winkeler, K. A., et al. (2019a). Overexpression of a serine hydroxymethyltransferase increases biomass production and reduces recalcitrance in the bioenergy crop *Populus*. *Sustain. Energy Fuels* 3, 195–207. doi: 10.1039/c8se00471d

Zhang, J., Xie, M., Li, M., Ding, J., Pu, Y., Bryan, A. C., et al. (2019b). Overexpression of a *Prefoldin* β subunit gene reduces biomass recalcitrance in the bioenergy crop *Populus*. *Plant Biotechnol. J.* 18, 859–871. doi: 10.1111/pbi.13254

Zhang, J., Xie, M., Tuskan, G. A., Muchero, W., and Chen, J. G. (2018a). 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

Zhang, J., Yang, Y., Zheng, K., Xie, M., Feng, K., Jawdy, S. S., et al. (2018b). 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, Q. A., Wang, H. Z., Yin, Y. B., Xu, Y., Chen, F., and Dixon, R. A. (2010). Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14496–14501. doi: 10.1073/pnas.1009170107

Zhao, Y., Sun, J., Xu, P., Zhang, R., and Li, L. (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., Lee, C., and Ye, Z. H. (2010a). 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., 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., and Ye, Z.-H. (2011). 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. Q., 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

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: 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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 May 2020
doi: 10.3389/fpls.2020.00651

Cell Wall Acetylation in Hybrid Aspen Affects Field Performance, Foliar Phenolic Composition and Resistance to Biological Stress Factors in a Construct-Dependent Fashion

*Marta Derba-Maceluch[1], Fariba Amini[2,3], Evgeniy N. Donev[1], Prashant Mohan-Anupama Pawar[1†], Lisa Michaud[2], Ulf Johansson[4], Benedicte R. Albrectsen[2] and Ewa J. Mellerowicz[1]**

[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, [3] Biology Department, Faculty of Science, Arak University, Arak, Iran, [4] Tönnersjöheden Experimental Forest, Swedish University of Agricultural Sciences, Simlångsdalen, Sweden

The production of biofuels and "green" chemicals from the lignocellulose of fast-growing hardwood species is hampered by extensive acetylation of xylan. Different strategies have been implemented to reduce xylan acetylation, resulting in transgenic plants that show good growth in the greenhouse, improved saccharification and fermentation, but the field performance of such plants has not yet been reported. The aim of this study was to evaluate the impact of reduced acetylation on field productivity and identify the best strategies for decreasing acetylation. Growth and biological stress data were evaluated for 18 hybrid aspen lines with 10–20% reductions in the cell wall acetyl content from a five year field experiment in Southern Sweden. The reduction in acetyl content was achieved either by suppressing the process of acetylation in the Golgi by reducing expression of *REDUCED WALL ACETYLATION* (*RWA*) genes, or by post-synthetic acetyl removal by fungal acetyl xylan esterases (AXEs) from two different families, CE1 and CE5, targeting them to cell walls. Transgene expression was regulated by either a constitutive promoter (*35S*) or a wood-specific promoter (*WP*). For the majority of transgenic lines, growth was either similar to that in WT and transgenic control (*WP:GUS*) plants, or slightly reduced. The slight reduction was observed in the AXE-expressing lines regulated by the *35S* promoter, not those with the *WP* promoter which limits expression to cells developing secondary walls. Expressing AXEs regulated by the *35S* promoter resulted in increased foliar arthropod chewing, and altered condensed tannins and salicinoid phenolic glucosides (SPGs) profiles. Greater growth inhibition was observed in the case of CE5 than with CE1 AXE, and it was associated with increased foliar necrosis and distinct SPG profiles, suggesting that CE5 AXE could

OPEN ACCESS

Edited by:
Zeng-Yu Wang,
Qingdao Agricultural University, China

Reviewed by:
Ajaya K. Biswal,
University of Georgia, United States
R. Glen Uhrig,
University of Alberta, Canada

***Correspondence:**
Ewa J. Mellerowicz
ewa.mellerowicz@slu.se

†Present address:
Prashant Mohan-Anupama Pawar,
Regional Centre for Biotechnology, NCR Biotech Science Cluster, Haryana, India

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *21 February 2020*
Accepted: *27 April 2020*
Published: *25 May 2020*

Citation:
Derba-Maceluch M, Amini F, Donev EN, Pawar PM-A, Michaud L, Johansson U, Albrectsen BR and Mellerowicz EJ (2020) Cell Wall Acetylation in Hybrid Aspen Affects Field Performance, Foliar Phenolic Composition and Resistance to Biological Stress Factors in a Construct-Dependent Fashion. Front. Plant Sci. 11:651. doi: 10.3389/fpls.2020.00651

be recognized by the pathogen-associated molecular pattern system. For each of three different constructs, there was a line with dwarfism and growth abnormalities, suggesting random genetic/epigenetic changes. This high frequency of dwarfism (17%) is suggestive of a link between acetyl metabolism and chromatin function. These data represent the first evaluation of acetyl-reduced plants from the field, indicating some possible pitfalls, and identifying the best strategies, when developing highly productive acetyl-reduced feedstocks.

Keywords: *Populus tremula* × *tremuloides*, transgenic trees, field trial, biotic resistance, salicinoid phenolic glucosides, condensed tannins, *HjAXE*, *AnAXE1*

INTRODUCTION

Plant cell walls (lignocellulose) constitute by far the most abundant carbon source on Earth available for the sustainable production of advanced biofuels and "green" chemicals (Bar-On et al., 2018). These products are made through saccharification which converts lignocellulose to fermentable sugars. The industrial saccharification and fermentation processes are challenged by, among other factors, the abundance of acetylation substituents (Jönsson and Martín, 2016) present in most cell wall polymers (Gille and Pauly, 2012; Donev et al., 2018). Dicotyledonous plants, including broadleaf trees (hardwoods), are particularly rich in *O*-acetyl substituents, the majority of which are associated with xylan (Pawar et al., 2013; Pauly and Ramírez, 2018). Biological role of xylan acetylation is not fully understood, but it is known to affect xylan solubility (Gröndahl et al., 2003), susceptibility to enzymatic degradation (Biely et al., 2016), interaction with cellulose (Grantham et al., 2017) and lignin (Giummarella and Lawoko, 2016). On the other hand, there is a considerable variation among different groups of plants in xylan acetylation, and some of them, like conifers, have no acetyl xylan substitution (Pawar et al., 2013). Several attempts have therefore been made to reduce acetyl content in dicotyledon species (Pogorelko et al., 2011; Xiong et al., 2015; Pawar et al., 2016) including hardwoods (Ratke et al., 2015; Pawar et al., 2017a,b; Wang et al., 2020). Based on the performance of greenhouse-grown plants, reductions in *O*-acetylation were found to be well tolerated by plants when the degree of xylan substitution was reduced by 30% or less (Pogorelko et al., 2011; Xiong et al., 2015; Pawar et al., 2016). Moreover, reducing acetylation was found to be one of the most promising strategies for improving plant cell walls for the purposes of saccharification and fermentation (Donev et al., 2018). These results were encouraging, but the performance of such acetylation-reduced lines also needs to be tested in the field.

Field conditions impose both biotic and abiotic stresses on plants, and therefore field performance may be very different from growth observed in the greenhouse (Strauss, 2003). Acetylation-challenged plants in particular could perform differently between these two sets of conditions, since such plants have been shown to react differently to both biotic and abiotic stresses. For example, *Arabidopsis* plants with a mutation in the *TRICHOME BIREFRINGENCY-LIKE 29 (TBL29)* gene encoding a key acetyl transferase involved in secondary wall xylan acetylation (Urbanowicz et al., 2014) were reported to be highly resistant to water deficit and freezing stress, and were thus named *eskimo1 (esk1)* (Xin and Browse, 1998; Xin et al., 2007; Lefebvre et al., 2011; Xu et al., 2014). Plants mutated in *TBL44/PMR5* from the same family are known to be resistant to powdery mildew (Vogel et al., 2004). Similarly, an acetylation deficit caused by mutations in *REDUCED WALL ACETYLATION* (*RWA*) genes that affect acetylation of all cell wall polysaccharides resulted in biotic resistance to biotrophic and necrotrophic fungi in *Arabidopsis* (Manabe et al., 2011; Pawar et al., 2016). Post-synthetic removal of acetic groups from the xylan backbone by transgenic expression of fungal acetyl xylan esterases (AXEs) has been shown to increase resistance to certain pathogenic fungi (Pogorelko et al., 2013; Pawar et al., 2016). Thus deacetylation of xylan appears to lead to better plant resistance to biotic and abiotic stresses. Naturally occurring deacetylation of pectin by the enzyme encoded by *PECTIN ACETYLESTERASE 9* (*AtPAE9*) has been shown to be required for proper basal levels of innate immunity and resistance to aphids (Kloth et al., 2019). A knock out *pae9* mutant with increased rhamnogalacturonan I (RGI) and homogalacturonan acetylation compared to wild-type plants (De Souza et al., 2014) exhibited decreased concentrations of JA, SA, ABA, and IAA, and initial facilitation of cell wall penetration by aphids (Kloth et al., 2019). Although the mechanism by which the cell wall acetylation level is communicated to the plant cell protoplast is at present not known (reviewed by Bacete et al., 2018), it is clear that modifying acetylation can impact plant biotic and abiotic resistance, which are key parameters affecting the field performance of plants.

To assess the field performance of acetylation-reduced plants we tested transgenic hybrid aspen (*Populus tremula* L. × *tremuloides* Michx.) lines in which the acetyl content was reduced by different means. These lines, in which xylan acetylation was post-synthetically reduced, included ones expressing AXEs of fungal origin from two Carbohydrate Esterase families, CE1 and CE5, *Aspergillus niger* AXE1 (*An*AXE1) and *Hypocrea jecorina* AXE (*Hj*AXE), respectively (Ratke et al., 2015; Pawar et al., 2017b; Wang et al., 2020).

Abbreviations: *35S*, Cauliflower mosaic virus *35S* promoter; ABA, abscisic acid; *An*AXE1, *Aspergillus niger* ACETYL XYLAN ESTERASE 1; CT, Condensed Tannin; *Hj*AXE, *Hypocrea jecorina* ACETYL XYLAN ESTERASE; IAA, indole-3-acetic acid; JA, jasmonic acid; PAE9, PECTATE ACETYLESTERASE 9; RGI, rhamnogalacturonan I; RWA, REDUCED WALL ACETYLATION; SA, salicylic acid; SPGs, salicinoid phenolic glucosides; TBL, TRICHOME BIREFRINGENCY-LIKE; *WP*, Wood Promoter.

AnAXE1 and *HjAXE* expressing lines grown in a greenhouse environment developed as well as wild type, and had superior saccharification properties. The lines in which acetyl content was reduced due to deficiencies in the biosynthetic acetylation machinery were those with reduced expression of native *REDUCED WALL ACETYLATION* (*RWA*) genes (Pawar et al., 2017a). The latter lines exhibited similar reductions in acetylation and improved saccharification properties, with good growth in the greenhouse, as the lines in which the xylan was post-synthetically deacetylated. Using these lines we were able to compare the impacts of reducing acetylation during the biosynthesis of xylan in the Golgi by suppressing *RWA* genes with those where the reduction was achieved post-synthetically in cell walls by expressing the fungal enzymes *An*AXE1 and *Hj*AXE targeted to the apoplast. We also addressed the question of the promoter to be used for genetic engineering. We compared the effects of the same transgenes expressed from either constitutive *35S* or wood-specific *WP* (Ratke et al., 2015) promoters. Finally, we compared the effects of two fungal enzymes belonging to the different CE families. We monitored growth over five years, and determined foliar biotic damage and foliar concentrations of phenylpropanoid compounds, which are indicators of stress induction and stress resistance (Dixon and Paiva, 1995; Papazian et al., 2019). In *Populus* spp., biotic stress has commonly been associated with levels of condensed tannins (CTs) (Bandau et al., 2015; Lindroth and Madritch, 2015; Lindroth et al., 2015) and salicinoid phenolic glucosides (SPGs) (Albrectsen et al., 2010; Robinson et al., 2012; Lindroth and St. Clair, 2013; Lindroth and Madritch, 2015), and these phenylpropanoid compounds are often related to environmental stress responses and performance (Lindroth et al., 2011; Robinson et al., 2012; Keefover-Ring et al., 2014; Bandau et al., 2015; Decker et al., 2016). We therefore measured foliar concentrations of these compounds in acetylation-compromised aspen lines. This is the first analysis of the field performance of plants with reduced acetylation.

MATERIALS AND METHODS

Biological Material

Hybrid aspen (*Populus tremula* L. × *tremuloides* Michx.) clone T89 was used as wild-type and all transgenic lines were made in this genetic background. Transgenic lines initially tested in the greenhouse included those expressing *35S:An*AXE1 (Pawar et al., 2017b), *35S:HjAXE* and *WP:Hj*AXE (Ratke et al., 2015; Wang et al., 2020), as well as lines with RNAi constructs targeting hybrid aspen *RWA-C* and *RWA-D* genes, denoted *35S:RWA-CD* (previously called *35S::CD-RWA RNAi*), and all four *RWA* genes, denoted *WP:RWA-ABCD* (previously called *pGT43B::RWA-ABCD RNAi*) (Pawar et al., 2017a). Two lines expressing a β-glucuronidase (GUS) gene under the control of the *WP* promoter (Ratke et al., 2015) were used as transgenic controls. Additional lines were generated that expressed *WP:AnAXE1* using the *pK-pGT34B-GW* destination vector (Ratke et al., 2015) and the *AnAXE1* cDNA as previously described (Pawar et al., 2017b). Each construct was represented by two to four lines selected from among approx. 20 independent lines, as described previously, based on the strength of transgene expression and superiority of greenhouse performance with regard to growth and saccharification properties (as described by Ratke et al., 2015; Pawar et al., 2017a,b; Wang et al., 2020), with the exception of *WP:AnAXE1*. The *WP:AnAXE1* lines were selected based on the strength of transgene expression in plants cultivated *in vitro*, as determined by RT-PCR analysis (**Supplementary Figure S1**). A list of the lines and constructs is given in **Table 1**.

TABLE 1 | Lines included in the field trial analysis.

Construct	Lines	References
35S:RWA-CD	10, 21, 22	Pawar et al., 2017a
WP:RWA-ABCD	11, 15	Pawar et al., 2017a
35S:AnAXE1	4, 8, 17	Pawar et al., 2017b
WP:AnAXE1	1, 5, 8, 10	This paper, **Supplementary Figure S1**
35S:HjAXE	9, 13, 22	Ratke et al., 2015; Wang et al., 2020
WP:HjAXE	11, 14B, 14C	Ratke et al., 2015; Wang et al., 2020
WP:GUS	25, 27	Ratke et al., 2015

Field Trial Establishment and Experimental Design

Transgenic trees were propagated *in vitro* at the Umeå Plant Science Centre transformation facility in Umeå, Sweden (Nilsson et al., 1992) and transplanted into soil for one month of acclimatization in the greenhouse in early spring 2014. Then, with permission from the Swedish Board of Agriculture (DNR. 4.6.18-761/14), the trees were moved to sheltered outdoor premises in Umeå for a two-week hardening period, before translocation to the field site (ca 1000 km south of Umeå) in Våxtorp, Laholm community, Sweden (56.42°N, 13.07°E). Between August 4 and 8, 2014, the trees were planted in the field, with a 3 m spacing, on abandoned farm land fenced according to the requirements for genetically modified plants. In total, 636 trees included in the current analysis, along with other trees not analyzed here, were arranged in a 14 block design, with two trees of each transgenic line randomly distributed within each block along with four wild-type (WT) trees (**Supplementary Figure S2**). For weed control, the field was harrowed twice a year during the first two years following planting and grass was mowed twice a year during subsequent years. All biosecurity and safety procedures for field trials with transgenic plants required by the Swedish Board of Agriculture have been adhered to.

Histochemical Analysis of GUS Expression

The stability of GUS gene expression was investigated in July of year four (2017) during the period of active cambial growth. The basal part of a two-year-old branch was hand sectioned and the sections were prefixed in acetone for 30 min, washed with water, placed in the reaction solution (1 mM X-GlcA (5-bromo-4-chloro-3-indolyl β-D-glucuronide), 50 mM Na-phosphate buffer pH 7, 0.1% (v/v) Triton, 1 mM $K_3[Fe(CN)_6]$, 1 mM $K_4[Fe(CN)_6]$), and incubated for 3 days in the dark at room temperature. Sections were then fixed in FAA (50% ethanol, 5% formaldehyde, 10% acetic acid, all v/v) overnight followed by clearing and dehydration in an ethanol series. Samples were

rehydrated, mounted in 50% (v/v) glycerol, and imaged with a Zeiss Axioplan 2 microscope using a 40× objective. Micrographs were taken with an Axiocam HRc camera and Axiovision V 4.8.2 Software (Carl Zeiss Light Microscopy, Göttingen, Germany). Images were combined into panoramas covering the section from bark to pith using the program Adobe Photoshop CS6.

Periodic Growth and General Damage Assessment

Growth parameters (plant height and root collar diameter) were assessed at the end of each growth season. Final tree height and diameter were also measured in June 2018, before harvesting. The height was assessed with a measuring stick and the stem diameter with a caliper (3 cm above ground level). Stem volume was calculated as: $V = 1/3\pi R^2 H$, where R – stem radius, H – stem height.

Standard assessments (Nilsson and Örlander, 1999) of damage were conducted four times a year; they included cause of damage (fungi, frost, drought, waterlogging, rodents, herbivores, insects, vegetation, unknown) and severity on a six-level scale (0 = undamaged, 1 = slight damage, 2 = uncertain or moderate damage, 3 = severe damage, 4 = life-threatening damage, and 5 = dead).

Detailed Biotic Stress Assessment and Leaf Collection

A detailed assessment of leaf damage was performed in July 2017 during the fourth growing season, according to previous methodology (Albrectsen et al., 2010; Robinson et al., 2012). Chewing damage by arthropods (chewing) was assessed as percentage of chewed leaves in the canopy. Evidence of other types of damage was scored in terms of presence (1) or absence (0). This was done for damage caused by arthropods including aphids, miners, gall-producing organisms, and pathogens including rust (*Melampsora spp.*) and venturia (*Venturia spp.*). Symptoms of chlorosis, necrosis, and hypersensitive response (HR) were recorded in the same way.

Assessment of Traits Related to Architecture and Leaf Chemical Profiling

Architectural traits, chlorophyll index, and leaf CT contents were assessed for a subsample consisting of 50% of trees having superior height selected from each transgenic line and wild type. The rationale for stratifying the samples by height rather than randomly selecting 50% of trees from each line was to avoid those trees that had been damaged by planting, field work or other types of random disturbance.

Architectural Traits

Branching was assessed according to Luquez et al. (2008). The apical dominance was scored on a scale from 0 to 8 as follows: clear leader (8)/good recovery after apex damage in 2017 (7)/good recovery after apex damage in 2016 (6)/good recovery after apex damage in 2015 (5)/main shoot lost in 2015 (4)/main shoot lost in 2016 (3)/main shoot lost in 2017 (2)/bushy growth (1)/dwarf (0).

Chlorophyll Content

Chlorophyll content was measured in fully developed leaves with no visible damage, collected from the upper part of the main stem between June 26 and 30, 2017, using a CCM-200 plus (Opti-Science, Huston, United States). A mean value from 18 measurements per tree was obtained (six leaves per tree and three measurements per leaf). The same leaves were collected for metabolite analyses and dry weight assessment. They were immediately frozen on dry ice and freeze-dried before being transported to Umeå for weighing, grinding and metabolic profiling.

Condensed Tannins

Six freeze-dried leaves per tree were ground together to a powder. Foliar CT contents were assessed based on the acid-butanol method of Porter et al. (1986). In short, 10.0 ± 2.0 mg leaf powder (exact weight) was extracted with 800 μl of a mix of acetone and 10 mM ascorbic acid solutions in a 70:30 (v:v) ratio, mixed by vortexing, sonicated, and centrifuged for 5 min at 3500 rpm on a bench top centrifuge. The absorbance of the extract (150 μl of supernatant) at 550 nm was measured with a spectrophotometer (Hitachi U-5100 UV/VIS, Hitachi High-Technologies, Tokyo, Japan). Results were compared to a standard curve of procyanidin B2 ($C_{30}H_{26}O_{12}$, Sigma-Aldrich®, St. Louis, MO, United States) and recalculated to give mg/g (d.w.) leaf powder.

Metabolite Analysis

Four trees per line were randomly selected from the set used for leaf CT determination. Ultra high performance liquid chromatography (UHPLC) with UV and electro-spray ionization time-of-flight mass spectrometry (ESI-TOF/MS) detectors was used as described by Abreu et al. (2011) and Keefover-Ring et al. (2014). In short, 10.00 ± 1.00 mg of ground leaf material was extracted in 1 ml of cold (4°C) methanol: chloroform: water, 60:20:20 (v:v:v), with deuterated SA as an internal standard. After centrifugation, 200 μl of the extract supernatant was dried in a speedvac. Before analysis, the samples were reconstituted with 20 μl of methanol and 20 μl of a 0.1% v/v aqueous formic acid solution. Compounds in the reconstituted plant extracts were separated on a C18 UPLC™ column (2.1 × 100 mm, 1.7 μm) and analyzed by an Acquity photodiode array detector coupled in line with a LCT Premier TOF/MS (all from Waters, Milford, MA, United States) as described by Abreu et al. (2011).

The MassLynx 4.1 software package (Waters Corp.) was used to extract single ion chromatograms (±0.15 exact mass unit) using the QuanLynx module to search for known and theoretical phenylpropanoids (using deprotonated ($[M-H]^-$) and formate adduct ($[M-H+FA]^-$) ions). QuanLynx software was used to obtain peak areas that were normalized with respect to internal standard peak area and sample weight; as described in Abreu et al. (2011) and Keefover-Ring et al. (2014). The phenylpropanoids salicortin, tremulacin, salicin, tremuloidin, salicyloylsalicin, HCH-salicortin, 2′-(E)-, and 2′-(Z)-cinnamoylsalicortin were determined using retention times and molecular weight information for purified standards. Other compounds

(2′-acetylsalicin, 2′-acetylsalicortin, acetyl tremulacin, HCH-2′-acetylsalicortin, HCH-tremulacin, and arachidonic acid) were tentatively identified based on LC-MS molecular weights and defragmentation patterns.

Statistical Analyses

All analyses were performed using the software package JMP 14.0.0 2018 (SAS Institute Inc.). The consequences of decreased acetylation for field performance were analyzed for a total of 19 growth related traits, 14 biotic stress related traits, and 23 foliar defense related chemicals (mainly phenylpropanoids). Effects of individual lines were tested by a one-way ANOVA with "line" used as fixed effect (**Supplementary Tables S1, S2**). Similarity of individual lines to WT was evaluated by a Dunnett's test, and consistent line effect within a construct were assessed by a contrast analysis (all lines for a given construct versus WT). These results were used to identify cases of reproducible construct effects among different transgenic lines.

A nested-ANOVA model design was used to answer questions about the impact of construct on phenotypic trait expression. "Line" nested in "construct" and block (random, considered when possible) effects were included (**Supplementary Tables S3, S4**). A multiple comparison Tukey test was used for cross-comparisons among different constructs.

Impacts of the promoter and the transgene (fixed effects) were analyzed by a two-way ANOVA with an interaction (**Supplementary Table S6**), and comparison of deacetylation strategy was carried out by a nested ANOVA model with "construct " nested in "pre- or post-synthetic strategy" and "line" nested within "construct" (**Supplementary Table S7**), all used as fixed effects.

RESULTS

Field Growth Analysis Identified Three Lines With Anomalies

Uniform growth and a stable survival rate of close to 100% (86–100%) characterized the majority of the transgenic lines throughout the period of testing at the field site (**Figure 1A** and **Supplementary Table S1**). However, three out of 20 transgenic lines for three different constructs (namely line 22 for construct *35S:RWA-CD*, line 15 for construct *WP:RWA-ABCD*, and line 11 for construct *WP:HjAXE*) exhibited distinct deviations from the general growth pattern. These lines were dwarf, reaching only 1, 26, and 2% of WT stem volume, respectively, and their apical dominance was reduced compared to that of WT and other lines with the same construct (**Figures 1B,C** and **Supplementary Table S1**). Line 11 of *WP:HjAXE* had significantly higher mortality than the other transgenic lines and WT (**Figure 1D**), and approx. 50% of trees of this line showed a striking variegated phenotype that suggested genomic instability (**Figures 1D,E**). None of these phenotypes were seen in greenhouse trials (Ratke et al., 2015; Pawar et al., 2017a,b; Wang et al., 2020). The three dwarf lines were also more affected by hare browsing and multiple injuries than the other lines or WT trees (**Supplementary Table S1**).

Since other lines with the *WP:HjAXE* and *35S:RWA-CD* constructs did not exhibit dwarf phenotypes, and moreover since line 11 of *WT:HjAXE* had lower transgene expression levels than the other lines carrying this construct (Wang et al., 2020), we do not consider these anomalies to have been caused by the respective transgenes. Rather these phenotypes should be attributed to mutations induced either by random transgene insertions in the genome or by somaclonal variation. However, the *WP:RWA-ABCD* construct was represented by only two lines that had very different phenotypes (**Figures 1B–D** and **Supplementary Table S1**). It was therefore not possible to infer any true construct effect from the data and both these lines were subsequently omitted from analyses testing construct effects. Thus, the subsequent analyses evaluate the effects of five constructs: *35S:RWA-CD*, *35S:AXE1*, *35S:HjAXE*, *WP:AnAXE1*, and *WP:HjAXE,* with the anomalous lines removed from analyses.

Effects of Constructs on Tree Growth and Development

Stem growth parameters (height and diameter) for the different constructs are shown in **Figure 2** and **Supplementary Tables S2, S3**. Trees expressing *WP:AXE1* and *WP:HjAXE* were slightly bigger than WT in the first year, but this advantage disappeared during the subsequent years. In contrast, those with two other constructs, *35S:AnAXE1* and *35S:HjAXE,* showed a small reduction in height and /or stem diameter after three and four years of growth (**Figures 2A,B**), resulting in a decrease in stem volume in the fourth year of 24 and 37%, respectively, as compared to WT (**Figure 2C**).

Shoot apical dominance and the branching pattern did not show any construct-related effects (**Supplementary Tables S2, S3**). Leaf dry weight and chlorophyll content were reduced in *35S:HjAXE* expressing trees as compared to WT, but their leaf morphology did not change (**Figure 3**). Interestingly, the chlorophyll content was slightly increased in *35S:RWA-CD* (**Figure 3**); this was not observed in any of the *WP:RWA-ABCD* lines (**Supplementary Table S1**). Unexpectedly, a small decrease in chlorophyll content was observed in *WP:GUS* lines.

To assess the suitability of the *35S* and *WP* promoters and the two *AXE* transgenes, *AnAXE1 and HjAXE* for transgenic expression, we analyzed the growth traits of a subset of transgenic lines (*35S:AXE1*, *35S:HjAXE*, *WP:AXE1*, *WP:HjAXE*) by a two-way ANOVA with promoter and transgene as fixed effects (**Supplementary Table S6**). This comparison revealed a positive effect of *WP*, compared to the *35S* promoter, on stem height, diameter, volume, and chlorophyll index (**Figures 2A–C**, **3B** and **Supplementary Table S6**). In contrast, there was hardly any effect of the transgene (*AnAXE1* vs. *HjAXE*) on growth, whereas the chlorophyll index was decreased in *HjAXE* compared with *AnAXE1* but only when combined with the *35S* promoter (**Figure 3B** and **Supplementary Table S6**).

Specificity of *WP* activity has been previously tested in greenhouse conditions (Ratke et al., 2015). We therefore investigated whether its activity and expression pattern are maintained in field conditions, using histochemical

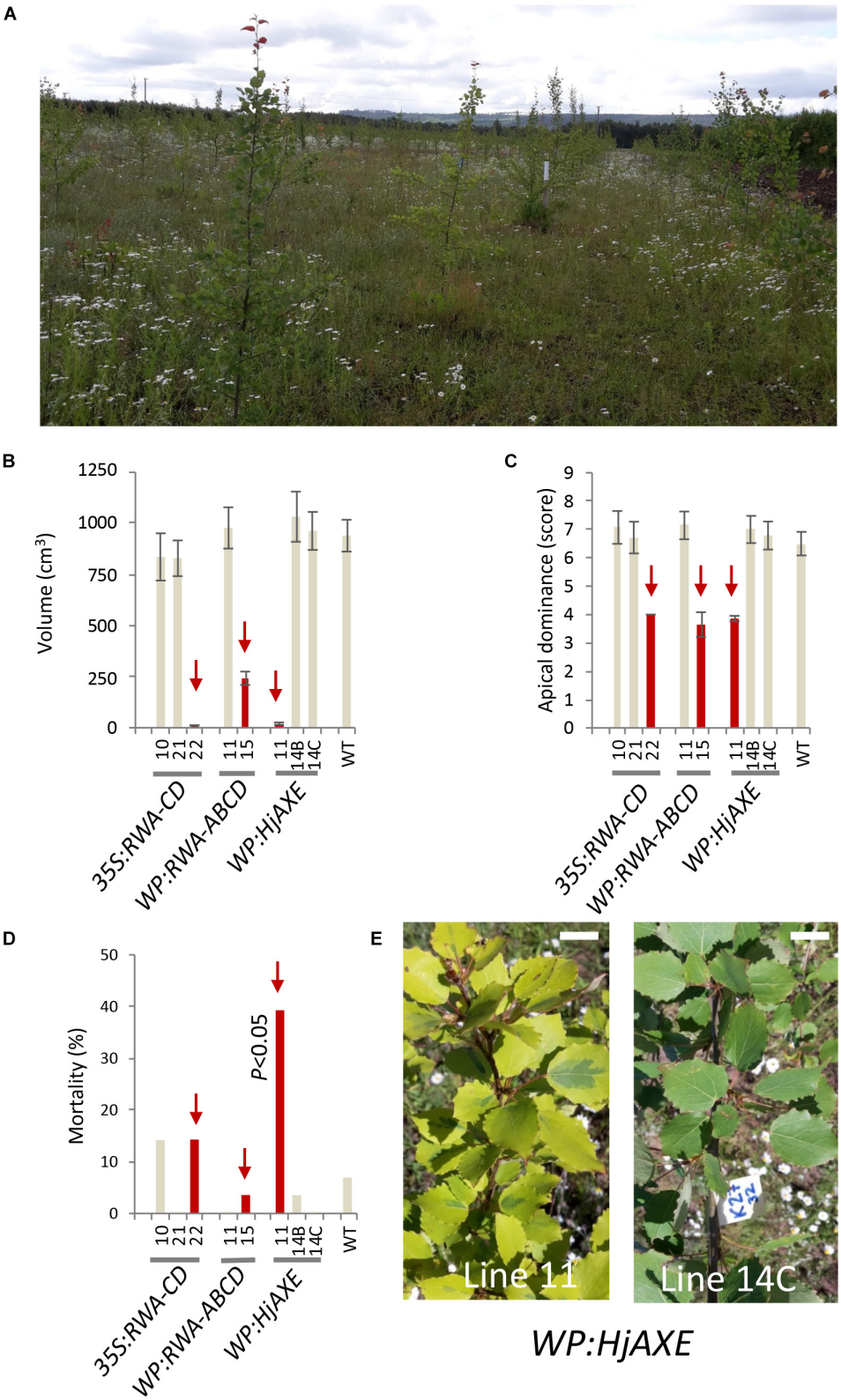

FIGURE 1 | Field testing revealed striking phenotypes in three out of 20 lines tested, effects which could not be related to transgenes. Overview of the field trial in July 2017 (fourth year) **(A)** and the corresponding data for anomalous lines: stem volume **(B)**, apical dominance **(C)** and mortality within the lines **(D)**. Lines marked in red showed aberrant morphology compared to other lines with the same construct and to WT. Line 11 with the *WP:HjAXE* construct had higher mortality than all other lines, and exhibited a variegated phenotype **(E)** not seen in other lines carrying this construct. Scale bar in E – 2 cm. Data in panels **(B)** and **(C)** are means ± SE.

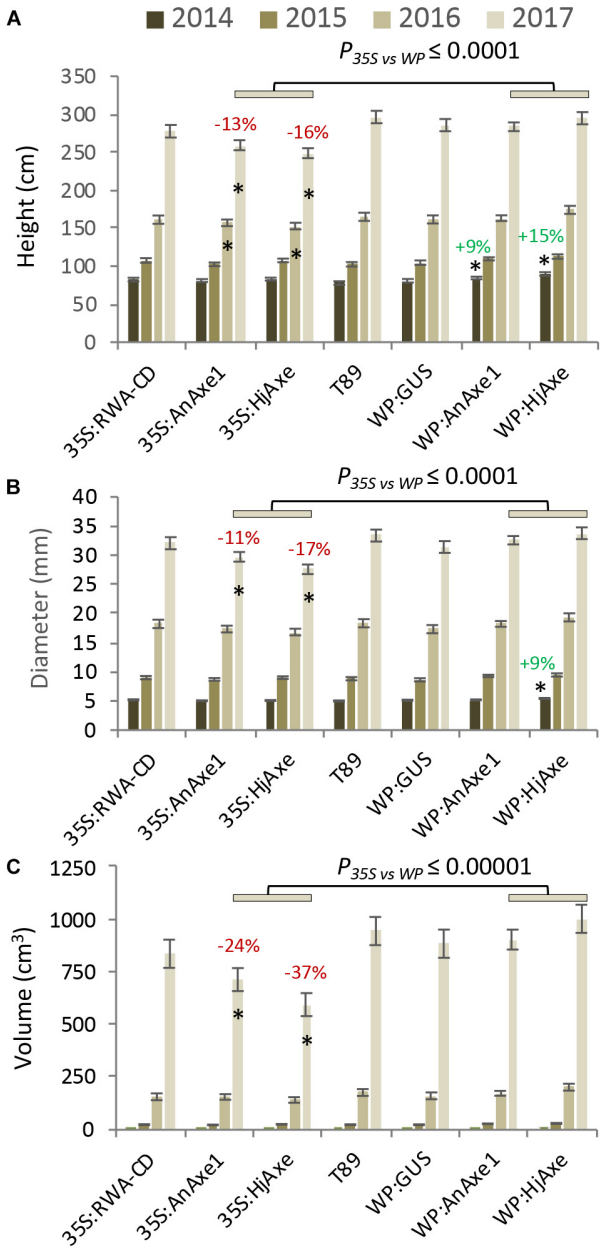

FIGURE 2 | Effects of different types of genetic modification on growth of transgenic lines during a four-year field trial. Height **(A)**, diameter **(B)**, and stem volume **(C)**. Means and SE, stars indicate means significantly different from WT, post-ANOVA contrast, $p \leq 0.05$). *P*-value indicates significance of the promoter effect based on a two-way ANOVA (**Supplementary Table S6**).

β-glucuronidase analysis of two *WP:GUS* lines during the fourth growing season. The test was carried out on branches with 1- and 2-year old cambia. In both samples and both lines, β-glucuronidase expression was detected in differentiating secondary xylem cells and secondary phloem fibers and sclereids depositing secondary walls (**Figure 4** and **Supplementary Figure S3**). This expression pattern was consistent with that previously observed in the same lines in the greenhouse (Ratke et al., 2015). We conclude that the activity and specificity of *WP* was not altered in the field.

We also investigated whether growth was differentially affected in plants in which xylan acetylation was modified during biosynthesis in the Golgi compared to plants in which post-synthetic xylan deacetylation was implemented (denoted *post-* vs. *synthetic* comparison). To answer this question, the lines with *RWA-CD* RNAi suppression driven by the *35S* promoter were compared with those where fungal *AXEs* were driven by *35S* using ANOVA (**Supplementary Table S7**). For these constructs, growth, assessed by stem height, diameter and volume, and leaf dry weight, was scarcely affected by the engineering strategy used; only the final height and diameter (measured in the middle of the fifth growing season) were slightly reduced in lines with the post-synthetic deacetylation strategy compared to those with synthetic reduction (**Supplementary Table S7**). This growth inhibition was preceded by a decrease in leaf dry weight and chlorophyll contents in the fourth year in lines with post-synthetic reduction (*35S:AnAXE1* and *35S:HjAXE*) (**Figure 3**).

Stress Response Traits

Stress-related responses were recorded regularly throughout the four-year field test period, and detailed mapping of necrosis, rust and chewing symptoms was additionally conducted in 2017 simultaneously with collection of leaf material for phenolic profiling. Some of these traits were related to construct identity and, in particular, to the choice of promoter (**Figure 5** and **Supplementary Tables S2, S3, S5–S7**). Necrosis was generally elevated for *35S:HjAXE* plants, potentially at the expense of rust symptoms (**Figure 5A**), which were generally reduced on plants belonging to the same construct. Moreover, both kinds of fungal AXE transgenes under the *35S* promoter (**Figure 5B**) suffered from an increase in the extent of chewing symptoms, with damage being increased by 187% (*35S:AnAXE1*) and 44% (*35S:HjAXE*) compared to wild type plants. Synthetic acetylation reduction (*35S:RWA-CD*), on the other hand, had no impact on chewing damage (**Figure 5B**).

Levels of phenolic compounds that are often associated with damage risk varied in a genotype-related way. For example, salicortin and HCH-acetyl-salicortin were elevated in *35S:RWA-CD* trees (**Table 2**). *35S:AnAXE1* and *35S:HjAXE* lines also expressed higher levels of certain SPGs, such as OH-tremuloidin, whereas acetyl-tremulacin and *p*-coumaric acid were greatly elevated in *35SAnAXE1*. Salicylic acid, on the other hand, was reduced in *35S:HjAXE* trees. Thus, use of the constitutive 35S promoter led to more alterations in concentrations of SPGs compared with transgenics created with the help of the *WP* promoter, which did not deviate significantly from the wildtype (WT, **Table 2**).

Catechin (the precursor of CTs) was elevated in *35S:AnAXE1*, which also had higher levels of CTs (**Figure 5C**). The elevated CT level characteristic of genotype *35S:AnAXE1* was accompanied by very much higher levels of damage caused by chewing herbivores, whereas a general decrease in CT levels for *35S:HjAXE* was also associated with elevated herbivory (by 44%). That the two fungal *35S* transformations were associated with varying responses in terms of CTs, but at the same time had similar susceptibility to

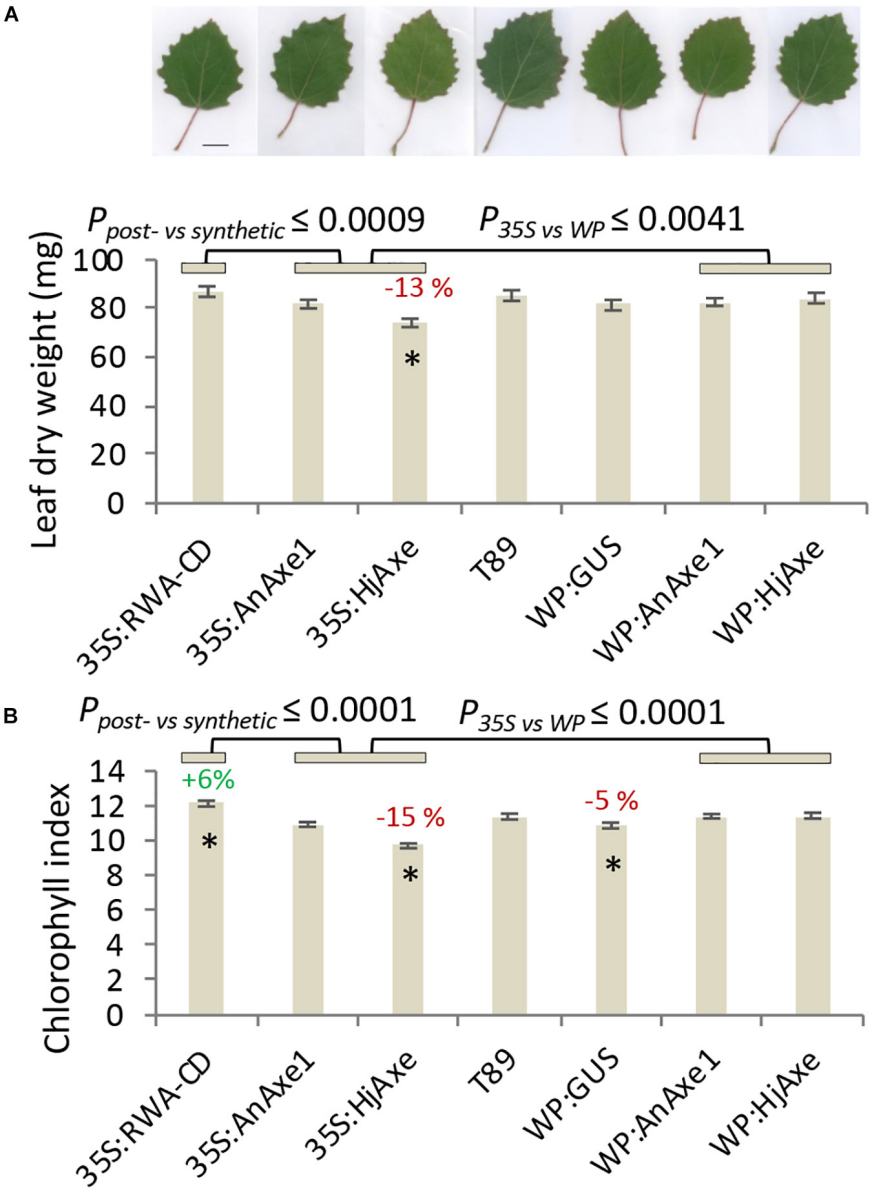

FIGURE 3 | Effects of different types of genetic modification on leaf traits. Leaf dry weight with representative images of leaves, size bar = 1 cm **(A)**. Leaf chlorophyll index **(B)**. Means and SE, stars indicate means significantly different from WT, post-ANOVA contrast, $p \leq 0.05$). *P*-value indicates significance of the promoter effect based on a two-way ANOVA (**Supplementary Table S6**) and synthetic vs. post-synthetic xylan modification (**Supplementary Table S7**).

herbivores, indicated a lack of any general relationship between CT levels and the risk of chewing damage.

In addition to SPGs, we monitored the foliar levels of arachidonic acid, a polyunsaturated fatty acid (20:4 Δ5,8,11,14) regulating different physiological and stress responses that is found in some plant species including poplars (Groenewald and van der Westhuizen, 1997). It has been shown to trigger different plant stress responses and induce resistance to fungal pathogens (Savchenko et al., 2010). Interestingly, the levels of arachidonic acid were decreased compared to WT in lines with all constructs except for those containing *AnAXE1* (**Table 2**). This suggests that the quality of lines carrying *AnAXE1* constructs is potentially superior compared to that of other transgenic lines with regard to biotic stress resistance.

DISCUSSION

Good Growth and Field Performance of Lines Carrying Constructs Targeting Acetylation

We have tested for the first time the field performance of trees with transgenically reduced cell wall acetyl content. We found

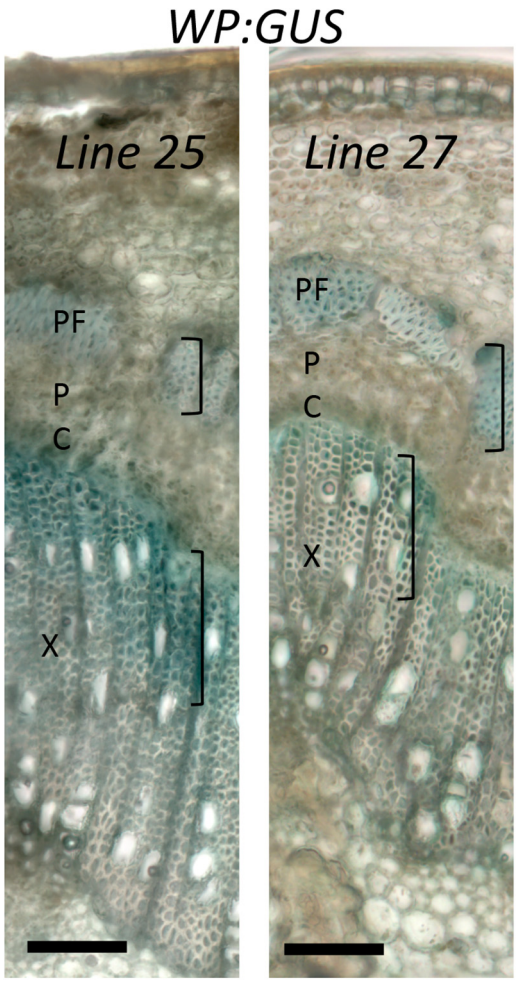

FIGURE 4 | Activity of wood-specific promoter (*WP*) in trees grown in the field. Histochemical β-glucuronidase analysis of a branch with one-year old cambium of *WP:GUS* trees (lines 25 and 27) in the fourth growing season in July, during the active wood production period. X – secondary xylem; P – secondary phloem, PF – phloem fibers; C-vascular cambium. Activity is seen in cells depositing secondary cell walls (marked with brackets). Scale bar = 100 μm.

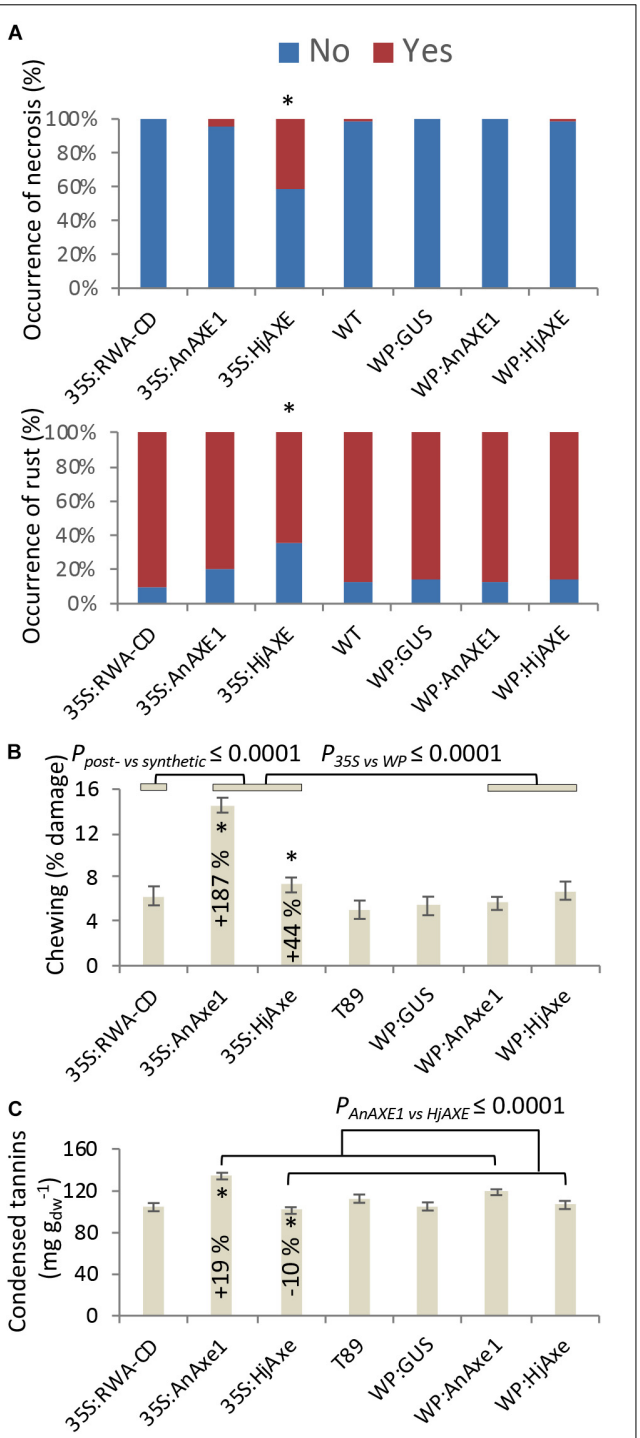

FIGURE 5 | Effects of different types of genetic modification on biotic stress responses. Instances of leaf damage recorded in the 2017 survey that showed significant effects of "construct." **(A)** Necrosis and rust. Constructs producing significantly different distributions are marked with *. Details of statistical analysis are provided in **Supplementary Table S5**. **(B)** Chewing damage. **(C)** Condensed tannin contents. Data in B and C are means and SE, stars indicate means significantly different from WT (post-ANOVA contrast, $p \leq 0.05$). *P*-value indicates significance of the transgene or promoter effect based on a two-way ANOVA (**Supplementary Table S6**) and synthetic vs. post-synthetic xylan modification strategy (**Supplementary Table S7**).

that, in general, apart from those lines with abnormal phenotypes, the reduction of acetylation in cell walls did not lead to either major detrimental effects or significant growth stimulation in the field (**Figures 2**, **3**). The level of reduction in acetyl content for the lines previously tested varied between 10 and 16% in *WP:HjAXE* (lines 11, 14B, 14C; Wang et al., 2020), and 13 and 16% in *35S:AnAXE1* (lines 4, 8, 17; Pawar et al., 2017a), and it was 20% in *35S:RWA-CD* (line 10; Pawar et al., 2017b). The reductions in acetyl content in *WP:AnAXE1* lines have not been analyzed, but based on previous comparisons between *35S* and *WP*-driven transgenes (Ratke et al., 2015), and the documented observation that the specific activity of the WP is stable (**Figure 4**), we expect a slightly stronger effect with *WP:AnAXE1* than with *35S:AnAXE1*. Thus, the mild reductions in cell wall acetylation levels (by 30% or less) are well supported by plants both, in the field (**Figure 2**), and

in the greenhouse conditions (discussed in Pawar et al., 2017b; Donev et al., 2018).

Better Field Performance of WP Compared to 35S Promoter

Analyses of growth and biotic stress responses in lines harboring *35S*- and *WP*-driven fungal AXEs point to a clear advantage offered by the specific transgene expression achieved with the *WP*. Both height and diameter growth were reduced in *35S* lines compared to *WP* lines, and the stem volume was reduced by as much as 21% and 41% in, respectively, *35S:AnAXE1* and *35S:HjAXE* compared to the corresponding *WP* constructs (**Figure 2**). These growth penalties in *35S* lines were seen only in the field; these lines did not exhibit growth defects in the greenhouse (Ratke et al., 2015; Pawar et al., 2017b; Wang et al., 2020). One factor which could have contributed to the growth penalty in the field is the alteration in interaction with herbivores as revealed by the more extensive foliar chewing damage (**Figure 5B**). The increase in susceptibility to herbivores is likely to be due to metabolic changes in the leaves in *35S* plants caused by AXE activity, changes which are largely avoided when *WP* targets the transgene expression to the developing wood. The fungal AXEs targeted to cell walls are expected to hydrolyze acetyl esters liberating acetic acid, thus changing the pH of the leaf, and since acetic acid can cross membranes in uncharged form, the reaction could contribute to the biosynthesis of acetyl-CoA. This compound is used in a variety of reactions, including the TCA cycle, glyoxylate cycle, lipid biosynthesis, mevalonate pathway, and it is considered an energy-status marker for a eukaryotic cell (Cai and Tu, 2011). It is therefore perhaps not surprising that ectopic manipulation of acetyl-CoA pools can have far-reaching consequences, and our results demonstrate that restricting the transgenic modification to specific tissues, such as developing wood, can prevent or attenuate such undesirable side-effects.

Synthetic Versus Post-synthetic Deacetylation Strategies

Assuming that the *RWA* genes encode Golgi-localized acetyl-CoA transporters (Gille and Pauly, 2012; Pauly and Ramírez, 2018), cytosolic accumulation of acetyl-CoA is expected when expression of these genes is suppressed. Apoplastic AXE expression which results in high acetic acid levels in the apoplast might eventually result in a similar outcome, assuming that acetic acid diffuses via membranes and is converted to acetyl-CoA by acetyl-CoA synthases. Even though the cytosolic accumulation of acetyl-CoA might be similar with the synthetic and the post-synthetic strategy, these two strategies might lead to differences in cell wall polymer structures. For example, increased glucuronosylation of xylan is expected when the xylan acetylation machinery is suppressed since the two processes apparently compete for the same substrate (Chong et al., 2014; Xiong et al., 2015; Grantham et al., 2017). Post-synthetic deacetylation might be also more specific toward a targeted polymer – xylan – in particular, *Hj*AXE, which was used in this study for the post-synthetic modification, has documented xylan specificity (Koutaniemi et al., 2013).

TABLE 2 | Metabolite (mostly phenylpropanoid) profiles affected by genetic transformation (construct) based on nested ANOVA.

	35S:RWA-CD	***35S:AnAXE1***	***35S:HjAXE***	**WT**	***WP:GUS***	***WP:AnAXE1***	***WP:HjAXE***
Acetyl-salicortin	458047*a* ± 27307	403962*ab* ± 22296	332471*b* ± 21540	387394*ab* ± 38618	419924*ab* ± 27307	408180*ab* ± 20098	442228*a* ± 27307
Acetyl-tremulacin	2182*b* ± 954	**109067*a*** ± 779	2293*b* ± 753	3158*b* ± 1349	2861*b* ± 954	3206*b* ± 702	2950*b* ± 954
Benzoic acid	9397*ab* ± 683	9340*a* ± 558	6842*b* ± 539	8803*ab* ± 966	7745*ab* ± 683	8337*ab* ± 503	7808*ab* ± 683
Catechin	18107820*b* ± 507156	**20517646*a*** ± 414091	19323366*ab* ± 400050	18648679*ab* ± 717226	18937764*ab* ± 507156	19594648*ab* ± 373256	18383082*b* ± 507156
Cinnamoyl-salicin	12097*ab* ± 987	11944*ab* ± 806	10358*b* ± 779	11808*ab* ± 1396	13966*a* ± 987	12987*ab* ± 727	14650*a* ± 987
HCH-Ac-salicortin	**30552*a*** ± 2279	15154*b* ± 1861	22385*ab* ± 1798	22058*ab* ± 3223	25912*a* ± 2279	23954*a* ± 1677	25120*a* ± 2279
HCH-salicortin	**2027209*a*** ± 117742	1941761*a* ± 96136	1590870*a* ± 92876	1819946*a* ± 166513	1814078*a* ± 117742	1891767*a* ± 86656	1690340*a* ± 117742
HCH-tremulacin	779420*ab* ± 54671	901451*a* ± 44639	702145*b* ± 43125	856067*ab* ± 77317	854546*ab* ± 54671	850986*ab* ± 40237	773174*ab* ± 54671
OH-tremuloidin	3083760*c* ± 22013	**379038*ab*** ± 17973	**404189*a*** ± 17364	295890*abc* ± 31131	295881*bc* ± 22013	312900*bc* ± 16201	285702*c* ± 22013
p-coumaric acid	72598*b* ± 5669	**117312*a*** ± 4628	69830*b* ± 4471	80959*b* ± 8017	83148*b* ± 5669	76046*b* ± 4172	81188*b* ± 5669
Salicoyl-tremuloidin	3393*ab* ± 536	3518*a* ± 438	1419*b* ± 423	2618*ab* ± 758	3566*a* ± 536	3503*a* ± 394	4019*a* ± 536
Salicortin	**31708303*** ± 1359824	27844354 ± 1110292	27399864 ± 1072644	26846382 ± 1923082	27934434 ± 1359824	30052206 ± 1000803	30601384 ± 1359824
Salicylic acid	1202017*a* ± 23487	1155301*ab* ± 19177	**1085672*b*** ± 18527	1179529*ab* ± 33216	1214922*a* ± 23487	1174446*a* ± 17286	1160759*ab* ± 23487
Arachidonic acid	**5687*b*** ± 252	6418*ab* ± 205	**5929*b*** ± 198	7235*a* ± 356	**6013*ab*** ± 252	6476*ab* ± 185	**6108*ab*** ± 252

Means ± SE and Tukey's test data (letters) are extracted from ***Supplementary Table S4****. Stars * and bold indicate means significantly different from WT (post-ANOVA contrast $P < 0.05$,* ***Supplementary Table S2****).*

Comparisons of the phenotypic effects of synthetic versus post-synthetic strategies (both using the *35S* promoter) revealed that some leaf-related traits were affected. Leaf weight, chlorophyll content, and chewing resistance appeared to be lower in the case of post-synthetic modification (**Figures 3**, **5**). Many SPGs were also affected by the deacetylation strategy (**Supplementary Table S7**). These foliar changes were not matched by stem growth during the first four years in the field, but in the final (fifth) year, both stem diameter and stem height were somewhat reduced by the post-synthetic modification strategy compared to the synthetic one (**Supplementary Table S7**). The lines available only allowed us to draw conclusions about the pre- vs. synthetic strategy in the case of ectopic modification using the *35S* promoter. It would be interesting to investigate whether the same conclusion applies to modification targeted specifically to developing wood.

Variability in Foliar Phenolics and Resistance Properties of the Transgenic Lines

Although the trees in this study were never exposed to an outbreak of severe attack by a particular herbivore or pathogen, the relatively low and variable relationships in the field between various kinds of biotic stressors and genotypes confirmed that the transgenic procedure in itself is unlikely to be associated with any systematic impact on surrounding organisms and *vice versa* (Strauss, 2003).

Leaf CTs of natural aspen populations are strongly tied to genotype (Lindroth et al., 2011; Robinson et al., 2012; Bandau et al., 2015; Decker et al., 2016), a feature that was also observed in our transformed genotypes. CTs are considered to be anti-oxidant phenolic polymers (Gourlay and Constable, 2019) that are expected to influence the presence and impact of plant-consuming microorganisms and herbivores (Mutikainen et al., 2000; Bailey et al., 2005; Barbehenn and Constabel, 2011), although they also express a high degree of plasticity in response to environmental factors such as nitrogen addition (Bandau et al., 2015) and they may be equally important and indicative of the extent of internal recovery and the mode of growth (Harding et al., 2013; Lindroth and Madritch, 2015; Decker et al., 2016). However, the two lines in this experiment that suffered from elevated chewing symptoms varied in tannin content, with *35S:HjAXE* giving lower and *35SAnAXE1* higher foliar CT concentrations when compared to WT (**Figures 5B,C**). No consistent relationship between CTs and chewing damage caused could therefore be deduced from this study, supporting the hypothesis that the potential defensive role of CTs in plants is indeed complex.

Salicinoid phenolic glucosides have often been investigated as suggested markers of innate resistance to herbivore damage in woody species (Mutikainen et al., 2000; Albrectsen et al., 2004; Philippe and Bohlmann, 2007; Witzell and Martin, 2008; Fabisch et al., 2019) and in particular as constitutive markers in aspen (Albrectsen et al., 2010; Robinson et al., 2012; Bernhardsson et al., 2013; Lindroth and St. Clair, 2013; Lindroth et al., 2015). The *35S*-driven transgenes appeared to be more affected with respect to their SPG profiles compared to the *WP*-driven transgenes, although the low number of lines per construct tested in the present experiment (between two and four) resulted in few significant changes in SPG contents among the constructs (**Table 2** and **Supplementary Tables S2, S4**). Despite this deficiency, the greater impact of the *35S* promoter compared to the *WP* promoter on SPG profiles was obvious. The synthesis of phenolic compounds belonging to the SPG group is still unresolved due to reticulate pathways with no apparent direct connection to the most simple salicinoid, salicin (Babst et al., 2010; Fellenberg et al., 2020), although it is increasingly accepted that the specialized metabolism of phenolic compounds is tightly linked to primary metabolism (Harding et al., 2013), and our study further suggests that cell wall acetylation may indeed alter, and determine levels of, phenolic compounds in aspen.

CE1 and CE5 AXEs Induce Distinct Foliar Phenotypes

The selection of enzymes appropriate for transgenic modification was addressed in this study by comparing two fungal enzymes, a CE1 representative, *An*AXE1 and a CE5 representative, *Hj*AXE, expressed from either *35S* or *WP* promoters, for their effects on several traits related to growth, biotic stress resistance and foliage characteristics. Stem growth and leaf weight were not affected by the enzyme used. In contrast, the occurrence of necrosis was associated with *HjAXE*, and there was a higher incidence of chewing with *AnAXE1*, although both transgenes induced more chewing than was seen in WT (**Figure 5** and **Supplementary Table S6**). CTs were more characteristic of *HjAXE* than *AnAXE1* expressing plants, and several SPGs accumulated differentially in the leaves of transgenic plants with the two transgenes (**Figure 5** and **Table 2**). These differences were primarily seen in the lines with *35S*-driven transgenes; effects were negligible in lines with the WP promoter. These data indicate that each transgene induced different susceptibilities to specific biotic stresses, associated with different patterns of accumulation of some stress-related SPGs and CTs. Previous greenhouse studies with *35S:AnAXE1* and *35S:HjAXE* expressing plants did not reveal any major morphological differences between plants with the two transgenes (Ratke et al., 2015; Pawar et al., 2017a), highlighting the importance of field testing.

The physiological background behind the contrasting phenotypes observed in *35SAnAXE1* and *35SHjAXE* expressing plants is not known, and it could encompass many factors. Beside the difference in enzymatic specificities and mode of action in the cell wall (discussed by Wang et al., 2020), the two proteins could be differentially perceived by the pathogen-associated molecular patterns (PAMP) recognition system (Bellincampi et al., 2014). The induction of necrosis by ectopically expressed *HjAXE* seen in our trial (**Figure 5**) is reminiscent of the effects of several fungal xylanases from family 11, including *Hypocrea jecorina* xylanase II, which induce ethylene and hypersensitive responses in plants, leading to necrosis (Noda et al., 2010). A conserved amino acid motif, TEIGSVTSDGS, has been identified as being involved in induction of necrosis. Amino

acid alignments of sequences used in the two constructs reveal that *Hj*AXE, but not *An*AXE1, includes a similar motif, 183-VGTCTTQG-190, and it would be interesting to test this for necrosis-inducing activity.

Imbalance in Cellular Acetyl Levels Could Lead to Genomic Instability

In this trial, three out of 18 transgenic lines with reduced acetylation exhibited dwarfism (**Figure 1B**) and growth abnormalities (**Figure 1C**), and one of them showed increased mortality (**Figure 1D**). These detrimental effects could not be associated with the transgenes introduced, and they are likely to have been caused by somaclonal variation or positional effects. Such a high (17%) incidence of dwarfism in acetylation-modified lines is, however, remarkable, and it was not predicted on the basis of the growth observed during greenhouse trials with the same transgenic lines. We also observed no dwarfism other than dwarfism related to transgenes among another set of 48 transgenic lines modified for other qualities that were grown in nearby transgenic fields. Reports from previous American field trials support the conclusion that somaclonal variations or detrimental positional effects are rare in transgenic poplars. For example, in a long-term trial with 948 lines engineered for sterility, not a single incident of detectable somaclonal variation was reported (Klocko et al., 2018); similarly, in a survey of field trial studies in United States covering a period of over 20 years and more than 100 transgenic poplar lines, only 0.1–1% dwarfism that could putatively be linked to positional effects or somaclonal variation was detected (Strauss et al., 2016).

The exceptionally high occurrence of random dwarfism and abnormalities among our transgenic lines with reduced acetyl content suggests a potential link between the acetylation status and genomic stability. Indeed, studies in other eucaryotes including mammals and yeasts showed that cellular levels of acetyl-CoA are directly associated with histone acetylation, which in turn regulates chromatin epigenetic state (Cai and Tu, 2011; Etchegaray and Mostoslavsky, 2016). In plants, epigenetic changes in chromatin state have been linked to the activation of transposable elements under stress conditions, thus contributing to somaclonal variation (Kaeppler et al., 2000). Moreover, in mammalian cells, the cellular ability to repair double strand breaks in DNA requires histone acetylation (Sivanand et al., 2017). These data support the hypothesis that the higher levels of acetyl-CoA expected to be induced by our engineering strategies could indeed lead to increased rates of mutation, especially when combined with stress. This hypothesis could be addressed by field testing and whole-genome sequencing of different acetyl-modified plants.

CONCLUDING REMARKS

The field growth of genetically modified plants is highly controversial in large parts of the World, but it is also warranted so that their potential can be carefully validated (Strauss, 2003; Viswanath et al., 2012; Strauss et al., 2016). On the one hand, genetic modifications promise to make it possible to tailor plants to perform better and produce higher quality products. On the other hand, genetic modification is a contentious topic among the public. The only way to assess the benefits and drawbacks of applying genetic modification techniques to crops, including forest trees, is to perform thorough characterization of transgenic lines. Here we present the results of the first field test of transgenic plants that target xylan-acetylation in cell walls, carried out to assess the consequences for growth, environmental stress resistance and biotic stress resistance in conditions similar to those used in short-rotation plantation forestry. Our data revealed novel plant phenotypes, not seen in the previous greenhouse experiments, as well as novel traits concerning the interaction of the genetically modified trees with their environment. The results highlight the need for early field testing in order to evaluate transgenic strategy and to assess the potential benefits and drawbacks expected when transgenic crops are used compared to their non-transgenic commercial counterparts.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MD-M, BA, and EM designed the research. MD-M, FA, ED, LM, UJ, BA, and EM carried out field the work and sample preparation. PP produced the transgenic lines. FA conducted the tannin analyses and prepared the leaves for metabolomics analyses. MD-M, ED, BA, and EM analyzed the data. BA and EM wrote the manuscript with contributions from all authors.

FUNDING

This work was supported by the Bio4Energy and the SSF program ValueTree RBP14-0011.

ACKNOWLEDGMENTS

We are grateful to Dr. Johanna Witzell, SLU, for help with conservation of field samples before transport to Umeå and the Swedish Metabolomic Centre for assistance with metabolomic analyses and to Erasmus student Kacper Dziewit for help with histochemical GUS analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00651/full#supplementary-material

REFERENCES

Abreu, I. N., Ahnlund, M., Moritz, T., and Albrectsen, B. R. (2011). UHPLC-ESI/TOFMS determination of salicylate-like phenolic glycosides in *Populus tremula* leaves. *J. Chem. Ecol.* 37, 857–870. doi: 10.1007/s10886-011-9991-7

Albrectsen, B. R., Gardfjell, H., Orians, C. M., Murray, B., and Fritz, R. S. (2004). Slugs, willow seedlings and nutrient fertilization: intrinsic vigor inversely affects palatability. *Oikos* 105, 268–278. doi: 10.1111/j.0030-1299.2004.12892.x

Albrectsen, B. R., Witzell, J., Robinson, K. M., Wulff, S., Luquez, V. M. C., Ågren, R., et al. (2010). Large scale geographic clines of parasite damage to *Populus tremula L. Ecography* 33, 483–493.

Babst, B. A., Harding, S. A., and Tsai, C.-J. (2010). Biosynthesis of phenolic glycosides from phenylpropanoid and benzenoid precursors in *Populus*. *J. Chem. Ecol.* 36, 286–297. doi: 10.1007/s10886-010-9757-7

Bacete, L., Melida, H., Miedes, E., and Molina, A. (2018). Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses. *Plant J.* 93, 614–636. doi: 10.1111/tpj.13807

Bailey, J. K., Deckert, R., Schweitzer, J. A., Rehill, B. J., Lindroth, R. L., Gehring, C. A., et al. (2005). Host plant genetics affect hidden ecological players: links among *Populus,* condensed tannins, and fungal endophyte infection. *Can. J. Bot.* 83, 356–361. doi: 10.1139/b05-008

Bandau, F., Decker, V. H., Gundale, M. J., and Albrectsen, B. R. (2015). Genotypic tannin levels in *Populus tremula* impact the way nitrogen enrichment affects growth and allocation responses for some traits and not for others. *PLoS ONE* 10:e0140971. doi: 10.1371/journal.pone.0140971

Barbehenn, R. V., and Constabel, P. C. (2011). Tannins in plant-herbivore interactions. *Phytochemistry* 72, 1551–1565. doi: 10.1016/j.phytochem.2011.01.040

Bar-On, Y. M., Phillips, R., and Milo, R. (2018). The biomass distribution on Earth. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6506–6511.

Bellincampi, D., Cervone, F., and Lionetti, V. (2014). Plant cell wall dynamics and wall-related susceptibility in plant–pathogen interactions. *Front. Plant Sci.* 5:228. doi: 10.3389/fpls.2014.00228

Bernhardsson, C., Robinson, K. M., Abreu, I. N., Jansson, S., Albrectsen, B. R., and Ingvarsson, P. K. (2013). Geographic structure in metabolome and herbivore community co-occurs with genetic structure in plant defence genes. *Ecol. Lett.* 16, 791–798. doi: 10.1111/ele.12114

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

Cai, L., and Tu, B. P. (2011). On acetyl-CoA as a gauge of cellular metabolic state. *Cold. Spring Harb. Symp. Quant. Biol.* 76, 195–202. doi: 10.1101/sqb.2011.76.010769

Chong, S.-L., Virkki, L., Maaheimo, H., Juvonen, M., Derba-Maceluch, M., Koutaniemi, S., et al. (2014). *O*-acetylation of glucuronoxylan in *Arabidopsis thaliana* wild type and its change in xylan biosynthesis mutants. *Glycobiology* 24, 494–506. doi: 10.1093/glycob/cwu017

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

Decker, V. H. G., Bandau, F., Gundale, M. J., Cole, C., and Albrectsen, B. R. (2016). Aspen phenylpropanoid genes' expression levels correlate with genets' tannin richness and vary both in responses to soil nitrogen and associations with phenolic profiles. *Tree Physiol.* 37, 270–289.

Dixon, R. A., and Paiva, N. L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell* 7, 1085–1097. doi: 10.1105/tpc.7.7.1085

Donev, E., Gandla, M. L., Jönsson, L. J., and Mellerowicz, E. J. (2018). Engineering non-cellulosic polysaccharides of wood for the biorefinery. *Front. Plant Sci.* 9:1537. doi: 10.3389/fpls.2018.01537

Etchegaray, J.-P., and Mostoslavsky, R. (2016). Interplay between metabolism and epigenetics: a nuclear adaptation to environmental changes. *Mol. Cell* 62, 695–711. doi: 10.1016/j.molcel.2016.05.029

Fabisch, T., Gershenzon, J., and Unsicher, S. (2019). Specificity of herbivore defense responses in a woody plant, black poplar (*Populus nigra*). *J. Chem. Ecol.* 45, 162–177. doi: 10.1007/s10886-019-01050-y

Fellenberg, C., Corea, O., Yan, L.-H., Archinuk, F., Piirtola, E.-M., Gordon, H., et al. (2020). Discovery of salicyl benzoate UDP-glycosyltransferase, a central enzyme in poplar salicinoid phenolic glycoside biosynthesis. *Plant J.* 102, 99–115. doi: 10.1111/tpj.14615

Gille, S., and Pauly, M. (2012). *O*-Acetylation of plant cell wall polysaccharides. *Front. Plant Sci.* 3:12. doi: 10.3389/fpls.2012.00012

Giummarella, N., and Lawoko, M. (2016). Structural basis for the formation and regulation of lignin–xylan bonds in birch. *ACS Sustain. Chem. Eng.* 4, 5319–5326. doi: 10.1021/acssuschemeng.6b00911

Gourlay, G., and Constable, C. P. (2019). Condensed tannins are inducible antioxidants and protect hybrid poplar against oxidative stress. *Tree Physiol.* 39, 345–355. doi: 10.1093/treephys/tpy143

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

Groenewald, E. G., and van der Westhuizen, A. J. (1997). Prostaglandins and Related Substances in Plants. *Bot. Rev.* 63, 199–220. doi: 10.1007/bf02857948

Gröndahl, M., Teleman, A., and Gatenholm, P. (2003). Effect of acetylation on the material properties of glucuronoxylan from aspen wood. *Carbohydr. Polym.* 52, 359–366. doi: 10.1016/s0144-8617(03)00014-6

Harding, S. A., Xue, L.-J., Du, L., Nyamdari, B., Lindroth, R. L., Sykes, R., et al. (2013). Condensed tannin biosynthesis and polymerization synergistically condition carbon use, defense, sink strength and growth in *Populus*. *Tree Physiol.* 34, 1240–1251. doi: 10.1093/treephys/tpt097

Jönsson, L. J., and Martín, C. (2016). Pretreatment of lignocellulose: formation of inhibitory by-products and strategies for minimizing their effects. *Bioresour. Technol.* 199, 103–112. doi: 10.1016/j.biortech.2015.10.009

Kaeppler, S. M., Kaeppler, H. F., and Rhee, Y. (2000). Epigenetic aspects of somaclonal variation in plants. *Plant Mol. Biol.* 43, 179–188.

Keefover-Ring, K., Ahnlund, M., Abreu, I. N., Jansson, S., Moritz, T., and Albrectsen, B. R. (2014). No evidence of geographical structure of salicinoid chemotypes within *Populus tremula*. *PLoS ONE* 9:e107189. doi: 10.1371/journal.pone.0107189

Klocko, A. L., Lu, H., Magnuson, A., Brunner, A. M., Ma, C., and Strauss, S. H. (2018). Phenotypic expression and stability in a large-scale field study of genetically engineered poplars containing sexual containment transgenes. *Front. Bioeng. Biotechnol.* 6:100. doi: 10.3389/fbioe.2018.00100

Kloth, K. J., Abreu, I. N., Delhomme, N., Petøík, I., Villard, C., Ström, C., et al. (2019). PECTIN ACETYLESTERASE9 affects the transcriptome and metabolome and delays aphid feeding. *Plant Physiol.* 181, 1704–1720. doi: 10.1104/pp.19.00635

Koutaniemi, S., van Gool, M. P., Juvonen, M., Jokela, J., Hinz, S. W., Schols, H. A., et al. (2013). Distinct roles of carbohydrate esterase family CE16 acetyl esterases and polymer-acting acetyl xylan esterases in xylan deacetylation. *J. Biotechnol.* 168, 684–692. doi: 10.1016/j.jbiotec.2013.10.009

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

Lindroth, R. L., and Madritch, M. D. (2015). Condensed tannins increase nitrogen recovery by trees following insect defoliation. *New Phytol.* 208, 410–420. doi: 10.1111/nph.13444

Lindroth, R. L., Rubert-Nason, K. F., Couture, J. J., Major, I. T., and Constabel, C. P. (2015). Influence of genotype, environment, and gypsy moth herbivory on local and systemic chemical defenses in trembling aspen (*Populus tremuloides*). *J. Chem. Ecol.* 41, 651–661. doi: 10.1007/s10886-015-0600-z

Lindroth, R. L., Scioneaux, A. N., Schmidt, M. A., Moore, M. A., Wooley, S. C., and Hagerman, A. E. (2011). Qualitative variation in proanthocyanidin composition of *Populus* species and hybrids: genetics is the key. *J. Chem. Ecol.* 37, 57–70. doi: 10.1007/s10886-010-9887-y

Lindroth, R. L., and St. Clair, S. B. (2013). Adaptations of quaking aspen (*Populus tremuloides* Michx.) for defense against herbivores. *For. Ecol. Manage.* 299, 14–21. doi: 10.1016/j.foreco.2012.11.018

Luquez, V., Hall, D., Albrectsen, B. R., Karlsson, J., Ingvarsson, P., and Jansson, S. (2008). Natural phenological variation in aspen (*Populus tremula*): the

SwAsp collection. *Tree Genet. Genomes* 4, 279–292. doi: 10.1007/s11295-007-0108-y

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

Mutikainen, P., Walls, M., Ovaska, J., Keina, M., Julkunen-Tiitto, R., and Vapaavuori, E. (2000). Herbivore resistance in *Betula pendula*: effect of fertilization, defoliation, and plant genotype. *Ecology* 81, 49–65. doi: 10.1890/0012-9658(2000)081%5B0049:hribpe%5D2.0.co;2

Nilsson, O., Aldén, T., Sitbon, F., Little, C. H. A., Chalupa, V., Sandberg, G., et al. (1992). Spatial pattern of cauliflower mosaic virus *35S* promoter-luciferase expression in transgenic hybrid aspen trees monitored by enzymatic assay and non-destructive imaging. *Transgen. Res.* 1, 209–220. doi: 10.1007/bf02524751

Nilsson, U., and Örlander, G. (1999). Vegetation management on grass-dominated clearcuts planted with Norway spruce in southern Sweden. *Can. J. For. Res.* 29, 1015–1026. doi: 10.1139/x99-071

Noda, J., Brito, N., and Gonzalez, C. (2010). The *Botrytis cinerea* xylanase Xyn11A contributes to virulence with its necrotizing activity, not with its catalytic activity. *BMC Plant Biol.* 10:38. doi: 10.1186/1471-2229-10-38

Papazian, S., Girdwood, T., Wessels, B. A., Poelman, E. H., Dicke, M., Moritz, T., et al. (2019). Leaf metabolic signatures induced by real and simulated herbivory in black mustard (*Brassica nigra*). *Metabolomics* 15:130.

Pauly, M., and Ramírez, V. (2018). New insights into wall polysaccharide O-acetylation. *Front. Plant Sci.* 9:1210. doi: 10.3389/fpls.2018.01210

Pawar, P. M.-A., Ratke, C., Balasubramanian, V. B., Chong, S. L., Gandla, M. L., Adriasola, M., et al. (2017b). Downregulation of RWA genes in hybrid aspen affects xylan acetylation and wood processing properties. *New Phytol.* 214, 1491–1505. doi: 10.1111/nph.14489

Pawar, P. M.-A., Derba-Maceluch, M., Chong, S.-L., Gómez, L. D., Miedes, E., Banasiak, A., et al. (2016). Expression of fungal acetyl xylan esterase in Arabidopsis thaliana improves saccharification of stem lignocellulose. *Plant Biotechnol. J.* 14, 387–397. doi: 10.1111/pbi.12393

Pawar, P. M.-A., 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.-A., Derba-Maceluch, M., Chong, S. L., Gandla, M. L., Bashar, S. S., Sparrman, T., et al. (2017a). *In muro* deacetylation of xylan affects lignin properties and improves saccharification of aspen wood. *Biotechnol. Biofuels* 10:98. doi: 10.1186/s13068-017-0782-4

Philippe, R. N., and Bohlmann, J. (2007). Poplar defense against insect herbivores. *Can. J. Bot.* 85, 1111–1126.

Pogorelko, G., Fursova, O., Lin, M., Pyle, E., Jass, J., and Zabotina, O. A. (2011). Post-synthetic modification of plant cell walls by expression of microbial hydrolases in the apoplast. *Plant Mol. Biol.* 77, 433–435.

Pogorelko, G., Lionetti, V., Fursova, O., Sundaram, R. M., Qi, M., Whitham, S. A., et al. (2013). Arabidopsis and *Brachypodium distachyon* transgenic plants expressing *Aspergillus nidulans* acetylesterases have decreased degree of polysaccharide acetylation and increased resistance to pathogens. *Plant Physiol.* 162, 9–23. doi: 10.1104/pp.113.214460

Porter, L. J., Hrstich, L. N., and Chan, B. G. (1986). The conversion of procyanidins and prodelphinidins to cyanidin and delphinidin. *Phytochemistry* 25, 223–230. doi: 10.1016/s0031-9422(00)94533-3

Ratke, C., Pawar, P. M., Balasubramanian, V. K., Naumann, M., Duncranz, M. L., Derba-Maceluch, M., et al. (2015). Populus GT43 family members group into distinct sets required for primary and secondary wall xylan biosynthesis and include useful promoters for wood modification. *Plant Biotechnol. J.* 13, 26–37. doi: 10.1111/pbi.12232

Robinson, K. M., Ingvarsson, P. K., Jansson, S., and Albrectsen, B. R. (2012). Genetic variation in functional traits influences arthropod community composition in aspen (*Populus tremula* L.). *PLoS ONE* 7:e37679. doi: 10.1371/journal.pone.0037679

Savchenko, T., Walley, J. W., Chehab, E. W., Xiao, Y., Kaspi, R., Pye, M. F., et al. (2010). Arachidonic acid: an evolutionarily conserved signaling molecule modulates plant stress signaling networks. *Plant Cell* 22, 3193–3205. doi: 10.1105/tpc.110.073858

Sivanand, S., Rhoades, S., Jiang, Q., Lee, J. V., Benci, J., Zhang, J., et al. (2017). Nuclear acetyl-CoA production by ACLY promotes homologous recombination. *Mol. Cell* 67, 252–265.

Strauss, S. H. (2003). Genomics, genetic engineering, and domestication of crops. *Science* 300, 61–62. doi: 10.1126/science.1079514

Strauss, S. H., Ma, C., Ault, K., and Klocko, A. L. (2016). "Lessons from two decades of field trials with genetically modified trees in the USA: biology and regulatory compliance," in *Biosafety of Forest Transgenic Trees*, eds C. Vettori, F. Gallardo, H. Häggman, V. Kazana, F. Migliacci, G. Pilate, et al. (Dordrecht: Springer), 101–124. doi: 10.1007/978-94-017-7531-1_5

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

Viswanath, V., Albrectsen, B. R., and Strauss, S. H. (2012). Global regulatory burden for field testing of genetically modified trees. *Tree Genet. Genomes* 8, 221–226. doi: 10.1007/s11295-011-0445-8

Vogel, J. P., Raab, T. K., Somerville, C. R., and Somerville, S. C. (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

Wang, Z., Pawar, P. M.-A., Derba-Maceluch, M., Hedenström, M., Chong, S.-L., Tenkanen, M., et al. (2020). Hybrid aspen expressing a carbohydrate esterase family 5 acetyl xylan esterase under control of a wood-specific promoter shows improved saccharification. *Front. Plant Sci.* 11:380. doi: 10.3389/fpls.2020.00380

Witzell, J., and Martin, J. A. (2008). Phenolic metabolites in the resistance of northern forest trees to pathogens — past experiences and future prospects. *Can. J. For. Res.* 38, 2711–2727. doi: 10.1139/x08-112

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., 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

Xu, F., Liu, Z., Xie, H., Zhu, J., Zhang, J., Kraus, J., et al. (2014). Increased drought tolerance through the suppression of *ESKMO1* gene and overexpression of CBF-related genes in *Arabidopsis*. *PLoS ONE* 9:e106509. doi: 10.1371/journal.pone.0106509

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 28 May 2020
doi: 10.3389/fpls.2020.00620

Insights of Molecular Mechanism of Xylem Development in Five Black Poplar Cultivars

Lei Zhang[1], Bobin Liu[2], Jin Zhang[3] and Jianjun Hu[1]**

[1] State Key Laboratory of Tree Genetics and Breeding, Key Laboratory of Tree Breeding and Cultivation of National Forestry and Grassland Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China, [2] College of Forestry, Fujian Agriculture and Forestry University, Fuzhou, China, [3] Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States

Black poplar (*Populus deltoides*, *P. nigra*, and their hybrids) is the main poplar cultivars in China. It offers interesting options of large-scale biomass production for bioenergy due to its rapid growth and high yield. Poplar wood properties were associated with chemical components and physical structures during wood formation. In this study, five poplar cultivars, *P. euramericana* 'Zhonglin46' (Pe1), *P. euramericana* 'Guariento' (Pe2), *P. nigra* 'N179' (Pn1), *P. deltoides* 'Danhong' (Pd1), and *P. deltoides* 'Nanyang' (Pd2), were used to explore the molecular mechanism of xylem development. We analyzed the structural differences of developing xylem in the five cultivars and profiled the transcriptome-wide gene expression patterns through RNA sequencing. The cross sections of the developing xylem showed that the cell wall thickness of developed fiber in Pd1 was thickest and the number of xylem vessels of Pn1 was the least. A total of 10,331 differentially expressed genes were identified among 10 pairwise comparisons of the five cultivars, most of them were related to programmed cell death and secondary cell wall thickening. *K*-means cluster analysis and Gene Ontology enrichment analysis showed that the genes highly expressed in Pd1 were related to nucleotide decomposition, metabolic process, transferase, and microtubule cytoskeleton; whereas the genes highly expressed in Pn1 were involved in cell wall macromolecule decomposition and polysaccharide binding processes. Based on a weighted gene co-expression network analysis, a large number of candidate regulators for xylem development were identified. And their potential regulatory roles to cell wall biosynthesis genes were validated by a transient overexpression system. This study provides a set of promising candidate regulators for genetic engineering to improve feedstock and enhance biofuel conversion in the bioenergy crop *Populus*.

Keywords: *Populus*, developing xylem, transcriptome, cell wall, transcriptional regulation

OPEN ACCESS

Edited by:
Agnieszka Ludwików, Adam Mickiewicz University, Poland

Reviewed by:
Guohua Chai, College of Resources and Environment, Qingdao Agricultural University, China
Naoki Takata, Forestry and Forest Products Research Institute, Japan
Zanmin Hu, Chinese Academy of Sciences, China

****Correspondence:***
Jin Zhang
zhangj1@ornl.gov
Jianjun Hu
hujj@caf.ac.cn

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *14 February 2020*
Accepted: *22 April 2020*
Published: *28 May 2020*

Citation:
Zhang L, Liu B, Zhang J and Hu J (2020) Insights of Molecular Mechanism of Xylem Development in Five Black Poplar Cultivars. Front. Plant Sci. 11:620. doi: 10.3389/fpls.2020.00620

INTRODUCTION

Energy issue is one of the major concerns of this century. As an important biomass energy, wood is expected to increase with the development of social economy. Biomass production as energy raw material accounts for about 14% of the world's primary energy sources (Parikka, 2004; El Kasmioui and Ceulemans, 2012). Poplar is used as short-rotation coppice (SRC) tree and main raw materials

of bioenergy because of its fast-growing, large biomass, and lower requirements for cultivation (Willebrand and Verwijst, 1993; Davis, 2008; Zhang et al., 2020). The biomass conversion rate of poplar wood is higher than that of other tree species due to its less fermentation inhibitory extract (Guerra et al., 2013).

Wood, the secondary xylem of trees, is mainly composed of cellulose, hemicellulose, and lignin. All xylem cell types first undergo secondary cell wall (SCW) thickening and experienced programmed cell death (PCD) in xylem (Courtois-Moreau et al., 2009; Zhong and Ye, 2015). While lignin content determines whether wood is used for pulp and the conversion efficiency of liquid biofuels (Wang et al., 2018). Lignin is a major phenolic polymer which is composed of 4-coumaryl alcohol (H-subunit), coniferyl alcohol (G-subunit), and sinapyl alcohol (S-subunit) (Freudenberg, 1965). Phenylalanine finally forms three monomers through the catalytic reactions of 10 enzyme families, including *PAL* (phenylalanine ammonia-lyase), *C4H* (cinnamate-4-hydroxylase), *4CL* (4-coumarate:CoA ligase), *HCT* (p-hydroxycinnamoyltransferase), *C3H* (4-coumarate 3-hydroxylase), *CCoAOMT* (caffeoyl-CoA *O*-methyltransferase), *CCR* (cinnamoyl-CoA reductase), *CAld5H* (coniferyl aldehyde 5-hydroxylase), *COMT* (caffeic acid/5-hydroxyconiferaldehyde *O*-methyltransferase), and *CAD* (cinnamyl alcohol dehydrogenase) (Freudenberg, 1965; Shi et al., 2009). And *LAC* (laccase) was involved in oxidative polymerization of lignin precursors and thus affected the process of vessel element and fiber lignification (Zhao et al., 2013). The changes of their expression can affect lignin content (Wagner et al., 2011; Lin et al., 2015; Wang et al., 2018).

The population genetic methods identified naturally occurring genetic variation for wood formation. Single-nucleotide polymorphism (SNP)-based association mapping, including quantitative trait locus (QTL) and genome-wide association studies (GWAS), has been used to identify SNPs related to wood properties in specific wood formation biosynthesis pathways in trees (Guerra et al., 2013; Zhang et al., 2018b), but only some of these associations were affiliated with genes that have *a priori* involvement in wood formation (Takata and Taniguchi, 2014; Zinkgraf et al., 2017). Transcriptional regulation is a primary mechanism that firstly responds to the environment and ultimately emits developmental signals during wood formation (Du and Groover, 2010; Zinkgraf et al., 2017; Zhang et al., 2018a). Transcriptomics has been widely used to compare and recognize specific regulatory networks in xylem development. It provides massive data for co-expression analysis, which can be used for potential gene mining and identify similar biological pathways or subject to similar regulatory pathways (D'haeseleer et al., 2000; Usadel et al., 2009). For example, Chano et al. (2017) analyzed the transcriptional profiles during the growing season in *Pinus canariensis*. Sundell et al. (2017) firstly established a high-spatial resolution transcriptome profile and revealed a gene expression module of wood formation in *P. tremula*. Subsequently, Seyfferth et al. (2018) distinguished the expression networks of ethylene-related genes in wood formation using this database.

Black poplar is widely used as the woody sources of fiber for the pulp, paper industry, biofuel production, and ecological shelter forest species in China. *P. euramericana* 'Zhonglin46', *P. euramericana* 'Guariento', *Populus nigra* 'N179', *Populus deltoides* 'Danhong', and *P. deltoides* 'Nanyang', are important poplar cultivars in China, and there were differences in growth and wood properties (Song et al., 2010; Yang et al., 2011; Hu et al., 2013; Zhang et al., 2020). They can represent *P. nigra*, *P. deltoides*, and their hybrids (*P. euramericana*), respectively. Wood formation mainly comes from the development of secondary xylem, which mainly refers to the deposition of lignin and thickening on the cell wall of xylem fibers and vessels (Zhang et al., 2014; Xu et al., 2017). It is great significance to explore the mechanism of cell wall formation for the study of wood formation. To obtain insights of molecular mechanism of xylem development in the five black poplar cultivars, we examined gene expression profiles of xylem and identified a large number of candidate regulators for xylem development. Three novel MYB transcription factors were identified and proved to be involved in the regulation of lignin biosynthesis. It provides new strategies and important resources for the exploration of xylem development of novel regulatory genes.

MATERIALS AND METHODS

Plant Materials

In this study, five black poplar cultivars, *P. euramericana* 'Zhonglin46' (Pe1, ♀), *P. euramericana* 'Guariento' (Pe2, ♀), *P. nigra* 'N179' (Pn1, ♂), *P. deltoides* 'Danhong' (Pd1, ♀), and *P. deltoides* 'Nanyang' (Pd2, ♂) were used as the plant materials. Poplar trees are grown in Jiaozuo, Henan Province, China (35°14′21″N, 113°18′40″E). The stem sample was collected from breast height of the stem in an area devoid of damage. The stems were debarked in 10 cm × 20 cm region. Then, the current year's xylem (1–2 mm) was scraped from 9-year-old trees using a sharp double-edged razor blade prior to August 2018. All the 20 samples (5 cultivars × 4 biological replicates) used for RNA sequencing (RNA-Seq) were immediately flash frozen in liquid nitrogen and then kept at −80°C until use. Then, the stem pieces, including bark, phloem, cambium, and xylem, were collected in the adjacent position by knife and reserved in formaldehyde-acetic acid-ethanol fixative (FAA) for anatomical observation.

Light Microscopy

The stem pieces were dehydrated in a graded ethanol series and embedded in steps of 25, 50, and 75% Spurr resin and finally in 100% a full day and polymerized overnight at 60°C as described by Samuels et al. (2002). Cross section of 4-μm thick was obtained from stem by Leica M205FA. Sections were stained by 0.05% toluidine blue O (TBO) and then washed with distilled water. Finally, all the sections were examined with microscope (Zeiss). The number and diameter of vessels in each sample were measured in the same area (860 μm × 940 μm). And we measured the wall thickness of developed fibers 12–20 layers away from the cambium. All data were measured using ImageJ software.

Illumina Sequencing and Mapping

Total RNA was isolated using the RNAprep Pure Plant Plus Kit (TIANGEN, China). Three micrograms of high-quality RNA per sample was used for the sequencing libraries preparation using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, United States). Then, 150 bp paired-end reads were generated on an Illumina Hiseq platform. We first cleaned the raw sequences and mapped to reference genome *P. trichocarpa* v3.0[1] using TopHat v2.0.12 (Trapnell et al., 2009). Gene expression was estimated as transcripts per million (TPM) (Li and Dewey, 2011). Sequencing data are available in NCBI SRA database (SRA number: SRP234303).

Differential Expression Genes and Functional Analysis

To identify the differential expression genes (DEGs) between the five black poplar cultivars, we performed pair-wise comparisons (Pe1 vs Pe2, Pe1 vs Pn1, Pe1 vs Pd1, Pe1 vs Pd2, Pe2 vs Pn1, Pe2 vs Pd1, Pe2 vs Pd2, Pn1 vs Pd1, Pn1 vs Pd2, and Pd1 vs Pd2) by DESeq2 R package. The parameters used to "call a gene" between conditions was determined at a false discovery rate (FDR)-adjusted P-value < 0.05. We computed gene expression based on the obtained clean reads using TPM values. Gene Ontology (GO) enrichment was performed based on FDR-adjusted P-value < 0.05. Principal component analysis (PCA) was performed using R package.

Clustering Analysis and Co-expression Network Construction

K-means clustering of the transcript expression patterns was performed using $\log_2$-transformed TPM in R package. Weighted gene co-expression network analysis (WGCNA) was performed according to Langfelder and Horvath (2008). The resulting network was visualized by Cytoscape 3.7.0 (Shannon et al., 2003).

Transient Expression Assay

The coding sequence (CDS) of three novel transcription factors, *PdMYB55*, *PdMYB74*, and *PdMYB160*, were amplified from Pd1 by special primers (**Supplementary Table S1**). Thermal cycler program was as follows: 95°C for 5 min followed by 35 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 50 s and a final at 72°C for 5 min. To analyze *PdMYB55*, *PdMYB74*, and *PdMYB160* transcriptional activity in yeast, the amplification products were cloned into pGBKT7 vector and transformed into the yeast strain Y2HGold containing *His3* reporter gene regulated by Gal4-responsive promoter (Liu et al., 2018). And the full-length *MYBs* were inserted in pCAMBIA2300-35S-OCS at *Acc*65I and *Sal*I sites. The empty vector was used as control. The recombinant expression vectors were introduced into *Nicotiana tabacum* leaf by transient *Agrobacterium*-mediated transformation method (Buschmann et al., 2011). After 3 days of infiltration, total RNA was extracted from infiltrated leaf region for the quantitative Real-Time PCR (qRT-PCR).

[1]https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Ptrichocarpa

The amino acid sequences of MYB and NAC transcription factors were obtained by BLAST searchers[2] Amino acid sequences were aligned and phylogenic analysis was performed by MEGA6.0 with the neighbor-joining method (Tamura et al., 2013).

Quantitative Real-Time PCR

qRT-PCR was used to verify the reliability of the RNA-Seq analyses and new transcription factor (TF) roles in SCW regulation. Eight DEGs were selected for qRT-PCR analysis. Primer pairs were designed using Primer 3[3]. qRT-PCR was performed using a TB Green® Premix Ex Taq™ II qPCR master mix (TaKaRa, Dalian, China) according to the manufacturer's instructions. *PtrActin* and *NtActin* were the internal controls of poplar and tobacco, respectively. The relative gene expression was calculated by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). All experiments were performed by using three biological replicates and three technical replicates. All the primers used in this study were listed in **Supplementary Table S1**.

RESULTS

The Microstructure of Five Cultivar Stems

To find the differences of main cultivars, two *P. deltoides* (Pd1 and Pd2), one *P. nigra* (Pn1), and two hybrids (Pe1 and Pe2) were selected for analysis. Firstly, we compared the longitudinal and latitudinal growths of the five poplar cultivars from 9-year-old trees. As is shown in **Figure 1A**, the diameter of the two hybrids was bigger than *P. deltoides* and *P. nigra*, but no significant changes were observed in height among the five poplar cultivars (**Figure 1B**).

The sections of five cultivars showed that the cambial region was composed of six to eight layers cell (**Figure 1C**). The number of cell layer and thickness of cambium in Pe1 and Pe2 are significantly greater than Pn1 (**Figures 1D,E**). In xylem, there were differences in cell wall thickness between developed fibers of different varieties of xylem, among which Pd1 cell wall was the thickest (**Supplementary Figure S1C**). The number of vessels in the same area showed a significant difference, with Pn1 having the lowest number but the largest size of vessels (**Figure 1F** and **Supplementary Figure S1**). This indicates that the breast diameter is associated with cambium and xylem secondary growth.

Transcriptome Sequencing and Alignment to the Reference Genome

To reveal the potential molecular mechanisms of cell wall thickening in developing xylem of the five poplar cultivars, the developing xylem was used for high-throughput RNA-Seq. A total of 10.03 billion high-quality reads were generated, of which 79.81% were successfully mapped to the *P. trichocarpa*

[2]http://www.phytozome.com

[3]http://primer3.ut.ee/

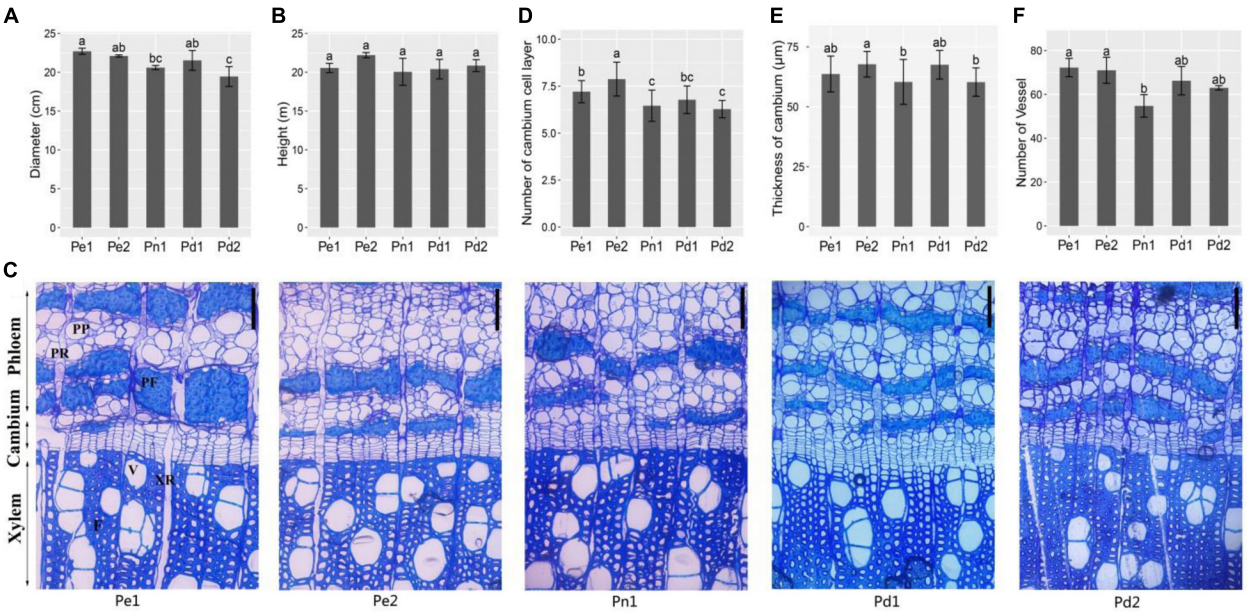

FIGURE 1 | Growth traits and microscopic analysis of stem from five black poplar cultivars. The diameter **(A)** and height **(B)** of the five cultivars from 9-year-old plants. **(C)** Cross sections of phloem–xylem region from 9-year-old trees. Scale bars = 100 μm. PP, phloem parenchyma; PR, phloem ray; PF, phloem fiber; V, xylem vessel; XR, xylem ray; F, xylem fiber. Statistical analysis of the number of cambium cell layer **(D)**, thickness of cambium **(E)**, and number of vessel cells in the same region (860 μm × 940 μm, **Supplementary Figure S1**) **(F)**. Means ± SD from four biological replicates. Lowercase letters (a, b, and c) indicate the results of Duncan's multiple range test (significant differences at $P < 0.05$). Pe1, *P. euramericana* 'Zhonglin46'; Pe2, *P. euramericana* 'Guariento'; Pn1, *P. nigra* 'N179'; Pd1, *P. deltoides* 'Danhong'; Pd2, *P. deltoides* 'Nanyang'.

reference genome, constituting 151 Gb of cDNA sequences. The GC content was 43.31%, and the Q30 was 93.33% (**Supplementary Table S2**).

A principal component analysis (PCA) plot of the whole data set revealed a sequential order of the different samples. The results showed that the five cultivars were divided into three clusters, and the biological replicates were projected closely. Two *P. deltoides* Pd1 and Pd2 were clustered together, hybrids Pe1 was clustered close to Pe2 in the middle of *P. deltoides* and *P. nigra*, highlighting the genetic relationship of five cultivars (**Figure 2A**).

Analysis and Functional Annotation of Differentially Expressed Genes

To identify the global transcriptional changes in varieties, we performed a pair-wise comparison with 10 comparable groups. In total, 10,331 differentially expressed genes (DEGs) were identified (**Figure 2B**). The largest DEG set was identified in comparison "Pn1 vs Pd1" (a total of 5,034 DEGs, including 2,101 upregulated and 2,933 downregulated genes), suggesting the difference of *P. nigra* and *P. deltoides* 'Danhong'. In contrast, the smallest DEG set was identified in comparison "Pd1 vs Pd2" (a total of 339 DEGs, including 81 upregulated and 258 downregulated genes) (**Figure 2C**).

To further characterize the biological role of DEGs, GO enrichment analysis was performed. The significant GO terms of the DEGs were classified into three major categories: 100 terms of biological process (BP), 65 terms of molecular function (MF), and seven terms of cellular component (CC) (**Figure 2D** and **Supplementary Table S3**). The most enriched terms were cell death, secondary cell wall, and lignin biosynthesis. In the BP category, subcategories of "apoptosis" (GO:0006915), "cell death" (GO:0008219), and "programmed cell death" (GO:0012501) were significantly enriched. The DEGs of "Pe1 vs Pd2" were enriched in "response to abiotic stimulus process" (GO:0009628), suggesting that there are differences in abiotic stress between the two cultivars. The GO terms "carbohydrate metabolic process" (GO:0005975) and "metabolic process" (GO:0008152) were specifically enriched in "Pd1 vs Pd2" DEGs. In the MF category, terms "ADP binding" (GO:0043531) and "receptor activity" (GO:0004872) were significantly enriched. The GO terms "nucleotide binding" (GO:0000166), "receptor activity" (GO:0004872), and "signal transducer activity" (GO:0004871) were mainly enriched in comparisons Pn1 with four other varieties. In the CC category, DEGs of "Pe1 vs Pd1" primarily belonged to "microtubule" (GO:0005874) and "cytoskeleton" (GO:0015630). "Pe2 vs Pd1" was significantly enriched in GO term "membrane" (GO:0016020).

K-Means Cluster of Five Cultivars

To further explore the functional diversity of DEGs from the five poplar cultivars, we performed a *K*-means clustering analysis and grouped the 10,331 DEGs into 20 clusters (**Figure 3**, **Supplementary Data Sheet 1**, and **Supplementary Table S4**). Three clusters (1, 8, and 15) showed a high expression level in Pd1. Genes in cluster 1 were mainly involved in "catabolic process", "metabolic process", "catalytic activity", and

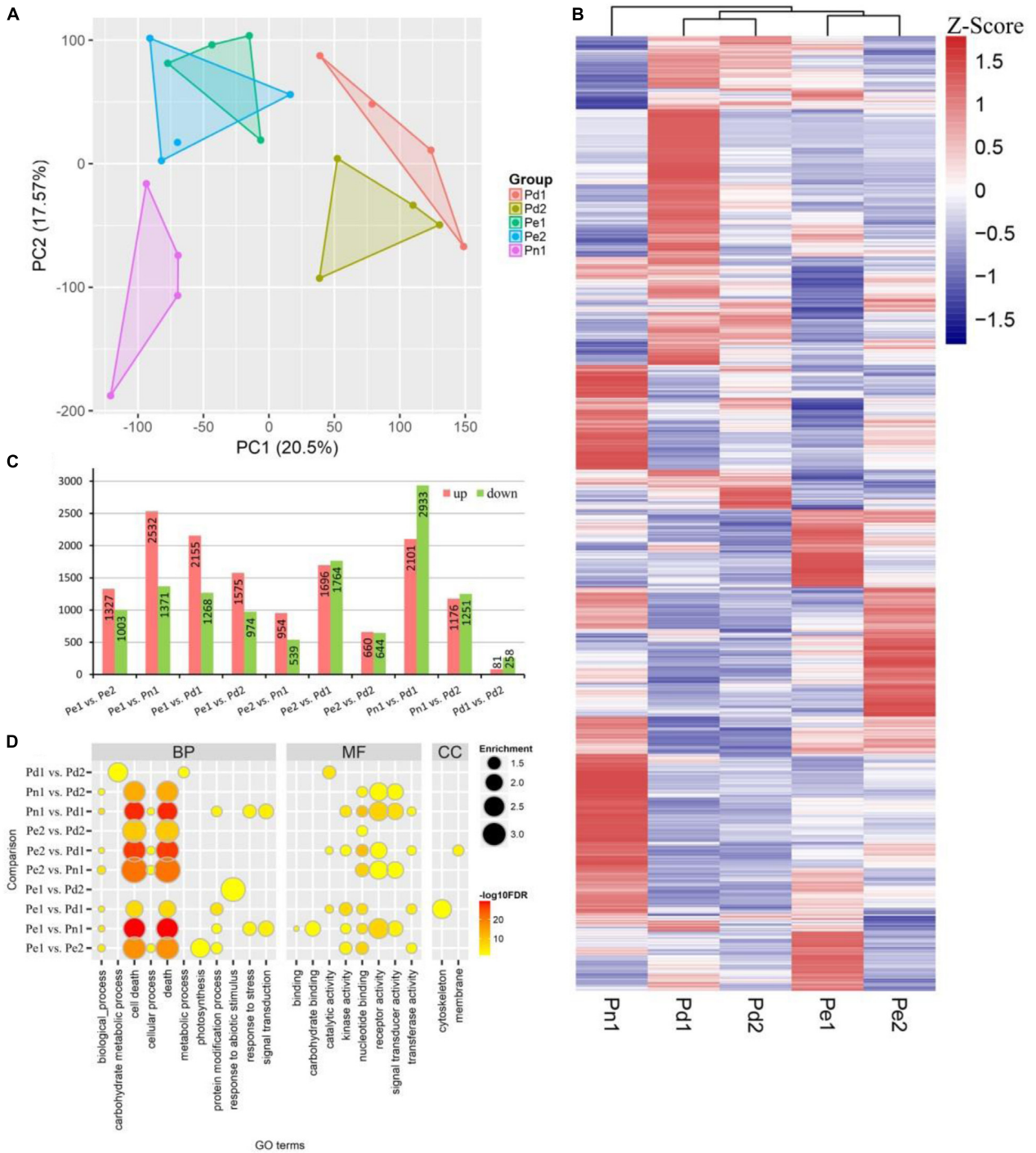

FIGURE 2 | Differentially expressed genes (DEGs) in xylem of the five poplar cultivars. **(A)** Principal component analysis (PCA) of the expressed genes showing sample separation. Principal component 1 (PC1) and PC2 explaining 20.5 and 17.57% of the total variance, respectively. **(B)** Heat map of transcriptome data for the DEGs in the five poplar cultivars. **(C)** The number of upregulated and downregulated genes in pair-wise comparisons between the five cultivars. **(D)** Gene Ontology (GO) enrichment analysis of genes in different comparisons. Node color represents -$\log_{10}$ transformed false discovery rate (FDR) corrected *P*-value. Node size represents rich factor. Full list of enriched GO terms was shown in **Supplementary Table S3**.

"cytoskeleton"; genes in clusters 8 and 15 were involved in "carbohydrate metabolic process", "membrane processes", and "catalytic activity". In addition, GO terms of "cellular amino acid and derivative metabolic process", "motor activity", and "cytoskeleton" were enriched in cluster 8. Genes in four clusters (4, 9, 10, and 14) were highly expressed in Pn1, and genes in

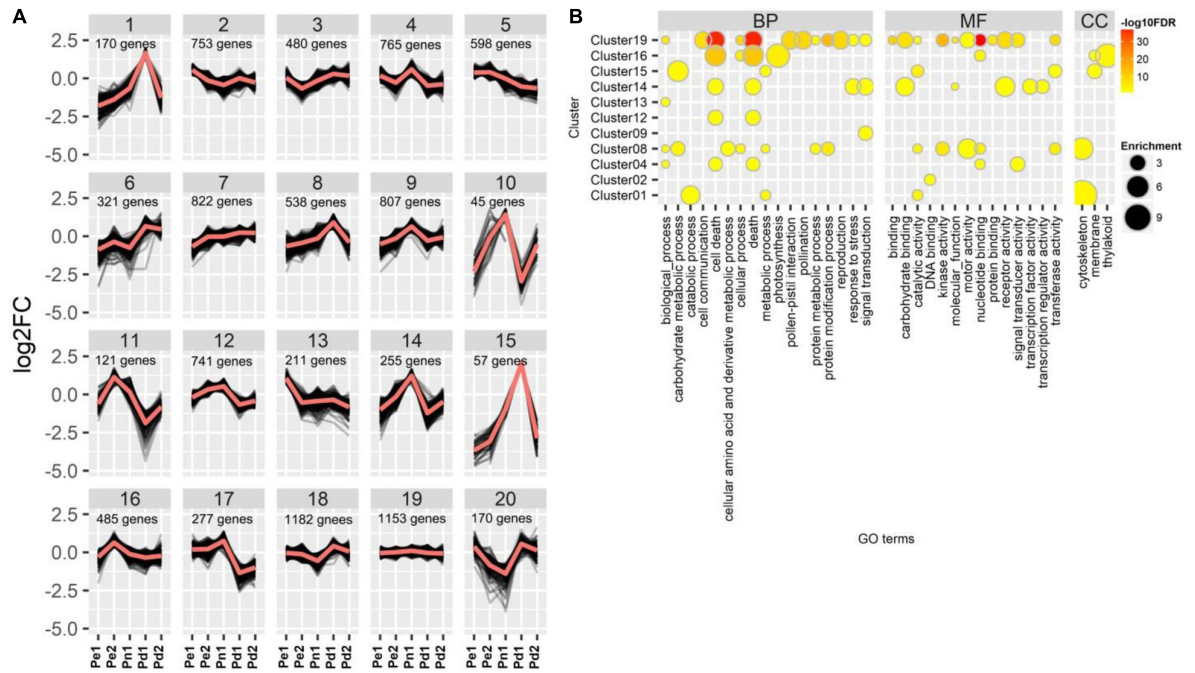

FIGURE 3 | *K*-means clustering and Gene Ontology (GO) classification of differentially expressed genes (DEGs). **(A)** Gene expression profiling of clusters, showing $\log_2$ of fold change variations among the five poplar cultivars. **(B)** GO enrichment analysis of genes in different clusters. Node color represents $-\log_{10}$ transformed false discovery rate (FDR) corrected *P*-value. Node size represents rich factor. Full list of enriched GO terms was shown in **Supplementary Table S4**.

clusters 10 and 14 showed low expression level in both Pe1 and Pd1. Genes in clusters 4 and 14 were both involved in "cell death" and "death" processes, whereas genes in cluster 14 were involved in "response to stress", "signal transduction", "carbohydrate binding", and "receptor activity". Cluster 16 which was enriched in "cell death", "photosynthesis", and "thylakoid" was highly expressed in Pe2. DEGs of cluster 19 did not show a difference among varieties; they were involved in fundamental categories of BP and MF.

TFs, binding to *cis*-acting elements in the promoters of target genes, as master regulators activate or repress a large number of functional genes (Gujjar et al., 2014; Yao et al., 2018). Of the 10,331 DEGs identified in this study, 671 differentially expressed TFs, including 73 bHLHs (basic helix-loop-helix), 70 MYBs, 63 NACs, and 56 ERFs (ethylene response factor), were identified in different clusters, except in clusters 10 and 15. The largest number of TFs was distributed in clusters 19 (79 TFs) and 7 (78 TFs). In addition, 10 MYBs and 5 NACs were enriched in cluster 8, which were related to cell wall biosynthesis and mainly expressed in Pd1 (**Supplementary Table S5**).

Construction of Gene Co-expression Network

To obtain a comprehensive understanding of gene expression and identify novel regulatory genes during poplar wood formation, we performed a weighted gene correlation network analysis (WGCNA) using DEGs. Modules were defined as clusters of highly interconnected genes, and genes within the same module have high correlation coefficients. A total of 26 distinct modules (labeled as different colors) were identified and shown in the dendrogram (**Supplementary Figure S2A**). We then compared the overlapped genes between WGCNA modules and *K*-means clusters. Module turquoise (1,309 DEGs) is highly correlated with clusters 1, 8, and 15, which genes were highly expressed in Pd1 (**Supplementary Figure S2B**). It mainly participated in lignin, cellulose, and secondary cell wall biosynthesis, including 65% of secondary cell wall biosynthesis module, 57.6% of lignin biosynthesis module, and 48.5% of S-lignin and xylan biosynthesis module in AspWood database (Sundell et al., 2017) (**Supplementary Table S6**). In turquoise module, 23 putative MYB and 10 NAC genes were identified including the master switches homologous of *MYB46*, *MYB83*, *NACSECONDARY WALL THICKENING PROMOTING FACTOR 1* (*NST1*), and *SECONDARY WALL-ASSOCIATED NAC DOMAINPROTEIN 2* (*SND2*) of SCW formation (Zhang et al., 2018a) (**Supplementary Figures S2D,F**). Module darkturquoise is major participated in cellular component organization progress and cell wall, external encapsulating structure, and extracellular region (**Supplementary Figure S2C**).

Lignin synthesis pathway was regulated by three layers of regulatory network in wood plants, including MYBs, NACs, miR397a, etc. (Lu et al., 2013; Zhang et al., 2018a). To further identify potential novel regulatory genes in lignin biosynthesis, we extracted the subnetwork of lignin biosynthetic genes from our co-expression dataset (**Figure 4A** and **Supplementary Table S7**). Many known SCW regulatory TFs were identified in this subnetwork, including *MYB4*, *MYB46*, *MYB83*, *MYB102*,

NST1, *SND2*, and *VASCULAR NAC DOMAIN 4* (*VND4*), etc. In addition, several functional unknown TFs were highly connected with those key regulators and the lignin biosynthetic genes, including four R2R3 MYB subfamily *MYB19* (Potri.009G096000), *MYB43* (Potri.011G041600), *MYB55* (Potri.014G111200), *MYB74* (Potri.015G082700), and one MYB3R4 subfamily *MYB160* (Potri.006G241700). *MYB74* directly co-expressed with genes related to lignin biosynthesis, including MYBs (*MYB46*, *MYB63*, *MYB4*, and *MYB85*), NACs (*NST1* and *SND2*), and structural genes (*PAL1*, *4CL*, *C4H*, and *CCoAOMT1*). *MYB55* indirectly co-expressed with lignin biosynthesis through a positive correlation with protein kinase and zinc finger. *MYB160* co-expressed with noyl-CoA hydratase, glutamine synthetase, and sinapine esterase, which were positive correlation with *4CL*, *C3H1*, and *CCoAMT* (**Figure 4A** and **Supplementary Table S7**). The phylogenetic relationship shows that *PtrSND2/3-B1* and *PtSND2* are the closest to *AtSND2*, *PtrVND6-A1* and *PtrVND6-B1* are the closest to *AtVND4*, *PtrVND6-C2* is the closest to *AtVND1/2*, and *PtrWND2A* is the closest to *AtNST1*. *MYB*55 and *MYB74* have the closest relationship with known SCW-associated R2R3-MYB transcription factors *PtrMYB121* and *PtoMYB170*, while *MYB160* as MYB3R4 type is the furthest from R2R3-MYB (**Supplementary Figure S3**).

We found 77 DEGs in our datasets that were involved in lignin biosynthesis. The highly expressed genes in Pd1 cover almost 10 enzyme families in monolignol biosynthesis and the most of *LAC*s (**Supplementary Figure S4**). Eight genes identified from DEG list were selected for qRT-PCR validation, which include three potential novel regulatory genes (*PdMYB55*, *PdMYB74*, and *PdMYB160*), three known MYBs (*MYB43*, *MYB63*, and *MYB83*), and two cell wall biosynthesis structural genes (*C3H* and *CesA4*). The high expression of the genes in Pd1 was consistent with RNA-seq, indicating the reliability of the RNA-seq results and the xylem of Pd1 is in active stage (**Figure 4B**).

Transient Expression Assay in *Nicotiana tabacum*

In order to verify whether these novel regulators identified in our study play potential roles in lignin biosynthesis, the three selected *MYB* genes were cloned from Pd1 and were transiently overexpressed in tobacco. Yeast cells expressing BD-MYB55, BD-MYB74, or BD-MYB160 but not BD alone grow in the absence of His (-His) on SD plates, suggesting that three *MYBs* possess the activity to promote HIS marker gene expression in yeast (**Figure 5A**). qRT-PCR analysis for three independent lines indicated that *PdMYB55, PdMYB74*, and *PdMYB160* can regulate the expression of lignin biosynthetic structural genes (**Figures 5B,C**). Similar to co-expression analysis *PdMYB74* can promote the expression of *PAL*, *CSE*, *HCT*, and *LAC*. The expression of genes in the lignin biosynthetic pathway, including *4CL*, *C4H*, *CCR*, and *CSE*, appeared strong downregulation in *PdMYB55* and *PdMYB160* transient overexpression lines compared to control plants (**Figure 5C**).

DISCUSSION

Typical poplar wood contains about 33% (vol/vol) vessel elements, 53%–55% fibers, 11–14% ray parenchyma, and about 1% axial parenchyma (Mellerowicz et al., 2001; Groover et al., 2010). In our study, there were significant differences in the number of vessels, the width of cambium region, and the xylem cell wall among five black poplar cultivars (**Figure 1**). Vessel, tracheary elements, transport water and soluble minerals from the roots throughout (Yamaguchi et al., 2011). Its size and number contribute to define wood density (Leal et al., 2011). Cell death is transcriptionally regulated as a part of an overall xylem maturation, which includes secondary cell wall formation (Bollhoner et al., 2012). The DEGs of Pn1 compared with the other four cultivars were enriched in cell death and death and also involved in molecular function such as kinase activity, nucleotide binding, and receptor activity (**Figure 2D**). And we recognized many related genes which influence the cell death and SCW of xylem vessels and fibers, such as *accelerated cell death2* (*ACD2*), *programmed cell death 4-like*, *XYLEM CYSTEINE PEPTIDASE1* (*XCP1*) and *XCP2*, *metacaspase9* (*MC9*), and *BIFUNCTIONAL NUCLEASE1* (*BFN1*) (**Supplementary Table S8**). VND and NST regulate vessel element and fiber differentiation (Mitsuda et al., 2005, 2007; Yamaguchi et al., 2011; Tan et al., 2018). VND6 and VND7 directly control PCD and autolysis in the element differentiation as transcriptional master switches (Escamez and Tuominen, 2014). XYLEM NAC DOMAIN1 (XND1) and VND-INTERACTING2 (VNI2) are NAC transcription factors that suppress secondary wall formation and cell death of vessel elements, suggesting they were negative regulators of xylem vessel formation (Grant et al., 2010; Yamaguchi et al., 2010). While in our study *XND1*and *VNI2* were highly expressed in *P. euramericana* and *P. deltoides*. We thought the difference of vessel development may be due to the interaction of NAC and PCD related genes, thus affecting the transport of nutrients and plant growth. We found *ERF1* (Potri.008G166200), *WRKY75* (Potri.012G101000), and disease resistance protein [CC-NBS-LRR class (Potri.T052300) and TIR-NBS-LRR class (Potri.011G014700 and Potri.019G114500)] highly expressed in Pn1 (**Supplementary Figure S5**), which participated in disease and defense response. And these genes were not expressed in the xylem of *P. tremula* by AspWood. The results suggest Pn1 should have stronger resistance and adaptability.

Phenotypic differences are often caused by the differential expression of genes. Only few number of DEGs (339) were identified between Pd1 and Pd2, suggesting their close relationship—they were progeny of *P. deltoides* '55/65' × *P. deltoides* '2KEN8' (Zhang et al., 2008; Hu et al., 2013). And all of them participated in the metabolic process, which might be the reason of radial growth differences between the two cultivars. DEGs between Pe1 and Pd1 are related with microtubule cytoskeleton (**Figure 2D**), which is a dynamic filamentous structure participating in nuclear and cell division, deposition of cell wall, cell expansion, organelle movement, and secretion processes in cell morphogenesis (Hussey et al., 2002).

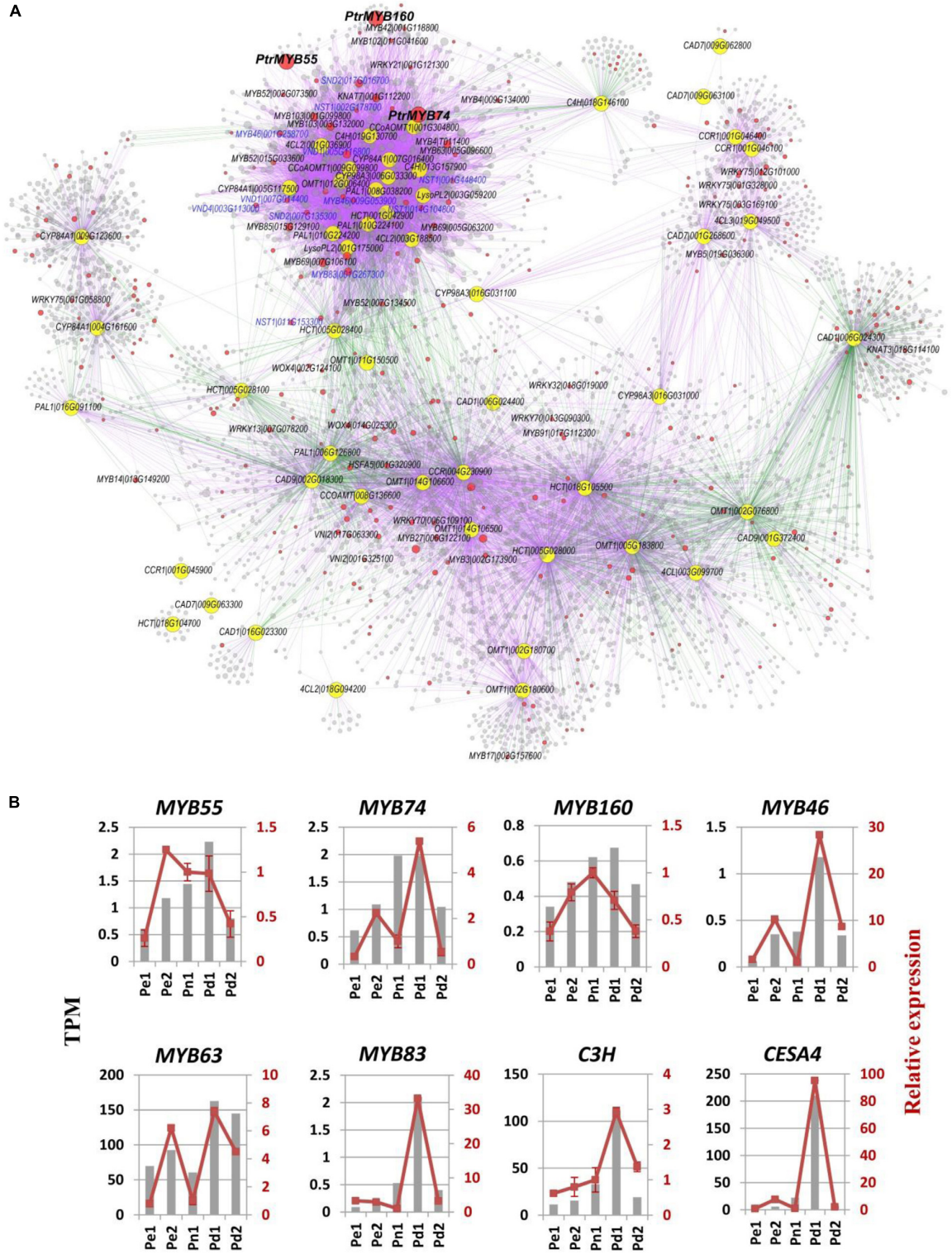

FIGURE 4 | Co-expression network of lignin biosynthetic genes based on our RNA sequencing (RNA-Seq) dataset. **(A)** The subnetwork was extracted from our RNA-Seq co-expression analysis. Yellow and red nodes represent lignin biosynthetic genes and transcription factors, respectively. Purple and green edges represent positive and negative correlation, respectively. The blue letter labeled the known key regulators in the first layer and the second layer of secondary cell wall formation regulatory network (Zhang et al., 2018a). See **Supplementary Table S7** for detailed node information. **(B)** Expression confirmation of eight critical genes using quantitative Real-Time PCR (qRT-PCR). Transcripts per million (TPM) values and relative expression of target genes by qRT-PCR of eight critical genes were shown. Each sample was conducted four biological replicates and four technical replicates.

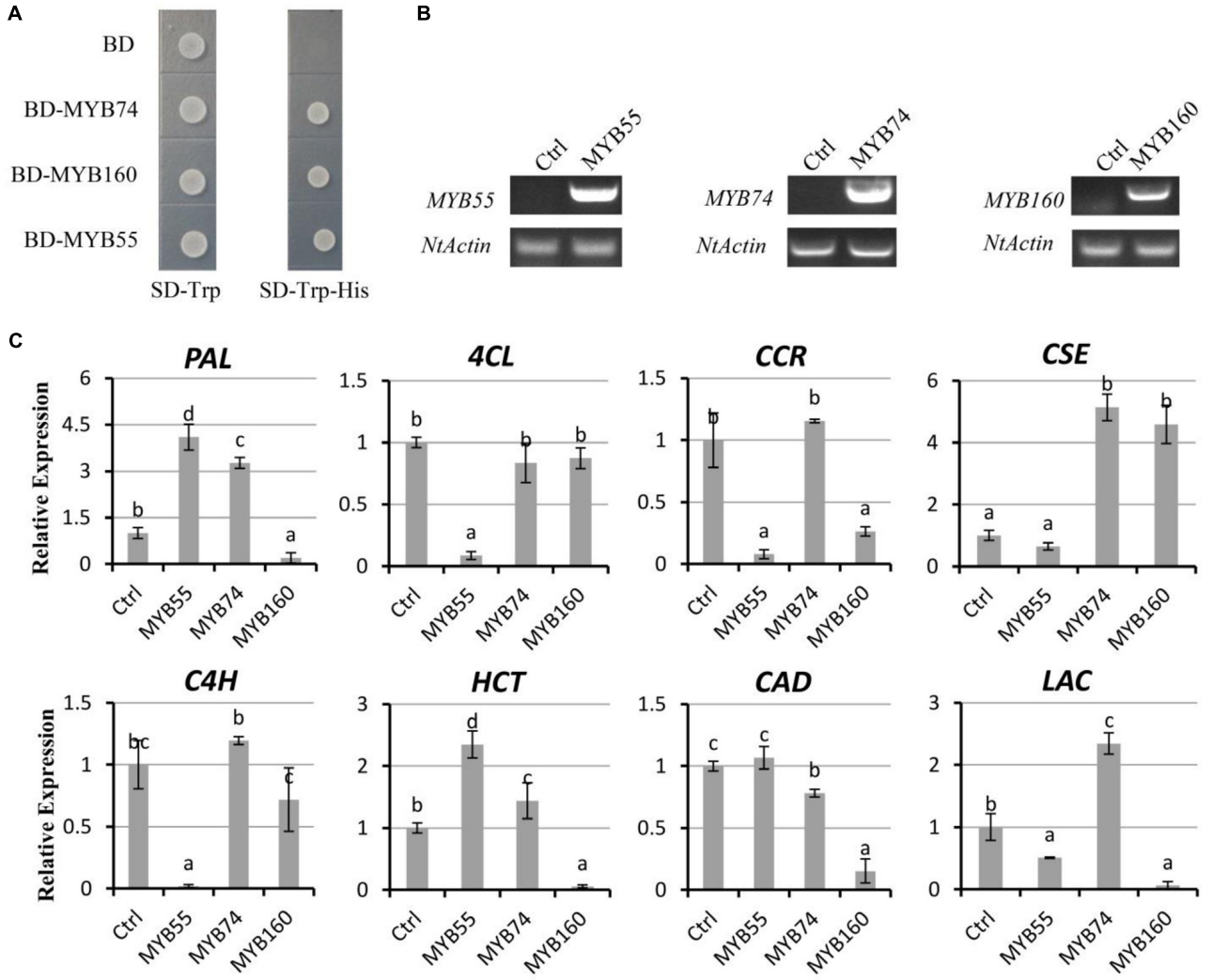

FIGURE 5 | Novel transcription factors regulate the expression of the secondary wall biosynthetic genes. **(A)** Transactivation analysis of MYB in yeast. Full-length *MYB55, MYB74*, and *MYB160* were fused with the GAL4 DNA binding domain and expressed in the yeast strain Y2HGold harboring the *His3* reporter gene under the control of GAL4-responsive promoter elements. **(B)** Semiquantitative RT-PCR analysis of *MYB55*, *MYB74*, and *MYB160* expression in tobacco. **(C)** The quantitative Real-Time PCR (qRT-PCR) of structural genes in lignin biosynthesis. Error bars ± SD from three biological replicates. Lowercase letters (a, b, c, and d) indicate the results of Duncan's multiple range test (significant differences at $P < 0.05$).

Plant cell walls are also a source of renewable biomass for conversion to biofuels and bioproducts (Li et al., 2012). Lignin impregnate with cellulose and hemicellulose simultaneously to provide additional mechanical strength, hardness, and hydrophobicity to the secondary wall (Zhong and Ye, 2015). Zhang et al. (2018a) systematically reviewed the complex regulatory network of SCW biosynthesis, which includes a series of NAC and MYB TFs. We identified a SCW-associated module (turquoise) by WGCNA. Most genes in this module are structural genes involved in the biosynthesis of lignin and cellulose, such as *PAL*, *4CL*, *CCR*, and *CesAs*, etc. In addition, we identified a large number of transcription factors, which are known as three layers of transcription factors in the regulatory network in secondary wall thickening and lignification in wood plants, including *VND*, *SND*, and *WND* in the first layer, master switches *MYB46* and *MYB83* in the second layer, and *MYB4*, *MYB61*, and *MYB103*, etc. in the third layer (**Figure 4**, **Supplementary Figures S2, S3**). The high expression of these genes in Pd1 is related to the development state of xylem and finally leaded the thickest wall of Pd1. Evolutionary trees show the relationship between known and novel TFs (**Supplementary Figure S3**). For example, the deposition of lignin and thickening of secondary walls were influenced in overexpressing *PtoVNS11* transgenic poplar (Yang et al., 2015). Splicing variants of *PtrVND6-C1*^{IR} and *PtrSND1-A2*^{IR} function together to cross-regulate the VND and SND families to maintain the wood formation and plant development (Lin et al., 2017). *PtrWND2B* and *PtrWND6B* influenced the expression of SCW-associated TFs and structural genes and, concomitantly, the ectopic deposition of cellulose, xylan, and lignin (Zhong et al., 2010). *PtoMYB156* and *PtoMYB189*

negatively regulate secondary cell wall biosynthesis during wood formation in poplar (Yang et al., 2017; Jiao et al., 2019). While *PtrMYB152* and *PtoMYB92* have been reported as activators of lignin biosynthesis (Wang et al., 2014; Li et al., 2015). In our study, SCW-associated modules were identified, including orthologs of *PtrSND2/3-B1*, *PtrSND1* and its target *PtrMYB021*, which influenced the thickness of secondary cell wall of xylem fiber and the content of cellulose and lignin in stem (Li et al., 2012; Wang et al., 2013). *PdMYB*55 and *PdMYB74* were clustered with *PtrMYB121*, *PtrMYB74*, and *PtoMYB170* which were identified as the downstream targets of wood-associated NAC domain TFs to influence wood formation (Zhong et al., 2011; Xu et al., 2017) (**Figure 4**, **Supplementary Figure S3**), so *PdMYB*55 and *PdMYB74* may also positively regulate lignin biosynthesis. Although *MYB160* has the furthest relationship with others, it may participate in the development of secondary wall as a member of SCW module. Those results suggested some uncharacterized NAC and MYB TFs may participate in the SCW biosynthesis.

PdMYB55, a homolog of *AtMYB55,* could influence the expression of key genes in lignin biosynthetic pathway in our transient expression assay. *AtMYB55*, as a brassinolide-inducible gene, participates in basal cell of mature leaves and downregulated by the Aux/IAA protein in an organ-specific manner (Nakamura et al., 2006; Schliep et al., 2010). *PtrMYB74* and *AtMYB50* as downstream genes of *NAC102* participate in the formation of secondary walls in xylem fiber and vessels (Ko et al., 2007; Zhong et al., 2011). The expression of structural genes may be upregulated by direct action in *PdMYB74*, suggesting it was a positive regulator of SCW. *PdMYB55* and *PdMYB74* are closely related in evolutionary relationship, but it is possible that their functions are not completely consistent because it regulates interaction with lignin pathway genes by protein kinase and zinc finger. *PdMYB160* belongs to c-myb-like MYB3R4 subfamily. MYB3R4 can bind to MSA motifs in promoters of B-type cyclins (CYCB) to regulate the cell cycle in *Arabidopsis* and tobacco (Haga et al., 2011; Kobayashi et al., 2015; Olszak et al., 2019). Although *PdMYB160* has transcriptional activity and represses the expression of structural genes, which may be due to the indirect effect of regulation and needs further study in the future. *PdMYB55*, *PdMYB74*, *PdMYB160*, and other SCW TFs are highly expressed in Pd1. Three MYBs and other TFs jointly regulate structural gene expression in lignin biosynthesis. These results indicate that three novel TFs are participated in the regulation of lignin biosynthetic pathway. The results of the case study prove that our dataset provides a great resource to discover novel regulators in the lignin biosynthetic pathway.

CONCLUSION

Secondary cell wall biosynthesis is a biological process of producing wood, which is an important renewable material and energy raw material. The chemical structure and the content of lignin directly affect the costs of pretreatment and conversion efficiency in biofuel production from cellulosic biomass. In this study, we compared the xylem anatomical structures of five poplar cultivars in China and analyzed the transcriptome-wide gene expression profiles of developing xylem. A large number of TFs co-expressed with lignin biosynthetic genes were identified by *K*-means clustering and co-expression analysis. Furthermore, transient expression showed that *MYB55*, *MYB74*, and *MYB160* may function as novel regulators in lignin biosynthesis pathway. This study provides a useful resource for future studies seeking for the molecular mechanisms of xylem development and utilization of bioenergy.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the sequencing data are available in NCBI SRA database (SRA number: SRP234303).

AUTHOR CONTRIBUTIONS

JZ and JH designed and conducted the experiments. LZ and BL performed the experiments. LZ and JZ conducted the data and wrote the manuscript. JZ and JH contributed to discussion and manuscript revision. All the authors were involved in the discussion of the data and approved the final manuscript.

FUNDING

This research was supported by the National Key Research and Development Program of China (2017YFD0600201), National Non-profit Institute Research Grant of CAF (CAFYBB2017ZY008), and National Key Program on Transgenic Research (2018ZX08020002) to JH and the National Natural Science Foundation of China (31870661) to BL.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00620/full#supplementary-material

FIGURE S1 | Microscopic analysis of xylem form five black poplar cultivars. **(A)** The region for vessel count. Scale bars = 200 μm. **(B)** Average diameter of vessel. **(C)** The cell wall thickness of fiber. At least 50 cells per sample were measured. Means ± SD from four biological replicates. Lowercase letters (a, b, c, and d) indicate the results of Duncan's multiple range test (significant differences at $P < 0.05$).

FIGURE S2 | Construction of co-expression modules by WGCNA. **(A)** The cluster dendrogram of genes. Each branch in the figure represents one gene, and every color below represents one co-expression module. **(B)** The percentage of module-to-cluster in module. Node color represents the percentage in module. Node size represents gene number. **(C)** Gene Ontology (GO) enrichment analysis of genes in different modules. Node color represents $-\log_{10}$ transformed FDR corrected *P*-value. Node size represents rich factor. The heatmap of MYBs **(D)** and NACs **(E)** in module turquoise.

FIGURE S3 | Phylogenetic trees of NACs **(A)** and MYBs **(B)** from different plant species by the neighbor-joining method with 1000 bootstraps. The bootstrap values are indicated as percentages at the nodes.

FIGURE S4 | The differentially expressed genes related to monolignol biosynthesis **(A)** and laccase phenoloxidases **(B)** in five cultivars.

FIGURE S5 | The differentially expressed genes related to defense response genes. The average TPM of four biological replicates. *ERF1*: Potri.008G166200, *WRKY75*: Potri.012G101000, *RPP4:* Potri.019G114500, *CC-NBS-LRR*: Potri.T052300 and *TIR-NBS-LRR*: Potri.011G014700.

DATA SHEET S1 | Summary of all differentially expressed genes (DEGs) TPM, cluster, modules and other information.

TABLE S1 | The primers used in this study.

TABLE S2 | Summary of the RNA-Seq results.

TABLE S3 | GO enrichment analysis of all differentially expressed genes (DEGs).

TABLE S4 | GO enrichment analysis of 20 clusters.

TABLE S5 | Identification of 671 transcription factors (TFs) representing 47 gene families in the 20 clusters.

TABLE S6 | Common genes between turquoise module in this study and SCW-associated modules in AspWood.

TABLE S7 | The genes in co-expression network of lignin biosynthetic.

TABLE S8 | Differentially expressed of vessel differentiation and PCD related genes.

REFERENCES

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

Buschmann, H., Green, P., Sambade, A., Doonan, J. H., and Lloyd, C. W. (2011). Cytoskeletal dynamics in interphase, mitosis and cytokinesis analysed through *Agrobacterium*-mediated transient transformation of tobacco BY-2 cells. *New Phytol.* 190, 258–267. doi: 10.1111/j.1469-8137.2010.03587.x

Chano, V., Collada, C., and Soto, A. (2017). Transcriptomic analysis of wound xylem formation in *Pinus canariensis*. *BMC Plant Biol.* 17:234. doi: 10.1186/s12870-017-1183-3

Courtois-Moreau, C. L., Pesquet, E., Sjödin, A., Muñiz, 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.1365-313X.2008.03777.x

Davis, J. M. (2008). "Genetic improvement of poplar (*Populus* spp.) as a bioenergy crop," in *Genetic Improvement of Bioenergy Crops*, ed. W. Vermerris (New York, NY: Springer), 397–419. doi: 10.1007/978-0-387-70805-8_14

D'haeseleer, P., Liang, S., and Somogyi, R. (2000). Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics* 16, 707–726. doi: 10.1093/bioinformatics/16.8.707

Du, J., and Groover, A. (2010). Transcriptional regulation of secondary growth and wood formation. *Chin. Bull. Bot.* 52, 17–22. doi: 10.1111/j.1744-7909.2010.00901.x

El Kasmioui, O., and Ceulemans, R. (2012). Financial analysis of the cultivation of poplar and willow for bioenergy. *Biomass Bioenergy* 43, 52–64. doi: 10.1016/j.biombioe.2012.04.006

Escamez, S., and Tuominen, H. (2014). Programmes of cell death and autolysis in tracheary elements: when a suicidal cell arranges its own corpse removal. *J. Exp. Bot.* 65, 1313–1321. doi: 10.1093/jxb/eru057

Freudenberg, K. (1965). Lignin: its constitution and formation from p-hydroxycinnamyl alcohols. *Science* 148, 595–600. doi: 10.1126/science.148.3670.595

Grant, E. H., Fujino, T., Beers, E. P., and Brunner, A. M. (2010). Characterization of NAC domain transcription factors implicated in control of vascular cell differentiation in *Arabidopsis* and *Populus*. *Planta* 232, 337–352. doi: 10.1007/s00425-010-1181-2

Groover, A., Nieminen, K., Helariutta, Y., and Mansfield, S. (2010). "Wood formation in *Populus*," in *Genetics and Genomics of Populus*, eds S. Jansson, R. Bhalerao, and A. Groover (New York, NY: Springer), 201–224. doi: 10.1007/978-1-4419-1541-2_10

Guerra, F. P., Wegrzyn, J. L., Sykes, R., Davis, M. F., Stanton, B. J., and Neale, D. B. (2013). Association genetics of chemical wood properties in black poplar (*Populus nigra*). *New Phytol.* 197, 162–176. doi: 10.1111/nph.12003

Gujjar, R., Akhtar, M., and Singh, M. (2014). Transcription factors in abiotic stress tolerance. *Indian J. Plant Physiol.* 19, 306–316. doi: 10.1007/s40502-014-0121-8

Haga, N., Kobayashi, K., Suzuki, T., Maeo, K., Kubo, M., Ohtani, M., et al. (2011). Mutations in MYB3R1 and MYB3R4 cause pleiotropic developmental defects and preferential down-regulation of multiple G2/M-specific genes in *Arabidopsis*. *Plant Physiol.* 157, 706–717. doi: 10.1104/pp.111.180836

Hu, J., Lu, M., Zhao, Z., Su, X., Li, X., Li, S., et al. (2013). An elite variety of *Populus deltoides* 'Nanyang'. *Sci. Silvae Sin.* 49:188. doi: 10.11707/j.1001-7488.20130728

Hussey, P. J., Hawkins, T. J., Igarashi, H., Kaloriti, D., and Smertenko, A. (2002). The plant cytoskeleton: recent advances in the study of the plant microtubule-associated proteins MAP-65, MAP-190 and the Xenopus MAP215-like protein, MOR1. *Plant Mol. Biol.* 50, 915–924. doi: 10.1023/A:1021236307508

Jiao, B., Zhao, X., Lu, W., Guo, L., and Luo, K. (2019). The R2R3 MYB transcription factor MYB189 negatively regulates secondary cell wall biosynthesis in *Populus*. *Tree Physiol.* 39, 1187–1200. doi: 10.1093/treephys/tpz040

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

Kobayashi, K., Suzuki, T., Iwata, E., Nakamichi, N., Suzuki, T., Chen, P., et al. (2015). Transcriptional repression by MYB3R proteins regulates plant organ growth. *EMBO J.* 34, 1992–2007. doi: 10.15252/embj.201490899

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Leal, S., Sousa, V. B., Knapic, S., Louzada, J. L., and Pereira, H. (2011). Vessel size and number are contributors to define wood density in cork oak. *Eur. J. For. Res.* 130, 1023–1029. doi: 10.1007/s10342-011-0487-3

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323

Li, C., Wang, X., Ran, L., Tian, Q., Fan, D., and Luo, K. (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, 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

Lin, C. Y., Wang, J. P., Li, Q., Chen, H. C., Liu, J., Loziuk, P., et al. (2015). 4-Coumaroyl and caffeoyl shikimic acids inhibit 4-Coumaric acid:coenzyme A ligases and modulate metabolic flux for 3-Hydroxylation in monolignol biosynthesis of *Populus trichocarpa*. *Mol. Plant* 8, 176–187. doi: 10.1016/j.molp.2014.12.003

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. U.S.A.* 114, E9722–E9729. doi: 10.1073/pnas.1714422114

Liu, B., Zhang, J., Yang, Z., Matsui, A., Seki, M., Li, S., et al. (2018). *PtWOX11* acts as master regulator conducting the expression of key transcription factors to induce de novo shoot organogenesis in poplar. *Plant Mol. Biol.* 98, 389–406. doi: 10.1007/s11103-018-0786-x

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta}$Ct Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

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

Mellerowicz, E., Baucher, M., Sundberg, B., and Boerjan, W. (2001). Unraveling cell wall formation in the woody dicot stem. *Plant Mol. Biol.* 47, 239–274. doi: 10.1023/A:1010699919325

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., 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

Nakamura, A., Nakajima, N., Goda, H., Shimada, Y., Hayashi, K.-I., Nozaki, H., et al. (2006). *Arabidopsis* Aux/IAA genes are involved in brassinosteroid-mediated growth responses in a manner dependent on organ type. *Plant J.* 45, 193–205. doi: 10.1111/j.1365-313X.2005.02582.x

Olszak, M., Truman, W., Stefanowicz, K., Sliwinska, E., Ito, M., Walerowski, P., et al. (2019). Transcriptional profiling identifies critical steps of cell cycle reprogramming necessary for Plasmodiophora brassicae-driven gall formation in *Arabidopsis*. *Plant J.* 97, 715–729. doi: 10.1111/tpj.14156

Parikka, M. (2004). Global biomass fuel resources. *Biomass Bioenergy* 27, 613–620. doi: 10.1016/j.biombioe.2003.07.005

Samuels, A. L., Rensing, K. H., Douglas, C. J., Mansdield, S. D., Dharmawardhana, D. P., and Ellis, B. E. (2002). Cellular machinery of wood production: differentiation of secondary xylem in *Pinus contorta* var. *latifolia*. *Planta* 216, 72–82. doi: 10.1007/s00425-002-0884-4

Schliep, M., Ebert, B., Simon-Rosin, U., Zoeller, D., and Fisahn, J. (2010). Quantitative expression analysis of selected transcription factors in pavement, basal and trichome cells of mature leaves from *Arabidopsis thaliana*. *Protoplasma* 241, 29–36. doi: 10.1007/s00709-009-0099-7

Seyfferth, 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

Shannon, P., Markiel, A., Ozier, O., Baliga, N., Wang, J., 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., Sun, Y.-H., Li, Q., Heber, S., Sederoff, R., and Chiang, V. L. (2009). Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes. *Plant Cell Physiol.* 51, 144–163. doi: 10.1093/pcp/pcp175

Song, X. L., Yao, C. L., Le, W., and Zou, X. (2010). Analysis of wood property of four species of *Populus euramericanas* and studies on their pulping performance. *China Puplp Pap. Ind.* 31, 33–37. doi: 10.3969/j.issn.1007-9211.2010.10.006

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

Takata, N., and Taniguchi, T. (2014). Expression divergence of cellulose synthase (*CesA*) genes after a recent whole genome duplication event in *Populus*. *Planta* 241, 29–42. doi: 10.1007/s00425-014-2217-9

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, T. T., Endo, H., Sano, R., Kurata, T., Yamaguchi, M., Ohtani, M., et al. (2018). Transcription factors VND1-VND3 contribute to cotyledon xylem vessel formation. *Plant Physiol.* 176, 773–789. doi: 10.1104/pp.17.00461

Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. doi: 10.1093/bioinformatics/btp120

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

Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Torr, K., Donaldson, L., et al. (2011). CCoAOMT suppression modifies lignin composition in *Pinus radiata*. *Plant J.* 67, 119–129. doi: 10.1111/j.1365-313X.2011.04580.x

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, 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, 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

Willebrand, E., and Verwijst, T. (1993). Population dynamics of willow coppice systems and their implications for management of short-rotation forests. *For. Chron.* 69, 699–704. doi: 10.5558/tfc69699-6

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

Yamaguchi, M., Mitsuda, N., Ohtani, M., Ohme-Takagi, M., Kato, K., and Demura, T. (2011). VASCULAR-RELATED NAC-DOMAIN7 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 transcrip-tion factor, negatively regulates xylem vessel formation in *Arabidopsis*. *Plant Cell* 22, 1249–1263. doi: 10.1105/tpc.108.064048

Yang, F., Wang, Y., Wang, J., Deng, W., Liao, L., and Li, M. (2011). Different eco-physiological responses between male and female *Populus deltoides* clones to waterlogging stress. *For. Ecolo. Manag.* 262, 1963–1971. doi: 10.1016/j.foreco.2011.08.039

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., Li, C., Fan, D., and Luo, K. (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

Yao, W., Zhao, K., Cheng, Z., Li, X., Zhou, B., and Jiang, T. (2018). Transcriptome analysis of poplar under salt stress and over-expression of transcription factor *NAC57* gene confers salt tolerance in transgenic *Arabidopsis*. *Front. Plant Sci.* 9:1121. doi: 10.3389/fpls.2018.01121

Zhang, C., Li, S., Zhao, Z., Hu, J., and Han, Y. (2008). A new poplar variety *Populus deltoides* 'Danhong'. *Scie. Silvae Sin.* 44:169. doi: 10.11707/j.1001-7488.20080127

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

Zhang, J., Song, X., Zhang, L., Jia, H., Peng, X., Zhao, Z., et al. (2020). Agronomic performance of 27 Populus clones evaluated after two 3-year coppice rotations in Henan. China. *GCB Bioenergy* 12, 168–181. doi: 10.1111/gcbb.12662

Zhang, J., Xie, M., Tuskan, G. A., Muchero, W., and Chen, J.-G. (2018a). 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

Zhang, J., Yang, Y., Zheng, K., Xie, M., Feng, K., Jawdy, S. S., et al. (2018b). 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, 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., 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., McCarthy, R. L., Lee, C., and Ye, Z.-H. (2011). Dissection of the transcriptional program regulating secondary wall biosynthesis during wood formation in poplar. *Plant Physiol.* 157, 1452–1468. doi: 10.1104/pp.111.181354

Zhong, R., and Ye, Z.-H. (2015). Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant Cell Physiol.* 56, 195–214. doi: 10.1093/pcp/pcu140

Zinkgraf, M., Liu, L., Groover, A., and Filkov, V. (2017). Identifying gene coexpression networks underlying the dynamic regulation of wood-forming tissues in Populus under diverse environmental conditions. *New Phytol.* 214, 1464–1478. doi: 10.1111/nph.14492

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 19 June 2020
doi: 10.3389/fpls.2020.00843

Silencing *Folylpolyglutamate Synthetase1* (*FPGS1*) in Switchgrass (*Panicum virgatum* L.) Improves Lignocellulosic Biofuel Production

Mitra Mazarei[1,2,3], Holly L. Baxter[1,2], Avinash Srivastava[2,4], Guifen Li[2,4], Hongli Xie[2,4], Alexandru Dumitrache[2,5], Miguel Rodriguez Jr.[2,3,5], Jace M. Natzke[2,5], Ji-Yi Zhang[2,4], Geoffrey B. Turner[2,6], Robert W. Sykes[2,6], Mark F. Davis[2,3,6], Michael K. Udvardi[2,3,4], Zeng-Yu Wang[2,4], Brian H. Davison[2,3,5], Elison B. Blancaflor[2,4], Yuhong Tang[2,4] and Charles Neal Stewart Jr.[1,2,3]**

[1] Department of Plant Sciences, The University of Tennessee, Knoxville, TN, United States, [2] BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [3] The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [4] Noble Research Institute, Ardmore, OK, United States, [5] Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [6] National Renewable Energy Laboratory, Golden, CO, United States

OPEN ACCESS

Edited by:
Mengzhu Lu,
Chinese Academy of Forestry, China

Reviewed by:
Mingyang Quan,
Beijing Forestry University, China
Bin Xu,
Nanjing Agricultural University, China

***Correspondence:**
Yuhong Tang
ytang@noble.org
Charles Neal Stewart Jr.
nealstewart@utk.edu

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: 26 March 2020
Accepted: 26 May 2020
Published: 19 June 2020

Citation:
Mazarei M, Baxter HL, Srivastava A, Li G, Xie H, Dumitrache A, Rodriguez M Jr, Natzke JM, Zhang J-Y, Turner GB, Sykes RW, Davis MF, Udvardi MK, Wang Z-Y, Davison BH, Blancaflor EB, Tang Y and Stewart CN Jr (2020) Silencing Folylpolyglutamate Synthetase1 (FPGS1) in Switchgrass (Panicum virgatum L.) Improves Lignocellulosic Biofuel Production. Front. Plant Sci. 11:843. doi: 10.3389/fpls.2020.00843

Switchgrass (*Panicum virgatum* L.) is a lignocellulosic perennial grass with great potential in bioenergy field. Lignocellulosic bioenergy crops are mostly resistant to cell wall deconstruction, and therefore yield suboptimal levels of biofuel. The one-carbon pathway (also known as C1 metabolism) is critical for polymer methylation, including that of lignin and hemicelluloses in cell walls. Folylpolyglutamate synthetase (FPGS) catalyzes a biochemical reaction that leads to the formation of folylpolyglutamate, an important cofactor for many enzymes in the C1 pathway. In this study, the putatively novel switchgrass *PvFPGS1* gene was identified and its functional role in cell wall composition and biofuel production was examined by RNAi knockdown analysis. The *PvFPGS1*-downregulated plants were analyzed in the field over three growing seasons. Transgenic plants with the highest reduction in *PvFPGS1* expression grew slower and produced lower end-of-season biomass. Transgenic plants with low-to-moderate reduction in *PvFPGS1* transcript levels produced equivalent biomass as controls. There were no significant differences observed for lignin content and syringyl/guaiacyl lignin monomer ratio in the low-to-moderately reduced *PvFPGS1* transgenic lines compared with the controls. Similarly, sugar release efficiency was also not significantly different in these transgenic lines compared with the control lines. However, transgenic plants produced up to 18% more ethanol while maintaining congruent growth and biomass as non-transgenic controls. Severity of rust disease among transgenic and control lines were not different during the time course of the field experiments. Altogether, the unchanged lignin content and composition in the low-to-moderate *PvFPGS1*-downregulated lines may suggest that partial downregulation of *PvFPGS1* expression did not impact lignin biosynthesis in switchgrass. In conclusion, the manipulation of *PvFPGS1* expression in bioenergy crops may be useful to increase biofuel potential with no growth penalty or increased susceptibility to rust in feedstock.

Keywords: folylpolyglutamate synthetase, switchgrass, RNAi-gene silencing, *PvFPGS1*, lignocellulosic, biofuel

INTRODUCTION

The addition or removal of one-carbon units (C1 metabolism) is required for the synthesis and regulation of many biological compounds and metabolic processes. The synthesis of basic compounds and a range of methylated molecules in all organisms requires C1 metabolism (Appling, 1991). In plants, a balanced supply of C1 units is required for the synthesis of numerous plant secondary metabolites (e.g., lignin and phytohormones) and hemicellulose (Hanson and Roje, 2001; Eckardt, 2007; Urbanowicz et al., 2012). As the major sink of methyl units, lignin biosynthesis is affected by changes of enzymes in the C1 pathway (Shen et al., 2002; Tang et al., 2014; Li et al., 2015). Folylpolyglutamate synthetase (FPGS) catalyzes a biochemical reaction that leads to the formation of folylpolyglutamate, an important cofactor for many enzymes in the C1 pathway (Shane, 1989). In plant cells, folates, mainly as polyglutamylated byproducts, are found in chloroplasts, mitochondria, and cytosol (Rébeillé et al., 2006; Hanson and Gregory, 2011). In Arabidopsis, there are three genes each encoding FPGS isoforms that are localized to chloroplasts (FPGS1), mitochondria (FPGS2), and cytosol (FPGS3) (Ravanel et al., 2001).

Switchgrass (*Panicum virgatum* L.) is a lignocellulosic perennial grass known for its high yield of biomass, wide adaptability, and ability to grow on marginal soil conditions. These characteristics have made switchgrass a promising bioenergy feedstock (van der Weijde et al., 2013). One of the major problems with lignocellulosic crops is the resistance of the cell wall to deconstruction for efficient conversion into biofuels (known as biomass recalcitrance) (Himmel and Bayer, 2009). Cell wall lignin is one of the main causes of recalcitrance, which limits efficient conversion of biomass into biofuels (Chen and Dixon, 2007; David and Ragauskas, 2010). Cell wall components have been key targets to reduce feedstock recalcitrance; manipulating cell wall biosynthesis gene expression has been the primary strategy (Nelson et al., 2017; Biswal et al., 2018; Brandon and Scheller, 2020). Several studies with transgenic plants have been conducted in greenhouses under tightly controlled environmental conditions and the actively growing green tissue has been analyzed most often (Fu et al., 2011a,b, 2012; Xu et al., 2011; Shen et al., 2012, 2013; Liu et al., 2018). In contrast, end-of-season senesced tissue is most often used for biofuel production. Thus, greenhouse experiments may not be predictive of transgenic plants performance and recalcitrance of field-grown plants. Field-grown plants are exposed to a wider range of biotic and abiotic stresses not present in the greenhouse. Therefore, field experiments are especially important for modified plants to better predict agronomic performance across multiple growing seasons.

We previously showed that reduced lignin content and improved cell wall digestibility was observed in an Arabidopsis mutant with a disrupted *FPGS1* gene (Srivastava et al., 2015). We also showed that disruption in both FPGS1 and caffeoyl-CoA-3-*O*-methyltransferase (CCoAOMT), a lignin biosynthetic enzyme, resulted in further reduction in lignin content and improvement in cell wall digestibility in Arabidopsis (Xie et al., 2019). These studies prompted us to examine the possible role of FPGS in improving biofuel production for switchgrass.

In the present study, a novel switchgrass *PvFPGS1* gene was identified and its functional role was examined by downregulation using RNAi technology in switchgrass. A field experiment using multiple transgenic switchgrass lines downregulating *PvFPGS1* was conducted for three field growing seasons (2014–2016) to evaluate (i) *PvFPGS1* transcript levels, (ii) growth traits and biomass production, (iii) cell wall composition, sugar release, and conversion into biofuel, and (iv) susceptibility to rust disease.

MATERIALS AND METHODS

Gene Identification

Using the *FPGS* cDNA sequence of Arabidopsis *AtFPGS1* (At5g05980), TBLASTN was used to identify the homologous gene sequences from switchgrass EST databases (Zhang et al., 2013) as well as from the draft switchgrass genome (*Panicum virgatum* v1.1 DOE-JGI) at Phytozome. A FPGS family gene tree was originally constructed using neighbor-joining in the software MEGA6 (Tamura et al., 2013) and a potential isolog of *AtFPGS1* was identified from switchgrass. A phylogeny tree for FPGS protein family was constructed by neighbor-joining using Geneious Prime 2019 software (www.geneious.com).

Vector Construction and Transgenic Plant Production

The RNAi construct was made using the switchgrass *PvFPGS1* gene sequence (Pavir.Ib00114). A 462 bp sequence (**Supplementary Figure S1**) was amplified by PCR from switchgrass cultivar 'Alamo' using primer pair PvDFB-RNAi-F: 5′-AAGCAGGGGCATAAGGACA-3′ and PvDFB-RNAi-R: 5′-ATCGATTTGTTCAGGCTCAGC-3′. The target fragment was cloned into pCR8 entry vector and confirmed by sequencing. The target fragment was then sub-cloned into pANIC-8A RNAi-vector (**Supplementary Figure S2**) (Mann et al., 2012) to be driven under the maize ubiquitin 1 (*ZmUbi1*) promoter. Transgenic plants were produced using NFCX01 clonal of switchgrass 'Alamo' via *Agrobacterium*-mediated transformation (Xi et al., 2009).

Greenhouse Plants

Plants were grown in greenhouse under the 16-h day/8-h night light at 28°C day/22°C night temperature.

Field-Grown Plants and Experimental Design

The T_0 generation of *PvFPGS1*-downregulated plants were used in field experiments. The plants included six independent transgenic lines and one non-transgenic control (wild type). The plants were transplanted onto a field on June 05, 2014. The field was located at the University of Tennessee Plant Sciences Unit of the East Tennessee Research and Education Center (ETREC). The field site was 24.2 m × 15.1 m. Three replicate plots for each transgenic and control lines were distributed throughout the field in a randomized complete block design (RCBD). Each

transgenic and control replicate plot contained four vegetatively propagated clones of each line. Replicate lines were spaced 152 cm apart with 76 cm between the four clonal plants within every single replicate. The experimental plots were surrounded by non-transgenic border plants (**Supplementary Figure S3**).

Field Maintenance

The field trial was conducted for three consecutive growing seasons, in which no environmental anomalies were observed. The soil fertility was in the range of switchgrass recommendations and no soil amendments were added. The plants were irrigated only for the first 2 months as needed after transplantation for establishment. Weeds were removed by tilling or hand. No herbicides were applied for the duration of the study. Following USDA APHIS BRS guidelines, plants were observed daily during the reproductive stage and emerging panicles were removed from all plants (transgenic, non-transgenic control and border plants) at R0-R1 developmental stage (Moore et al., 1991) by cutting the plant below the top node containing the inflorescence.

Analysis of *PvFPGS1* Transcript Levels

Baseline expression of *FPGS1* in different tissues of switchgrass at the R1 growth stage was tested in greenhouse-grown plants using quantitative reverse transcription-polymerase chain reaction (qRT-PCR). Total RNA was extracted from 1st, 2nd, 3rd, and 4th internodes; 1st, 2nd, 3rd, and 4th nodes; leaf blade; leaf sheath; crown; inflorescence; and root of non-trangenic lines or from the leaf blade, leaf sheath, and 2nd internode of transgenic lines at the R1 growth stage. For field-grown plants, samples were collected from green plants in August of each growing season. All samples were collected at the same date and time of each year analyzed. Tillers at the R0 developmental stage were chosen at random from two plants within each replicate. Each tiller was excised below the top internode. The resulting top portion of the tiller with the two intact top leaves was flash-frozen in liquid nitrogen and stored at −80°C for qRT-PCR analysis. Total RNA was isolated from the frozen tissues using Spectrum™ Plant Total RNA Kits (Sigma-Aldrich, St. Louis, MO, United States) following the manufacturer's instructions. RNA quality was checked with Agilent Bioanalyzer 2100 (Agilent, Palo Alto, CA, United States) and quantified using Qubit™RNA BR Assay Kit (Fisher Scientific, Santa Clara, CA, United States). Five micrograms of total RNA was treated with TURBO DNA-free™ Kit (Invitrogen, Carlsbad, CA, United States) to remove any potential genomic DNA contaminants. Two micrograms of DNA-free total RNA was used for first-strand cDNA synthesis using SuperScript™ III First-Strand Synthesis System (Invitrogen). qRT-PCR was performed with Power SYBR™ Green PCR Master Mix (Applied Biosystems, Foster City, CA, United States) using ABI PRISM 7900 HT Sequence Detection System (Applied Biosystems). The primer pair used for the *PvFPGS1* transcript analysis was PvDFB-RNAi-qRT-F1: 5′-CAAAGAGCTTCGGAGTTGG-3′ and PvDFB-RNAi-qRT-R1: 5′-GGTAGGGGATCAGTACGATTGA-3′. Data were collected and analyzed using the SDS 2.2.1 software (Applied Biosystems). The relative transcript quantification was normalized by the levels of switchgrass ubiquitin 1 (*PvUbi1*) transcripts (Shen et al., 2009) using primer pair SWUbi_F304: 5′-TTCGTGGTGGCCAGTAAGC-3′ and SWUbi_R367: 5′-AGAGACCAGAAGACCCAGGTACAG-3′.

Agronomic Performance

Growth measurements were recorded each December (end-of-season) of the three growing seasons in the field after all aboveground biomass was completely senesced. For tiller height, the tallest tiller from each individual plant was measured from soil level to the tip of the top leaf. For plant width, the circumference at the mid-section of each whole plant was measured. Tiller numbers were tallied for each plant. For biomass yield, whole aboveground senesced biomass was harvested. The biomass was oven-dried at 43°C for 96 h and weighed to determine total dry aboveground biomass. The dried biomass samples were chipped into 5–8 pieces prior to milling. The chipped samples were milled with a Wiley mill (Thomas Scientific, Model 4, Swedesboro, NJ, United States) through a 20-mesh screen (1.0 mm particle size). This milled biomass was used for cell wall characterization and bioconversion analyses.

Cell Wall Characterization

Lignin content and composition were determined by pyrolysis molecular beam mass spectrometry (py-MBMS) using the NREL high-throughput method wherein soluble extractive and starch were removed from the biomass samples (Sykes et al., 2009; Decker et al., 2012). Lignin content was estimated from the relative intensities of the lignin precursor peaks. S/G lignin monomer ratio was determined by dividing the sum of the intensity of syringyl peaks by the sum of the intensity of guaiacyl peaks. Sugar release by enzymatic hydrolysis was determined using the NREL high-throughput method as previously described (Selig et al., 2010). Cell wall residues prepared by removing soluble extractives and starch were subjected to a hot water pretreatment (180°C for 17.5 min) followed by a 72-h incubation at 40°C with hydrolyzing enzymes. Glucose and xylose release were determined by colorimetric assays, and total sugar release is the sum of glucose + xylose released (Studer et al., 2009).

Ethanol Yield

Ethanol yield was determined by separate hydrolysis and fermentation (SHF) as described previously (Dumitrache et al., 2017). Biomass samples were incubated at 50°C for 5 days with hydrolyzing enzymes. The resulting sugars were fermented at 35°C for 72 h with *Saccharomyces cerevisiae* D5α (ATCC 200062). Ethanol yield was determined at the endpoint by HPLC quantification (Bio-Rad, Hercules, CA, United States).

RNA-Seq Analysis

The whole tillers and internode of the greenhouse-grown plants at the R1 developmental stage were harvested for RNA-seq analysis as described previously (Rao et al., 2019). RNA-seq was conducted at Joint Genome Institute (JGI) using Illumina TruSeq technology. Four biological replicates for each sample group were included. For each sample of the *PvFPGS1*-downregulated and

control lines, a total of 40–50 million paired-end (PE) reads of 150 bp was generated. Paired-end Illumina reads after filtering and trimming treatment were mapped to the Switchgrass genome *Panicum virgatum* v3.1 using HISAT2 (Kim et al., 2015) with default parameters. The selected genes were annotated with switchgrass genome v5.1[1]. Genes whose expression was different from the control were selected through comparison between each transgenic line and its control, using differential analysis software such as DESeq (Anders and Huber, 2010) with default settings of adjusted P-value < 0.05. Genes were annotated against Arabidopsis, rice and other model species using blast search.

Rust Disease Evaluation

Plants were evaluated for rust disease caused by *Puccinia novopanici* (formerly known as *Puccinia emaculata*) infection. Disease severity was assessed at weekly time points during the second (2015) and third (2016) growing seasons between July and August as described by Baxter et al. (2018). Two plants within each replicate were selected at random. A single tiller from each plant was tagged and all leaves on the selected tillers were examined for rust severity. The coverage of the top leaf surface with rust uredia was visually evaluated using the following scale: 0 = 0%, 1 $\leq$ 5%, 2 $\leq$ 10%, 3 $\leq$ 25%, 4 $\leq$ 40%, 5 $\leq$ 55%, 6 $\leq$ 70%, and 7 $\leq$ 100%. Because of the severity of the rust, the whole field was treated with fungicide in late August of each growing season. All data reported were collected before fungicide treatments. Fungicides used included "Quilt" (Syngenta Canada Inc., Guelph, ON, Canada) at a rate of 0.21 ml/m^2, and "Heritage" (Syngenta Crop Protection, Greensboro, NC, United States) at a rate of 20 ml/m^2.

Statistics

Means were analyzed with one-way ANOVA using Fisher's least significant difference method in SAS version 9.4 (SAS Institute Inc., Cary, NC, United States). Differences were considered statistically significant where P-values were less than 0.05.

RESULTS

Identification of PvFPGS Homologs

Switchgrass FPGS (PvFPGS) was initially identified using the FPGS amino acid sequences from *Arabidopsis thaliana* (At5g05980). Analysis of the sequences showed that switchgrass assembly v1.1 has only three isoforms of FPGS: Pavir.Ib00114.1, Pavir.Ib03621.1, and Pavir.Ia04781.1. A phylogenetic tree was constructed using all three isoforms of switchgrass and Arabidopsis as well as FPGS variants from several other species. These species, *Populus trichocarpa*, *Medicago truncatula*, *Oryza sativa*, *Zea mays*, and *Panicum hallii* from Phytozome 12, served as reference to elucidate homologous relationship among different members of FPGSs. Based on homology analysis, Pavir.Ib00114 was identified and named PvFPGS1 (**Figure 1**). There are two other isoforms of FPGS in switchgrass genome: *PvFPGS2* (Pavir.Ia04781.1) and *PvFPGS3* (Pavir.Ib03621.1). Based on this whole genome level phylogeny analysis of FPGSs, the relationship of FPGSs between monocot and dicot is not as straightforward as one to one, especially AtFPGS3 had no isologs in other species. However, among all three FPGSs between Arabidopsis and switchgrass, AtFPGS1 is still closest to PvFPGS1 with an identity of 59.4%. Moreover, *PvFPGS1* is isolog of maize *brown midrib4* (*bm4*) which encode a functional FPGS and its loss-of-function leads to lower lignin content (Li et al., 2015).

[1]https://phytozome-next.jgi.doe.gov/

Expression Patterns of *PvFPGS1* in Switchgrass

Transcript abundance via qRT-PCR analysis indicated that *PvFPGS1* in non-transgenic plants is expressed in stems, leaves, crown, inflorescences, and roots at the R1 developmental stage. The level of *PvFPGS1* expression was highest in the crown and leaves, and lowest in the root (**Figure 2**). *PvFPGS1* transcripts were detected in all the tissue types tested. In the transgenic plants, depending on the transgenic event, *PvFPGS1* transcript abundance was reduced by 23–82% in the joint leaf blade, leaf sheath, and 2nd internode tissues of the RNAi-transgenic lines (**Supplementary Figure S4**). Regardless of the level of reduction in *PvFPGS1* expression, the growth and phenotype of transgenic lines was not apparently different from non-transgenic controls under greenhouse conditions (**Figure 3**).

Field Experiments

From greenhouse studies, we selected six independent transgenic lines for the field experiment, which had a range of decreased *PvFPGS1* expression. The plants included the transgenic lines T2 (decreased expression by 71%), T8 (decreased expression by 72%), T10 (decreased expression by 73%), T12 (decreased expression by 82%), T32 (decreased expression by 67%), T115 (decreased expression by 63%), and one non-transgenic control. The field study was conducted for three consecutive growing seasons (2014–2016) (**Figure 4**).

PvFPGS1 Gene Expression Under Field Conditions

Transgene expression in the field-grown *PvFPGS1*-downregulated switchgrass lines was studied by qRT-PCR on sequentially harvested tissue over the course of the field study. *PvFPGS1* transcript levels were reduced in the transgenic lines for field year one (2014), year two (2015), and year three (2016) compared to the control. The *PvFPGS1* expression was lowest in transgenic line T8 with 80–89% decrease in transcript level compared to the control in years one, two, and three. The decrease in *PvFPGS1* transcript level was followed by transgenic lines T2 (67–83%), T10 (76–81%), T12 (78–81%), T32 (72–77%), and T115 (by 58–86%) relative to the control during the three growing seasons (**Figure 5**).

Agronomic Performance

For each growing season, the following end-of-year growth characteristics were assessed: tiller height, plant width, tiller number, and aboveground dry biomass. Biomass production of transgenic lines T10, T12, and T32 was comparable to that of

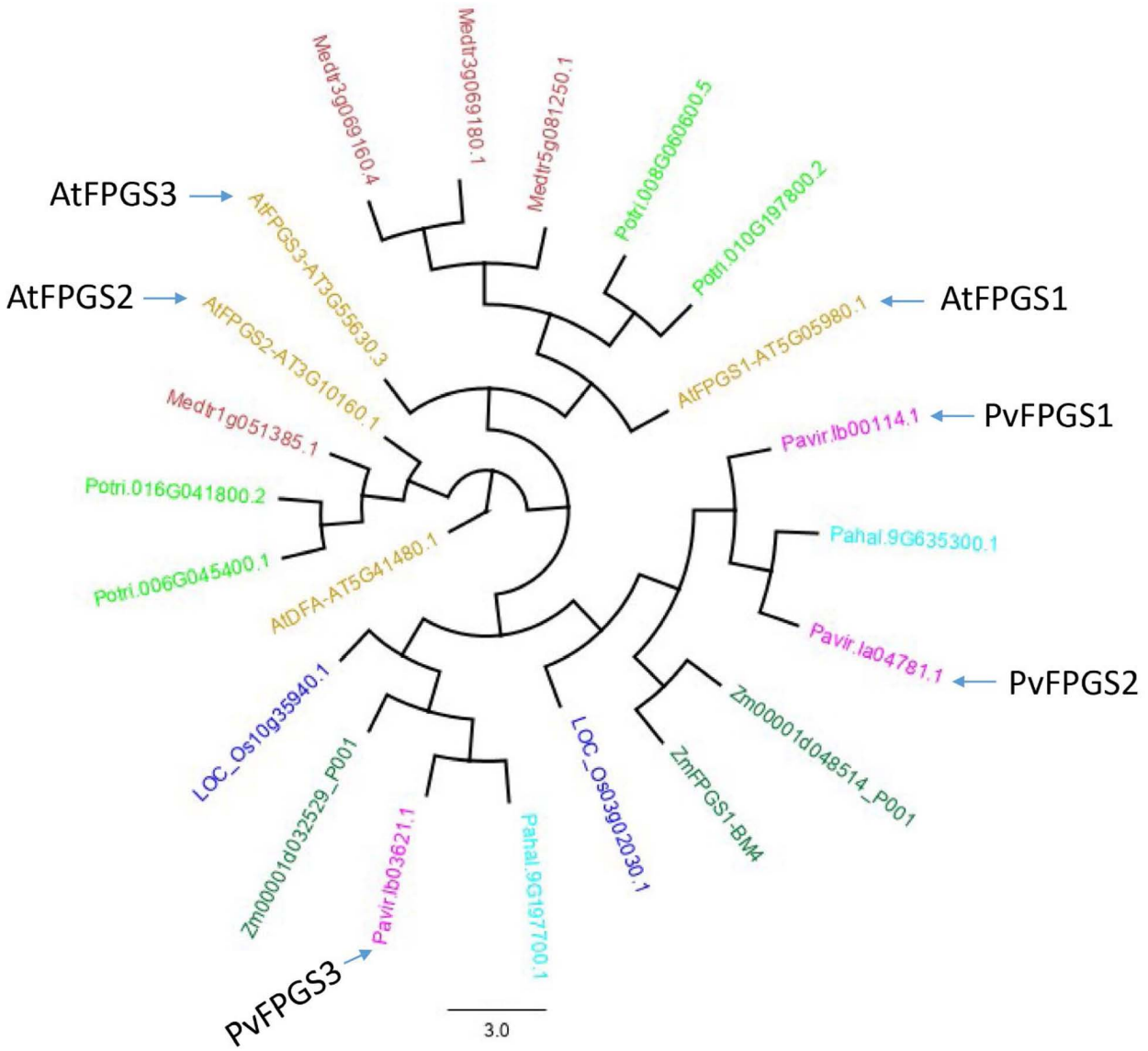

FIGURE 1 | The gene tree of the FPGS protein family. Various members are shown by plant species: *Arabidopsis thaliana* (gold), *Populus trichocarpa* (light green), *Medicago truncatula* (brown), *Oryza sativa* (blue), *Zea mays* (green), *Panicum hallii* (light blue), and *Panicum virgatum* (pink) from Phytozome 12 (https://phytozome.jgi.doe.gov/) showing relationship based on amino acid sequences. The phylogenetic analysis shows that switchgrass genome has three isoforms of FPGS: *PvFPGS1* (Pavir.Ib00114.1), *PvFPGS2* (Pavir.Ia04781.1), and *PvFPGS3* (Pavir.Ib03621.1). Based on this whole genome level comparison, the relationship of FPGSs between monocot and dicot is not as straightforward. However, among all three FPGSs in Arabidopsis and switchgrass, AtFPGS1 is still closest to PvFPGS1. In this graph, AtDFA encodes a dihydrofolate synthetase, homologous to FPGS but with different functions and serve as outgroup for this tree.

the control in the first year (2014), while lines T2 (decreased biomass by 45%), T8 (decreased biomass by 81%), and T115 (decreased biomass by 62%) exhibited yield reduction. The yield reductions in these lines were congruent with decreased tiller height, plant width, and tiller number by up to 56%. In the second year (2015), the transgenic lines showed similar biomass yield and growth traits to that of the control with the exception of line T8, in which only one plant (out of 12 plants) survived the first field winter. The sole T8 survival had decreased biomass by 95% accompanied by up to 69% decrease in tiller height, plant width, and tiller number compared to the control. Similar to the second year, transgenic lines did not differ from the control in biomass production in the third year (2016), except for the T8 surviving plant which had a 99% decrease in biomass yield accompanied by up to 94% reduction in tiller height, plant width, and tiller number relative to the control (**Table 1**).

Lignin Content and Composition

Cell wall lignin content and the S/G ratio were measured for aboveground biomass harvested at end-of-season of each year by pyrolysis molecular beam mass spectrometry (py-MBMS). Decrease in lignin content (14% reduction) and S/G ratio (18%

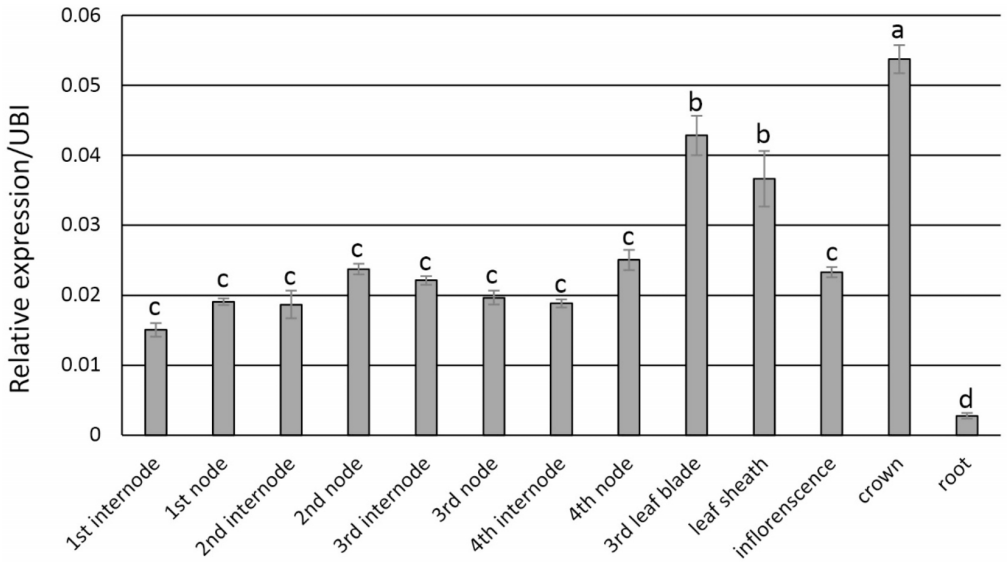

FIGURE 2 | Expression patterns of *PvFPGS1* in different plant tissues as determined by qRT-PCR. Plant samples for RNA extraction used in the qRT-PCR experiments were collected at R1 (reproductive stage 1) developmental stage. The relative levels of transcripts were normalized to the switchgrass ubiquitin 1 gene expression (UBI). Bars represent mean values of three biological replicates ± standard error. Means were compared by a one-way ANOVA and letter groupings were obtained using Fisher's least significant difference method. Bars with different letters are significantly different at the 5% level.

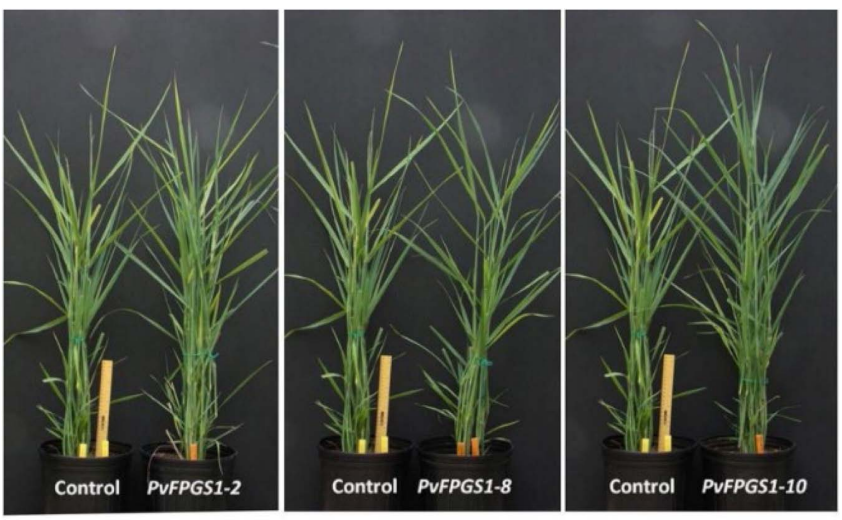

FIGURE 3 | Representative *PvFPGS1*-RNAi transgenic and non-transgenic control lines grown under greenhouse conditions at 3 months old.

reduction) were only observed in line T8 relative to the control in year one (2014). There were no significant differences in lignin content or S/G ratio between all the other transgenic lines and the control lines at the first (2014), second (2015), and third (2016) growing seasons (**Table 2**).

Sugar Release Efficiency

Enzymatic hydrolysis was used to determine the sugar release for aboveground biomass harvested at end of each season. In year one (2014), line T8 exhibited a decrease in xylose (13%) and total sugar (8%) release relative to the control, whereas line T12 had a 10% higher xylose and total sugar release than control. In year two (2015), a decrease in xylose (12%) and total sugar (10%) release was only observed in line T2. In year three (2016), there was an increase in xylose release in lines T10 (7%) and T12 (5%). There were no significant differences in sugar release between the other transgenic and the control lines at the first (2014), second (2015), and third (2016) growing seasons (**Table 2**).

Ethanol Yield

The fermentation potential of aboveground biomass harvested at end-of-season of each year was determined by separate hydrolysis

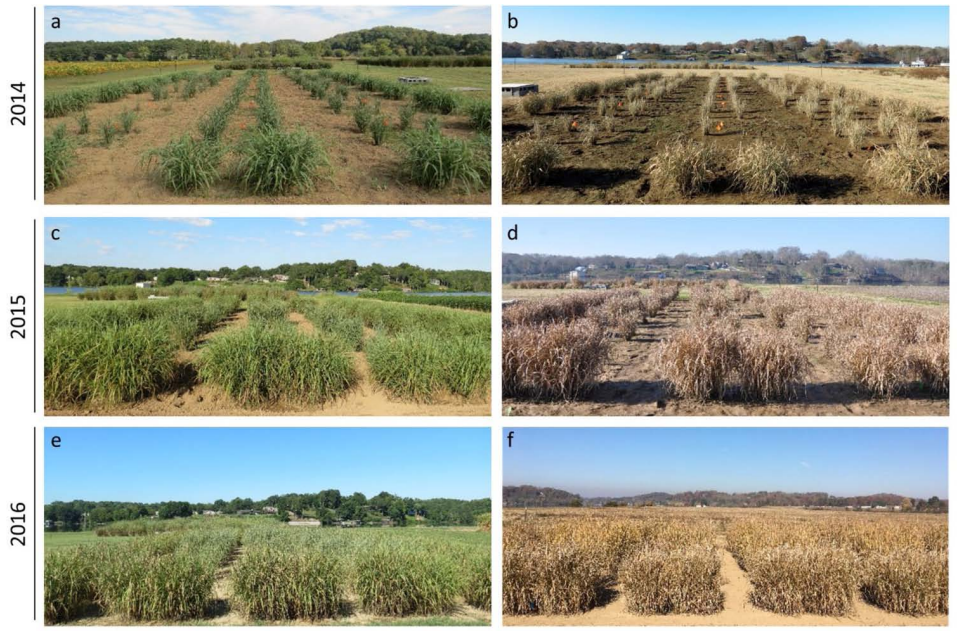

FIGURE 4 | Photos of *FPGS1*-downregulated switchgrass in the field. **(a)** October 01, 2014; **(b)** November 21, 2014; **(c)** August 26, 2015; **(d)** December 08, 2015; **(e)** August 16, 2016; **(f)** November 17, 2016.

and fermentation (SHF). Lines T2 and T10 with moderate levels of the decreased *PvFPGS1* expression were selected to evaluate bioconversion efficiency. Line T10 showed an increase in ethanol yield in year one (2014) by 7%, year two (2015) by 18%, and year three (2016) by 9% relative to the control. There were no differences in ethanol yield between line T2 and the control at the three growing seasons (**Figure 6**).

RNA-Seq Analysis

Transcriptome analysis by RNA-seq was conducted to identify underlying pathways through which modification may further improve biofuel production. In both tiller and internode, the number of putative genes of different expression in the *PvFPGS1*-downregulated line T10 compared with control was similar, with 148 higher and 145 lower in tiller while 134 higher and 167 lower in the internode (**Supplementary Table S1**). FPGS1 has two subtypes presented in the annotated switchgrass genome, one in K genome and one in the N genome, each was named based on their subgenome, namely FPGS1K (Pavir.9KG559800 V5.1, originally Pavir.Ib00114.1 in V1.1), and FPGS1N (Pavir.9NG789900 V5.1, originally Pavir.Ia04781.1 in V1.1). Sequence identity between these two subtypes is 96.7%, and both were significantly reduced to 20–30% of the control levels, indicating these target genes were successfully downregulated in line T10. Reads for switchgrass FPGS2 is not well mapped in this RNA-seq data and its expression pattern will be worth looking at in future study. In total, there were 565 genes that were differentially expressed in line T10. Based on gene annotation, 38 differentially expressed genes were putatively involved in activities related to plant cell wall (**Supplementary Table S2**). Among them, in addition to the FPGS1 target genes, seven genes encode enzymes in the flavonoid and aromatic amino acids synthesis pathways which directly interact with the phenylpropanoid pathway, the upstream of lignin biosynthesis pathway. These genes include *chalcone synthase* (*CHS*) (Pavir.8NG107400 and Pavir.8KG261900), *chalcone isomerase* (Pavir.9NG037000), *chorismate mutase* (Pavir.1KG086400), *isochorismate synthase* (Pavir.2KG294000), *anthocyanidin reductase* (Pavir.7NG368500), and *isoflavone 7-O-glucosyltransferase* (Pavir.9NG095000). Peroxidase and laccase are enzymes involved in lignin polymerization and deposition (Tobimatsu and Schuetz, 2019). Three peroxidase (Pavir.2KG513700, Pavir.2NG638600, and Pavir.7NG321900) and three laccase genes (Pavir.5KG623300, Pavir.5NG585100, and Pavir.5KG613400) are among the selected genes. One gene potentially involved in monolignol metabolism, monolignol beta-glucoside homolog without catalytic acid/base, also present in the selected gene list (Pavir.7NG350700). Ten genes involved in different steps of some type of sugar and wall polysaccharide synthesis and modification are also among the selected genes. At the upstream of the pathway, two genes involved in myo-inositol metabolism have reduced expression in line T10, *myo-inositol-1-phosphate synthase* (Pavir.9KG643000) and *myo-inositol 2-dehydrogenase* (Pavir.2NG256100). Following these steps, four UDP nucleotide sugar metabolism enzymes are also among the mostly downregulated gene list: *UDP-arabinose 4-epimerase* (Pavir.9KG554500), *UDP-arabinopyranose mutase 1 related* (Pavir.9NG214400), *UDP-glycosyltransferase 73B4* (Pavir.7NG278900), *UDP-glucuronic acid decarboxylase* (Pavir.9KG487800). Genes involved in pectin and hemicellulose also are among genes of changed expression in line T10. Among them are *polygalacturonate 4-alpha-galacturonosyltransferase*

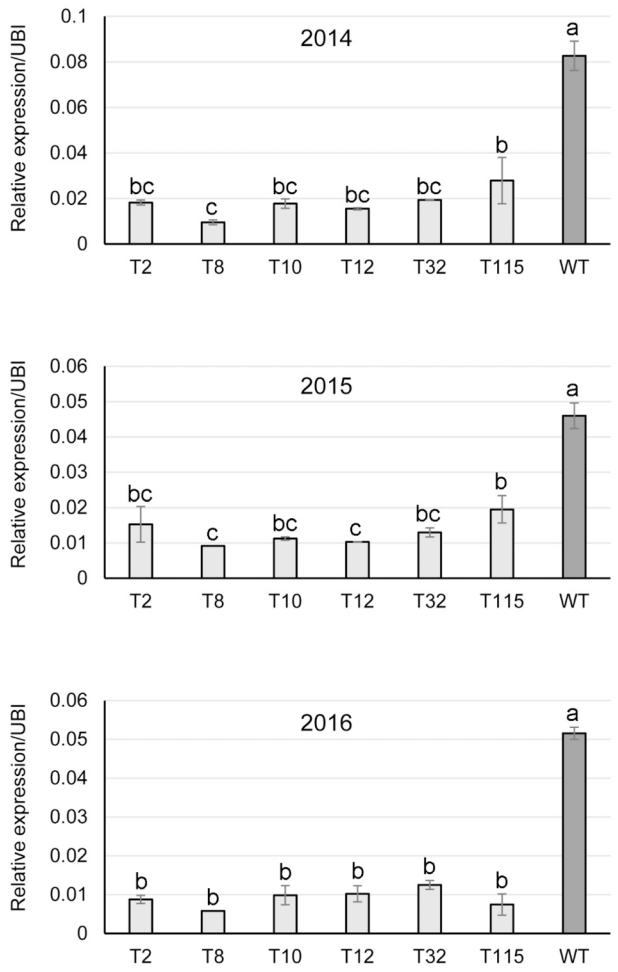

FIGURE 5 | Relative transcript levels of *PvFPGS1* in RNAi-transgenic lines as determined by qRT-PCR. Samples were collected in year one (2014), year two (2015) and year 3 (2016) of the field trial. WT: non-transgenic control. The relative levels of transcripts were normalized to the switchgrass ubiquitin 1 gene expression (UBI). Bars represent mean values of three biological replicates ± standard error, with the exception of line T8 in 2015 and 2016, which is from one surviving plant during the two growing seasons. Means within each year were compared by a one-way ANOVA and letter groupings were obtained using Fisher's least significant difference method. Bars with different letters are significantly different at the 5% level.

(*GAUT4*) (Pavir.2KG409000), *cellulose synthase* (*CESA8*) (Pavir.2KG167600), *licheninase/mixed linkage beta-glucanase* (Pavir.3NG120500); three glycotransferase (Pavir.5NG133600, Pavir.1KG262700, and Pavir.4KG337800) are all downregulated in line T10; the last group of genes on the list that are directly involved in cell wall structure are five wall associated proteins: four hydroxyproline-rich glycoprotein family protein (Pavir.2KG047700, Pavir.2KG047500, Pavir.3KG332300, and Pavir.9NG522300); the other one is *wall-associated kinase 2* (Pavir.7NG318000). One interesting gene that is downregulated in the internode of line T10 encodes an EamA-like transporter family member (Pavir.5KG129300), homology to nodulin (MtN21) and walls are thin 1 (WAT1), which is a tonoplast-localized protein required for secondary wall formation in fibers (Ranocha et al., 2010).

Disease Susceptibility

All plants were evaluated for rust disease severity during the second (2015) and third (2016) growing seasons. Rust disease was detected in late July and infection was advanced largely through late August for both growing seasons. In order to maintain the field for downstream analyses, the field was sprayed with fungicides. Disease severity of rust infection was rated weekly prior to fungicide treatments based on the percentage of the leaf area coverage with rust uredia. The disease severity ranged from 0 to 9% during the second (2015) and 4–18% during the third (2016) growing seasons. There were no significant differences in rust susceptibility between transgenic and control lines for both growing seasons (**Supplementary Figures S5**, **S6**).

DISCUSSION

A clear understanding of cell wall enzymology is needed to engineer reduced recalcitrance in bioenergy crops. We previously reported that the loss of function of Arabidopsis FPGS1 resulted in lignin reduction and improved saccharification efficiency in Arabidopsis (Srivastava et al., 2015; Xie et al., 2019). Given the link between the FPGS and lignin biosynthetic pathways, our findings gave rise to the hypothesis that FPGS plays a functional role in reducing recalcitrance in switchgrass, a leading lignocellulosic bioenergy feedstock. The present study describes identification of switchgrass *FPGS* gene (*PvFPGS1*). Our study strengthens the FPGS-recalcitrance hypothesis.

In Arabidopsis, there are three genes that each encode FPGS isoforms: FPGS1, FPGS2, and FPGS3 (Ravanel et al., 2001). The homologous sequence in switchgrass discovered using AtFPGS1 was named PvFPGS1. At whole genome level, PvFPGS1 has two subtypes, each belong to the N and K subgenome, they show similarity closer to AtFPGS1 and AtFPGS3, farther from AtFPGS2.

Endogenous expression of *PvFPGS1* transcripts was highest in the crown of plants. Transcript abundance decreased sequentially in leaves, internodes, inflorescences, and roots. The expression profile of *PvFPGS1* is slightly different from Arabidopsis as the expression profile of Arabidopsis *AtFPGS1* showed the highest expression in stems compared to other tissues (Srivastava et al., 2015). Secondary cell walls consisted of high amount of lignin provide much of the rigidity in stem tissue, in contrast to the flexible organs such as roots. Consistently, expression of secondary cell wall-related genes has been shown to be higher in stems (Zhao and Bartley, 2014; Mazarei et al., 2018; Rao et al., 2019). Noteworthy, other than inflorescence, the tissues of high *PvFPGS1* expression in switchgrass are all significantly lignified (Crowe et al., 2017).

All transgenic *PvFPGS1*-downregulated lines had reduced targeted transcript levels over the three growing seasons of the field trial. These results confirm that the *PvFPGS1*-downregulated lines grown under field conditions sustained the reduction in *PvFPGS1* transcript levels compared to the control.

TABLE 1 | Morphology and dry weight yield of *PvFPGS1*-downregulated lines at the first (2014), second (2015), and third (2016) growing seasons per plant.

Year	Line	Tiller height (cm)	Plant width (cm)	Tiller number	Dry weight yield (g/plant)
2014	T2	**44.5 ± 14.3^{c}**	**36.7 ± 12.1^{b}**	**31.0 ± 2.6^{b}**	**36.0 ± 2.4^{b}**
	T8	**56.0 ± 0.6^{bc}**	**41.7 ± 2.4^{b}**	**28.8 ± 5.7^{b}**	**12.3 ± 2.4^{c}**
	T10	80.0 ± 1.7^{a}	78.5 ± 1.5^{a}	58.4 ± 7.4^{a}	59.7 ± 2.9^{a}
	T12	82.3 ± 2.3^{a}	76.2 ± 6.5^{a}	54.8 ± 2.4^{a}	67.3 ± 7.5^{a}
	T32	86.6 ± 2.0^{a}	91.7 ± 1.8^{a}	58.1 ± 3.8^{a}	83.7 ± 3.4^{a}
	T115	69.7 ± 10.0^{ab}	**46.1 ± 12.1^{b}**	**22.4 ± 8.3^{b}**	**25.0 ± 8.5^{b}**
	WT	78.4 ± 0.8^{a}	83.4 ± 5.1^{a}	60.4 ± 3.1^{a}	65.8 ± 6.3^{a}
2015	T2	101.8 ± 1.3^{ab}	106.4 ± 14.6^{a}	80.4 ± 16.1^{a}	178.6 ± 36.3^{a}
	T8*	**63.5 ± 0.0^{c}**	**35.6 ± 0.0^{b}**	**34.0 ± 0.0^{b}**	**10.0 ± 0.0^{b}**
	T10	104.7 ± 6.6^{ab}	114.7 ± 7.4^{a}	118.1 ± 12.7^{a}	204.2 ± 20.4^{a}
	T12	116.2 ± 12.0^{a}	128.9 ± 22.3^{a}	110.3 ± 20.8^{a}	311.7 ± 114.8^{a}
	T32	107.1 ± 6.7^{a}	133.4 ± 8.6^{a}	113.8 ± 6.5^{a}	308.9 ± 64.0^{a}
	T115	88.9 ± 2.9^{b}	99.1 ± 22.6^{a}	77.1 ± 22.6^{a}	142.5 ± 48.5^{a}
	WT	100.4 ± 3.5^{ab}	116.3 ± 1.0^{a}	96.4 ± 13.2^{a}	206.1 ± 11.5^{a}
2016	T2	**127.7 ± 9.0^{b}**	189.2 ± 56.5^{a}	96.9 ± 7.8^{a}	446.8 ± 98.0^{a}
	T8*	**34.3 ± 0.0^{c}**	**15.2 ± 0.0^{b}**	**11.0 ± 0.0^{b}**	**4.5 ± 0.0^{b}**
	T10	152.5 ± 4.5^{a}	208.1 ± 8.8^{a}	121.5 ± 6.0^{a}	593.3 ± 68.8^{a}
	T12	162.8 ± 6.3^{a}	220.8 ± 34.7^{a}	125.4 ± 21.1^{a}	741.3 ± 211.5^{a}
	T32	153.1 ± 3.1^{a}	226.6 ± 5.9^{a}	135.7 ± 5.5^{a}	782.4 ± 38.5^{a}
	T115	**136.3 ± 4.2^{b}**	185.0 ± 29.5^{a}	97.2 ± 25.6^{a}	378.5 ± 65.4^{a}
	WT	161.8 ± 5.7^{a}	242.7 ± 13.1^{a}	119.5 ± 12.5^{a}	752.4 ± 68.8^{a}

*Values are the mean of three biological replicates for each transgenic line (T2, T8, T10, T12, T32, and T115) and wild-type control (WT) ± standard error, with the exception of line T8 in 2015 and 2016, which is from one surviving plant during the two growing seasons. Means within each year were compared by a one-way ANOVA and letter groupings were obtained using Fisher's least significant difference method. Values followed by different letters are significantly different at the 5% level. Bold values are significantly different from controls. *Only one surviving plant.*

For most of transgenic lines, plant growth was comparable to the control plants. However, in the field, the transgenic *PvFPGS1*-downregulating line T8, which had the highest decrease in *PvFPGS1* transcript levels (up to 89%) among lines either did not survive or had up to 99% reduction in biomass production over the control. The reduced biomass yield was accompanied with significant decrease in tiller height, plant width, and tiller number relative to the control. Yet, the transgenic line T8, which had the highest reduction in *PvFPGS1* expression had normal growth under greenhouse conditions. These observations that the transgenic lines with normal plant growth and development under greenhouse conditions could not survive or had substantial biomass loss under field conditions emphasize the importance of performing field studies. Likewise, an association between levels of transgene expression and biomass production for switchgrass grown under field conditions has been shown for transgenic *PvMYB4*-overexpressing lines, where the transgenic lines with higher expression levels of *PvMYB4* did not survive and/or had substantial reduction in biomass relative to control when grown under field conditions (Baxter et al., 2015). Yet, transgenic lines with low to moderate decreases in *PvFPGS1* expression levels produced biomass yield comparable to that of the controls. These observations highlight the significance of an optimized level of gene expression in transgenic plants.

Lignin content and S/G ratios were not significantly different in low-to-moderate downregulated *PvFPGS1* transgenic lines. Similarly, these transgenic lines showed similar sugar release efficiency compared with the control lines. Although our results are in contrast to Arabidopsis FPGS1 findings in which a loss of function was associated with reduced lignin content, coupled with an increase in sugar release efficiency (Srivastava et al., 2015; Xie et al., 2019), they used homozygous null-mutants in their studies. Similar observations were shown in maize where disruption in a gene encoding the isolog of PvFPGS1 resulted in low lignin (Li et al., 2015). We observed similar reductions in lignin in the T8 line where expression of FPGS1 was reduced to 80%. Cell wall lignin is one of the barriers to lignocellulosic biofuel production limiting digestion of cellulose into fermentable sugars. The manipulation of cell walls to decrease lignin content has been shown to improve bioconversion efficiency (Chen and Dixon, 2007; Baxter et al., 2014, 2015; Bonawitz et al., 2014; Eudes et al., 2014; Wilkerson et al., 2014; Hu et al., 2018). Since we did not find any noticeable difference in lignin content and composition in the selected *PvFPGS1*-downregulated lines, it is possible that the difference is very subtle and it is noticeable when we significantly disrupt the FPGS gene and associated supplies to the methylation pathway. This may also explain why there were not significant differences in sugar release

TABLE 2 | Lignin content, S/G ratios, and sugar release of *PvFPGS1*-downregulated lines at the first (2014), second (2015), and third (2016) growing seasons.

Year	Line	Lignin content (% CWR)	*S/G* ratio	Glucose release (mg/g CWR)	Xylose release (mg/g CWR)	Total sugar release (g/g CWR)
2014	T2	20.4 ± 0.3^{b}	0.59 ± 0.01^{ab}	0.219 ± 0.00^{a}	0.178 ± 0.00^{b}	0.396 ± 0.00^{b}
	T8	**18.3 ± 0.4^{c}**	**0.49 ± 0.02^{c}**	0.205 ± 0.00^{a}	**0.154 ± 0.00^{c}**	**0.359 ± 0.01^{c}**
	T10	20.8 ± 0.2^{b}	0.56 ± 0.01^{b}	0.228 ± 0.00^{a}	0.183 ± 0.00^{b}	0.410 ± 0.00^{ab}
	T12	21.2 ± 0.2^{ab}	0.58 ± 0.02^{ab}	0.236 ± 0.01^{a}	**0.195 ± 0.00^{a}**	**0.430 ± 0.01^{a}**
	T32	22.0 ± 0.2^{a}	0.61 ± 0.01^{a}	0.216 ± 0.01^{a}	0.185 ± 0.01^{ab}	0.401 ± 0.01^{b}
	T115	20.4 ± 0.4^{b}	0.57 ± 0.02^{ab}	0.238 ± 0.01^{a}	0.177 ± 0.00^{b}	0.414 ± 0.02^{ab}
	WT	21.1 ± 0.3^{ab}	0.59 ± 0.01^{ab}	0.215 ± 0.00^{a}	0.177 ± 0.00^{b}	0.392 ± 0.00^{b}
2015	T2	20.8 ± 0.2^{a}	0.68 ± 0.02^{a}	0.190 ± 0.00^{c}	**0.183 ± 0.01^{c}**	**0.374 ± 0.00^{c}**
	T8*	20.8 ± 0.0^{a}	0.64 ± 0.00^{a}	0.229 ± 0.00^{a}	0.205 ± 0.00^{ab}	0.434 ± 0.00^{a}
	T10	21.0 ± 0.1^{a}	0.67 ± 0.00^{a}	0.216 ± 0.01^{ab}	0.211 ± 0.01^{a}	0.427 ± 0.02^{a}
	T12	21.5 ± 0.5^{a}	0.65 ± 0.02^{a}	0.205 ± 0.00^{bc}	0.209 ± 0.01^{a}	0.414 ± 0.00^{ab}
	T32	22.3 ± 0.5^{a}	0.70 ± 0.01^{a}	0.201 ± 0.01^{bc}	0.209 ± 0.00^{a}	0.409 ± 0.01^{ab}
	T115	21.3 ± 0.4^{a}	0.71 ± 0.02^{a}	0.202 ± 0.00^{bc}	0.191 ± 0.01^{bc}	0.393 ± 0.01^{bc}
	WT	21.5 ± 0.4^{a}	0.68 ± 0.02^{a}	0.205 ± 0.01^{bc}	0.209 ± 0.01^{a}	0.414 ± 0.01^{ab}
2016	T2	22.8 ± 0.3^{a}	0.78 ± 0.02^{a}	0.151 ± 0.01^{a}	0.167 ± 0.00^{ab}	0.318 ± 0.01^{a}
	T8*	ND	ND	ND	ND	ND
	T10	22.3 ± 0.4^{a}	0.75 ± 0.02^{a}	0.155 ± 0.00^{a}	**0.174 ± 0.00^{a}**	0.329 ± 0.00^{a}
	T12	22.7 ± 0.1^{a}	0.71 ± 0.01^{a}	0.149 ± 0.01^{a}	**0.172 ± 0.00^{a}**	0.321 ± 0.01^{a}
	T32	23.0 ± 0.2^{a}	0.77 ± 0.02^{a}	0.140 ± 0.00^{a}	0.168 ± 0.00^{ab}	0.308 ± 0.00^{a}
	T115	22.7 ± 0.6^{a}	0.79 ± 0.05^{a}	0.142 ± 0.00^{a}	0.162 ± 0.00^{b}	0.304 ± 0.00^{a}
	WT	23.4 ± 0.2^{a}	0.78 ± 0.01^{a}	0.142 ± 0.01^{a}	0.163 ± 0.01^{b}	0.306 ± 0.01^{a}

*Values are the mean of the biological replicates (n = 3) for each transgenic line (T2, T8, T10, T12, T32, and T115) and wild-type control (WT) ± standard error, with the exception of line T8 in 2015 and 2016, which is from one surviving plant during the two growing seasons. Means within each year were compared by a one-way ANOVA and letter groupings were obtained using Fisher's least significant difference method. Values followed by different letters are significantly different at the 5% level. Bold values are significantly different from controls. *Only one surviving plant. ND, not done due to unavailability of material needed for the analysis. CWR: cell wall residue.*

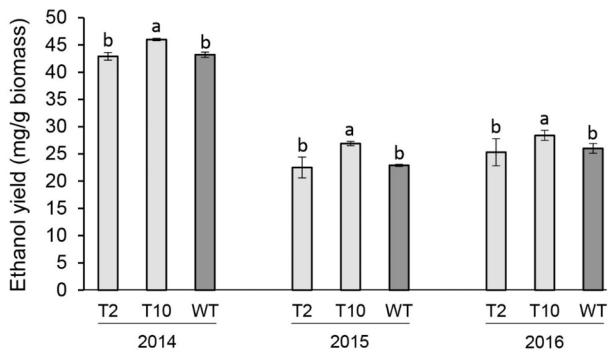

FIGURE 6 | Ethanol yields of *FPGS1*-downregulated switchgrass in the first (2014), second (2015), and third (2016) growing seasons. Bars are the mean value of three biological replicates for each transgenic line (T2 and T10) and wild-type control (WT) ± standard error. Means within each year were compared by a one-way ANOVA and letter groupings were obtained using Fisher's least significant difference method. Bars with different letters are significantly different at the 5% level.

efficiency between the selected *PvFPGS1*-downregulated lines and the controls.

Nevertheless, the *PvFPGS1*-downregulated line T10 produced up to 18% more ethanol than controls over the course of three field seasons. These observations may suggest that *PvFPGS1* expression is associated with improved biofuel production. Transcriptome data provides some clues to the increased ethanol production in the transgenic line T10. In grass plant, lignin can be synthesized from both phenylalanine and tyrosine and it is tightly linked to other secondary metabolite pathways, and tightly regulated through metabolites flux (Recent review by Vanholme et al., 2019). The secondary metabolites that derived from the same upper biochemical pathways but branched off to produce different type of flavonoids including anthocyanidin[2,3]. Genes in several steps leading to the synthesis of anthocyanidin show affected expression in line T10 indicating that this metabolite pathway could be affected. Work has shown that disturbance in flavone synthase II in rice (Lam et al., 2017) and CHS in maize (Eloy et al., 2017) have changed lignin content and digestibility. Increased level of *CHS* in line T10 with increased ethanol production is in agreement with this evidence. Even though the lignin content was not significantly different in line T10, decreased transcript levels in several laccases and peroxidases hint that the level of polymerization of lignin could be reduced, which could contribute to its increased ethanol production.

Furthermore, among the enzymes involved in wall simple sugar to polysaccharide synthesis and wall modifications, a number of the genes show affected expression levels in line T10. Wall polysaccharide formed a network of

[2]https://www.genome.jp/kegg/kegg2.html

[3]https://en.wikipedia.org/wiki/Anthocyanin#Biosynthesis

cellulosic, hemicellulosic, and pectic polysaccharides and protein (Bashline et al., 2014) and changes are dynamic. Although not all expression changes will have consequential effects as the plants maintain normal growth and have similar sugar release, however, changes in some steps have led accumulated changes that have contributed to the increased ethanol productions in line T10. In rice, mutants of a putative glycosyltransferase in a grass-specific subfamily of GT61, are deficient in ferulic acid, coumaric acid, and aromatic compounds and exhibit an increased saccharification efficiency (Chiniquy et al., 2012). It is worth noting that one of the genes annotated as glycosyltransferase family 61 protein among the selected gene list. Lower expression of this gene could indirectly affect the level of these small metabolites and hence could contribute to the higher level of ethanol in line T10. The other gene that could be of significant effect is gene encoding licheninase/mixed linkage beta-glucanase. This enzyme could cause release of smaller oligosaccharides (DP $<$ 6) from graminaceous hemicelluloses (Yoshida and Komae, 2006). Some smaller oligosaccharides were shown to be recalcitrant to fermentation (Jonathan et al., 2017). In the line T10, expression level of licheninase is reduced and thus could add another factor that potentially contribute to its higher ethanol yield.

Besides, phenolic composition of switchgrass has been shown to be another important factor affecting recalcitrance (Tschaplinski et al., 2012; Yee et al., 2012; Shen et al., 2013; Baxter et al., 2015). Studies with *PvCOMT*-downregulated switchgrass plants had higher levels of certain phenolic compounds which inhibited the microbial fermentation (Tschaplinski et al., 2012; Yee et al., 2012). In contrast, *PvMYB4*-overexpressing switchgrass plants grown in the greenhouse and field had lower amounts of phenolic compounds that inhibit microbial fermentation (Shen et al., 2013; Baxter et al., 2015). Interestingly, one *PvMYB4*-overexpressing line showed no increase in sugar release efficiency but still resulted in higher levels of ethanol production when grown in the field (Baxter et al., 2015). These studies suggest that there are factors other than lignin, e.g., phenolic compounds, that play roles in improved biofuel production observed in transgenic switchgrass overexpressing *PvMYB4* (Shen et al., 2013; Baxter et al., 2015). Given that C1 metabolism pathway is involved in synthesis of variety of polymers and plant secondary metabolites, it is tempting to speculate that the increased ethanol production in the *PvFPGS1*-downregulated line may be caused by changes in phenolic fermentation inhibitors. Elucidating of this mechanism will be the subject of future study.

Successful establishment and sustainability of bioenergy feedstocks are key factors for production of fuel from biomass (Stewart and Cromey, 2011). Performing field studies of transgenic plants is crucial to examine consequences of their genetic modifications on plant defenses. Of particular significance is the rust disease caused by fungal pathogen *P. novopanici* that is identified as potentially damaging to switchgrass fields (Uppalapati et al., 2013). During the last two growing seasons, rust severity in *PvFPGS1*-downregulated lines was not different from that of control plants. Levels of *PvFPGS1* expression appears to not be a factor in switchgrass rust disease.

In conclusion, we have shown that genetic manipulation of *PvFPGS1* could lead to improved biofuel production without negatively impacting plant growth and biomass yield. These results provide further insights into the effect of knockdown expression of *PvFPGS1* on improving biofuel production in switchgrass. As of interest is to use *PvFPGS1* in complementation studies of the Arabidopsis and maize mutants with loss of function of native FPGS orthologs. This could further elucidate functionality of FPGS and provide information on potential strategies to enhance productivity in bioenergy crops. Further research on the FPGS genes enhance understanding of the factors associated with reducing recalcitrance. Our study provides a starting point for a more rigorous exploration of the role of *PvFPGS1* in the bioenergy field.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

MM designed the experiments, participated in characterization of field-grown plants and preparation of plant samples for cell wall analysis, analyzed the data, and wrote the manuscript. HB performed rust disease phenotyping and statistical analysis, and participated in preparation of plant samples for cell wall analysis. AS performed transgenic lines generation and phylogenetic tree work. GL and HX performed gene expression analysis. AD, MR, JN, and BD performed ethanol yield analysis. J-YZ and MU performed cloning of the target gene. GT, RS, and MD performed lignin and sugar release analyses. Z-YW produced the transgenic plants. YT performed RNA-seq analysis. EB and YT conceived of the experimental approach and made significant intellectual contributions about the target gene and cell wall biology. CS conceived of the field study and its design and coordination, and assisted with interpretation of results and revisions to the manuscript. All authors contributed to text and data analysis. All authors read and approved the final manuscript.

FUNDING

This work was primarily supported by BioEnergy Science Center, and secondarily by the Center for Bioenergy Innovation. The BioEnergy Science Center and the Center for Bioenergy Innovation are United States Department of Energy Bioenergy Research Centers, supported by the Office of Biological and Environmental Research in the Department of Energy's Office of Science. This manuscript has been partially authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the DOE. Funding was also provided by the Ivan Racheff Endowment and a USDA Hatch grant to CS.

ACKNOWLEDGMENTS

We greatly thank Joint Genome Institute (JGI) for RNA-seq data generation and Xin Chen and Chunman Zuo for assistance with RNA-seq data analysis. We thank Crissa Doeppke and Melvin Tucker for assistance with cell wall characterization. We also thank Ben Wolfe, Marcus Laxton, and the UT field staff for assistance with data collection and general field maintenance, and Reggie Millwood for assistance with the USDA APHIS BRS permit regulations. We also thank 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.2020.00843/full#supplementary-material

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11:R106.

Appling, D. R. (1991). Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. *FASEB J.* 5, 2645–2651. doi: 10.1096/fasebj.5.12.1916088

Bashline, L., Lei, L., Li, S., and Gu, Y. (2014). Cell wall, cytoskeleton, and cell expansion in higher plants. *Mol. Plant* 7, 586—600. doi: 10.1093/mp/ssu018

Baxter, H. L., Mazarei, M., Dumitrache, A., Natzke, J., Rodriguez, M., Gou, J., et al. (2018). Transgenic miR156 switchgrass in the field: growth, recalcitrance and rust susceptibility. *Plant Biotechnol. J.* 6, 39–49. doi: 10.1111/pbi.12747

Baxter, H. L., Mazarei, M., Labbe, 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. Jr., 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., and Pu, Y. (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

Brandon, A. G., and Scheller, H. V. (2020). Engineering of bioenergy crops: dominant genetic approaches to improve polysaccharide properties and composition in biomass. *Front. Plant Sci.* 11:282. doi: 10.3389/fpls.2020.00282

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

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

Crowe, J. D., Feringa, N., Pattathil, S., Merritt, B., Foster, C., and Dines, D. (2017). Identification of developmental stage and anatomical fraction contributions to cell wall recalcitrance in switchgrass. *Biotechnol. Biofuels* 10:184. doi: 10.1186/s13068-017-0870-5

David, K., and Ragauskas, A. J. (2010). Switchgrass as an energy crop for biofuel production: a review of its ligno-cellulosic chemical properties. *Energy Environ. Sci.* 3, 1182–1190. doi: 10.1039/B926617H

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

Dumitrache, A., Natzke, J., Rodriguez, M. Jr., Yee, K. L., Thompson, O. A., Poovaiah, C. R., et al. (2017). Transgenic switchgrass (*Panicum virgatum* L.) targeted for reduced recalcitrance to bioconversion: a two-year comparative analysis of field-grown lines modified for target gene or genetic element expression. *Plant Biotechnol. J.* 15, 688–697. doi: 10.1111/pbi.12666

Eckardt, N. A. (2007). Gibberellins are modified by methylation in planta. *Plant Cell* 19, 3–6. doi: 10.1105/tpc.107.050955

Eloy, N. B., Voorend, W., Lan, W., Saleme, M. L., Cesarino, I., Vanholme, R., et al. (2017). Silencing CHALCONE SYNTHASE in maize impedes the incorporation of tricin into lignin and increases lignin Content. *Plant Physiol.* 173, 998—1016. doi: 10.1104/pp.16.01108

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

Fu, C., Mielenz, J. R., Xiao, X., Ge, Y., Hamilton, C. Y., Rodriguez, M., et al. (2011a). 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

Fu, C., Sunkar, R., Zhou, C., Shen, H., Zhang, J. Y., Matts, J., et al. (2012). Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotechnol. J.* 10, 443–452. doi: 10.1111/j.1467-7652.2011.00677.x

Fu, C., Xiao, X., Xi, Y., Ge, Y., Chen, F., Bouton, J., et al. (2011b). Downregulation of cinnamyl alcohol dehydrogenase (CAD) leads to improved saccharification efficiency in switchgrass. *BioEnergy Res.* 4, 153–164. doi: 10.1007/s12155-010-9109-z

Hanson, A. D., and Gregory, J. F. (2011). Folate biosynthesis, turnover, and transport in plants. *Annu. Rev. Plant Biol.* 62, 105–125. doi: 10.1146/annurev-arplant-042110-103819

Hanson, A. D., and Roje, S. (2001). One-carbon metabolism in higher plants. *Annu. Rev. Plant Biol.* 52, 119–137. doi: 10.1146/annurev.arplant.52.1.119

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

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

Jonathan, M. C., DeMartini, J., Van Stigt Thans, S., Hommes, R., and Kabel, M. A. (2017). Characterisation of non-degraded oligosaccharides in enzymatically hydrolysed and fermented, dilute ammonia-pretreated corn stover for ethanol production. *Biotechnol. Biofuels* 10:112. doi: 10.1186/s13068-017-0803-3

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Lam, P. Y., Tobimatsu, Y., Takeda, Y., Suzuki, S., Yamamura, M., Umezawa, T., et al. (2017). Disrupting flavone synthase II alters lignin and improves biomass digestibility. *Plant Physiol.* 174, 972—985. doi: 10.1104/pp.16.01973

Li, L., Skinner, S. H., Liu, S., Beuchle, D., Tang, H. M., Yeh, C.-T., et al. (2015). The maize *brown midrib4 (bm4)* gene encodes a functional folylpolyglutamate synthase. *Plant J.* 81, 493–504. doi: 10.1111/tpj.12745

Liu, W., Mazarei, M., Ye, R., Peng, Y., Shao, Y., Baxter, H. L., et al. (2018). Switchgrass (*Panicum virgatum* L.) promoters for green tissue-specific expression of the *MYB4* transcription factor for reduced-recalcitrance transgenic switchgrass. *Biotechnol. Biofuels* 11:122. doi: 10.1186/s13068-018-1119-7

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

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 or RNAi-mediated gene silencing. *Front. Plant Sci.* 9:1114. doi: 10.3389/fpls.2018.01114

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

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

Ranocha, P., Denancé, N., Vanholme, R., Freydier, A., Martinez, Y., Hoffmann, L., et al. (2010). *Walls are thin 1* (*WAT1*), an *Arabidopsis* homolog of *Medicago truncatula NODULIN21*, is a tonoplast-localized protein required for secondary wall formation in fibers. *Plant J.* 63, 469−483. doi: 10.1111/j.1365-313X.2010.04256.x

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

Ravanel, S., Cherest, H., Jabrin, S., Grunwald, D., Surdin-Kerjan, Y., Douce, R., et al. (2001). Tetrahydrofolate biosynthesis in plants: molecular and functional characterization of dihydrofolate synthetase and three isoforms of folylpolyglutamate synthetase in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 98, 15360–15365. doi: 10.1073/pnas.261585098

Rébeillé, F., Ravanel, S., Jabrin, S., Douce, R., Storozhenko, S., and Van Der Straeten, D. (2006). Folates in plants: biosynthesis, distribution, and enhancement. *Physiol. Plant.* 126, 330–342. doi: 10.1111/j.1399-3054.2006.00587.x

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

Shane, B. (1989). Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism. *Vitam. Horm.* 45, 263–335. doi: 10.1016/S0083-6729(08)60397-0

Shen, B., Li, C., and Tarczynski, M. C. (2002). High free-methionine and decreased lignin content result from a mutation in the *Arabidopsis S-adenosyl-l-methionine synthetase 3* gene. *Plant J.* 29, 371–380. doi: 10.1046/j.1365-313X.2002.01221.x

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. *BioEnerg Res.* 2, 233–245. doi: 10.1007/s12155-009-9058-6

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

Shen, H., Poovaiah, C. R., Ziebell, A., Tschaplinski, T. J., Pattathil, S., Gjersing, E., et al. (2013). Enhanced characteristics of genetically modified switchgrass (*Panicum virgatum* L.) for high biofuel production. *Biotechnol. Biofuels* 6:71. doi: 10.1186/1754-6834-6-71

Srivastava, A. C., Chen, F., Ray, T., Pattathil, S., Peña, M. J., Avci, U., et al. (2015). Loss of function of folylpolyglutamate synthetase 1 reduces lignin content and improves cell wall digestibility in *Arabidopsis*. *Biotechnol. Biofuels* 8:224. doi: 10.1186/s13068-015-0403-z

Stewart, A., and Cromey, M. (2011). Identifying disease threats and management practices for bio-energy crops. *Curr. Opin. Environ. Sustain.* 3, 75–80. doi: 10.1016/j.cosust.2010.10.008

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

Tang, H. M., Liu, S., Hill-Skinner, S., Wu, W., Reed, D., Yeh, C. T., et al. (2014). The maize *brown midrib2 (bm2)* gene encodes a methylenetetrahydrofolate reductase that contributes to lignin accumulation. *Plant J.* 77, 380–392. doi: 10.1111/tpj.12394

Tobimatsu, Y., and Schuetz, M. (2019). Lignin polymerization: how do plants manage the chemistry so well? *Curr. Opin. Biotechnol.* 56, 75–81. doi: 10.1016/j.copbio.2018.10.001

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

Uppalapati, S. R., Serba, D. D., Ishiga, Y., Szabo, L. J., Mittal, S., Bhandari, H. S., et al. (2013). Characterization of the rust fungus, *Puccinia emaculata*, and evaluation of genetic variability for rust resistance in switchgrass populations. *BioEnergy Res.* 6, 458–468. doi: 10.1007/s12155-012-9263-6

Urbanowicz, B. R., Pena, M. J., Ratnaparkhe, S., Avci, U., Backe, J., Steet, H. F., et al. (2012). 4-*O*-methylation of glucuronic acid in *Arabidopsis* glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14253–14258. doi: 10.1073/pnas.1208097109

van der Weijde, T., Alvim Kamei, C. L., Torres, A. F., Vermerris, W., Dolstra, O., Visser, R. G., et al. (2013). The potential of C4 grasses for cellulosic biofuel production. *Front. Plant Sci.* 4:1. doi: 10.3389/fpls.2013.00107

Vanholme, R., De Meester, B., Ralph, J., and Boerjan, W. (2019). Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.* 56, 230−239. doi: 10.1016/j.copbio.2019.02.018

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. *Bioenerg. Res.* 2, 275–283. doi: 10.1007/s12155-009-9049-7

Xie, H., Engle, N. L., Venketachalam, S., Yoo, C. G., Barros, J., Lecoultre, M., et al. (2019). Combining loss of function of FOLYLPOLYGLUTAMATE SYNTHETASE1 and CAFFEOYL-COA 3-O-METHYLTRANSFERASE1 for lignin reduction and improved saccharification efficiency in *Arabidopsis thaliana*. *Biotechnol. Biofuels* 12:108. doi: 10.1186/s13068-019-1446-3

Xu, B., Escamilla-Treviño, L. L., Sathitsuksanoh, N., Shen, Z., Shen, H., Zhang, Y. H., et al. (2011). Silencing of 4-coumarate:coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. *New Phytol.* 192, 611–625. doi: 10.1111/j.1469-8137.2011.03830.x

Yee, K. L., Rodriguez, M. Jr., Tschaplinski, T. J., Engle, N. L., Martin, M. Z., Fu, C., et al. (2012). Evaluation of the bioconversion of genetically modified switchgrass using simultaneous saccharification and fermentation and a consolidated bioprocessing approach. *Biotechnol. Biofuels* 5:81. doi: 10.1186/1754-6834-5-81

Yoshida, K., and Komae, K. (2006). A rice family 9 glycoside hydrolase isozyme with broad substrate specificity for hemicelluloses in type II cell walls. *Plant Cell Physiol.* 47, 1541−1554. doi: 10.1093/pcp/pcl020

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

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

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.

frontiers in Plant Science

REVIEW
published: 02 July 2020
doi: 10.3389/fpls.2020.00914

Regulation of Lignin Biosynthesis by Post-translational Protein Modifications

*Daniel B. Sulis and Jack P. Wang**

Forest Biotechnology Group, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, United States

Post-translational modification of proteins exerts essential roles in many biological processes in plants. The function of these chemical modifications has been extensively characterized in many physiological processes, but how these modifications regulate lignin biosynthesis for wood formation remained largely unknown. Over the past decade, post-translational modification of several proteins has been associated with lignification. Phosphorylation, ubiquitination, glycosylation, and S-nitrosylation of transcription factors, monolignol enzymes, and peroxidases were shown to have primordial roles in the regulation of lignin biosynthesis. The main discoveries of post-translational modifications in lignin biosynthesis are discussed in this review.

Keywords: lignin, lignification, proteins, PTMs, SCW, trees, wood formation

OPEN ACCESS

Edited by:
Jaime Barros-Rios, University of North Texas, United States

Reviewed by:
Chang-Jun Liu, Brookhaven National Laboratory (DOE), United States
Jeongim Kim, University of Florida, United States
Xiaolan Rao, University of North Texas, United States

***Correspondence:**
Jack P. Wang
jpwang@ncsu.edu

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *21 March 2020*
Accepted: *04 June 2020*
Published: *02 July 2020*

Citation:
Sulis DB and Wang JP (2020) Regulation of Lignin Biosynthesis by Post-translational Protein Modifications. Front. Plant Sci. 11:914. doi: 10.3389/fpls.2020.00914

INTRODUCTION

Wood is a biological composite and a valuable source of feedstock for construction timber, pulp and paper, and biofuels (Wang et al., 2018). Wood is composed of secondary xylem tissue formed by secondary cell walls (SCW) (Fromm, 2013). The secondary xylem consists of two types of cells: fibers, which provide mechanical support, and tracheary elements composed of vessels (not found in gymnosperm wood) and tracheids (found in both angiosperm and gymnosperm) for water and solutes transport (Demura and Fukuda, 2007). SCW is mainly composed of cellulose, hemicelluloses, and lignin, in different ratios of these constituents (Li et al., 2011; Ye and Zhong, 2015).

Lignin is a phenolic polymer formed by phenylpropanoid monomeric units, 4-coumaryl alcohol (H-unit), coniferyl alcohol (G-unit), and sinapyl alcohol (S-unit), also known as monolignols. The biosynthesis of lignin occurs in several consecutive reactions involving up to 11 different enzyme families and 24 metabolites (Wang et al., 2019b). The pathway is complex and regulated by a network of substrates and inhibitors in the conversion of phenylalanine or tyrosine to monolignols (Wang et al., 2014). The monolignols are then transported to the lignifying zone and oxidized by peroxidases and laccases to phenoxy radicals (Li et al., 2014). Many efforts have been made to elucidate the regulation of the lignin biosynthetic pathway. Transcription factors (TFs) associated with lignification have been identified for *Pinus* (Patzlaff et al., 2003a,b), *Eucalyptus* (Goicoechea et al., 2005; Legay et al., 2010), and *Populus* (Karpinska et al., 2004; Zhong et al., 2010b; Ohtani et al., 2011; Li et al., 2012, 2015; Tian et al., 2013; Wang S. et al., 2015; Yang et al., 2017; Chen et al., 2019; Gui et al., 2019; Zheng et al., 2019) tree species. These TFs impart transactivation and transrepression of monolignol genes and other TFs in complex hierarchical transcriptional regulatory networks. Some TFs associated with monolignol biosynthesis form protein-protein

(TF-TF) interactions that may affect TF-target DNA binding (Chen et al., 2019). Lignification is also regulated at the enzyme level. Monolignol biosynthetic enzymes can directly interact with each other (e.g., Ptr4CL3-Ptr4CL5, PtrC3H-PtrC4H, At4CL1-AtC3H, At4CL1-AtC4H, and AtCCR1-AtC4H interactions) (Chen et al., 2011; Chen H. -C. et al., 2014; Gou et al., 2018; Wang et al., 2019a) or indirectly through common mediators such as AtMSPBs (e.g., AtC3H-AtMSPBs, AtC4H-AtMSPBs, AtF5H-AtMSPBs interactions) (Gou et al., 2018). In addition, monolignol biosynthetic enzymes can interact with proteins in other biological pathways (e.g., OsCCR1-OsRac1, ZmCCoAOMT/ZmHCT-ZmRp1 interactions) (Kawasaki et al., 2006; Wang G. H. et al., 2015; Wang and Balint-Kurti, 2016). These interactions may modulate enzyme stability, activity, and metabolic flux in lignin biosynthesis, increase the biosynthesis of lignin in response to pathogen infection, and suppress plant hypersensitive response (Kawasaki et al., 2006; Chen et al., 2011; Chen H. -C. et al., 2014; Wang G. H. et al., 2015; Wang et al., 2019a; Wang and Balint-Kurti, 2016; Gou et al., 2018). However, the role of post-translational modifications (PTMs) in SCW formation remains poorly understood.

PTMs are covalent processes that alter the properties of proteins by proteolytic cleavage or addition of modifying groups to one or more amino acids (Mann and Jensen, 2003). Over 200 different types of PTMs have been identified, ranging from small chemical modifications (e.g., phosphorylation and acetylation) to the addition of complete proteins (e.g., ubiquitination) (Beltrao et al., 2013; Spoel, 2018). PTMs are essential for many proteins and can affect the localization, stability, structure, activity, and molecular interactions of proteins (Nørregaard Jensen, 2004; Wang J. P. et al., 2015). In plants, PTMs are associated with plant growth and development, biotic and abiotic stress response, and metabolism (Catala et al., 2007; Mazzucotelli et al., 2008; Stulemeijer and Joosten, 2008; Miura and Hasegawa, 2010; Friso and van Wijk, 2015). Phosphorylation of cellulose synthases and cellulose synthase-like proteins were identified to have essential roles in regulating the activity and distribution of cellulose synthase complexes (CSCs) along microtubules (Speicher et al., 2018). The ubiquity of PTMs and their diverse regulatory functions is indicative of the complexity of secondary cell wall biosynthesis in general, and lignin biosynthesis in particular (**Table 1**).

PHOSPHORYLATION OF MONOLIGNOL BIOSYNTHETIC ENZYMES

Lignin biosynthesis is typically perceived as a matrix of linear or branched enzymatic reactions that successively modify the aromatic ring of the phenylpropane units and conversion of the side-chain carboxyl to an alcohol moiety. The involvement of many enzymes suggests that the pathway is well coordinated to mediate precise control of the rate and ratios of monolignol biosynthesis for polymerization. PTMs provide an efficient way to impart spatiotemporal activation and inactivation of monolignol enzyme activities in plants. However, the identification of PTMs involved in lignin biosynthesis remains challenging. Some studies based on phosphoproteomics could not reliably detect phosphopeptides for monolignol enzymes in wood forming cells (Mauriat et al., 2015). The difficulty of detecting protein phosphorylation is in part due to the highly dynamic nature of the regulatory mechanism.

PAL PHOSPHORYLATION

Phenylalanine ammonia-lyase (PAL) is a family of enzymes that catalyzes the deamination of phenylalanine to cinnamic acid, representing the first reaction step in the phenylpropanoid pathway (Shi et al., 2013; Zhang and Liu, 2015). PAL enzymes have been extensively characterized for their protein structure, functionality, and primordial role in wood formation (Ritter and Schulz, 2004; Shi et al., 2013). Some evidence regarding the chemical modification of PAL has been proposed. In *Phaseolus vulgaris* (French Bean) suspension-cultured cells, a phosphopeptide derived from PAL2 was detected (Allwood et al., 1999). The authors suggested that the kinase responsible for the PAL2 phosphorylation belongs to the calmodulin-like domain protein kinase (CDPK) family (Allwood et al., 1999). Consistently, a recombinant PAL2 protein from *P. trichocarpa* × *P. deltoides* was phosphorylated by CDPK derived from French bean suspensions and Arabidopsis (Allwood et al., 1999; Cheng et al., 2001). In PAL of *Phyllostachys edulis*, *in silico* functional analysis indicated a likely presence of nine casein kinase phosphorylation sites, eight protein kinase C phosphorylation sites, and two N-glycosylation sites (Gao et al., 2012). The high degree of PAL peptide sequence similarity across plant species and their conserved phosphorylation by CDPK may suggest a common phosphoregulatory mechanism for PAL in phenylpropanoid biosynthesis (Allwood et al., 1999; Cheng et al., 2001; Zhang and Liu, 2015).

The functional role of post-translational phosphorylation in PAL remains to be clarified. Allwood et al. (1999) showed that one hour after co-incubation of PAL2 and CDPK, the Km and V_{max} of PAL2 is slightly reduced. On the other hand, longer co-incubation of up to 4 h reduced the protein stability of the phosphorylated isoform of PAL2 and result in a 3-fold reduction in the V_{max} when compared to the non-phosphorylated isoform (Allwood et al., 1999). The phosphorylation of PAL2 likely limits the rate of phenylalanine conversion to cinnamic acid, thereby influencing the metabolic flux for lignification.

The CDPK family of kinases has been associated with the regulation of biological processes encompassing plant growth, development, and response to biotic and abiotic stresses (Romeis, 2001; Schulz et al., 2013). Pathogen infections induce the expression and activity of PAL (Jones, 1984; Shoresh and Harman, 2008), which in turn promotes immune response mediated by salicylic acid production, enhanced physical barriers (lignin), and accumulation of antimicrobial compounds (e.g., phytoalexins) (Jones, 1984; Solecka, 1997; Chen et al., 2009; Adams-Phillips et al., 2010; Hamann, 2012; Yan and Dong, 2014). Since PAL transcript expression is upregulated under biotic stress, the CDPK-mediated phosphorylation of PAL may mark it for turnover, thus reducing PAL activity

TABLE 1 | PTMs of proteins involved in lignification.

Protein	Organism	Types of PTM	Detection of PTM	Effect of PTM	References
PAL2	*P. vulgaris* (French bean)	Phosphorylation	*In vitro*	Unknown	Allwood et al., 1999
PAL2	*P. trichocarpa x P. deltoides*	Phosphorylation	*In vitro*	Phosphorylation of PAL2 decreases Km and V_{max}	Allwood et al., 1999; Cheng et al., 2001
AldOMT2	*P. trichocarpa*	Phosphorylation	*In vitro*	Phosphorylations of AldOMT2 reduce enzyme activity	Wang J. P. et al., 2015
MYB4	*P .taeda*	Phosphorylation	*In vitro*	Phosphorylation of MYB4 reduces the MYB4-transactivation activity over the lignin target genes	Morse et al., 2009
RAI1	*O. sativa*	Phosphorylation	*In vitro*	Phosphorylation activates RAI1 and increases the expression of PAL1.	Kawasaki et al., 2006; Kim et al., 2012; Akamatsu et al., 2013; Nasir et al., 2018
LTF1	*P. trichocarpa*	Phosphorylation	*In vivo* and *In vitro*	Phosphorylation reduces the LTF1 stability via 26S proteasome and reduces the LTF1-transrepression activity over the lignin target genes	Gui et al., 2019
PAL1-4	*A. thaliana*	Ubiquitination	*In vivo*	Ubiquitination reduces the protein stability via 26S proteasome	Zhang et al., 2013
CCR	*O. sativa*	Ubiquitination	Untested	Interaction of OsCCR with SCF^{OsFBK1} reduces the OsCCR stability via 26S proteasome	Borah and Khurana, 2018
MYB156	*P. tomentosa*	Ubiquitination	Untested	Interaction of MYB156 with UBC34 reduces the MYB156-transactivation over the lignin genes possibly decreasing the MYB156 stability via 26S proteasome	Zheng et al., 2019
MYB221	*P. tomentosa*	Ubiquitination	Untested	Interaction of MYB221 with UBC34 reduces the MYB221-transactivation over the lignin genes possibly decreasing the MYB221 stability via 26S proteasome	Zheng et al., 2019
VND7	*A. thaliana*	S-nitrosylation	*In vitro*	S-nitrosylation of VND7 decreases the VND7-transactvation activity over the SCW genes	Kawabe et al., 2018
VND7	*A. thaliana*	Ubiquitination	Untested	VND7 protein accumulates upon treatment with MG−132 in transformed tobacco BY−2 cells	Yamaguchi et al., 2008
PXP1-6	*P. trichocarpa*	Glycosylation	*In vivo*	Unknown	Christensen et al., 1998
PRX	*Z. elegans*	Glycosylation	*In vitro*	Glycosylations change the PRX catalytic efficiencies.	Gabaldón et al., 2005, 2006, 2007

(Allwood et al., 1999) to maintain homeostasis (**Figure 1A**). The phosphorylation of PAL has also been suggested to translocate the proteins to membranes, which may contribute to metabolic channeling in lignin biosynthesis (Allwood et al., 1999; Rasmussen and Dixon, 1999).

AldOMT2 PHOSPHORYLATION

5-Hydroxyconiferaldehyde *O*-methyltransferase 2 (AldOMT2) catalyzes the 3/5-methylation of caffeoyl- and 5-hydroxyferuloyl–containing acids, aldehydes, and alcohols for monolignol biosynthesis. In *P. trichocarpa*. *PtrAldOMT2* is the highest transcribed gene in the monolignol biosynthetic pathway and the third-highest transcribed gene in stem differentiating xylem (SDX). The protein abundance of PtrAldOMT2 is also the highest of all monolignol enzymes, accounting for 5.9% of the SDX proteome (Shuford et al., 2012; Lin et al., 2013). Given the abundance in transcripts and proteins of PtrAldOMT2, regulation of its activity by transcriptional control is energy-intensive. In contrast, regulation of PtrAldOMT2 activity by protein phosphorylation removes the need to synthesize/degrade RNA and protein, thus providing a rapid and energetically efficient mode of regulation for O-methyltransferase activity (Wang J. P. et al., 2015). Phosphoproteomic analysis by mass-spectrometry in SDX revealed two phosphopeptides in PtrAldOMT2 that contain either a phosphoserine at peptide position-123 or position-125. Concurrent phosphorylation of both Serine123 and Serine125 was not detected *in vivo*. The phosphorylation of PtrAldOMT2 is mediated by kinases in the SDX protein extract and could be reversed by phosphatase treatment (Wang J. P. et al., 2015). Phosphorylation reduced the enzymatic activity of recombinant PtrAldOMT2 considerably, and this reduction in activity was not due to

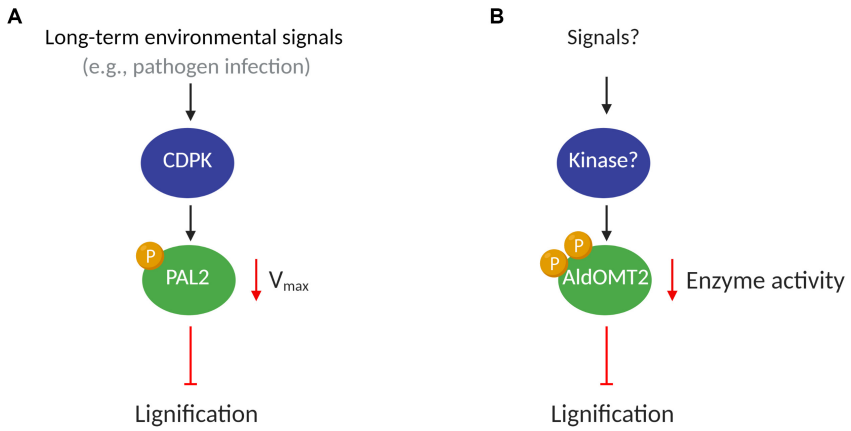

FIGURE 1 | A putative model of kinases regulating monolignol enzymes activity. **(A)** Under potential stimulus such as pathogen infections, the CDPK proteins phosphorylate PAL2, decreasing the V_{max} of the enzyme and reducing lignification in plants; **(B)** Kinase(s) protein(s), after an endogenous or environmental stimulus, phosphorylates AldOMT2 at either of the two phosphorylation sites and negatively regulates the AldOMT2 activity.

protein degradation (Wang J. P. et al., 2015). Moreover, site-directed mutagenesis by replacing either Serine123 or Serine125 with asparagine, a negatively charged amino acid that mimics the properties of phosphoserine, showed that Serine123 is more critical for phosphoregulation of enzyme activity. Serine123 is conserved in 43 of 46 (93%) AldOMTs across diverse plant species (Wang J. P. et al., 2015), suggesting a conserved evolutionary phosphoregulation of the enzyme. The presence of two phosphoserine residues in PtrAldOMT2 and their distinct enzymatic regulations provide a basis for further exploring mechanistic insights to phosphoregulation in monolignol biosynthesis. Whether different kinases and signaling pathways regulate the two phosphorylation sites or if the sites are functionally redundant, remains to be determined (Wang J. P. et al., 2015). Lignification is regulated by many environmental and developmental stimuli that combinatorially modulate the rate and ratios of monolignol biosynthesis. The understanding of how PTMs transduce the regulation to changes in metabolic flux for monolignol biosynthesis would provide valuable knowledge of plant adaptation and cell wall biosynthesis (**Figure 1B**).

PHOSPHORYLATION OF TFS IN LIGNIN BIOSYNTHESIS

R2R3-MYBs are TFs that contain a DNA binding domain at the N-terminus, which is composed of two imperfect helix-turn-helix repeats of approximately 50 amino acids (R2 and R3) (Jin and Martin, 1999; Patzlaff et al., 2003a). Protein members from the R2R3-MYB family of TFs are known to bind at AC rich cis-elements present in the promoters of many genes in the lignin biosynthetic pathway (Lois et al., 1989; Joos and Hahlbrock, 1992; Leyva et al., 1992; Hauffe et al., 1993; Hatton et al., 1995; Logemann et al., 1995; Bell-Lelong et al., 1997; Séguin et al., 1997; Lacombe et al., 2000; Lauvergeat et al., 2002; Patzlaff et al., 2003a). Some R2R3-MYBs have been identified and characterized as key regulators of plant cell wall biosynthesis (Patzlaff et al., 2003a). PtMYB4 is an ortholog of AtMYB56 and AtMYB83 (Patzlaff et al., 2003a; Yang et al., 2016; Gui et al., 2019), which function as second layer transactivators of SCW biosynthesis in Arabidopsis (Yang et al., 2016). The overexpression of *PtMYB4* in tobacco induced monolignol biosynthesis and promoted lignin accumulation in non-lignified cell types, suggesting that this TF regulates positively the lignin biosynthesis genes during the wood formation process (Patzlaff et al., 2003a).

Using a cDNA library from *Pinus taeda* SDX, Morse et al. (2009) screened MAPKs (PtMAPK6 and PtMAPK13) potentially involved in wood formation. *In vitro* phosphorylation assays demonstrated that PtMAPK6 could phosphorylate PtMYB1 and PtMYB4, transregulators of lignin biosynthesis in SDX. Replacement of Lysine68 for arginine abolished the kinase activity of PtMAPK6, demonstrating that the lysine is essential for mediating target phosphorylation. Serine236 was identified to be the phosphorylation site in PtMYB4. Replacement of Serine236 for a glutamate residue (mimicking constitutive phosphorylation of PtMYB4) did not interfere with the affinity for DNA binding but significantly reduced the transactivation of the target gene (Morse et al., 2009).

MAPKs are components of signal transduction networks that trigger a variety of biological processes in plants (Nühse et al., 2000). MAPK6 is associated with response to pathogen infection, ethylene response, and cell wall biosynthesis and remodeling (Andreasson and Ellis, 2010; Seifert and Blaukopf, 2010; Bacete et al., 2018). All these processes affect secondary cell wall biosynthesis for wood formation (Adams-Phillips et al., 2010; Fromm, 2013; Li et al., 2015; Ye and Zhong, 2015). In *Oriza sativa*, OsMAPK3 and OsMAPK6 are involved in the phosphorylation of OsRAI1 through a complex signaling cascade upon chitin elicitation. The activation of OsRAI1 by phosphorylation induces the transcript expression of defense-associated genes including *PAL1* (Kawasaki et al., 2006; Kim et al., 2012; Akamatsu et al., 2013; Nasir et al., 2018). In *P. taeda*, PtMAPK6 is expressed in cell differentiation stages 1, 2 (division and expansion), and 3, 4 (SCW synthesis)

of SDX, but its kinase activity for PtMYB4 phosphorylation was mainly observed in the initial stages of 1 and 2. The phosphorylated isoform of PtMAPK6, representing the active form of the enzyme, was only found in stages 1 and 2 (Morse et al., 2009). PtMAPK6 may be autophosphorylated or activated by other kinase-mediated phosphorylation during early xylogenesis. The activated PtMAPK6 then phosphorylates PtMYB4, reducing the TF transregulation of cell wall biosynthetic genes (**Figure 2A**). PtMAPK6 activity during early xylogenesis may function to suppress the expression of PtMYB4-regulated genes from resulting in premature deposition of lignin in dividing and expanding cells. The absence of PtMAPK6 activity in the late stages of differentiating xylem allows PtMYB4 to transactivate lignin biosynthetic genes in these cells (Morse et al., 2009).

LTF1 is the closest ortholog of AtMYB4 in *P. deltoides* × *P. euramericana* (Gui et al., 2019). Arabidopsis *AtMYB4* mutants showed increased expression of *C4H* and reduced expression of *CCoAOMT*. Overexpression of *AtMYB4* in tobacco increased the expression of *CCoAOMT* and reduced the expression of *C4H, 4CL*, and *CAD* genes (Jin et al., 2000). In *P. deltoides* × *P. euramericana*, LTF1 is mainly expressed in the SDX as a transrepressor of monolignol biosynthetic

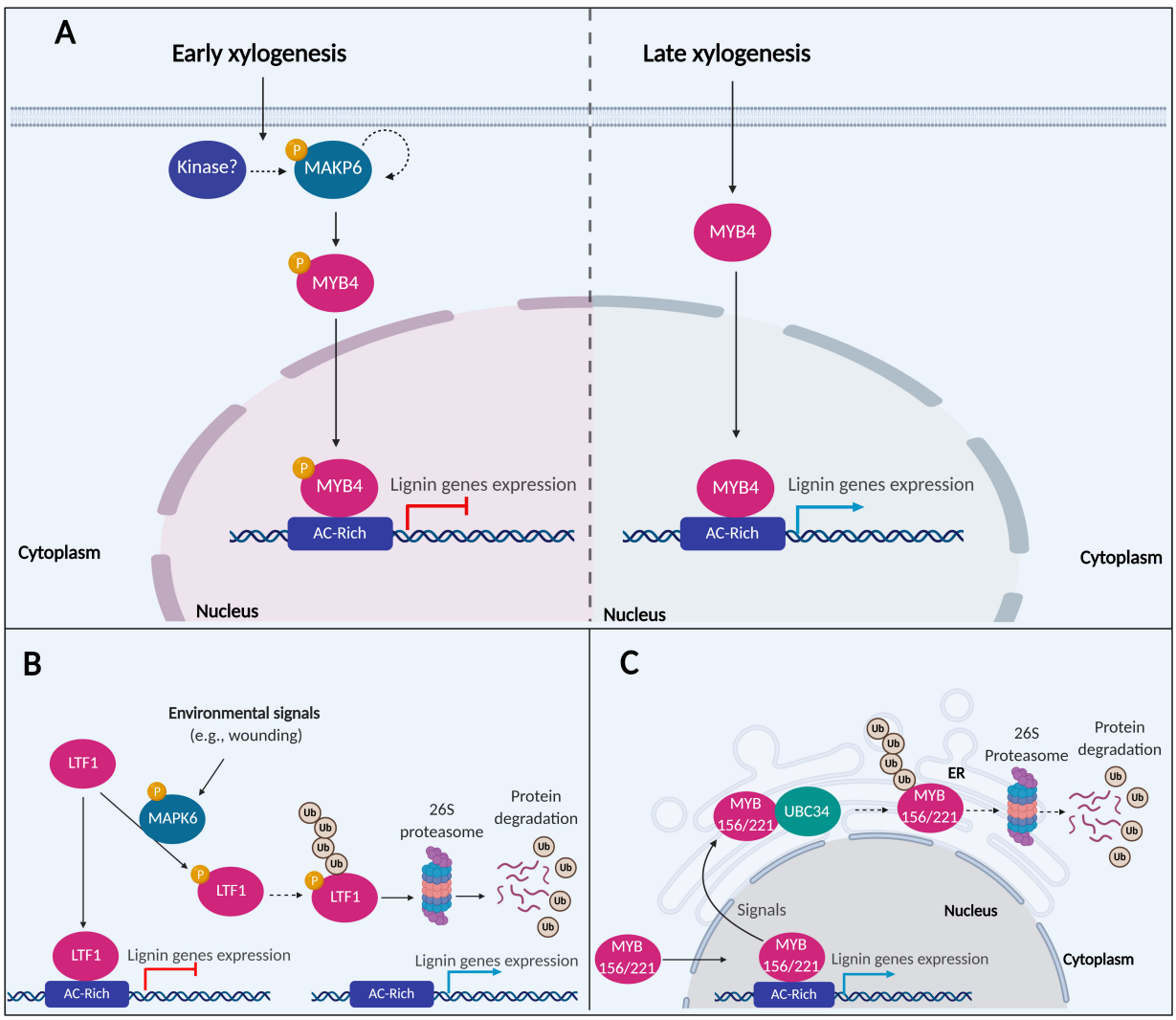

FIGURE 2 | A putative model of MAPK6 regulating the activation of lignin biosynthesis. **(A)** In the early xylogenesis stage in *P. taeda*, MAPK6 proteins are autophosphorylated or activated by other kinase-mediated phosphorylation. MAPK6 becomes activated and phosphorylates the TF MYB4. Phosphorylation inactivates MYB4, resulting in the repression of lignin genes. In the late xylogenesis stage, MAPK6 is no longer active, and MYB4 induces the expression of lignin genes **(B)** In *P. trichocarpa*, LTF1 is a repressor of lignin genes under normal conditions. After environmental stimuli such as wounding, MAPK6 can interact and phosphorylate LTF1. Phosphorylation destabilizes LTF1 in the cells and promotes its degradation via 26S proteasome and attenuating the repression of lignin genes mediated by LRF1. **(C)** A putative model of MYB156 and MYB221 regulating the activation of lignin biosynthesis in *P. tomentosa*. MYB156 and MYB221 are repressors TFs of lignin genes. UBC34 ubiquitin-conjugating enzyme interacts with MYB156 and MYB221 and alters the subcellular localization of the TFs from the nucleus to the ER. The expression of the lignin genes are attenuated either by the TFs trapping into ER or translocation of the MYB156 and MYB221 either traps or degradation via ubiquitin and 26S proteasome pathway. The dotted arrows indicated hypothesized events not confirmed experimentally.

genes, including *4CL*. *LTF1* mutants in poplar showed elevated transcript expression of lignin genes and increased lignin content. The overexpression of *LTF1* in transgenic poplar showed opposite effects, decreasing the expression of lignin genes and reducing the overall lignin content. *LTF1* overexpression induced dwarfism in 2-month old poplar (grown in phytotron), but the severity of dwarfism was reduced when the plants reached eight months old (Gui et al., 2019). The protein abundance of LTF1 was decreased significantly in 8-month-old plants. Phosphoproteomic analysis identified two phosphopeptides corresponding to LTF1 in 8-month old plants. These phosphopeptides were not detected in the 2-month-old *LTF1*-overexpressed plants. Disruption of the phosphorylation sites in transgenic plants exhibited reduced plant height, stem diameter, and internode length, and reduced lignin deposition in the fiber cells of these plants, and xylem vessel collapse (Gui et al., 2019). Recombinant LTF1 proteins were rapidly degraded when co-incubated with SDX extracts, and the addition of proteasome inhibitor MG–132 re-establishes the initial levels LTF1. On the other hand, even in the absence of MG–132, the levels of LTF1 protein containing mutations in the phosphorylation sites remained stable after the incubation with the SDX extracts.

PdMAPK6 in *P. deltoides* × *P. euramericana* is a homolog of the Arabidopsis and *P. taeda* MAPK6, identified by protein-protein interactions to be associated with LTF1 (Morse et al., 2009; Gui et al., 2019). *In vitro* phosphorylation assays showed that PdMAPK6 could phosphorylate LTF1 at the Threonine146 and Threonine178 positions, while mutagenesis of these phosphorylation sites prevented LTF1 phosphorylation. A higher level of phosphorylated LTF1 and other PdMAPK6 activated proteins were detected in 8-month old greenhouse-grown poplar, compared to 2-month old plants grown in indoor growth chambers. Greenhouse-grown plants are more susceptible to environmental signals like biotic and abiotic stress (Gui et al., 2019), consistent with the functions of MAPK6 in pathogenic response (Andreasson and Ellis, 2010; Bacete et al., 2018). In this context, after mechanical wounding, increased expression of lignin genes and improved lignin deposition were detected in plants overexpressing *LTF1*. However, these effects were significantly lower in plants overexpressing *LTF1* without the phosphorylation sites. Consistently, the expression of Arabidopsis ortholog *AtMYB4* is repressed by environmental stimuli, including UV-B and wounding (Jin et al., 2000). These data suggest that LTF1 is a repressor of lignin genes, reducing the lignin content in plant cells. Under environmental stimuli, LTF1 is phosphorylated, decreasing its stability by degradation via 26S proteasome (Gui et al., 2019; **Figure 2B**).

The high peptide sequence identity of PdMAKP6 and PtMAKP6 (84%) and the similar phosphoregulation of MYB TFs suggest an important role of MAPK in the regulation of lignin biosynthesis in both angiosperm and gymnosperm species. Future studies focusing on MYB phosphorylation by MAPK6 and their interaction with external stimuli would better inform the phosphoregulation of lignin biosynthesis. For instance, is mechanical wounding the only stimulus of LTF1 phosphorylation, or are other signals such as ethylene response, cell wall biosynthesis, and cell remodeling also involved? PTMs are found in diverse classes of proteins, highlighting the regulatory interplay among PTMs as a common strategy to regulate protein functions (Yang, 2005; Nussinov et al., 2012; Lothrop et al., 2013; Pejaver et al., 2014). Other types of PTMs could also be involved in the regulation of MYB TFs. Identifying the upstream regulators of MAPK6 and other wood TFs phosphorylated by MAPK6 are needed to determine the network of regulatory interactions by kinases in the lignification of tree species.

UBIQUITINATION OF PAL

Ubiquitination is a common regulatory mechanism in all eukaryotes that target proteins for degradation via the 26S proteasome, thereby maintaining protein turnover in the cells. The ubiquitin attachment process involves three enzymes: the E1 ubiquitin-activating enzyme, the E2 ubiquitin-conjugating enzyme, and an E3 ubiquitin ligase (Guo and Yang, 2017). Proteins containing a well-conserved 40–50 amino acid F-box domain in the N-terminal region are referred to as F-box proteins that are one part of the ubiquitin E3 ligase complex. F-box proteins are involved in many physiological processes in plants, such as flowering, pathogen defenses, circadian cycle, and phytohormones signaling (Kuroda et al., 2002). PAL enzymes from Arabidopsis (PAL1, PAL2, PAL3, and PAL4) have been shown to interact with Kelch domain-containing F-box (KFB) proteins (KFB01, KFB20, KFB39, and KFB50) (Zhang et al., 2013; Zhang et al., 2015). The F-box domain in KFBs that confers ubiquitin ligase activity is associated with reducing PAL protein abundance by decreasing protein stability (Zhang et al., 2013). Co-overexpression of *PALs-GFP* and *KFBs* in Arabidopsis reduced the conversion efficiency of phenylalanine to t-cinnamic acid by up to 80% (Zhang et al., 2013). More recently, a protein called Small and Glossy Leaves 1 (SAGL1) was identified to be involved in the PTM regulation PAL1. Phylogenetic analysis using 99 Arabidopsis KFB homologous proteins revealed that SAGL1 is closely related to the PAL-regulating KFBs (KFB01, KFB20, KFB39, and KFB50), but located in a separate clade from these KFBs (Yu et al., 2019). SAGL1 can interact and reduce the stability of PAL1, leading to reduced PAL activity for monolignol biosynthesis. Similar to other KFB-mediated PALs regulation, the level of PAL1 was fully restored after MG–132 treatment for proteasome inhibition (Yu et al., 2019).

KFB and SAGL1 have similar regulatory roles in PAL protein stability. Incubation of recombinant PAL in protein extracts of 6-week old Arabidopsis stems led to a significant reduction in the abundance of PAL proteins. The addition of the MG–132 proteasome inhibitor to the assay could maintain PAL protein abundance by preventing ubiquitination-based protein degradation (Zhang et al., 2013). Double and triple mutants in Arabidopsis for the *KFB01*, *KFB20,* and *KFB50* genes showed an increased amount of PAL proteins, and consequently, more acetyl-bromide lignin in the plant cell walls. Overexpression of *KFBs* genes, in contrast, caused a 2 to 70% lignin reduction in the transgenics (Zhang et al., 2013). *SAGL1* mutants in Arabidopsis showed a strong purple color in rosette leaves,

leaf petioles, and inflorescence stems, typically found where high amounts of anthocyanin are accumulated (Nakatsuka and Nishihara, 2010; Yu et al., 2019). Quantification of lignin content in the mature stems of the mutants showed a 2-fold increase in lignin content compared to wild-type. The *SAGL1* mutants also had a 60% increased conversion rate of L-phenylalanine to trans-cinnamate, compared to wild-type. PAL enzyme activity increased in *SAGL1* mutants without changing the transcript levels of the *PAL* genes (Yu et al., 2019). Transgenic lines overexpressing *SAGL1* showed significant reductions in the conversion rate of L-phenylalanine to trans-cinnamate and lignin content in the mature stems compared to the wild-type (Yu et al., 2019).

The reduced abundance of PAL proteins in Arabidopsis overexpressing *KFBs* or *SAGL1* was not due to reduced transcript levels of *PAL* genes. In contrast, increased transcript abundance of *PALs* were observed in the transgenic lines overexpressing *KFBs*, possibly to compensate for the rapid turnover of PAL proteins in the cells (Zhang et al., 2013). The discrepancy between PAL transcript and protein abundances in the transgenics is indicative of cross-talks between transcriptional regulation and PTM to maintain homeostasis for lignin biosynthesis. PAL mediates the entry step in the phenylpropanoid pathway and is extensively regulated for metabolic control (Zhang and Liu, 2015). In *O. sativa*, an auxin-responsive Kelch-domain-containing F-box protein (OsFBK1) interacts with cinnamoyl-CoA reductase (OsCCR) to mediate its protein degradation via the proteasome pathway, thereby regulating lignin content in the cell walls (Borah and Khurana, 2018). This suggests that the ubiquitin E3 ligase complex proteins may interact with other lignin biosynthetic enzymes. In *P. trichocarpa,* 337 F-box proteins were identified using a bioinformatics approach (Schumann et al., 2011), suggesting a possible role of these proteins in the regulation of turnover rate in monolignol enzymes. The involvement of this class of protein in regulating lignin biosynthesis and wood formation in tree species remains to be validated.

S-NITROSYLATION AND UBIQUITINATION OF TFS

TFs containing the NAM/ATAF/CUC (NAC) domain are key regulators of plant development, senescence, SCW formation, and biotic and abiotic stress responses (Olsen et al., 2005; Nakashima et al., 2007; Puranik et al., 2012; Chen X. et al., 2014). Several members of the NAC family of TFs have been identified and characterized in Arabidopsis, rice, soybean, wheat, poplar, and citrus (Puranik et al., 2012). In *O. sativa*, pathogen infection induces the phosphorylation of OsNAC4, leading to its accumulation in the cell nucleus. OsNAC4 is involved in the transregulation of 139 genes and hypersensitive cell death (Kaneda et al., 2009), but its association with SWC biosynthesis has not been reported. Seven *NAC* genes are highly transcribed during vessel cell differentiation in Arabidopsis (Kubo, 2005). *VND7* has been extensively characterized as a transregulator of protoxylem vessel development (Kubo, 2005; Yamaguchi et al., 2010). Repression of *VND7* and *VND6* using plant repression domain (SRDX) resulted in defects in Arabidopsis growth and vessel formation, mediated by repressed central metaxylem vessel formation and central protoxylem vessel formation (Kubo, 2005). Induced expression of *VND7* accelerated vessel cell differentiation in all cell types encompassing cotyledon, leaf, hypocotyl, and root cells (Yamaguchi et al., 2010; Kawabe et al., 2018).

The knockout of S-nitrosoglutathione reductase 1 (*GSNOR1*) in transgenic Arabidopsis expressing *VND7* resulted in mutant seedlings lacking xylem vessel differentiation (mutant "suppressor of ectopic vessel cell differentiation induced by VND7" or *seiv*) (Yamaguchi et al., 2010; Kawabe et al., 2018). Indeed, the inducible-*VND7* expression in the *seiv* mutants did not show any evidence of cell wall deposition. These mutant seedlings showed a dwarf phenotype, suggesting that the *GSNOR1* gene may be involved in plant growth and development, in addition to regulating VND7 for xylem trans-differentiation (Kawabe et al., 2018). VND7 directly or indirectly regulates the expression of genes involved in SCW formation during xylem vessel differentiation, encompassing the biosynthesis of cellulose, hemicellulose, and lignin, and programed cell death (Yamaguchi et al., 2011). The expression of these VND7-downstream genes is induced by *VND7* overexpression, and suppressed in *seiv* mutants, demonstrating that the knockout of *GSNOR1* disrupts VND7-mediated regulation (Kawabe et al., 2018). Biotin switch assays showed that the VND7 was S-nitrosylated *in vitro*, and amino acid substitution of two Cysteine264 and Cysteine320 abolished the S-nitrosylation signals. The presence of S-nitrosoglutathione reduced VND7 activity in Arabidopsis protoplasts. VND7 activity was reduced when either Cysteine264 or Cysteine320 were substituted for tryptophan, which mimics a constitutive S-nitrosylated isoform of VND7. Thus, S-nitrosylation plays an essential role in negatively regulating VND7 activity for vessel development. However, replacing the Cysteine264 or Cysteine320 for a serine residue, which mimics a non-S-nitrosylated form of VND7, also caused a reduction in the transactivation activity of VND7 (Kawabe et al., 2018). Cysteine residues in plant proteins are modified by different types of PTMs in addition to S-nitrosylation, including S-sulfenylation, S-glutathionylation, sulfinylation, sulfonylation, and the formation of intra-/intermolecular disulfide bridges (Waszczak et al., 2015; Kawabe et al., 2018). Therefore, other PTMs may be regulating VND7 and causing the non-S-nitrosylated isoform of VND7 to reduce transactivation activity in protoplasts (Kawabe et al., 2018).

NAC-domain proteins have been demonstrated to interact with many cellular components involved in different cellular processes (Yamaguchi et al., 2008), making them a potential target for protein modifications. For instance, the RING domain protein SINA of Arabidopsis 5 (SINAT5) interacts with NAC1 and promotes its degradation via the ubiquitin/26S proteasome pathway (Olsen et al., 2005). Similarly, VND7

is regulated by 26S proteasome-mediated degradation (Yamaguchi et al., 2008). Protein extracts from transformed tobacco BY-2 cells expressing *VND7* had low levels of detected VND7 protein after 24 h, while the addition of proteasome inhibitor increased the abundance of VND7 (Yamaguchi et al., 2008).

Although herbaceous plants such as Arabidopsis have a limited secondary growth compared to tree species, some similarities may be found in the general organization of their tissues (Chaffey et al., 2002; Demura and Fukuda, 2007; Zhong et al., 2010b). Repeated removal of inflorescent stems in Arabidopsis can induce some secondary xylem production that has been used in developmental studies of secondary growth (Oh et al., 2003). Searching for NAC TFs in the *P. trichocarpa* genome, Zhong et al. (2010b) identified 16 homologous genes of Arabidopsis *VND*, *NST,* and *SMB* involved in secondary wall biosynthesis or xylem differentiation. These poplar genes were named *PtrWNDs* for "wood-associated NAC domain transcription factors" (Zhong et al., 2010b; Ohtani et al., 2011). Of the 16 genes, 12 showed a high level of transcript expression in vessel and fiber cells of developing woody tissue. Moreover, they were localized in the nucleus of plant cells and exhibited transactivation activities, consistent with their putative function as TFs (Zhong et al., 2010b). Expression of *PtrWNDs* in Arabidopsis *snd1 (nst3) nst1* double mutants rescued the pendent inflorescence stem phenotype and restored stem strength and deposition of secondary walls in interfascicular fibers. Overexpression of *PtrWND2B* and *PtrWND6B* (sub members of the *SMB* and *VND* class of TF, respectively) in Arabidopsis increased the transcript expression of cellulose synthases (*CesA4*, *CesA7*, and *CesA8*), xylan biosynthetic genes (*FRA8*, *IRX8*, and *IRX9*), and lignin biosynthetic genes (*4CL1* and *CCoAOMT1*) (Zhong et al., 2010b; Ohtani et al., 2011). PtrWND2B and PtrWND6B also transactivated wood-associated TFs involved in the biosynthesis of lignin, hemicelluloses, and cellulose (Zhong et al., 2010b). Overexpression of 12 *PtVNS*/*PtrWND* in *P. trichocarpa* caused SCW formation surrounding the transgenics' vascular tissue, similarly to which was seen for the induced expression of *VND7* in Arabidopsis and poplar epidermal cells (Ohtani et al., 2011). Expression of *VND7* in poplar induced the expression of many genes related to SCW formation, including enzyme-encoding genes and TFs (Ohtani et al., 2011). Taken together, NAC TFs in Arabidopsis and *P. trichocarpa* are functionally important for SCW formation (Zhong et al., 2010a; Ohtani et al., 2011). Orthologous proteins are typically conserved in their PTMs, and the level of conservation is higher in closely related species (Remmerie et al., 2011; Venne et al., 2014). PTMs identified for Arabidopsis NAC TFs, such as S-nitrosylation and ubiquitination, may also occur in PtrWNDs in the regulation of wood formation.

In *P. tomentosa,* PtoMYB156 and PtoMYB221 from the subgroup 4 of the R2R3-MYB TF family work as transrepressors of lignin genes (Zheng et al., 2019). Overexpression of *PtoMYB156* dramatically retarded plant growth and reduced secondary xylem formation, while thickening of SCWs in the vascular stem tissue was observed in *PtoMYB156* knockout plants (Yang et al., 2017). The regulation of PtoMYB156 and PtoMYB221 is mediated by interactions with an E2 ubiquitin-conjugating enzyme 34 (PtoUBC34), which translocate the TFs to the ER (Zheng et al., 2019). In mesophyll protoplasts of *P. tomentosa* overexpressing PtoMYB156 and PtoMYB221, the TFs are localized in the cell nucleus and could suppress the transcript expression of target genes in the absence of PtoUBC34. When PtoUBC34 is co-expressed in the same system, PtoMYB156 and PtoMYB221 are translocated to the ER in a dose-depend way, and their repression activity is reduced (Zheng et al., 2019). Whether the reduced repression activity of PtoMYB156 and PtoMYB221 is due to the trapping of these TFs in the ER or degradation via the 26S proteasome pathway remains unknown (**Figure 2C**). Lignin biosynthesis responds to various developmental and environmental signals such as light, sugar content, circadian rhythms, plant hormones, and wounding, where these signals are converted to a physiological process by hierarchical transcriptional regulation of lignin pathway genes (Zheng et al., 2019). The translocation to cellular compartments or possibly ubiquitination and degradation via the 26S proteasome pathway may be a pivotal point in regulating the activity of these TFs. Future studies will help to clarify the cross-talking of lignin biosynthesis and other physiological processes (Zheng et al., 2019).

GLYCOSYLATION IN PEROXIDASES

Class III peroxidases are present as large multigenic families in all land plants (Mathé et al., 2010). Generally, this class of proteins is heme-containing enzymes, which may oxidize a wide range of substrates using hydrogen peroxide as a reduction agent. Peroxidases are associated with multiple cellular processes, encompassing primary and secondary metabolism, hormone catabolism, pathogen defense, phenol oxidation, cross-linking of cell wall-structural proteins, and polysaccharides, and lignin polymerization (Christensen et al., 1998). As aforementioned, peroxidases oxidize the monolignols in the plant cell walls to free radicals for conjugating to the growing lignin polymer (Li et al., 2014). The oxidation reaction mediated by lignin peroxidases is characterized by a three-step cycle involving a two-electron enzyme oxidation: FeIII + H_2O_2 → CoI + H_2O (*k1*); followed by two one-electron reductions: (i) CoI + RH → CoII + $R^{\cdot}$; (*k2*) and (ii) CoII + RH → FeIII + $R^{\cdot}$; (*k3*), where FeIII, CoI, CoII, and RH are the native enzyme, compound I, compound II, and the monolignol, respectively (Gabaldón et al., 2006).

Most of the peroxidases are glycosylated, and the carbohydrate content may constitute up to 25% of the protein (Christensen et al., 1998; Veitch, 2004) Glycosylation sites are normally found in the sequence motif asn-x-ser/thr (where x means any amino acid residue), and they occur within surface turns or loop regions connecting helices (Veitch, 2004). Six anionic stem peroxidases (PXP1-6) were isolated from xylem tissues of *P. trichocarpa,* and all were heavily glycosylated, exhibiting N-glucosamine and ten

putative glycosylation sites in their structures (Christensen et al., 1998). Similarly, two isoforms of ZePrx, where one was fully glycosylated (ZePrx 34.70) and another partially glycosylated (ZePrx 33.44) were identified in *Zinnia elegans* suspensions cells (Gabaldón et al., 2005; Gabaldón et al., 2007). The two isoforms of this protein showed different biochemical properties. For instance, the *k1* (CoI formation constant – which monitors the reactivity of the enzyme with H_2O_2) was higher in the ZePrx 34.70 compared to ZePrx 33.44 in the presence of either *p*-coumaryl alcohol or coniferyl alcohol substrates. On the other hand, the opposite effects were observed when coniferaldehyde, sinapyl alcohol, and sinapaldehyde were used as substrates. Moreover, the *k3*, (CoII reduction constant – which monitors the ability of the oxidized form of the enzyme, CoII, to oxidize phenolics) was higher for ZePrx 33.44 comparing to the ZePrx 34.70 in the presence of coniferyl alcohol. Again, the opposite effects were observed in the presence of sinapyl alcohol, where the reactivity was higher for the ZePrx 34.70 (Gabaldón et al., 2006). The different reactivity rates suggest that glycans modulate the catalytic activity of the peroxidases (Gabaldón et al., 2006). Since multiple glycosylation sites are found in these enzymes, the different glycosylation patterns may determine the substrate specificity for monolignol oxidation, leading to changes in the composition of the lignin. Glycosylation may have a role beyond the regulation of peroxidase substrate specificity. Protein folding and stabilization, and protein-cell wall interactions have also been predicted as putative functions of this PTM in peroxidases (Schuller et al., 1996; Gabaldón et al., 2007; Cesarino et al., 2012).

FINAL CONSIDERATIONS

Despite the technical challenges in the detection and characterization of PTMs *in vivo*, recent discoveries have provided insights into their importance in regulating lignin biosynthesis and wood formation. The lignin biosynthetic pathway is complex and controlled by many external and developmental stimuli. The complexity of the regulations suggests that TFs, monolignol enzymes, and peroxidases could respond rapidly and specifically to different types of stimuli in the modulation of lignin biosynthesis. Phosphorylation (e.g., PAL2 and AldOMT2) and ubiquitination (e.g., PALs and CCR) of monolignol enzymes decrease the enzymatic activity and protein stability (Allwood et al., 1999; Zhang et al., 2013; Wang J. P. et al., 2015; Borah and Khurana, 2018). More comprehensive identification of PTMs associated with lignin biosynthesis will further improve our understanding of the spatiotemporal regulation of the pathway, and how such regulations coordinately affect the properties of the cell walls. MAPKs mediated MYB4 and LTF1 phosphorylation (Morse et al., 2009; Gui et al., 2019) have provided valuable mechanistic insights on how PTM modulates the SCW but the complete signaling pathway and the interactions between MAPKs and other MAPK-cascade components remain to be elucidated. Further studies focusing on how developmental and environmental signals trigger the MAPK-cascade to activate or repress TFs regulating lignin biosynthesis would be valuable to determine the network of regulatory interactions that control phenotypic alterations.

Protein degradation via 26S proteasome (e.g., VND7, LTF1, and possibly MYB156 and MYB221) plays an essential role in regulating TFs for lignin biosynthesis (Kawabe et al., 2018; Gui et al., 2019; Zheng et al., 2019). However, there are knowledge gaps to be filled for a better understand of how TF protein turnover is modulated by PTMs. Ubiquitination, which is related to the 26S proteasome pathway, remains to be identified in LTF1, MYB156, and MYB221. Moreover, the combinatorial effects of multiple PTMs and their coordinated control of lignin biosynthesis remain poorly understood. For example, the glycosylation of lignin peroxidases (e.g., ZePrx) controls the enzyme substrate specificity (Gabaldón et al., 2006), and further studies may elucidate how multiple glycosylations, and glycosylation together with other PTMs may combinatorially affect lignin composition and structure. Moreover, the relationship between VND7 S-nitrosylation and other types of cysteine modifications in the transregulation of SCW genes is unknown (Kawabe et al., 2018). One PTM can influence the modification of other types of PTMs, which can result in a broad variation of possible proteoforms (Friso and van Wijk, 2015). The combinatorial effect of multiple PTMs and cross-talk between PTMs in the same protein, or on different proteins within complexes, is crucial to defining the interrelationships of multiple PTMs in lignin biosynthesis.

Much of the early work on plant PTMs have focused on model organisms such as Arabidopsis. These PTM in woody species, particularly in the regulation of lignin biosynthesis has remained largely unknown. Due to the large variation in lignin content, composition, and structure between Arabidopsis and perennial woody species, the extent of similarity or difference in the SCW-associated PTMs across species should be evaluated. Systematic co-expression analysis of SCW protein abundances with PTM enzymes during SDX formation, or genome-scale protein interactomics could reveal putative PTM regulatory interactions controlling lignin biosynthesis. The resulting interactions may guide future studies to verify whether these PTMs occur *in vitro* and *in vivo* and how they regulate lignin biosynthesis and wood formation. Comprehensive identification of PTMs in trees will be crucial to understanding the transduction of complex regulatory information that mediates wood formation.

AUTHOR CONTRIBUTIONS

DS and JW conceptualized and literary reviewed and wrote the manuscript.

FUNDING

This work was supported by the United States Department of Agriculture (Grant# 2017-06529) and the NC State University Chancellor's Innovation Fund (Grant# 190549MA).

REFERENCES

Adams-Phillips, L., Briggs, A. G., and Bent, A. F. (2010). Disruption of poly(ADP-ribosyl)ation mechanisms alters responses of *Arabidopsis* to biotic stress. *Plant Physiol.* 152, 267–280. doi: 10.1104/pp.109.148049

Akamatsu, A., Wong, H. L., Fujiwara, M., Okuda, J., Nishide, K., Uno, K., et al. (2013). An OsCEBiP/OsCERK1-OsRacGEF1-OsRac1 module is an essential early component of chitin-induced rice immunity. *Cell Host Microbe* 13, 465–476. doi: 10.1016/j.chom.2013.03.007

Allwood, E. G., Davies, D. R., Gerrish, C., Ellis, B. E., and Bolwell, G. P. (1999). Phosphorylation of phenylalanine ammonia-lyase: evidence for a novel protein kinase and identification of the phosphorylated residue. *FEBS Lett.* 457, 47–52. doi: 10.1016/S0014-5793(99)00998-9

Andreasson, E., and Ellis, B. (2010). Convergence and specificity in the *Arabidopsis* MAPK nexus. *Plant Sci.* 15, 106–113. doi: 10.1016/j.tplants.2009.12.001

Bacete, L., Mélida, H., Miedes, E., and Molina, A. (2018). Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses. *Plant J.* 93, 614–636. doi: 10.1111/tpj.13807

Bell-Lelong, D. A., Cusumano, J. C., Meyer, K., and Chapple, C. (1997). Cinnamate-4-hydroxylase expression in *Arabidopsis*. Regulation in response to development and the environment. *Plant Physiol.* 113, 729–738. doi: 10.1104/pp.113.3.729

Beltrao, P., Bork, P., Krogan, N. J., and Noort, V. (2013). Evolution and functional cross-talk of protein post-translational modifications. *Mol. Syst. Biol.* 9:714. doi: 10.1002/msb.201304521

Borah, P., and Khurana, J. P. (2018). The OsFBK1 E3 ligase subunit affects anther and root secondary cell wall thickenings by mediating turnover of a cinnamoyl-CoA reductase. *Plant Physiol.* 176, 2148–2165. doi: 10.1104/pp.17.01733

Catala, R., Ouyang, J., Abreu, I. A., Hu, Y., Seo, H., Zhang, X., et al. (2007). The *Arabidopsis* E3 SUMO ligase SIZ1 regulates plant growth and drought responses. *Plant Cell* 19, 2952–2966. doi: 10.1105/tpc.106.049981

Cesarino, I., Araújo, P., Sampaio Mayer, J. L., Paes Leme, A. F., and Mazzafera, P. (2012). Enzymatic activity and proteomic profile of class III peroxidases during sugarcane stem development. *Plant Physiol. Biochem.* 55, 66–76. doi: 10.1016/j.plaphy.2012.03.014

Chaffey, N., Cholewa, E., Regan, S., and Sundberg, B. (2002). Secondary xylem development in *Arabidopsis*: a model for wood formation. *Physiol. Plant.* 114, 594–600. doi: 10.1034/j.1399-3054.2002.1140413.x

Chen, H.-C., Li, Q., Shuford, C. M., Liu, J., Muddiman, D. C., Sederoff, R. R., et al. (2011). Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 21253–21258. doi: 10.1073/pnas.1116416109

Chen, H.-C., Song, J., Wang, J. P., Lin, Y.-C., Ducoste, J., Shuford, C. M., et al. (2014). Systems biology of lignin biosynthesis in *Populus trichocarpa*: heteromeric 4-coumaric acid:coenzyme A ligase protein complex formation, regulation, and numerical modeling. *Plant Cell* 26, 876–893. doi: 10.1105/tpc.113.119685

Chen, X., Wang, Y., Lv, B., Li, J., Luo, L., Lu, S., et al. (2014). The NAC family transcription factor OsNAP confers abiotic stress response through the ABA pathway. *Plant Cell Physiol.* 55, 604–619. doi: 10.1093/pcp/pct204

Chen, H.-C., Wang, J. P., Liu, H., Li, H., Lin, Y.-C. J., and Shi, R. (2019). Hierarchical transcription factor and chromatin binding network for wood formation in *Populus trichocarpa*. *Plant Cell* 31, 602–626. doi: 10.1105/tpc.18.00620

Chen, Z., Zheng, Z., Huang, J., Lai, Z., and Fan, B. (2009). Biosynthesis of salicylic acid in plants. *Plant Signal. Behav.* 4, 493–496. doi: 10.4161/psb.4.6.8392

Cheng, S.-H., Sheen, J., Gerrish, C., and Bolwell, G. P. (2001). Molecular identification of phenylalanine ammonia-lyase as a substrate of a specific constitutively active *Arabidopsis* CDPK expressed in maize protoplasts. *FEBS Lett.* 503, 185–188. doi: 10.1016/S0014-5793(01)02732-6

Christensen, J. H., Bauw, G., Gjesing Welinder, K., Van Montagu, M., and Boerjan, W. (1998). Purification and characterization of peroxidases correlated with lignification in poplar xylem. *Plant Physiol.* 118, 125–135. doi: 10.1104/pp.118.1.125

Demura, T., and Fukuda, H. (2007). Transcriptional regulation in wood formation. *Trends Plant Sci.* 12, 64–70. doi: 10.1016/j.tplants.2006.12.006

Friso, G., and van Wijk, K. J. (2015). Update: post-translational protein modifications in plant metabolism. *Plant Physiol.* 169, 1469–1487. doi: 10.1104/pp.15.01378

Fromm, J. (2013). "Xylem development in trees: from cambial divisions to mature wood cells," in *cellular Aspects of Wood Formation*, ed. J. Fromm (Berlin: Springer Berlin Heidelberg), 3–39.

Gabaldón, C., Gómez-Ros, L. V., Núñez-Flores, M. J. L., Esteban-Carrasco, A., and Barceló, A. R. (2007). Post-translational modifications of the basic peroxidase isoenzyme from *Zinnia elegans*. *Plant Mol. Biol.* 65, 43–61. doi: 10.1007/s11103-007-9197-0

Gabaldón, C., López-Serrano, M., Pedreño, M. A., and Barceló, A. R. (2005). Cloning and molecular characterization of the basic peroxidase isoenzyme from *Zinnia elegans*, an enzyme involved in lignin biosynthesis. *Plant Physiol.* 139, 1138–1154. doi: 10.1104/pp.105.069674

Gabaldón, C., López-Serrano, M., Pomar, F., Merino, F., Cuello, J., Pedreño, M. A., et al. (2006). Characterization of the last step of lignin biosynthesis in *Zinnia elegans* suspension cell cultures. *FEBS Lett.* 580, 4311–4316. doi: 10.1016/j.febslet.2006.06.088

Gao, Z. M., Wang, X. C., Peng, Z. H., Zheng, B., and Liu, Q. (2012). Characterization and primary functional analysis of phenylalanine ammonia-lyase gene from *Phyllostachys edulis*. *Plant Cell Rep.* 31, 1345–1356. doi: 10.1007/s00299-012-1253-9

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: EgMYB2, a regulator of lignification. *Plant J.* 43, 553–567. doi: 10.1111/j.1365-313X.2005.02480.x

Gou, M., Ran, X., Martin, D. W., and Liu, C.-J. (2018). The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. *Nat. Plants* 4, 299–310. doi: 10.1038/s41477-018-0142-9

Gui, J., Luo, L., Zhong, Y., Sun, J., Umezawa, T., and Li, L. (2019). Phosphorylation of LTF1, an MYB transcription factor in *Populus*, acts as a sensory switch regulating lignin biosynthesis in wood cells. *Mol. Plant* 12, 1325–1337. doi: 10.1016/j.molp.2019.05.008

Guo, T., and Yang, X. (2017). Ubiquitination and its function in defense pathways. *Eur. J. Biomed. Res.* 3, 1–4. doi: 10.18088/ejbmr.3.2.2017

Hamann, T. (2012). Plant cell wall integrity maintenance as an essential component of biotic stress response mechanisms. *Front. Plant Sci.* 3:77. doi: 10.3389/fpls.2012.00077

Hatton, D., Sablowski, R., Yung, M. H., Smith, C., Schuch, W., and Bevan, M. (1995). Two classes of cis sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. *Plant J.* 7, 859–876. doi: 10.1046/j.1365-313X.1995.07060859.x

Hauffe, K. D., Lee, S. P., Subramaniam, R., and Douglas, C. J. (1993). Combinatorial interactions between positive and negative cis-acting elements control spatial patterns of 4CL-1 expression in transgenic tobacco. *Plant J.* 4, 235–253. doi: 10.1046/j.1365-313X.1993.04020235.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

Jin, H., and Martin, C. (1999). Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol. Biol.* 41, 577–585. doi: 10.1023/A:1006319732410

Jones, D. H. (1984). Phenylalanine ammonia-lyase: regulation of its induction, and its role in plant development. *Phytochemistry* 23, 1349–1359. doi: 10.1016/S0031-9422(00)80465-3

Joos, H.-J., and Hahlbrock, K. (1992). Phenylalanine ammonia-lyase in potato (*Solanum tuberosum* L.). Genomic complexity, structural comparison of two selected genes and modes of expression. *Eur. J. Biochem.* 204, 621–629. doi: 10.1111/j.1432-1033.1992.tb16675.x

Kaneda, T., Taga, Y., Takai, R., Iwano, M., Matsui, H., Takayama, S., et al. (2009). The transcription factor OsNAC4 is a key positive regulator of plant hypersensitive cell death. *EMBO J.* 28, 926–936. doi: 10.1038/emboj.2009.39

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

Kawasaki, T., Koita, H., Nakatsubo, T., Hasegawa, K., Wakabayashi, K., Takahashi, H., et al. (2006). Cinnamoyl-CoA reductase, a key enzyme in lignin biosynthesis, is an effector of small GTPase Rac in defense signaling in rice. *Proc. Natl. Acad. Sci. U.S.A.* 103, 230–235. doi: 10.1073/pnas.0509875103

Kim, S.-H., Oikawa, T., Kyozuka, J., Wong, H. L., Umemura, K., Kishi-Kaboshi, M., et al. (2012). The bHLH Rac immunity1 (RAI1) is activated by OsRac1 via OsMAPK3 and OsMAPK6 in rice immunity. *Plant Cell Physiol.* 53, 740–754. doi: 10.1093/pcp/pcs033

Kubo, M. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19, 1855–1860. doi: 10.1101/gad.1331305

Kuroda, H., Takahashi, N., Shimada, H., Seki, M., Shinozaki, K., and Matsui, M. (2002). Classification and expression analysis of *Arabidopsis* F-Box-containing protein genes. *Plant Cell Physiol.* 43, 1073–1085. doi: 10.1093/pcp/pcf151

Lacombe, E., Van Doorsselaere, J., Boerjan, W., Boudet, A. M., and Grima-Pettenati, J. (2000). Characterization of cis-elements required for vascular expression of the cinnamoyl-CoA reductase gene and for protein-DNA complex formation. *Plant J.* 23, 663–676. doi: 10.1046/j.1365-313x.2000.00838.x

Lauvergeat, V., Rech, P., Jauneau, A., Guez, C., Coutos-Thevenot, P., and Grima-Pettenati, J. (2002). The vascular expression pattern directed by the *Eucalyptus gunnii* cinnamyl alcohol dehydrogenase EgCAD2 promoter is conserved among woody and herbaceous plant species. *Plant Mol. Biol.* 50, 497–509.

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

Leyva, A., Liang, X., Pintor-Toro, J. A., Dixon, R. A., and Lamb, C. J. (1992). Cis-element combinations determine phenylalanine ammonia-lyase gene tissue-specific expression patterns. *Plant Cell* 4, 263–271. doi: 10.1105/tpc.4.3.263

Li, C., Wang, X., Ran, L., Tian, Q., Fan, D., and Luo, K. (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, 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, Q., Min, D., Wang, J. P.-Y., Peszlen, I., Horvath, L., Horvath, B., et al. (2011). Down-regulation of glycosyltransferase 8D genes in *Populus trichocarpa* caused reduced mechanical strength and xylan content in wood. *Tree Physiol.* 31, 226–236. doi: 10.1093/treephys/tpr008

Li, Q., Song, J., Peng, S., Wang, J. P., Qu, G.-Z., Sederoff, R. R., et al. (2014). Plant biotechnology for lignocellulosic biofuel production. *Plant Biotechnol. J.* 12, 1174–1192. doi: 10.1111/pbi.12273

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

Logemann, E., Parniske, M., and Hahlbrock, K. (1995). Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. *Proc. Natl. Acad. Sci. U.S.A.* 92, 5905–5909. doi: 10.1073/pnas.92.13.5905

Lois, R., Dietrich, A., Hahlbrock, K., and Schulz, W. (1989). A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive cis-acting elements. *EMBO J.* 8, 1641–1648. doi: 10.1002/j.1460-2075.1989.tb03554.x

Lothrop, A. P., Torres, M. P., and Fuchs, S. M. (2013). Deciphering post-translational modification codes. *FEBS Lett.* 587, 1247–1257. doi: 10.1016/j.febslet.2013.01.047

Mann, M., and Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 21, 255–261. doi: 10.1038/nbt0303-255

Mathé, C., Barre, A., Jourda, C., and Dunand, C. (2010). Evolution and expression of class III peroxidases. *Arch. Biochem. Biophys.* 500, 58–65. doi: 10.1016/j.abb.2010.04.007

Mauriat, M., Leplé, J.-C., Claverol, S., Bartholomé, J., Negroni, L., Richet, N., et al. (2015). Quantitative proteomic and phosphoproteomic approaches for deciphering the signaling pathway for tension wood formation in poplar. *J. Proteome Res.* 14, 3188–3203. doi: 10.1021/acs.jproteome.5b00140

Mazzucotelli, E., Mastrangelo, A. M., Crosatti, C., Guerra, D., Stanca, A. M., and Cattivelli, L. (2008). Abiotic stress response in plants: when post-transcriptional and post-translational regulations control transcription. *Plant Sci.* 174, 420–431. doi: 10.1016/j.plantsci.2008.02.005

Miura, K., and Hasegawa, P. M. (2010). Sumoylation and other ubiquitin-like post-translational modifications in plants. *Trends Cell Biol.* 20, 223–232. doi: 10.1016/j.tcb.2010.01.007

Morse, A. M., Whetten, R. W., Dubos, C., and Campbell, M. M. (2009). Post-translational modification of an R2R3-MYB transcription factor by a MAP kinase during xylem development. *New Phytol.* 183, 1001–1013. doi: 10.1111/j.1469-8137.2009.02900.x

Nakashima, K., Tran, L.-S. P., Van Nguyen, D., Fujita, M., Maruyama, K., Todaka, D., et al. (2007). Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice: rice OsNAC6 functions in stress responses. *Plant J.* 51, 617–630. doi: 10.1111/j.1365-313X.2007.03168.x

Nakatsuka, T., and Nishihara, M. (2010). UDP-glucose:3-deoxyanthocyanidin 5-O-glucosyltransferase from *Sinningia cardinalis*. *Planta* 232, 384–392. doi: 10.1007/s00425-010-1175-0

Nasir, F., Tian, L., Chang, C., Li, X., Gao, Y., Tran, L.-S. P., et al. (2018). Current understanding of pattern-triggered immunity and hormone-mediated defense in rice (*Oryza sativa*) in response to *Magnaporthe oryzae* infection. *Semin. Cell Dev. Biol.* 83, 95–105. doi: 10.1016/j.semcdb.2017.10.020

Nørregaard Jensen, O. (2004). Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.* 8, 33–41. doi: 10.1016/j.cbpa.2003.12.009

Nühse, T. S., Peck, S. C., Hirt, H., and Boller, T. (2000). Microbial Elicitors Induce Activation and Dual Phosphorylation of the *Arabidopsis thaliana* MAPK6. *J. Biol. Chem.* 275, 7521–7526. doi: 10.1074/jbc.275.11.7521

Nussinov, R., Tsai, C.-J., Xin, F., and Radivojac, P. (2012). Allosteric post-translational modification codes. *Trends Biochem. Sci.* 37, 447–455. doi: 10.1016/j.tibs.2012.07.001

Oh, S., Park, S., and Han, K.-H. (2003). Transcriptional regulation of secondary growth in *Arabidopsis thaliana*. *J. Exp. Bot.* 54, 2709–2722. doi: 10.1093/jxb/erg304

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: NAC domain protein family regulates wood formation. *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

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. W., 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

Pejaver, V., Hsu, W.-L., Xin, F., Dunker, A. K., Uversky, V. N., and Radivojac, P. (2014). The structural and functional signatures of proteins that undergo multiple events of post-translational modification: structural and functional signatures of PTM crosstalk. *Protein Sci.* 23, 1077–1093. doi: 10.1002/pro.2494

Puranik, S., Sahu, P. P., Srivastava, P. S., and Prasad, M. (2012). NAC proteins: regulation and role in stress tolerance. *Trends Plant Sci.* 17, 369–381. doi: 10.1016/j.tplants.2012.02.004

Rasmussen, S., and Dixon, R. A. (1999). Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. *Plant Cell* 11, 1537–1551. doi: 10.1105/tpc.11.8.1537

Remmerie, N., De Vijlder, T., Laukens, K., Dang, T. H., Lemière, F., Mertens, I., et al. (2011). Next generation functional proteomics in non-model plants: a survey on techniques and applications for the analysis of protein complexes and

post-translational modifications. *Phytochemistry* 72, 1192–1218. doi: 10.1016/j.phytochem.2011.01.003

Ritter, H., and Schulz, G. E. (2004). Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase. *Plant Cell* 16, 3426–3436. doi: 10.1105/tpc.104.025288

Romeis, T. (2001). Calcium-dependent protein kinases play an essential role in a plant defence response. *EMBO J.* 20, 5556–5567. doi: 10.1093/emboj/20.20.5556

Schuller, D. J., Ban, N., van Huystee, R. B., McPherson, A., and Poulos, T. L. (1996). The crystal structure of peanut peroxidase. *Structure* 4, 311–321. doi: 10.1016/S0969-2126(96)00035-4

Schulz, P., Herde, M., and Romeis, T. (2013). Calcium-dependent protein kinases: hubs in plant stress signaling and development. *Plant Physiol.* 163, 523–530. doi: 10.1104/pp.113.222539

Schumann, N., Navarro-Quezada, A., Ullrich, K., Kuhl, C., and Quint, M. (2011). Molecular evolution and selection patterns of plant F-Box proteins with C-terminal kelch repeats. *Plant Physiol.* 155, 835–850. doi: 10.1104/pp.110.166579

Séguin, A., Laible, G., Leyva, A., Dixon, R. A., and Lamb, C. J. (1997). Characterization of a gene encoding a DNA-binding protein that interacts in vitro with vascular specific cis elements of the phenylalanine ammonia-lyase promoter. *Plant Mol. Biol.* 35, 281–291. doi: 10.1023/A:1005853404242

Seifert, G. J., and Blaukopf, C. (2010). Irritable walls: the plant extracellular matrix and signaling. *Plant Physiol.* 153, 467–478. doi: 10.1104/pp.110.153940

Shi, R., Shuford, C. M., Wang, J. P., Sun, Y.-H., Yang, Z., Chen, H.-C., et al. (2013). Regulation of phenylalanine ammonia-lyase (PAL) gene family in wood forming tissue of *Populus trichocarpa*. *Planta* 238, 487–497. doi: 10.1007/s00425-013-1905-1

Shoresh, M., and Harman, G. E. (2008). The molecular basis of shoot responses of maize seedlings to *Trichoderma harzianum T22* inoculation of the root: a proteomic approach. *Plant Physiol.* 147, 2147–2163. doi: 10.1104/pp.108.123810

Shuford, C. M., Li, Q., Sun, Y.-H., Chen, H.-C., Wang, J., Shi, R., et al. (2012). Comprehensive quantification of monolignol-pathway enzymes in *Populus trichocarpa* by protein cleavage isotope dilution mass spectrometry. *J. Proteome Res.* 11, 3390–3404. doi: 10.1021/pr300205a

Solecka, D. (1997). Role of phenylpropanoid compounds in plant responses to different stress factors. *Acta Physiol. Plant.* 19, 257–268. doi: 10.1007/s11738-997-0001-1

Speicher, T., Li, P., and Wallace, I. (2018). Phosphoregulation of the plant cellulose synthase complex and cellulose synthase-like proteins. *Plants* 7:52. doi: 10.3390/plants7030052

Spoel, S. H. (2018). Orchestrating the proteome with post-translational modifications. *J. Exp. Bot.* 69, 4499–4503. doi: 10.1093/jxb/ery295

Stulemeijer, I. J. E., and Joosten, M. H. A. J. (2008). Post-translational modification of host proteins in pathogen-triggered defence signalling in plants. *Mol. Plant Pathol.* 9, 545–560. doi: 10.1111/j.1364-3703.2008.00468.x

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

Veitch, N. C. (2004). Structural determinants of plant peroxidase function. *Phytochem. Rev.* 3, 3–18. doi: 10.1023/B:PHYT.0000047799.17604.94

Venne, A. S., Kollipara, L., and Zahedi, R. P. (2014). The next level of complexity: Crosstalk of posttranslational modifications. *Proteomics* 14, 513–524. doi: 10.1002/pmic.201300344

Wang, G.-F., and Balint-Kurti, P. J. (2016). Maize homologs of CCoAOMT and HCT, two key enzymes in lignin biosynthesis, form complexes with the NLR Rp1 protein to modulate the defense response. *Plant Physiol.* 171, 2166–2177. doi: 10.1104/pp.16.00224

Wang, G.H., He, Y., Strauch, R., Olukolu, B., Nielsen, D., Li, X., et al. (2015). Maize homologs of HCT, a key enzyme in lignin biosynthesis, bind the NLR Rp1 proteins to modulate the defense response. *Plant Physiol.* 169, 2230–2243. doi: 10.1104/pp.15.00703

Wang, J. P., Chuang, L., Loziuk, P. L., Chen, H., Lin, Y.-C., Shi, R., et al. (2015). Phosphorylation is an on/off switch for 5-hydroxyconiferaldehyde O -methyltransferase activity in poplar monolignol biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8481–8486. doi: 10.1073/pnas.1510473112

Wang, S., Li, E., Porth, I., Chen, J.-G., Mansfield, S. D., and Douglas, C. J. (2015). Regulation of secondary cell wall biosynthesis by poplar R2R3 MYB transcription factor PtrMYB152 in *Arabidopsis*. *Sci. Rep.* 4:5054. doi: 10.1038/srep05054

Wang, J. P., Liu, B., Sun, Y., Chiang, V. L., and Sederoff, R. R. (2019a). Enzyme-enzyme interactions in monolignol biosynthesis. *Front. Plant Sci.* 9:1942. doi: 10.3389/fpls.2018.01942

Wang, J. P., Matthews, M. L., Naik, P. P., Williams, C. M., Ducoste, J. J., Sederoff, R. R., et al. (2019b). Flux modeling for monolignol biosynthesis. *Curr. Opin. Biotechnol.* 56, 187–192. doi: 10.1016/j.copbio.2018.12.003

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. P., Naik, P. P., Chen, H.-C., Shi, R., Lin, C.-Y., Liu, J., et al. (2014). Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa*. *Plant Cell* 26, 894–914. doi: 10.1105/tpc.113.120881

Waszczak, C., Akter, S., Jacques, S., Huang, J., Messens, J., and Van Breusegem, F. (2015). Oxidative post-translational modifications of cysteine residues in plant signal transduction. *J. Exp. Bot.* 66, 2923–2934. doi: 10.1093/jxb/erv084

Yamaguchi, M., Goué, 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

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: direct target genes of VND7. *Plant J.* 66, 579–590. doi: 10.1111/j.1365-313X.2011.04514.x

Yan, S., and Dong, X. (2014). Perception of the plant immune signal salicylic acid. *Curr. Opin. Plant Biol.* 20, 64–68. doi: 10.1016/j.pbi.2014.04.006

Yang, L., Zhao, X., Ran, L., Li, C., Fan, D., and Luo, K. (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

Yang, X.-J. (2005). Multisite protein modification and intramolecular signaling. *Oncogene* 24, 1653–1662. doi: 10.1038/sj.onc.1208173

Ye, Z.-H., and Zhong, R. (2015). Molecular control of wood formation in trees. *J. Exp. Bot.* 66, 4119–4131. doi: 10.1093/jxb/erv081

Yu, S., Kim, H., Yun, D.-J., Suh, M. C., and Lee, B. (2019). Post-translational and transcriptional regulation of phenylpropanoid biosynthesis pathway by kelch repeat F-box protein SAGL1. *Plant Mol. Biol.* 99, 135–148. doi: 10.1007/s11103-018-0808-8

Zhang, X., Gou, M., Guo, C., Yang, H., and Liu, C.-J. (2015). Down-regulation of kelch domain-containing F-Box protein in *Arabidopsis* enhances the production of (poly)phenols and tolerance to ultraviolet radiation. *Plant Physiol.* 167, 337–350. doi: 10.1104/pp.114.249136

Zhang, X., Gou, M., and Liu, C.-J. (2013). *Arabidopsis* kelch repeat F-Box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. *Plant Cell* 25, 4994–5010. doi: 10.1105/tpc.113.119644

Zhang, X., and Liu, C.-J. (2015). Multifaceted regulations of gateway enzyme phenylalanine ammonia-lyase in the biosynthesis of phenylpropanoids. *Mol. Plant* 8, 17–27. doi: 10.1016/j.molp.2014.11.001

Zheng, L., Chen, Y., Ding, D., Zhou, Y., Ding, L., Wei, J., et al. (2019). Endoplasmic reticulum–localized UBC34 interaction with lignin repressors MYB221 and MYB156 regulates the transactivity of the transcription factors in *Populus tomentosa*. *BMC Plant Biol.* 19:97. doi: 10.1186/s12870-019-1697-y

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

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 September 2020
doi: 10.3389/fpls.2020.545748

Genome-Wide Association Study of Wood Anatomical and Morphological Traits in *Populus trichocarpa*

*Hari B. Chhetri[1], Anna Furches[2,3], David Macaya-Sanz[1], Alejandro R. Walker[4], David Kainer[2], Piet Jones[2,3], Anne E. Harman-Ware[5], Timothy J. Tschaplinski[2], Daniel Jacobson[2,3], Gerald A. Tuskan[2] and Stephen P. DiFazio[1]**

[1] Department of Biology, West Virginia University, Morgantown, WV, United States, [2] Biosciences Division, and The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [3] The Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee, Knoxville, TN, United States, [4] Department of Oral Biology, College of Dentistry, University of Florida, Gainesville, FL, United States, [5] Biosciences Center, and National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO, United States

OPEN ACCESS

Edited by:
Mengzhu Lu,
Chinese Academy of Forestry, China

Reviewed by:
Nathaniel Robert Street,
Umeå University, Sweden
Alexander Andrew Myburg,
University of Pretoria, South Africa

***Correspondence:**
Stephen P. DiFazio
spdifazio@mail.wvu.edu

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *26 March 2020*
Accepted: *21 August 2020*
Published: *09 September 2020*

Citation:
Chhetri HB, Furches A, Macaya-Sanz D, Walker AR, Kainer D, Jones P, Harman-Ware AE, Tschaplinski TJ, Jacobson D, Tuskan GA and DiFazio SP (2020) Genome-Wide Association Study of Wood Anatomical and Morphological Traits in Populus trichocarpa. Front. Plant Sci. 11:545748. doi: 10.3389/fpls.2020.545748

To understand the genetic mechanisms underlying wood anatomical and morphological traits in *Populus trichocarpa*, we used 869 unrelated genotypes from a common garden in Clatskanie, Oregon that were previously collected from across the distribution range in western North America. Using GEMMA mixed model analysis, we tested for the association of 25 phenotypic traits and nine multitrait combinations with 6.741 million SNPs covering the entire genome. Broad-sense trait heritabilities ranged from 0.117 to 0.477. Most traits were significantly correlated with geoclimatic variables suggesting a role of climate and geography in shaping the variation of this species. Fifty-seven SNPs from single trait GWAS and 11 SNPs from multitrait GWAS passed an FDR threshold of 0.05, leading to the identification of eight and seven nearby candidate genes, respectively. The percentage of phenotypic variance explained (PVE) by the significant SNPs for both single and multitrait GWAS ranged from 0.01% to 6.18%. To further evaluate the potential roles of candidate genes, we used a multi-omic network containing five additional data sets, including leaf and wood metabolite GWAS layers and coexpression and comethylation networks. We also performed a functional enrichment analysis on coexpression nearest neighbors for each gene model identified by the wood anatomical and morphological trait GWAS analyses. Genes affecting cell wall composition and transport related genes were enriched in wood anatomy and stomatal density trait networks. Signaling and metabolism related genes were also common in networks for stomatal density. For leaf morphology traits (leaf dry and wet weight) the networks were significantly enriched for GO terms related to photosynthetic processes as well as cellular homeostasis. The identified genes provide further insights into the genetic control of these traits, which are important determinants of the suitability and sustainability of improved genotypes for lignocellulosic biofuel production.

Keywords: *Populus*, wood anatomy, leaf morphology, GWAS, networks, lignin

INTRODUCTION

It is of increasing interest to identify the molecular variants underlying adaptive and morphological trait variation in plant populations. Loci highlighted by such analyses have great potential for optimizing the trait of interest through genetic engineering or breeding, thereby producing trees with increased productivity, enhanced abiotic stress tolerance, and/or improved quality of end products. Because of their wide geographic distribution and climatic gradients, large effective population sizes, and high genetic variation, forest trees are excellent model systems for understanding local adaptation and the genetic architecture of complex traits (Neale and Savolainen, 2004; González-Martínez et al., 2006; Neale and Kremer, 2011; Street and Ingvarsson, 2011; Ingvarsson et al., 2016). In this regard, efforts have been made to optimize the ecologically and economically important tree *Populus* for lignocellulosic biofuel production. Vast amounts of genomic and phenotypic resources are available for the genus. Several large-scale genome-wide association studies have identified the underlying genetic architecture related to morphological, physiological, wood property and chemistry, salinity tolerance, and disease resistance traits (Ma et al., 2013; McKown et al., 2014b; Muchero et al., 2015; Zhang et al., 2018; Bdeir et al., 2019; Quan et al., 2019; Jia et al., 2020; Lu et al., 2020). Furthermore, the biology of wood formation, cell wall ultrastructure and composition, and cell wall recalcitrance are fairly well studied (Groover et al., 2010; Wegrzyn et al., 2010; Studer et al., 2011; Du et al., 2013; Porth et al., 2013a; Porth et al., 2013b; Muchero et al., 2015; Porth et al., 2015; Allwright et al., 2016; Du et al., 2016; Escamez et al., 2017; Fahrenkrog et al., 2017; Johnson et al., 2017; Xi et al., 2017; Du et al., 2018; Gandla et al., 2018; Du et al., 2019) and a few studies have also recognized the role of microRNA in controlling tree growth and wood property traits in *Populus* (Quan et al., 2016; Chen B. et al., 2018). However, the genetic architecture underlying wood anatomical traits such as vessel size and density are relatively uncharacterized, despite the importance of these traits for cell wall composition and the overall performance of the tree.

Wood anatomy not only contributes to the structural integrity of the tree, but it is also critical for transport and storage processes (Sperry, 2003; Hietz et al., 2017). Anatomical structures like vessel size and density are related to cell wall structure and composition and wood density. These traits together affect long-distance axial transport of nutrients and hydraulic conductivity. Lignified cell walls and fibers add strength and living parenchyma cells provide radial transport and storage. There are often trade-offs among vessel properties, wood density and hydraulic conductivity (Preston et al., 2006). Wood traits are generally heritable (Carlquist, 2012) and serve as useful traits in phylogenetic analyses (Hietz et al., 2017). Radial variation in wood anatomical properties in the stem affects wood functional traits such as hydraulic conductivity and often scales with tree size and leaf characteristics (Lachenbruch et al., 2011). Genome-wide association studies of wood anatomy traits undoubtedly complement the current understanding of genetic architecture of other structural and functional traits in *Populus* and will serve as an important step for optimizing traits suitable for lignocellulosic biofuels.

Nevertheless, the limitation of power due to sample size in detecting the genetic variants associated with complex traits is a major hurdle in trees and other non-model organisms. The recent technological revolution in genomics has made comprehensive genotyping routine, but challenges remain due to the cost and logistical difficulties inherent in the establishment, management, and intensive phenotyping of large common gardens of trees. Furthermore, SNP loci and candidate genes identified thus far explain only a small proportion of the genetic variation in complex traits in general (Visscher et al., 2017). Genes controlling complex traits do not work in isolation, but instead are interconnected in networks of hundreds to thousands of genes, each of which may contribute incrementally to the variation in complex traits (Boyle et al., 2017). Recently, methods such as multitrait GWAS and meta analyses using summary statistics (e.g., principal components) have become increasingly popular due to their role in enhancing the power of GWAS and the identification of potentially pleiotropic loci (Porter and O'Reilly, 2017). Performing multitrait GWAS directly on raw phenotypic data can enhance power and lead to discovery of novel associations (Chhetri et al., 2019).

Multi-omics network methods can be used to boost the signal of true positives and reduce noise in GWAS results by demonstrating concurrent support for hypotheses from multiple lines of evidence (LOE) (Pendergrass et al., 2013; Verma et al., 2018; Weighill et al., 2019). This can be particularly helpful for studies in which sample size or power are limiting factors (Genovese et al., 2006). Furthermore, LOE methods help identify GWAS hits that are strong gene candidates for experimental validation by examining the gene across multiple contexts in parallel. By considering expression patterns in conjunction with phenotypic associations, genes that have a higher probability of being feasibly validated in lab, greenhouse, and common garden settings can be prioritized.

Here we focus on *Populus trichocarpa*, a targeted species for lignocellulosic biofuel production that has a distribution spanning from northern California to northern British Columbia. Tremendous resources including whole-genome resequencing data, multiple common gardens for association mapping, transcriptome and metabolite data and expression networks are available for this species (Chhetri et al., 2019; Weighill et al., 2019). We present a genome-wide association study of wood anatomical traits for the first time for this species. We also present a GWAS for important morphological traits from the same trees that together with anatomical traits affect overall plant productivity. We complemented the single trait GWAS with multitrait analyses, followed by multi-omic LOE and functional enrichment analyses. Since this study is based on data collected from a common garden in Clatskanie, Oregon this allowed for the direct comparison of GWAS genes identified for some of the same traits from other plantations (Evans et al., 2014; Chhetri et al., 2019).

MATERIALS AND METHODS

Phenotypic Data Collection

Wood anatomical, and morphological trait data were collected from a field trial of 1,100 *P. trichocarpa* genotypes that was

established in Clatskanie, Oregon in 2009 (**Figure 1**). These genotypes were previously collected from across the natural range of *P. trichocarpa* from northern California to northern British Columbia and were clonally replicated and planted in a randomized block design with three replicates of each genotype at 2 m x 3 m spacing in the field trial (Slavov et al., 2012; Evans et al., 2014). Phenotyping was performed on clonal replicates for all traits as described in **Table 1**. In some cases replicate observations were recorded for a subset of the genotypes, with the number of replicated genotypes ranging from 59 to 860.

Wood Anatomy

In June 2012, 557 trees were sampled for wood anatomical traits. Wood cores of 5.5 mm diameter were extracted from bark to pith from the southern face of the main trunk of the tree using an increment borer. Free hand cross-sections were made from the previous year's growth ring and fixed in 70% alcohol. The tissue sections were stained in 1% Safranin O solution for 30 s before preparing the slides for imaging. Images at 100x magnification were taken to sample the early, intermediate and late wood from the growth ring. These images were used for measuring vessel density, size and number (**Figure 2**). All images were processed using the software *imageJ* to extract quantitative measurements—vessel count, density and size (Schindelin et al., 2015).

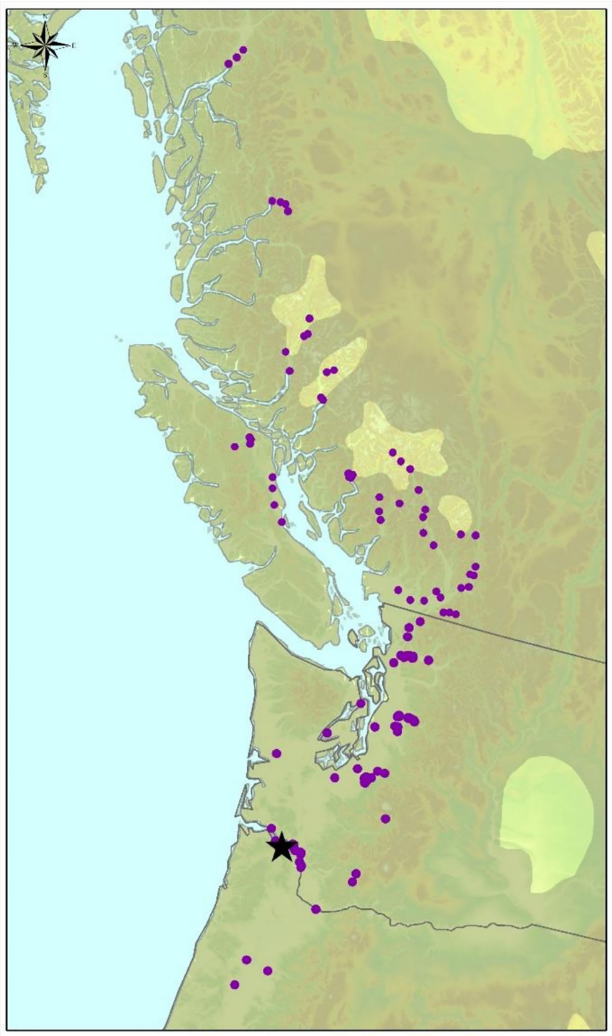

FIGURE 1 | Source locations of 869 P. *trichocarpa* genotypes sampled in this study (purple dots). The trees were grown in a common garden in Clatskanie, Oregon, USA (black star).

Leaf Anatomy

Leaf characteristics were measured for 676–687 trees (**Table 1**). The first and second fully expanded leaf (counting from the apex) were collected from a branch receiving full sunlight. One of the leaves was used for measuring petiole length and minimum and maximum petiole diameter with a caliper and then scanned using a hand-held scanner. The software *imageJ* (Schindelin et al., 2015) was used to estimate leaf area, leaf circularity, leaf length, leaf width, and leaf perimeter. Dry and wet weights were measured for the same leaf and leaf area and leaf dry weight were used to estimate the specific leaf area (SLA). Leaf chlorophyll content (SPAD) was assessed using a SPAD 502 Plus meter (Spectrum Technologies) with an average of three replicate measures on a leaf section. The second leaf was used for measuring the abaxial stomatal density. Clear nail polish was applied to the broadest part of the leaf close to the midrib. A clear piece of tape was then used to capture an imprint of the epidermal leaf surface and stored on microscope slides. The number of stomata in 1 mm^2 area in four random microscopic fields at 400x magnification were counted.

Leaf Metabolites

Leaf material was collected from 851 P. *trichocarpa* genotypes. Metabolites were extracted from tissues, measured by GCMS, and peaks were characterized (Weighill et al., 2018). For each sample, a Median Absolute Deviation (MAD) threshold of six was used to remove phenotypic values that were more than six MADs from the population median for that phenotype. In addition, phenotypes with greater than 20% missing data were removed prior to analysis, resulting in a total of 816 metabolite phenotypes. A Benjamini-Hochberg false-discovery rate cutoff of 0.1 was applied to the adjusted *p*-values.

Wood Composition

The wood cores were ground, processed and analyzed by pyrolysis molecular beam mass spectrometry (py-MBMS) as previously described (Muchero et al., 2015; Weighill et al., 2018). py-MBMS data were TIC normalized to account for variation in sample mass. Phenotypes with > 20% missing data or zeroes were removed prior to analysis (one phenotype, "m/z 76"). For each sample, a Median Absolute Deviation (MAD) threshold of five was used to remove phenotype values that were more than five MADs from the population median for that phenotype, resulting in the removal of an average of 1.49 py-MBMS phenotype values per sample (in total, 625 phenotype values).

TABLE 1 | Broad-sense heritability estimates (H^2) and the number of SNP-trait associations for wood anatomical and morphological traits in *P. trichocarpa*.

Trait	H^2	Genotypes	Total trees[a]	SNPs with FDR<0.1[b]	Chip_H^2 (+/- CI)[c]
Wood anatomy					
Early wood vessel area	0.335	489	548 (59)	0	0.542 (+/- 0.545)
Early wood vessel count	0.38	489	548 (59)	0	0.611 (+/- 0.582)
Early wood vessel size	0.392	489	548 (59)	0	0.199 (+/- 0.325)
Intermediate wood vessel area	0.196	570	633 (63)	0	0.475 (+/- 0.300)
Intermediate wood vessel count	0.233	570	633 (63)	0	0.562 (+/- 0.484)
Intermediate wood vessel size	0.218	570	633 (63)	1	0.301 (+/- 0.406)
Late wood vessel area	0.114	557	636 (79)	0	0.373 (+/- 0.359)
Late wood vessel count	0.146	557	636 (79)	0	0.327 (+/- 0.388)
Late wood vessel size	0.22	557	636 (79)	0	0.373 (+/- 0.359)
Morphology					
Diameter (breast height)	0.227	869	2438 (860)	9 (3)	0.929 (+/- 0.278)
Height	0.294	869	2438 (860)	0	0.965 (+/- 0.257)
Leaf area	0.434	676	813 (137)	0	0.87 (+/- 0.349)
Leaf aspect ratio	0.251	676	813 (137)	0	0.439 (+/- 0.320)
Leaf circularity	0.285	676	813 (137)	22 (2)	0.313 (+/- 0.318)
Leaf dry weight	0.452	685	836 (151)	27 (21)	0.99 (+/- 0.325)
Leaf length	0.477	676	813 (137)	0	0.679 (+/- 0.320)
Leaf perimeter	0.459	676	813 (137)	0	0.786 (+/- 0.357)
Leaf wet weight	0.393	687	842 (155)	33 (27)	0.985 (+/- 0.351)
Leaf width	0.386	676	813 (137)	0	0.905 (+/- 0.355)
Maximum petiole diameter	0.426	687	840 (153)	0	0.928 (+/- 0.323)
Minimum petiole diameter	0.249	683	834 (151)	0	0.843 (+/- 0.345)
Petiole length	0.397	685	839 (154)	0	0.987 (+/- 0.335)
SPAD	0.214	687	843 (156)	0	0.232 (+/- 0.282)
Specific leaf area	0.122	667	797 (130)	0	0.038 (+/- 0.318)
Stomatal density	0.477	721	884 (163)	20 (4)	0.597 (+/- 0.294)

[a]*Number of genotypes, with clonal replicates in parentheses.*
[b]*Number of SNPs at 10% FDR; Number in parentheses represents SNPs with FDR<0.05.*
[c]*Mean chip heritability values for phenotypes, with confidence interval (CI).*

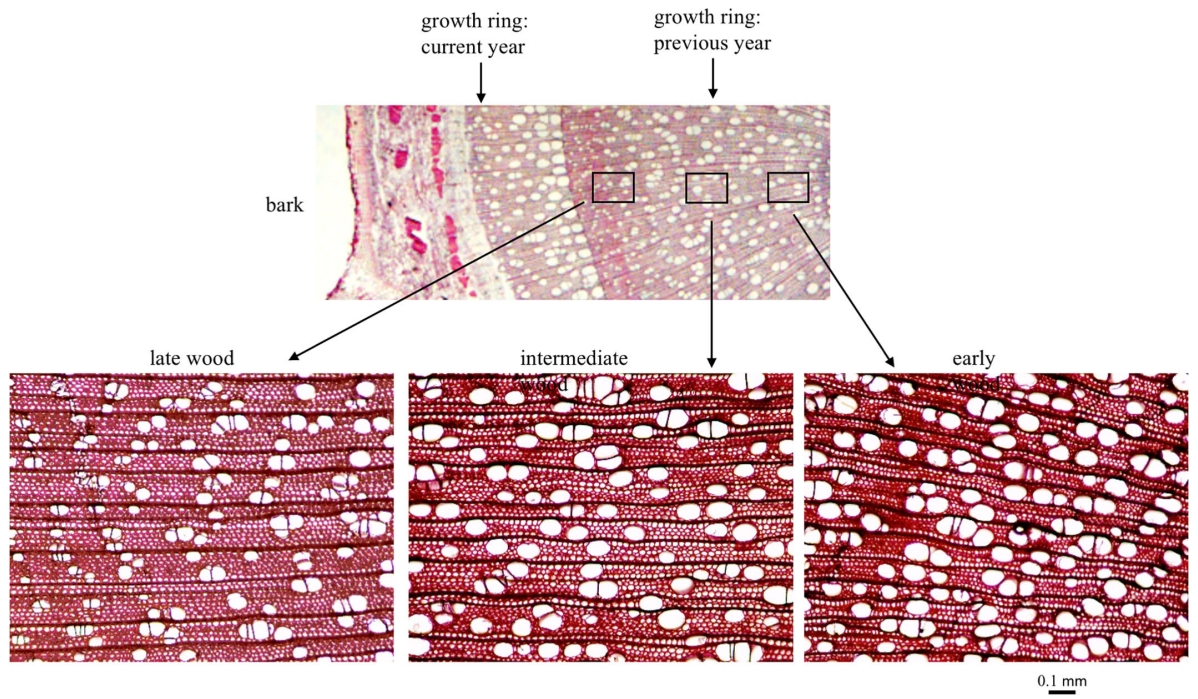

FIGURE 2 | *Populus trichocarpa* wood cross section from the previous year's growth ring showing variation in early, intermediate and late wood types.

Statistical Analyses

To estimate the genetic control of quantitative traits, broad-sense heritability (H^2) was estimated for all traits using the genotypes with replicate clonal measurements using the following formula:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_E^2},$$

where, σ_G^2 is genotypic variance due to clonal differences and σ_E^2 is environmental variance.

Outliers were removed, and the data were evaluated for normality. Variance components were estimated employing the linear regression model with the *lmer* and *ranef* functions of the *lme4* package implemented in R. Genotype and the position of the tree (i.e. row and column) in the garden was used as a random effect in the model. Error was estimated from the residuals of the model. Genetic correlation between the traits was estimated using the Best Linear Unbiased Predictors (BLUPs) from the same model (**Supplementary Figure S1**). Phenotypic BLUPs were also used for estimating the correlation between the phenotypic traits and 26 geoclimate variables and their first four principal components of the source location of the trees. Genetic correlations were also calculated between phenotypic traits collected at this site to those collected in a clonally replicated trial with the same genotypes in Corvallis, Oregon (Chhetri et al., 2019).

Genotypic Data

Methods for obtaining genotypic data were as described previously (Weighill et al., 2018; Chhetri et al., 2019). Briefly, whole genome re-sequencing data was obtained from 1,053 trees using Illumina genetic analyzers at the DOE Joint Genome Institute. After removing trees related more closely than first cousins and highly differentiated California trees, 869 trees were left, which were used for all analyses. A genetic relationship matrix was estimated using GEMMA and used as a covariate in the GWAS analyses. Furthermore, SNPs with minor allele frequency ≤0.05 were removed.

Association Analysis

The tests for statistical association for all phenotypic traits were conducted using the Genome-wide Efficient Mixed Model Association package (GEMMA, Zhou & Stephens, 2012; Zhou & Stephens, 2014), except for the py-MBMS data, which were analyzed using the efficient mixed-model association eXpedited (EMMAX) package (Kang et al., 2010). Phenotypic BLUPs, a genetic relationship matrix and 6,741,160 SNPs were used for the association test. Single trait GWAS was run for 25 phenotypes (**Table 1**). The tested model was:

$$y = x\beta + u + \epsilon,$$

where y is an n-vector of phenotypic BLUP values, where n is the number of individuals tested; x is an n-vector of marker genotypes, β is the effect size of the marker, u is an n-vector of random effects that includes a relatedness matrix and ϵ is an n-vector of errors. We also tested Principal Components as covariates in the model, as described previously (Evans et al., 2014), but we determined that inclusion of PCs resulted in overcorrection, as determined by a preponderance of points below the 1:1 line in Q-Q plots (not shown).

Multitrait GWAS was run for 9 sets of phenotype combinations (**Table 2**). The same procedure was used for selecting trait sets as described in Chhetri et al. (2019). Briefly, traits were combined if there was reason to believe that they could be functionally related. For example, leaf traits were combined with wood anatomy traits due to the expected interactive roles of both sets of traits in water stress tolerance (Sperry, 2000). Multitrait association was conducted with GEMMA using the same model as for single trait associations, except y is an $n \times d$ matrix of d phenotypes for n individuals.

We used an FDR cutoff of 0.05 and a more liberal FDR cutoff of 0.1 to identify suggestive associations (Storey and Tibshirani, 2003). For the purpose of summarizing the results, significant SNPs within 10 kb of one another were merged into peaks. Gene models that were closest to significant SNPs were identified based on v3 of the *P. trichocarpa* genome. Annotation information including gene expression level in different plant tissues and annotation of putative gene function was obtained from Phytozome (Goodstein et al., 2012). Percentage of variance explained (PVE) by SNPs was estimated using the formula in Shim et al. (2015).

Multiple Lines of Evidence (LOE) Analysis

A network was created using five data layers generated from *P. trichocarpa* genetic and phenotypic analyses: pairwise gene coexpression, pairwise gene comethylation, metabolite GWAS, rare variant regional metabolite GWAS, and py-MBMS GWAS. The first four of these layers were previously described (Furches et al., 2019). The py-MBMS GWAS analysis was performed for this study. For each layer, a network was constructed that consisted of gene pairs (nodes) linked by a correlation value or association p-value (edge).

Coexpression Network

P. trichocarpa (Nisqually-1 reference) RNA-seq data was obtained from the DOE Joint Genome Institute Plant Gene Atlas (available from Phytozome at: https://phytozome.jgi.doe.gov/pz/portal.html). For each sample, expression was measured in leaf, stem, root and bud tissues at multiple developmental stages and four different conditions (three nitrogen treatments and an untreated control). Plant cultivation, tissue collection, and RNA extraction and sequencing were previously described in Weighill et al. (2018), as were RNA-seq read trimming, read alignment, TPM calculations, Spearman correlation analysis, and network construction. A correlation value threshold of 0.85 was set as criteria for retention of gene pairs in the coexpression network. As previously demonstrated, this threshold minimized false positives, and the resulting network was significantly different from random (Furches et al., 2019).

Comethylation Network

P. trichocarpa MeDIP-seq data (Vining et al., 2012) was obtained from Phytozome. This data set contained MeDIP-seq reads from ten tissues: bud, callus, male and female catkins, internode, regenerated internode, leaf, phloem, xylem, and root. Read

TABLE 2 | List of traits used for multitrait associations and significant SNPs identified in *P. trichocarpa*.

Trait combination	Abbreviation	Trait name	PopN[a]	SNPs with FDR<0.1[b,c]
Early wood vessel area, early wood vessel count	EWva_EWvc	wood anatomy multitrait 1	489	0
Early wood vessel area, early wood vessel count, intermediate wood vessel area, intermediate wood vessel count, late wood vessel area, late wood vessel count	EWva_EWvc_IWva_IWvc_LWva_LWvc	wood anatomy multitrait 7	411	0
Early wood vessel area, intermediate wood vessel area, late wood vessel area	EWva_IWva_LWva	wood anatomy multitrait 4	411	0
Early wood vessel area, leaf area, stomatal density	EWva_LA_SD	morphology and wood anatomy multitrait	425	4
Early wood vessel count, intermediate wood vessel count, late wood vessel count	EWvc_IWvc_LWvc	wood anatomy multitrait 5	411	0
Early wood vessel size, intermediate wood vessel size, late wood vessel size	EWvs_IWvs_LWvs	wood anatomy multitrait 6	411	0
Intermediate wood vessel area, intermediate wood vessel count	IWva_IWvc	wood anatomy multitrait 2	570	3
Leaf area, leaf dry weight, leaf length, lead width	LA_LD_LL_LW	leaf morphology multitrait	674	9 (8)
Late wood vessel area, late wood vessel count	LWva_LWvc	wood anatomy multitrait 3	557	3 (3)

[a]*Number of unique genotypes.*
[b]*Number of significant SNPs in multi-trait GWAS.*
[c]*Numbers in parentheses represent SNPs with FDR <0.05.*

mapping, calculation of transcripts per million reads, Spearman correlation analysis, and network construction were conducted as previously described (Weighill et al., 2018). Gene pairs with correlation values greater than or equal to 0.95 were retained in the comethylation network (Furches et al., 2019).

Rare Variant Regional Metabolite GWAS Networks

Leaf metabolite data were analyzed using a rare variant regional GWAS analysis as described elsewhere (Weighill et al., 2018). This analysis jointly tested rare SNPs (MAF ≤ 0.01) within each gene's boundary and the 2-kb up- and downstream flanking regions of that gene using a linear mixed model as implemented by SKAT in RVtest (Zhan et al., 2016). An FDR cutoff of 0.1 was applied to the resulting p-values.

LOE Scoring of Candidate Genes

The purpose of this LOE method was to explore the networks of genes underlying the traits analyzed in the primary leaf and wood anatomy GWAS analyses to assess support for the functional hypotheses suggested by the gene annotations. For each of the significant (FDR<0.1) gene models identified by the single trait and multitrait GWAS analyses of wood and leaf anatomy traits (hereafter referred to as "anchor genes"), a merged network of anchor genes and their nearest one-hop neighbors in six data layers (including the gene-trait associations detected in this study) was created. Network statistics were calculated, including breadth (the number of layers to which a gene is connected), depth per layer (the number of nodes to which the gene is connected), and total depth. Anchor genes were directly scored based on the number of shared edges in each layer rather than scoring the nearest neighbors of anchor genes. These scores were then used to prioritize genes for further analysis.

Functional Enrichment of Coexpression Nearest Neighbors

The purpose of this analysis is to test whether functional categories represented in the one-hop networks occur more frequently than expected from a random draw of the same number of genes from the genome. For each anchor gene, two separate Gene Ontology (GO) functional enrichment analyses were performed on positively and negatively coexpressed one-hop neighbors, respectively using the software package Fuento (Weichselbaum et al., 2017). Fisher's Exact tests were performed for all available namespaces (molecular function, cellular compartment, biological process, aberrant functions) using 100 tries and an FDR adjusted p-value cutoff of p ≤ 0.1.

RESULTS

Heritabilities of Wood Anatomy and Morphology Traits

Overall the broad-sense heritabilities estimated for wood anatomy and morphology related traits in this study were low to moderate (**Table 1**). Broad-sense heritabilities for wood anatomy traits ranged from 0.114 for late wood vessel area to 0.392 for early wood vessel size. Late wood anatomy (late wood vessel count, vessel area and vessel size) had low heritabilities compared to other wood traits. For morphology related traits, broad-sense heritabilities ranged from 0.122 for specific leaf area to 0.477 for stomatal density.

Heritability estimates derived from the SNP data ("chip heritabilities") for wood anatomy traits ranged from 0.199 for early wood vessel size to 0.611 for early wood vessel count whereas for morphology related traits it ranged from 0.038 for specific leaf area to 0.965 for tree height (**Table 1**). Although there was no correlation between broad-sense and chip heritabilities, low chip heritability was generally reflective of low broad-sense heritability and there was a significant correlation between the sample size (number of genotypes used for GWAS) and chip heritability values (r=0.441, *P*=0.021) for the traits measured in this study (**Table 1**). GWAS results for traits with low chip heritabilities should be treated with caution.

Genetic Correlation of Phenotypic Traits Within and Between the Common Gardens

As expected, most morphological traits measured in the Clatskanie common garden were correlated to each other—leaf traits such as leaf area, leaf length, leaf dry weight, leaf wet weight, petiole diameter and length, leaf perimeter, and leaf aspect ratio were highly significantly correlated to each other and SPAD and stomatal density had low, but significant correlations with most leaf traits (**Supplementary Figure S2**). Similarly, most wood traits were correlated to each other. Wood anatomy traits such as vessel area and size and vessel count within each of the growth ring regions (early, intermediate or late wood indicating a different maturation stages) were correlated to each other. Wood anatomy traits were also significantly correlated to each other across growth ring areas (**Supplementary Figure S2**). Furthermore, some morphological traits such as tree height and diameter were significantly correlated with wood traits such as, early wood vessel size and area, intermediate wood vessel size and count, and late wood vessel size and count (**Supplementary Figure S2**).

Genetic correlation of the same phenotypic traits measured between the Clatskanie and Corvallis common gardens showed that most traits were weak to moderately, but significantly correlated to each other (**Supplementary Figure S3**). Pairwise genetic correlations of tree height (r=0.436, $p<0.001$) and stomatal density (0.345, $p<0.001$) between the sites showed the strongest correlations among all traits compared (**Supplementary Figure S3**).

Phenotypic Trait Correlations With Climate Variables

As seen for the phenotypic traits in the Corvallis common garden (Chhetri et al., 2019), most phenotypic traits had significant correlations with latitude, and therefore the correlations between the phenotypic traits and other geoclimate variables cannot easily be disentangled. Nevertheless, most morphological traits had significant (although weak) correlations with most geoclimatic variables (**Supplementary Figure S4**). Similarly, wood traits such as late wood vessel area had significant (although weak) correlations with most geoclimatic variables (**Supplementary Figure S4**).

Genes Identified From Single Trait and Multitrait GWAS

We performed single trait GWAS with 6.741 million SNPs for 25 morphological and wood anatomical traits. Fifty-seven SNPs passed the FDR threshold of $p<0.05$ (**Table 1**). However, we identified a total of 112 SNPs that passed suggestive association FDR *P*-value cutoff of 0.1 (**Table 1**, **Figures 3** and **4**, **Supplementary Figures S5** and **S6**, **Supplementary Table S1**). These associated SNPs belonged to 12 separate SNP peaks, and 21 P. *trichocarpa* gene models were identified as nearest neighbors to the significant SNPs (**Table 3**, **Supplementary Table S2**). Percentage of variance explained (PVE) for significant SNPs ranged from 3.30% to 5.65% (**Supplementary Table S1**).

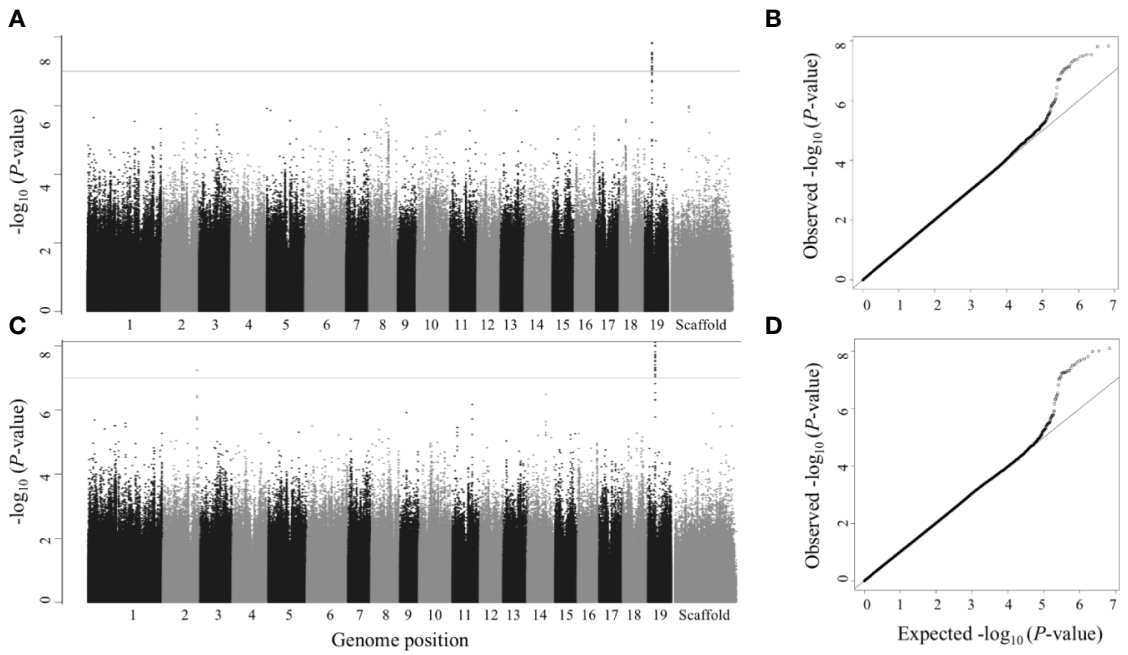

FIGURE 3 | Single trait GWAS – Manhattan (left) and QQ plots (right). Numbers 1 to 19 represent chromosomes; scaffolds are the reads that did not align to any of the 19 chromosomes. **(A, B)** Leaf dry weight; **(C, D)** Leaf wet weight. SNPs above gray line have $p<1\times10^{-7}$, which is roughly equivalent to FDR ≤ 0.1 and SNPs above black line (the top line) have $p<7.417\times10^{-9}$ (the Bonferroni correction threshold), which is roughly equivalent to FDR ≤ 0.05 in this study.

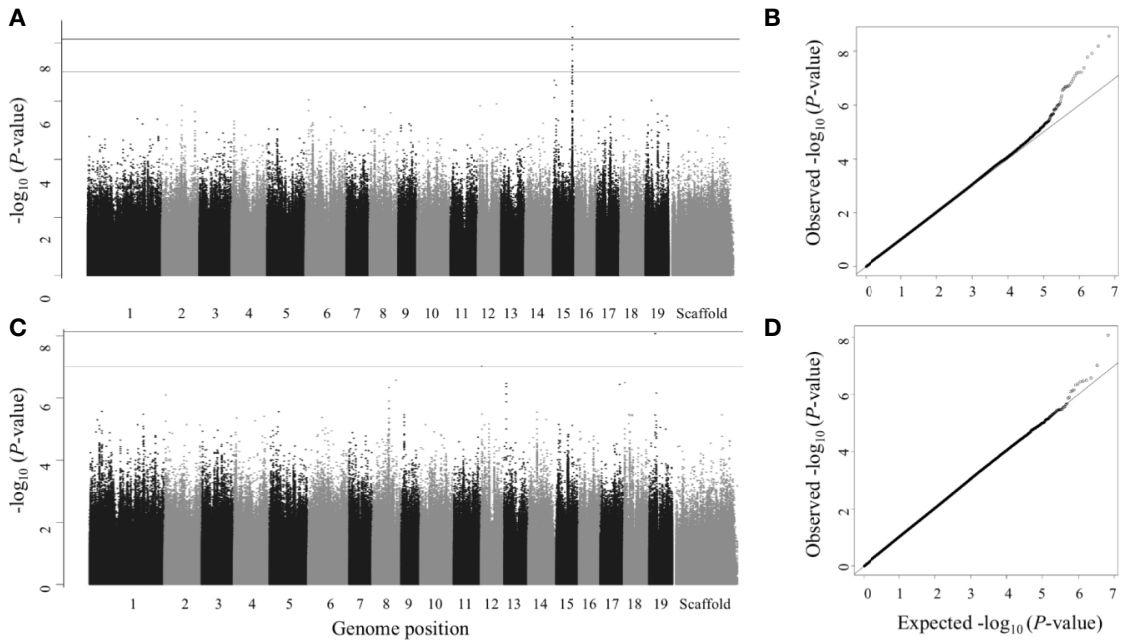

FIGURE 4 | Single trait GWAS – Manhattan (left) and QQ plots (right). Numbers 1 to 19 represent chromosomes; scaffolds are the reads that did not align to any of the 19 chromosomes. **(A, B)** Stomatal density; **(C, D)** Vessel size (intermediate wood). SNPs above gray line have $p<1\times10^{-7}$, which is roughly equivalent to FDR ≤ 0.1 and SNPs above black line (the top line) have $p<7.417\times10^{-9}$ (the Bonferroni correction threshold), which is roughly equivalent to FDR ≤ 0.05 in this study.

TABLE 3 | Genes identified from *P. trichocarpa* single trait GWAS.

Gene model[a]	Trait	P-value[b]	FDR[c]	Functional annotation
Potri.014G117300	Diameter (breast height)	1.95E-08	0.046	Gibberellin 2-beta-dioxygenase 2
Potri.014G117400	Diameter (breast height)	1.88E-08	0.046	similar to MYB family transcription factor
Potri.018G141400	Diameter (breast height)	5.67E-08	0.051	ABC transporter
Potri.019G037900	Intermediate wood vessel size	8.44E-09	0.057	Staufen and related double-stranded-RNA-binding proteins
Potri.005G231500	Leaf circularity	2.46E-07	0.094	Similar to expressed protein in Arabidopsis thaliana; co-ortholog of At1g75400, At1g19680
Potri.005G247200	Leaf circularity	3.27E-07	0.100	PF01535/PF13041 - PPR repeat (PPR)
Potri.005G247400	Leaf circularity	9.20E-08	0.094	similar to prolyl oligopeptidase
Potri.005G247600	Leaf circularity	4.63E-08	0.094	similar to maoC-like dehydratase domain-containing protein
Potri.005G247700	Leaf circularity	1.37E-08	0.047	multi-copper oxidase type 1 family protein
Potri.005G247900	Leaf circularity	9.01E-08	0.094	ubiquitin family protein
Potri.019G042700	Leaf dry weight, leaf wet weight	9.73E-09	0.018	similar to expressed protein in *Arabidopsis* thaliana; co-orthologof At4g27620, At4g27610
Potri.019G042600	Leaf dry weight, leaf wet weight[d]	1.49E-08	0.018	MITOCHONDRIAL OUTER MEMBRANE PROTEIN 25
Potri.002G238700	Leaf wet weight	5.74E-08	0.018	similar to hypothetical protein; co-ortholog of At5g48890
Potri.014G159800	Leaf wet weight	3.32E-07	0.077	Lysosomal acid lipase/cholesteryl ester hydrolase (LIPA)
Potri.015G015200	Stomatal density	1.95E-07	0.084	PF02362 - B3 DNA binding domain
Potri.015G028600	Stomatal density	2.81E-07	0.095	KOG0472 - Leucine-rich repeat protein
Potri.015G113800[e]	Stomatal density	1.62E-07	0.084	PF04578/PF13968 – Protein of unknown function; DUF594, DUF4220
Potri.015G117300	Stomatal density	6.49E-09	0.022	NA
Potri.015G117400	Stomatal density	6.11E-08	0.056	PROTEIN ARGONAUTE 2-RELATED
Potri.015G117500	Stomatal density	2.77E-09	0.019	Peptide-O-fucosyltransferase/GDP-L-fucose:polypeptide fucosyltransferase
Potri.015G117600	Stomatal density	8.31E-08	0.062	zinc finger (C3HC4-type RING finger) family protein;

[a]*Gene models are annotated using v3 of the P. trichocarpa genome.*
[b]*SNP p values <1 × 10^{-7}.*
[c]*FDR at 10% level of significance.*
[d]*smallest p-value reported (leaf wet weight).*

Multitrait GWAS performed on 9 sets of traits identified 11 SNPs that passed the FDR *p*-value cutoff of 0.05 and 19 SNPs that passed the suggestive association FDR *p*-value cutoff of 0.1. (**Table 2**, **Figure 5**, **Supplementary Figure S7**, **Supplementary Table S3**). These SNPs belonged to 13 separate SNP peaks and were within or close to 13 P. *trichocarpa* gene models (**Table 4**, **Supplementary Table S4**). PVE of these SNPs ranged from 0.01% to 6.18% for the individual traits comprising the multitrait set (**Supplementary Table S3**).

Out of the single trait GWAS for nine wood anatomical traits, we identified only one gene model belonging to intermediate wood vessel size (**Figures 4C, D**). A total of three gene models for seven wood anatomy multitrait sets (**Figures 5C, D**) and three gene

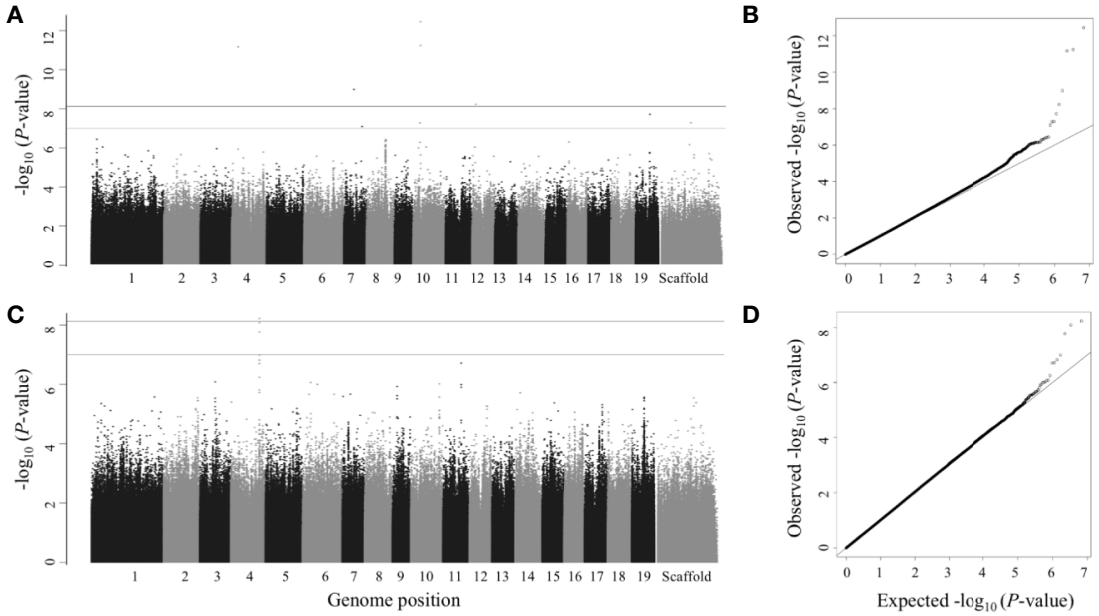

FIGURE 5 | Multitrait GWAS—Manhattan (left) and QQ plots (right). Numbers 1 to 19 represent chromosomes; scaffolds are the reads that did not align to any of the 19 chromosomes. **(A**, **B)** Leaf area, leaf dry weight, leaf length and leaf width; **(C**, **D)** Vessel area and vessel count (late wood). SNPs above gray line have $P<1\times10^{-7}$, which is roughly equivalent to FDR ≤ 0.1 and SNPs above black line (the top line) have $P<7.417\times10^{-9}$ (the Bonferroni correction threshold), which is roughly equivalent to FDR ≤ 0.05 in this study.

TABLE 4 | Genes identified from *P. trichocarpa* multitrait GWAS.

Gene model[a]	Trait	P-value[b]	FDR[c]	Functional annotation
Potri.016G013400	Early wood vessel area, leaf area, stomatal density	3.20E-08	0.054	K15223— upstream activation factor subunit UAF30, SPP27
Potri.004G002600	Early wood vessel area, leaf area, stomatal density	9.57E-09	0.054	UBIQUINOL OXIDASE 4, CHLOROPLASTIC/CHROMOPLASTIC
Potri.015G113800[d]	Early wood vessel area, leaf area, stomatal density	2.22E-08	0.054	Protein of unknown function: DUF594, DUF4220
Potri.006G275800	Intermediate wood vessel area, intermediate wood vessel count	2.37E-10	0.093	LEUCINE-RICH REPEAT-CONTAINING PROTEIN
Potri.001G058100	Intermediate wood vessel area, intermediate wood vessel count	4.15E-08	0.093	similar to expressed protein in Arabidopsis thaliana (co-ortholog of At1g27290)
Potri.004G183900	Late wood vessel area, late wood vessel count	5.91E-09	0.027	SERINE/THREONINE-PROTEIN KINASE
Potri.004G056300	Leaf area, leaf dry weight, leaf length, leaf width	6.88E-12	0.000	Fruit bromelain
Potri.007G061600	Leaf area, leaf dry weight, leaf length, leaf width	1.03E-09	0.002	vacuolar protein sorting-associated protein 35 (VPS35)
Potri.007G099700	Leaf area, leaf dry weight, leaf length, leaf width	8.07E-08	0.059	NB-ARC domain (NB-ARC)/TIR domain (TIR_2)/Leucine rich repeat (LRR_8)
Potri.010G031900	Leaf area, leaf dry weight, leaf length, leaf width	5.23E-08	0.043	similar to RUB-activating enzyme (Ubiquitin activating enzyme E1 like protein)
Potri.010G032600	Leaf area, leaf dry weight, leaf length, leaf width	3.62E-13	0.000	small subunit ribosomal protein S11 (RP-S11, MRPS11, rpsK)
Potri.012G028700	Leaf area, leaf dry weight, leaf length, leaf width	5.90E-09	0.008	Ras suppressor protein (contains leucine-rich repeats)
Potri.019G067300	Leaf area, leaf dry weight, leaf length, leaf width	1.90E-08	0.021	BETA-1,3-GALACTOSYLTRANSFERASE 8-RELATED

[a]*Gene models are annotated using v3 of the P. trichocarpa genome.*
[b]*SNP p values <1 × 10^{-7}.*
[c]*FDR at 10% level of significance.*
[d]*Potri.015G113800 is shared between single and multitrait GWAS.*

models for one morphology and wood anatomy multitrait set (**Figures S6A, B**) were identified.

LOE Scores and Rankings

Breadth scores ranged from one to four (**Supplementary Table S5**, **Supplementary Figure S8**). Two anchor genes had breadth scores of four: Potri.019G067300 and Potri.010G031900. Both of these genes were associated with the multitrait GWAS for leaf morphology (LA_LD_LL_LW). They have connections to coexpressed and comethylated genes, metabolite, and py-MBMS phenotypes (**Supplementary Tables S6** and **S7**). Two anchor genes had breadth scores of three, both of which were identified through single trait GWAS analysis: Potri.019G037900 (wood anatomy), and Potri.014G159800 (morphology). Total depth scores ranged from one to 2,014. The ranges of layer specific depth scores were: zero to 1,115 for coexpression, zero to 1,914 for comethylation, zero to one for metabolite GWAS, zero to one for RV metabolite GWAS, and zero to one for py-MBMS GWAS. A total of 7630 unique coexpressed and comethylated genes, four metabolites and three py-MBMS GWAS were found to be associated with 33 GWAS (defined as anchor) genes for single and multitrait GWAS (**Figures 6–8**, **Supplementary Tables S6** and **S7**, **Supplementary Figures S9–S13**).

Functional Enrichment Results

Thirteen anchor genes had significantly enriched functions among positively coexpressed one-hop neighbors. However, after FDR correction, only five anchor genes had enriched neighborhoods (**Supplementary Table S8**). Seven anchor genes had significantly enriched functions among negatively coexpressed neighbors, but only a single enrichment for one anchor gene remained after FDR correction. Overall there was an abundance of signaling, transport, growth and development, RNA regulation, hormone metabolism, cell wall related and stress related genes for the coexpressed and comethylated genes in the network for all anchor genes (**Supplementary Figure S14**).

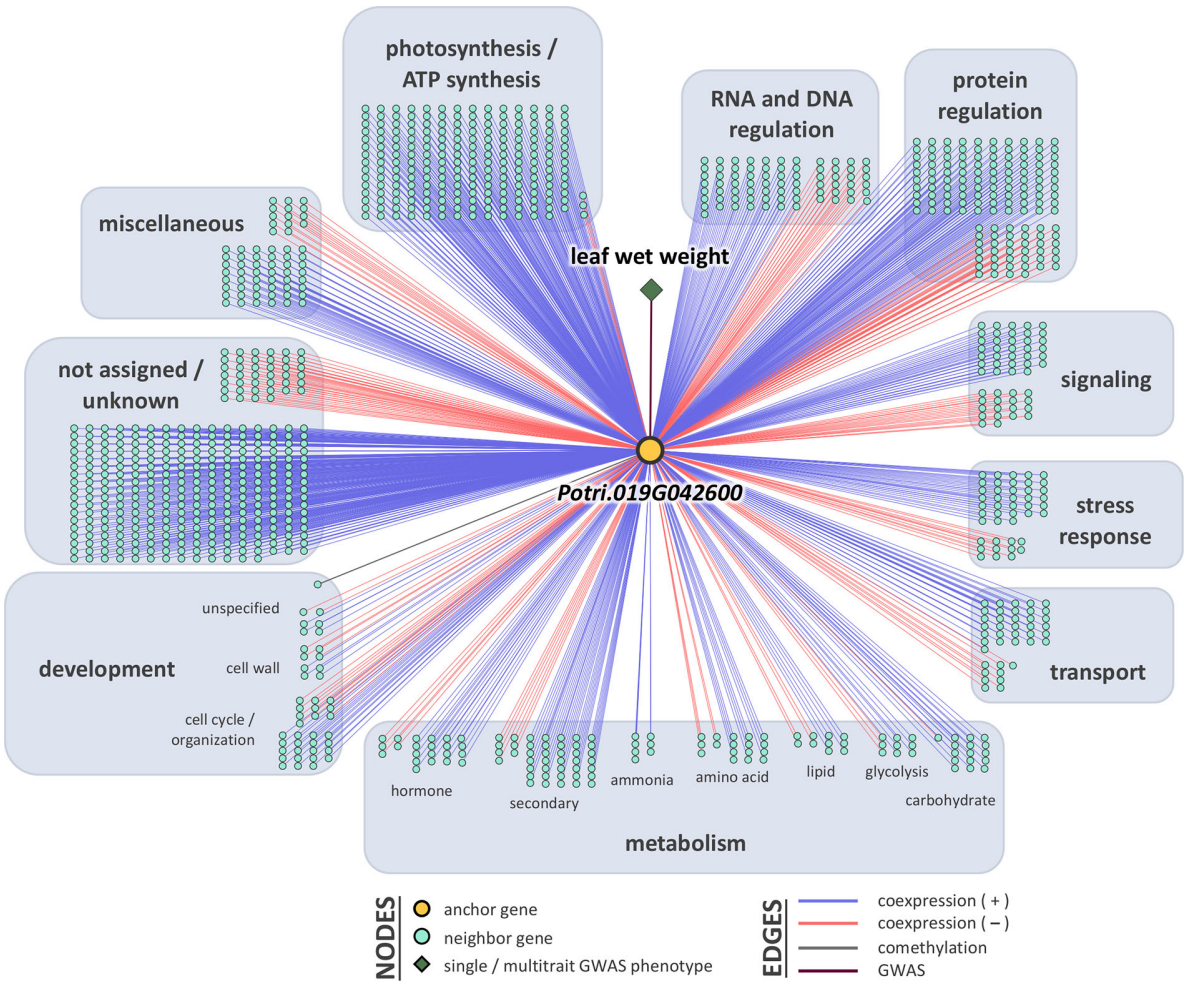

FIGURE 6 | One-hop multi-omic network for the gene Potri.019G042600, which was associated with leaf wet weight.

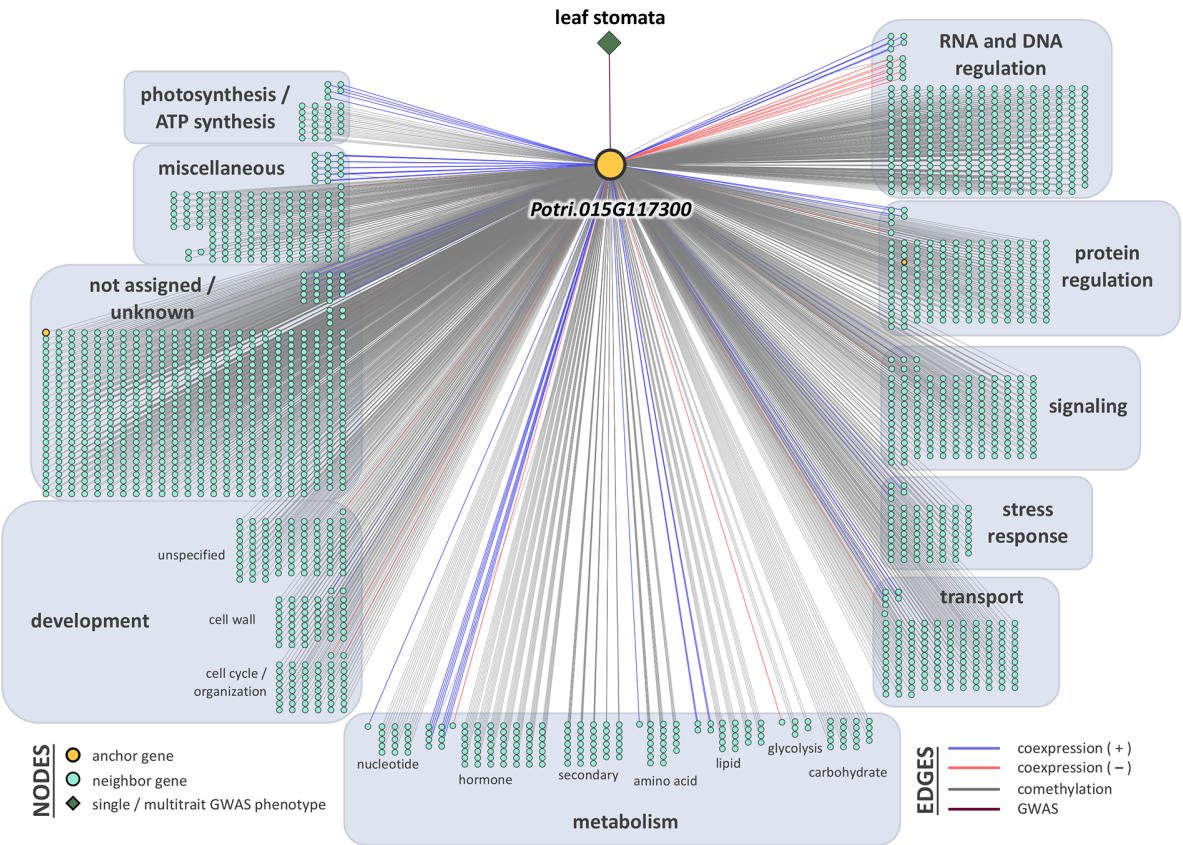

FIGURE 7 | One-hop multi-omic network for the gene Potri.015G117300, which was associated with leaf abaxial stomatal density.

Shared LOE for Single Trait and Multi-Trait Associations

Only one gene, Potri.015G113800, was significant for both a multitrait set (early wood vessel area-leaf area-stomatal density) and for one of the traits making up the set (stomatal density). We also searched for overlaps between the one-hop networks for multi-traits and the individual traits, but again this was the only case where there were shared genes in the one-hop networks (**Supplemental Figure S15**).

DISCUSSION

Patterns of Genetic Variation

Broad-sense heritabilities for most traits in this study ranged from 0.2 to 0.4, suggesting that the traits were under moderate genetic control. Heritability estimates for morphological traits were comparable to similar studies (Evans et al., 2014; Mckown et al., 2014; Chhetri et al., 2019). However, some wood related traits and SLA had low heritabilities (<0.2), which may be due to relatively small sample size compared to morphological traits. Low heritability may also be due to micro-environmental variation in the common garden, which can have a strong effect on the growth ring patterns and the anatomical traits studied here. Furthermore, some imprecision in wood trait phenotypes may result from the high throughput preparation method that we were forced to adopt due to the large sample numbers. Wood samples were prepared *via* free-hand sectioning, which results in less uniform images than can be obtained with embedded sections prepared with a microtome. In any case, the results for these low heritability traits should be interpreted with caution.

As reported in previous studies geography plays a major role in shaping adaptive trait variation in *P. trichocarpa*. Correlation of these adaptive traits with latitude can obscure the actual between-trait and trait-climate relationships (Chhetri et al., 2019). Most morphological traits such as tree height and leaf characteristics had strong significant correlations with latitude compared to previous studies (Mckown et al., 2014; Chhetri et al., 2019). In contrast to Chhetri et al. (2019), tree height including diameter at breast height had a very low or no correlation with leaf traits as in McKown et al. (2014c). However, tree height and diameter had significant negative correlations with intermediate and late wood vessel counts and significant positive correlations with vessel size of all maturation stages (early, intermediate and late wood). Interestingly, vessel density was positively correlated with latitude and vessel size was

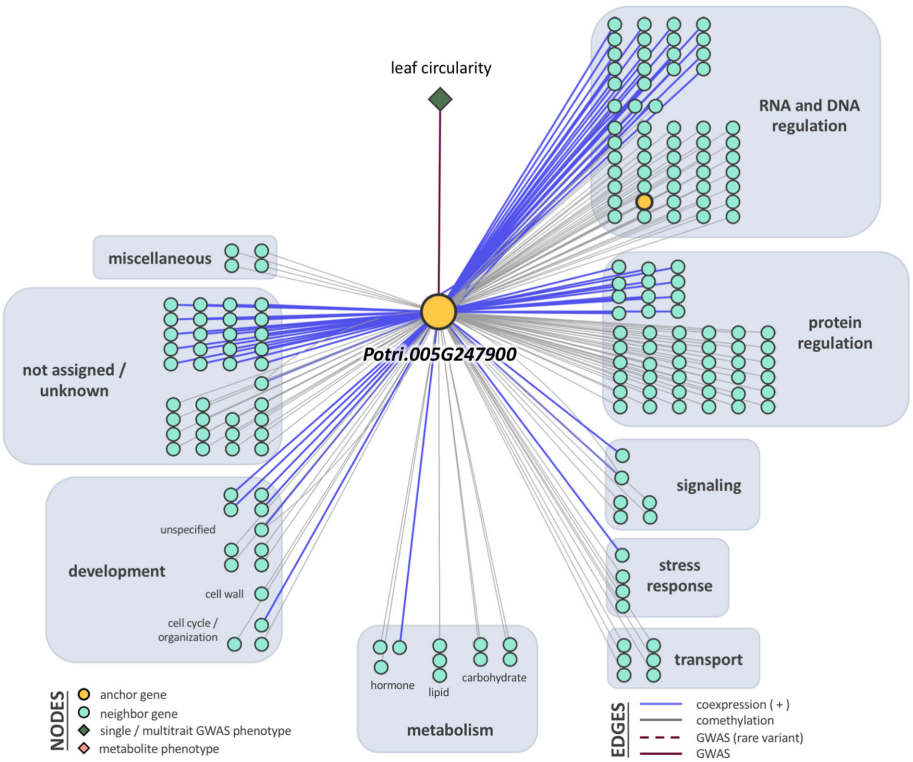

FIGURE 8 | One-hop multi-omic network for the gene Potri.019G067300, which was associated with the multitrait phenotype leaf area, leaf dry weight, leaf length, and leaf width (LA_LD_LL_LW).

negatively correlated with latitude. It is therefore difficult to deconvolute the indirect effects of latitudinal variation from the direct effects of wood anatomy on productivity in this study.

Further corroboration for the effects of environment shaping these traits were evident when comparing data from the two common gardens. Genetic correlation of the directly comparable traits between the Clatskanie and Corvallis common garden showed very low or no correlations except for stomatal density and tree height (**Supplementary Figure S3**). Furthermore, different loci seem to be controlling phenotypic traits in the two gardens, based on a complete lack of overlap in significant loci discovered by GWAS in the Corvallis, Oregon common garden (Chhetri et al., 2019). This might be due to differential effects of local environments on phenotypes in each of the respective common gardens, or may simply indicate that our analysis is underpowered.

As temperature decreases, and moisture and precipitation increase with latitude, relationships of the phenotypic traits with climate variables corroborates the trait relationships explained above. Leaf characteristics including leaf dry weight and petiole length and diameter had significant positive correlation with latitude, mean annual precipitation and relative humidity, but significant negative correlation with mean annual temperature. Similarly, abaxial stomatal density was negatively correlated with precipitation (**Supplementary Figure S4**). Taken together, southern trees had larger but less dense vessels, smaller leaves with higher abaxial stomatal density and low chlorophyll content, and larger tree height and diameter. Late wood vessel area (similar to the relationship of height and diameter with latitude and temperature) decreased with latitude and increases with temperature.

GWAS Genes

The GWAS analysis for wood anatomical traits we reported here is the first such study in *Populus*. While we found none to only a few significantly associated or suggestive SNPs for most traits we studied here, we think that this is mainly due to lack of power due to sample size although this is one of the most comprehensive GWAS analyses to date for *P. trichocarpa* in terms of the number of SNPs used. We did not find any gene model controlling more than one trait from our single trait GWAS analyses, but we identified 13 gene models with potential pleiotropic effects controlling sets of morphological and wood anatomical traits (**Table 4**).

Single Trait GWAS Genes

For single trait GWAS, we identified eight genes belonging to four traits based on 57 SNPs that passed the FDR threshold of 0.05 (**Table 3**). The top five genes identified were for two morphological traits, leaf wet weight (three genes) and stomatal density (two genes).

The gene model identified based on the top GWAS hit for both leaf wet weight and leaf dry weight (**Figure 3**) was Potri.019G042600 which is a mitochondrial outer membrane protein based on PANTHER annotation (https://phytozome.jgi.doe.gov/pz/portal.html). The gene is 6.1 kb downstream of the GWAS SNP (FDR = 0.015, p-value = 7.18×10^{-9}). Outer membrane proteins in mitochondria (and chloroplasts) play a role in important cellular functions such as protein import and ion exchange (Lee et al., 2014). The Arabidopsis homolog of this gene, AT5G17170 encodes rubredoxin family protein located in the plant chloroplast and is involved in electron transport and metal ion binding activity in Arabidopsis (https://www.arabidopsis.org/). Expression of rubredoxin family protein in a grass, *Puccinellia tenuiflora* (*PutRUB)* was shown to be induced by abiotic stresses due to NaCl, $NaHCO_3$, $CuCl_2$, and H_2O_2 suggesting a role of *PutRUB* in stress tolerance. Overexpression of *PutRUB* in Arabidopsis plants under stressed conditions showed an increase in seedling weight in overexpressed plants compared to wild type (Li et al., 2016). The GWAS gene has a breadth score of 2 and a total depth score of 1,116 (3rd highest among all single and multitrait anchor genes, **Figure 6**, **Supplementary Table S5**). It is highly expressed in late winter apical bud and fully open bud, stem, young and first fully expanded leaf and stem treated with ammonia, nitrate and urea (**Supplementary Table S7**). The anchor gene is coexpressed with 1,115 other genes (**Figure 6**) of which 915 are positively and 200 are negatively coexpressed. The coexpressed genes have functions related to photosystem I and II in photosynthesis (155 genes), oxidation reduction reaction (26), RNA regulation and binding (62), signaling (53, including light, G-protein, calcium, receptor kinases and sugar and nutrient physiology), abiotic and biotic stresses (25), transport (36), glycolysis (11), and hormone metabolism (25). All of the top 10 coexpressed genes are related to the photosynthesis electron transport chain including Photosystem I and Photosystem II complexes (**Supplementary Table S6**). Similarly, Potri.019G042600 showed significant functional enrichment for GO term Biological Processes (photosynthesis and respiration related) with functional annotations including ATP synthesis, oxidation reduction, glycolysis and photosynthesis. The gene also showed enrichment for GO term Molecular Function that included hydrolase, oxidoreductase, calcium ion binding, antioxidant activity and proton transport (**Supplementary Table S8**). In sum, network and functional enrichment analyses suggest that the function of this GWAS gene is related to photosynthetic and respiration processes and responses to biotic and abiotic stresses. One of the possible mechanisms by which plants respond to abiotic stresses, especially nutrient and water stress, is by changing morphology (e.g, leaf area and weight) and physiological processes (e.g., respiration) (Rymaszewski et al., 2017), thereby raising the expectation for coordinated responses of genes related to photosynthesis, oxidation reduction reactions, glycolysis, signaling, hormone metabolism, and stress responses, along with those controlling leaf morphology.

Potri.015G117500 was the top gene associated with stomatal density (p-value = 2.77×10^{-9}, **Figures 4A, B**). The top GWAS SNP is a synonymous substitution harbored within the gene. Potri.015G117500 encodes peptide-O-fucosyltransferase, which belongs to a family of genes involved in post-translational modification (Lira-Navarrete et al., 2011). The gene had a breadth score of 1 and the total depth score of 17 (comethylated with 17 other genes, **Supplementary Table S5**). Notable comethylated genes include a HSP20-like chaperone super family protein (Potri.009G153200), photosystem II oxygen evolving PsbQ complex subunit protein (Potri.011G031300) and farnesylcysteine lyase (Potri.007G101700) that indirectly has a negative regulatory effect in the ABA mediated signaling pathway in Arabidopsis (Huizinga et al., 2010). Another coexpressed gene, Potri.019G006900 encodes kunitz trypsin inhibitor and is involved in apoptosis during plant-pathogen interactions and responding to salt stress and salicylic acid. Kunitz trypsin inhibitors are known to have a function in pest resistance in poplar (Major and Peter Constabel, 2008). Similarly, another comethylated gene, Potri.001G045500, encodes cinnamoyl CoA reductase which is involved in lignin biosynthesis (Lacombe et al., 1997; Kawasaki et al., 2006). Lignin biosynthetic genes are known to affect growth and defense mechanisms in plants (Xie et al., 2018).

In *Arabidopsis* the O-fucosyltransferase SPINDLY (*SPY*, similar to the protein encoded by Potri.015G117500) affects DELLA and associated regulators such as brassinazole-resistant1 (*BZR1*) phytochrome-interacting-factor3 (*PIF3*) and *PIF4* in the brassinosteroid- and light-signaling pathways (Zentella et al., 2017). DELLA proteins are important integrators of multiple signaling pathways in flowering plants. DELLAs are repressors of phytohormone (GA) signaling, master growth repressors that restrict plant growth by affecting genes in cell division, expansion and differentiation (Zentella et al., 2007). In *Arabidopsis* O-fucosyltransferase modifies the DELLA protein RGA (repressor of *ga1-3*). The GA-deficient mutant *ga1-3* is a nongerminating, extreme dwarf that flowers late and produces male-sterile flowers. This protein is also known to be involved in cell-to-cell adhesion (deficiency in pectin biosynthesis pathway leading to loss of cell adhesion) together with *GAUT*8 (QUASIMODO1, see below) in *Arabidopsis* (Verger et al., 2016). McKown et al. (2014a) identified a similar gene, Potri.004G059000, encoding BRASSINOSTEROID-INSENSITIVE 2 for stomatal density based on the GWAS using a 34K SNP array data for 464 P. *trichocarpa* trees. BRASSINOSTEROID-INSENSITIVE 2 potentially affects the brassinosteroid signaling pathway and the regulation of stomatal development (Gudesblat et al., 2012; Khan et al., 2013). Brassinosteroids also play important role in plant growth, development and stress responses (Nolan et al., 2020).

The other top GWAS hit for stomatal density was within a gene model Potri.015G117300, a gene of unknown function (**Figures 4A, B**). The anchor gene has a breadth score of 2 and a total depth score of 2014 (highest breadth score among all single and multitrait anchor genes, **Figure 7**, **Supplemental Table S5**). The gene is comethylated and coexpressed with 1914 and 100 other genes, respectively, that encode proteins related to hormone signaling, stress response and transport functions. The anchor gene is coexpressed or comethylated with the genes related to diverse

functions including photosynthesis (22), cell wall (44), development (84), glycolysis (6), hormone metabolism (56), UDP glucosyl and glucoronyl transferases (6), redox (12), signaling (153, including receptor kinase, MAP kinase, light, sugar, G-proteins and calcium), stress (64, including abiotic and biotic, PR-proteins, respiratory burst), transport (123). For example, Potri.015G117300 is comethylated with two other genes, Potri.001G139700 and Potri.005G109300, both of which encode for plasma membrane proteins that belong to the subfamily of plant aquaporins (AQPs) that play important roles in plant physiological processes (Chaumont and Tyerman, 2014). Membrane proteins have been shown to have an involvement in regulating plant-water homeostasis and abiotic stress responses (Sade et al., 2009). Similarly, the anchor gene is comethylated with Potri.005G123200, that encodes for Sec14p-like phosphatidylinositol transfer family protein that has important role in lipid signaling pathways in plants (Huang et al., 2016). Another comethylated gene, Potri.005G134900, encodes a LATERAL ORGAN BOUNDARIES (*LOB*) domain protein. Proteins in this family play a role in organ development and stress responses (Grimplet et al., 2017), and also in regulating the expression of *VND7*, a master regulator of tracheary elements in plants (Ohashi-Ito et al., 2018). Overexpression of *VND7* represses *KNOX* gene activity, which in turn is responsible for meristem and organ development including the development of leaf primordia in plants (Hake et al., 2004) The anchor gene is coexpressed with germin-like (GLP) protein 1 (Potri.013G141900). *GLPs* are involved in a wide range of functions including growth and development, and biotic and abiotic stress responses (Dunwell et al., 2008). Another coexpressed gene, Potri.009G140900, encodes Small Auxin-Up RNA (SAUR)-like protein is involved in plant growth and development (cell division in particular) and the expression of the gene is regulated by auxin (Markakis et al., 2013). The coexpressed gene, Potri.010G112800 encodes auxin efflux carrier family protein, PIN-FORMED 3 (*PIN3*) which plays a role in seedling elongation and shade avoidance in Arabidopsis (Keuskamp et al., 2010). Another coexpressed gene, Potri.011G067500, encodes CBL-interacting protein kinase which is involved in drought tolerance mechanisms and the regulation of stomatal aperture (Wang et al., 2016).

Stomatal density in leaves facilitates gas exchange, and variation in the density of stomata can have a significant effect on overall plant fitness as it can change the dynamics of photosynthetic rate, respiration and growth (Franks and Farquhar, 2007; Franks and Beerling, 2009). Our GWAS and network analyses highlight the importance of genes involved in hormonal signaling and stress responses. Other recent studies in *Populus* have suggested a relationship between stomatal patterning and the trade-off between growth and disease susceptibility (and presence of defense compounds) (McKown et al., 2014a; McKown et al., 2019).

Multitrait GWAS Genes

We identified a total of seven genes based on the 11 SNPs that passed an FDR threshold of 0.05 in the multitrait GWAS, of which 6 genes were detected for the leaf morphology multitrait, (leaf area, leaf dry weight, leaf length and leaf width) and 1 gene was detected for wood anatomy multitrait (late wood vessel area and vessel count) (**Figure 5**)

The first anchor gene for the leaf morphology multitrait set was Potri.019G067300 (breadth score 4, total depth score 395, FDR=0.021, *p*-value=1.9 x 10^{-8}; **Figures 5A, B** and **8**, **Supplementary Table S5**). The gene encodes a protein related to galactosyltransferase-8 (*GAUT*8). The gene is comethylated with 392 other genes, positively coexpressed with one other gene, and related to one metabolite and linked to a py-MBMS (wood chemistry) GWAS. The anchor gene is comethylated with genes with various functions including cell wall (11 genes), growth and development related (11), glycolysis and Krebs cycle related (6), photosystem I and II related (4), oxidation reduction related (6), RNA regulation & binding (37), signaling (31, includes receptor kinase, light, photosynthesis, G-proteins and calcium signaling), stress (16, abiotic such as drought/salt and heat, biotic such as PR-proteins, respiratory burst), and transport (15) (**Figure 8**). The coexpressed gene Potri.006G049600 encodes concanavalin A-like lectin protein kinase family protein that is involved in amino acid phosphorylation. Lectin receptor-like kinases (*LecRLK)* are known to regulate abiotic and biotic stresses, and development in plants (Vaid et al., 2013; Yang et al., 2016; Labbe et al., 2019). A comethylated gene with a similar function as the anchor genes is Potri.006G137100, which encodes a pectin lyase-like superfamily protein that functions in pectin methylesterase (*PME*) activity. *PMEs* are involved in regulation of plant growth and development *via* structural modeling of the cell wall and also play an important role in heat stress tolerance (Wu et al., 2018). Another gene with a similar function, Potri.004G111000 (*GAUT*9) was detected for the same multitrait GWAS from the Corvallis common garden (Chhetri et al., 2019). *GAUT*8 and *GAUT*9 belonged to the same clade B1 in a phylogenetic tree based on the *GAUT* protein family of *Arabidopsis thaliana* and *P. trichocarpa* (Biswal et al., 2018). *GAUT*9 affects leaf size in *Populus deltoides* (Chhetri et al., 2019) whereas *GAUT*8 (also known as QUASIMODO1) has high expression in stems in *Arabidopsis* (Caffall et al., 2009) and is thought to be involved in cell wall pectic homogalacturonan (HG) and xylan biosynthesis (Orfila et al., 2005; Verger et al., 2016). Genes encoding glycosyltransferases, a large family of enzymes mainly involved in biosynthesis of polysaccharides and glycoproteins in the plant cell wall (Hansen et al., 2012), were also detected in single trait GWAS for stomatal density in this study (**Table 3**).

Another gene associated with the leaf morphology multi-trait set was Potri.010G031900 (FDR = 0.043, *P*-value = 5.23 x 10^{-8}) is coexpressed and comethylated with 12 and 67 other genes, respectively (breadth score 4, total depth score 81, **Figures 5A, B**, **Supplementary Table S5**). The gene encodes E1 C-terminal protein, which is a RUB-activating enzyme that indirectly triggers auxin response by interacting with the cullin *AtCUL1* in Arabidopsis (Del Pozo and Estelle, 1999). The anchor gene is coexpressed or comethylated with genes with diverse functions including signaling (9, including G-proteins, MAP kinase, receptor kinase and calcium), RNA regulation (9), transport (3), stress (5, abiotic- drought/salt and heat, biotic—

PR-proteins), hormone metabolism (4, auxin 2, ABA 2), and lipid metabolism (2) (**Supplementary Figure S13**).

A few of the top coexpressed genes in the network are related to the RING/U-box superfamily protein (e.g., Potri.018G064400). U-box (PUB) proteins have diverse functions in plants, including hormonal, defense and abiotic stress responses (Yee and Goring, 2009). RNA binding protein (Potri.019G017900) and mitogen-activated protein kinase kinase 3 (*MKK3*, Potri.001G345500)—*MKK3* belong to a bigger group of mitogen-activated protein kinases (*MAPKs*) that play a role in activating downstream hormonal signaling including auxin, jasmonic acid, salicylic acid, brassinosteroid, ethylene and ABA signaling targets in plants (Jagodzik et al., 2018). The gene is also comethylated with genes with similar roles such as Potri.009G015800 encoding IQ-domain 5 (IQD5) that functions in calmodulin binding. Calmodulin and calmodulin-like proteins are involved in responding to abiotic stresses in plants (Zeng et al., 2015; Badmi et al., 2018). There are other comethylated genes in the network with functions in stress response, and growth and development. Potri.009G001200 encodes RING-H2 group F2A (*RHF2A*) protein. RING-H2 belong to a group of E3 ubiquitin protein complexes and are involved in defense response, regulation of growth and development, and apoptosis (Guzmán, 2012). Potri.008G010800 encodes ABA-responsive element binding protein 3. Potri. 002G152100 encodes phospholipase D beta 1 (*PLDBETA1*). PLD functions in regulating plant growth and development, hormonal signaling and stress responses (Li et al., 2007). The anchor gene is also associated with a metabolite, caffeoyl-quercetin glycoside and one py-MBMS GWAS, m/z 414). Quercetin glycosides are known to play a role in osmotic adjustment during drought stress (Tschaplinski et al., 2019). Caffeoyl-quercetin glycoside was found to be associated with Potri.001G411800 (EF-hand Calcium-Binding Domain protein) in a multitrait GWAS for leaf area, stomatal density and carbon isotope composition in a previous study (Chhetri et al., 2019). Ethylene and ABA play important roles in stress responses with prominent effects in plant tissues (Cramer et al., 2011). Reduction in Leaf water potential and thus leaf wet weight is possible due to abiotic (e.g. drought) and biotic stress (e.g. pathogen) (Pandey et al., 2017). Drost et al. (2015) provided evidence for the effect of *PtARF1* on leaf morphology in a hybrid population of *P. deltoides* and *P. trichocarpa* that is mediated by auxin signaling such that the PIN polarization in leaf cells possibly plays a role in differential expression of *PtARF1*, ultimately causing the expansion of leaf in the length direction with the decrease in leaf width in *P. trichocarpa* and vice versa in *P. deltoides* (Kalluri et al., 2007). Various signaling genes including G-protein signaling (Lease et al., 2001), and hormonal signaling genes (McSteen and Zhao, 2008) are known to have an effect on plant growth and development. Calcium signaling regulates the function of stress related genes ultimately affecting the plant phenotype (Tuteja and Mahajan, 2007).

Multitrait GWAS Genes: Wood Anatomy

The gene model Potri.001G058100 was associated with the wood anatomy multitrait (**Figure 5**). This gene has no known function but is coexpressed and comethylated with 9 and 972 other genes (breadth score 2, total depth score 981, 4th highest depth score among all single and multitrait GWAS genes), respectively (**Supplementary Tables S5** and **S9**). The gene is comethylated with 3 RNA binding genes (Potri.009G152900, Potri.013G094400, Potri.017G063700) that are involved in biogenesis of small RNA in Arabidopsis (Eamens et al., 2009; Eamens et al., 2012). Small RNAs are known to regulate the expression of genes related to growth and development, and wood formation in plants (Tang et al., 2016; Chen C. et al., 2018). The anchor gene is coexpressed or comethylated with diverse genes that have functions in transport (61 genes), biotic and abiotic stress tolerance (35), signaling mechanism (84, hormonal, receptor kinase, light and calcium, sugar and G-protein signaling), cell wall related (20), developmental mechanisms (39), hormone metabolism (25) and RNA regulation (106) (**Figures 4C, D**; **Supplementary Table S9**). Notable top 10 coexpressed genes include Potri.001G043100 that encodes Chaperone DnaJ-domain superfamily protein which is a heat shock protein with a function in stress tolerance, Potri.002G169400 (encodes RCAR1, regulatory components of ABA receptor), Potri.009G033300 (encodes SCARECROW-like 14, *SCL14*) and Potri.017G134800 (encoding zinc finger protein 7, *ZFP7*). *SCL14* is a member of GRAS family of transcription factors that interacts with the TGA2 transcription factors and affects the transcription of stress-responsive genes (Fode et al., 2008). *ZFP7* is a member of *C2H2* zinc finger protein related to *ZFP3*. *ZFP3* and related *ZFP* proteins including *ZFP7* acts as a negative regulator of ABA signaling during seedling development, affect vegetative growth and regulate light signaling in Arabidopsis (Joseph et al., 2014). ABA receptors are involved in transmitting ABA signals that allow plants to respond to drought and other abiotic stresses (Fan et al., 2016; Duarte et al., 2019). The anchor gene is comethylated with Leucine-rich repeat protein kinase family protein (Potri.019G001800) that functions in ATP binding activity and regulating hormonal signaling and stress response in plants (Máthé et al., 2019). Taken together, this suggests a role of this gene in growth and development including wood formation by maintaining cellular homeostasis, signaling and stress response.

Comparison With Previous Studies

We compared our GWAS results with previous GWAS studies for similar traits in *P. trichocarpa* (Wegrzyn et al., 2010; Porth et al., 2013a; Porth et al., 2013b; Evans et al., 2014; McKown et al., 2014b; Chhetri et al., 2019) and other *Populus* species including *P. deltoides* (Fahrenkrog et al., 2017), *P. euphratica* (Ma et al., 2013; Jia et al., 2020) and *P. tomentosa* (Quan et al., 2016; Chen B. et al., 2018; Quan et al., 2019; Lu et al., 2020). We found very minimal overlap in the genes and/or gene families detected. Ma et al. (2013) identified multiple gene families that control adaptation in saline environments. Similar to their results, heat shock protein families and some ABA signaling proteins were found in the co-expression networks in our study. Jia et al. (2020) identified 82 genes related to seed salinity tolerance in *P. euphratica*, including a zinc finger protein similar to one that we detected for leaf abaxial density that regulates gas exchange in plants. Lu et al. (2020) identified 23 genes involved in three growth and six wood property traits and Quan et al. (2019) identified 203 lignin biosynthetic genes, 81 TF genes, 36 microRNAs and 71 long non-coding RNA genes for 10 growth and

wood property traits in the same *P. tomentosa* population. However, these studies have no apparent overlap in gene functions for similar morphology or wood property traits in our study.

Comparing our findings with previous *P. trichocarpa* studies we identified some non-overlapping gene models with similar functions in shared gene families. For example, we identified a gene Potri.019G067300 in a multi-trait GWAS for leaf morphology that encodes the GAUT8 protein. In a previous multi-trait GWAS study of leaf morphology at a different site, we identified GAUT9 (Potri.004G111000), which affects leaf size and is involved in pectin and xylan biosynthesis (Chhetri et al., 2019). Likewise, we identified a gene (Potri.005G247700) encoding "SKU5 similar 5" (sks5) that functions in oxidoreductase activity and copper ion binding in this study. Porth et al. (2013a) also identified a gene with similar function for soluble lignin (Potri.001G000500) encoding "SKU5 similar 12" (sks12) based on 34K SNP genotyping array data. McKown et al. (2019) identified several genes related to plant growth, physiological processes, biotic and abiotic stresses, and defense and immunity for abaxial and adaxial stomata related traits (stomatal density and pore length). We did not find any of the genes they identified in their study, despite analyzing similar traits here. The discrepancies in the gene models identified across the studies may be due to a) differential environmental effects on the traits due to plantation sites, b) traits measured in different seasons of the year and different years were subjected to different environmental conditions, c) variability in the microenvironment within the sites as reflected by the different broad-sense heritability values, d) differences in the developmental stages of the traits across the studies, e) difference in the genotypes and the sequence data types (e.g. SNP array vs whole genome sequence data) and f) the lack of power in this and previous studies. Furthermore, we were not able to replicate any significant SNP hits or the gene models for the same traits from another *P. trichocarpa* common garden in Corvallis, Oregon in our study, although several single traits such as stomatal density, petiole diameter and tree height and the multitrait set containing leaf area, leaf dry weight, leaf length and leaf width were common between the two common gardens, and measurements were performed using the same methods. This might very well reflect the differential influence of the environment on genotypes in the two populations. Nevertheless, we believe that our comprehensive GWAS study with genome-wide sequencing data highlights the underlying genes controlling adaptive traits in *P. trichocarpa* and complements the findings from other similar studies. Moreover, we have reported here potential genes underlying complex wood anatomy traits in *Populus* for the first time.

CONCLUSION

We presented here the first comprehensive GWAS for wood anatomical traits in *Populus* that provides insights into the type of genes controlling structural and functional properties important for plant development, function and stress tolerance. We complemented this with GWAS for important morphological traits that have functional relationships with wood anatomical traits. As shown in a previous *P. trichocarpa* GWAS study in another common garden (Chhetri et al., 2019), we identified additional genes with multitrait GWAS in this study as well. Furthermore, the multitrait sets formed based on the genetic correlations and functional relationships of the traits within and among the wood anatomical and morphological traits provided insight into pleiotropic genes controlling these traits. Some of the genes we identified in this study had no known functions. However, the LOE network analyses we performed here provided additional information not only for suggesting the role of the unknown genes, but also added an extra layer of support for the genes with known functions. Genes identified here, especially the genes controlling wood anatomical traits, can be good targets for biotechnological approaches toward optimizing wood traits for biofuel production. However, as is the case with most GWAS studies, despite the use of large scale whole genome resequencing data and much broader sampling compared to most previous studies in trees, a very small percent of the variation in the traits were explained by the significant SNPs. It is possible that we may simply be underpowered to detect SNPs with small effect sizes, given that we only detected a handful of genes or none for most of the complex traits that might potentially be controlled by hundreds or thousands of genes (Boyle et al., 2017). A more robust approach to GWAS is required, with large and homogeneous sampling from across the range of distribution. Alternatively, more populations with controlled structure can be used to take full advantage of the range of variation present while explicitly controlling for underlying structure, as is the case in nested association mapping (Yu et al., 2008). Moreover, none of the genetic variants detected in the GWAS from another common garden were validated in this study. This may be because of the differential effects of the environment between the common gardens, which is somewhat reflected by very low correlation of the traits between the two common gardens. Furthermore, additional variation due to the difference in the timing (different years) of data collection might have influenced the GWAS study. Nevertheless, by integrating data from multiple sources we were able to apply a LOE approach that provided further corroboration for relatively weak associations, while providing insights into gene function.

DATA AVAILABILITY STATEMENT

Populus trichocarpa genome sequence, annotation, and Gene Atlas expression and methylation data sets are available on Phytozome (http://phytozome.jgi.doe.gov). Populus trichocarpa variant data (DOI 10.13139/OLCF/1411410) is available from https://doi.ccs.ornl.gov/ui/doi/55.

AUTHOR CONTRIBUTIONS

HC, AF, DM-S, AW, DK, PJ, and DJ performed statistical and computational analyses. AH-W and TT performed metabolite analyses and analyzed the data. DJ, GT, and SD oversaw the

study. HC, AF, and SD drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Center for Bioenergy Innovation (CBI) and the Bioenergy Science Center. CBI is supported by the Office of Biological and Environmental Research in the DOE Office of Science. This manuscript has been coauthored 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). 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. This research described herein was supported by an award of computer time provided by the INCITE program and used resources of the Oak Ridge Leadership Computing Facility (OLCF) at the Oak Ridge National Laboratory. This work was authored in part by Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding was provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Bioenergy Technologies Office.

ACKNOWLEDGMENTS

We thank the multitude of researchers from the Bioenergy Science Center and the DOE Joint Genome Institute who provided invaluable logistical support for this work. In particular, we would like to thank Kat Haiby, Brian Stanton, Rich Shuren, Carlos Gantz, and Austin Himes of Greenwood Resources for their work in establishing and maintaining the plantation, for facilitating our work at the site, and for the many insights that they have provided about *Populus* biology and silviculture. We would also like to thank Crissa Doeppke and Robert Sykes for their help with py-MBMS analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.545748/full#supplementary-material

REFERENCES

Allwright, M. R., Payne, A., Emiliani, G., Milner, S., Viger, M., Rouse, F., et al. (2016). Biomass traits and candidate genes for bioenergy revealed through association genetics in coppiced European *Populus nigra* (L.). *Biotechnol. Biofuels* 9, 1–22. doi: 10.1186/s13068-016-0603-1

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.* 871:1669. doi: 10.3389/fpls.2018.01669

Bdeir, R., Muchero, W., Yordanov, Y., Tuskan, G. A., Busov, V., and Gailing, O. (2019). Genome-wide association studies of bark texture in *Populus trichocarpa*. *Tree Genet. Genomes* 15, 14. doi: 10.1007/s11295-019-1320-2

Biswal, A. K., Atmodjo, M. A., Pattathil, S., Amos, R. A., Yang, X., Winkeler, K., et al. (2018). Working towards recalcitrance mechanisms: Increased xylan and homogalacturonan production by overexpression of *GAlactUronosylTransferase12 (GAUT12)* causes increased recalcitrance and decreased growth in *Populus*. *Biotechnol. Biofuels* 11, 1–26. doi: 10.1186/s13068-017-1002-y

Boyle, E. A., Li, Y.II, 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

Caffall, K. H., Pattathil, S., Phillips, S. E., Hahn, M. G., and Mohnen, D. (2009). *Arabidopsis thaliana* T-DNA mutants implicate GAUT genes in the biosynthesis of pectin and xylan in cell walls and seed testa. *Mol. Plant* 2, 1000–1014. doi: 10.1093/mp/ssp062

Carlquist, S. (2012). How wood evolves: a new synthesis. *Botany* 90, 901–940. doi: 10.1139/b2012-048

Chaumont, F., and Tyerman, S. D. (2014). Aquaporins: Highly regulated channels controlling plant water relations. *Plant Physiol.* 164, 1600–1618. doi: 10.1104/pp.113.233791

Chen, B., Chen, J., Du, Q., Zhou, D., Wang, L., Xie, J., et al. (2018). Genetic variants in microRNA biogenesis genes as novel indicators for secondary growth in *Populus*. *New Phytol.* 219, 1263–1282. doi: 10.1111/nph.15262

Chen, C., Zeng, Z., Liu, Z., and Xia, R. (2018). Small RNAs, emerging regulators critical for the development of horticultural traits. *Hortic. Res.* 5, 6–8. doi: 10.1038/s41438-018-0072-8

Chhetri, H. B., Macaya-sanz, D., Kainer, D., Biswal, A. K., Evans, L. M., Chen, J., et al. (2019). Multitrait genome-wide association analysis of *Populus trichocarpa* identifies key polymorphisms controlling morphological and physiological traits. *New Phytol.* 223, 293–309. doi: 10.1111/nph.15777

Cramer, G. R., Urano, K., Delrot, S., Pezzotti, M., and Shinozaki, K. (2011). Effects of abiotic stress on plants: A systems biology perspective. *BMC Plant Biol.* 11, 163. doi: 10.1186/1471-2229-11-163

Del Pozo, J. C., and Estelle, M. (1999). The Arabidopsis cullin AtCUL1 is modified by the ubiquitin-related protein RUB1. *Proc. Natl. Acad. Sci. U. S. A.* 96, 15342–15347. doi: 10.1073/pnas.96.26.15342

Drost, D. R., Puranik, S., Novaes, E., Novaes, C. R. D. B., Dervinis, C., Gailing, O., et al. (2015). Genetical genomics of *Populus* leaf shape variation. *BMC Plant Biol.* 15, 1–10. doi: 10.1186/s12870-015-0557-7

Du, Q., Pan, W., Xu, B., Li, B., and Zhang, D. (2013). Polymorphic simple sequence repeat (SSR) loci within cellulose synthase (PtoCesA) genes are associated with growth and wood properties in *Populus tomentosa*. *New Phytol.* 197, 763–776. doi: 10.1111/nph.12072

Du, Q., Gong, C., Wang, Q., Zhou, D., Yang, H., Pan, W., et al. (2016). 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

Du, Q., Lu, W., Quan, M., Xiao, L., Song, F., Li, P., et al. (2018). Genome-Wide Association Studies to Improve Wood Properties: Challenges and Prospects. *Front. Plant Sci.* 9, 1912. doi: 10.3389/fpls.2018.01912

Du, Q., Yang, X., Xie, J., Quan, M., Xiao, L., Lu, W., et al. (2019). Time-specific and pleiotropic quantitative trait loci coordinately modulate stem growth in *Populus*. *Plant Biotechnol. J.* 17, 608–624. doi: 10.1111/pbi.13002

Duarte, K. E., de Souza, W. R., Santiago, T. R., Sampaio, B. L., Ribeiro, A. P., Cotta, M. G., et al. (2019). Identification and characterization of core abscisic acid (ABA) signaling components and their gene expression profile in response to

abiotic stresses in *Setaria viridis*. *Sci. Rep.* 9, 1–16. doi: 10.1038/s41598-019-40623-5

Dunwell, J. M., Gibbings, J. G., Mahmood, T., and Saqlan Naqvi, S. M. (2008). Germin and germin-like proteins: Evolution, structure, and function. *CRC. Crit. Rev. Plant Sci.* 27, 342–375. doi: 10.1080/07352680802333938

Eamens, A. L., Smith, N. A., Curtin, S. J., Wang, M. B., and Waterhouse, P. M. (2009). The *Arabidopsis thaliana* double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes. *Rna* 15, 2219–2235. doi: 10.1261/rna.1646909

Eamens, A. L., Kim, K. W., Curtin, S. J., and Waterhouse, P. M. (2012). DRB2 is required for microRNA biogenesis in *Arabidopsis thaliana*. *PloS One* 7, e35933. doi: 10.1371/journal.pone.0035933

Escamez, S., Latha Gandla, M., Derba-Maceluch, M., Lundqvist, S. O., Mellerowicz, E. J., Jönsson, L. J., et al. (2017). A collection of genetically engineered *Populus* trees reveals wood biomass traits that predict glucose yield from enzymatic hydrolysis. *Sci. Rep.* 7, 1–11. doi: 10.1038/s41598-017-16013-0

Evans, L. M., Slavov, G. T., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., et al. (2014). Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nat. Genet.* 46, 1089–1096. doi: 10.1038/ng.3075

Fahrenkrog, A. M., Neves, L. G., Resende, M. F. R., Dervinis, C., Davenport, R., Barbazuk, W. B., et al. (2017). Population genomics of the eastern cottonwood *(Populus deltoides)*. *Ecol. Evol.* 7, 9426–9440. doi: 10.1002/ece3.3466

Fan, W., Zhao, M., Li, S., Bai, X., Li, J., Meng, H., et al. (2016). Contrasting transcriptional responses of PYR1/PYL/RCAR ABA receptors to ABA or dehydration stress between maize seedling leaves and roots. *BMC Plant Biol.* 16, 1–14. doi: 10.1186/s12870-016-0764-x

Fode, B., Siemsen, T., Thurow, C., Weigel, R., and Gatz, C. (2008). The arabidopsis GRAS protein SCL14 interacts with class II TGA transcription factors and is essential for the activation of stress-inducible promoters. *Plant Cell* 20, 3122–3135. doi: 10.1105/tpc.108.058974

Franks, P. J., and Beerling, D. J. (2009). Maximum leaf conductance driven by CO2 effects on stomatal size and density over geologic time. *Proc. Natl. Acad. Sci. U. S. A.* 106, 10343–10347. doi: 10.1073/pnas.0904209106

Franks, P. J., and Farquhar, G. D. (2007). The mechanical diversity of stomata and its significance in gas-exchange control. *Plant Physiol.* 143, 78–87. doi: 10.1104/pp.106.089367

Furches, A., Kainer, D., Weighill, D., Large, A., Jones, P., Walker, A. M., et al. (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

Gandla, M. L., Martín, C., and Jönsson, L. J. (2018). Analytical Enzymatic Saccharification of Lignocellulosic Biomass for Conversion to Biofuels and Bio-Based Chemicals. *Energies* 11, 2936. doi: 10.3390/en11112936

Genovese, C. R., Roeder, K., and Wasserman, L. (2006). False discovery control with p-value weighting. *Biometrika*. 93, 509–524. doi: 10.1093/biomet/93.3.509

González-Martínez, S. C., Ersoz, E., Brown, G. R., Wheeler, N. C., and Neale, D. B. (2006). DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. *Genetics* 172, 1915–1926. doi: 10.1534/genetics.105.047126

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* 40, 1178–1186. doi: 10.1093/nar/gkr944

Grimplet, J., Pimentel, D., Agudelo-Romero, P., Martinez-Zapater, J. M., and Fortes, A. M. (2017). The LATERAL ORGAN BOUNDARIES Domain gene family in grapevine: Genome-wide characterization and expression analyses during developmental processes and stress responses. *Sci. Rep.* 7, 1–18. doi: 10.1038/s41598-017-16240-5

Groover, A. T., Nieminen, K., Helariutta, Y., and Mansfield, S. D. (2010). "Wood Formation in Populus," in *Genetics and Genomics of Populus*. Eds. S. Jansson, R. Bhalerao and A. T. Groover (New York: Springer), 201–224.

Gudesblat, G. E., Schneider-Pizoń, J., Betti, C., Mayerhofer, J., Vanhoutte, I., Van Dongen, W., et al. (2012). SPEECHLESS integrates brassinosteroid and stomata signalling pathways. *Nat. Cell Biol.* 14, 548–554. doi: 10.1038/ncb2471

Guzmán, P. (2012). The prolific ATL family of RING-H2 ubiquitin ligases. *Plant Signal. Behav.* 7, 1014–1021. doi: 10.4161/psb.20851

Hake, S., Smith, H. M. S., Holtan, H., Magnani, E., Mele, G., and Ramirez, J. (2004). The Role of Knox Genes in Plant Development. *Annu. Rev. Cell Dev. Biol.* 20, 125–151. doi: 10.1146/annurev.cellbio.20.031803.093824

Hansen, S. F., Harholt, J., Oikawa, A., and Scheller, H. V. (2012). Plant Glycosyltransferases Beyond CAZy: A Perspective on DUF Families. *Front. Plant Sci.* 3, 59. doi: 10.3389/fpls.2012.00059

Hietz, P., Rosner, S., Hietz-Seifert, U., and Wright, S. J. (2017). Wood traits related to size and life history of trees in a Panamanian rainforest. *New Phytol.* 213, 170–180. doi: 10.1111/nph.14123

Huang, J., Ghosh, R., and Bankaitis, V. A. (2016). Sec14-like phosphatidylinositol transfer proteins and the biological landscape of phosphoinositide signaling in plants. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* 1861, 1352–1364. doi: 10.1016/j.bbalip.2016.03.027

Huizinga, D. H., Denton, R., Koehler, K. G., Tomasello, A., Wood, L., Sen, S. E., et al. (2010). Farnesylcysteine lyase is involved in negative regulation of abscisic acid signaling in Arabidopsis. *Mol. Plant* 3, 143–155. doi: 10.1093/mp/ssp091

Ingvarsson, P. K., Hvidsten, T. R., and Street, N. R. (2016). Towards integration of population and comparative genomics in forest trees. *New Phytol.* 212, 338–344. doi: 10.1111/nph.14153

Jagodzik, P., Tajdel-Zielinska, M., Ciesla, A., Marczak, M., and Ludwikow, A. (2018). Mitogen-activated protein kinase cascades in plant hormone signaling. *Front. Plant Sci.* 9, 1387. doi: 10.3389/fpls.2018.01387

Jia, H., Liu, G., Li, J., Zhang, J., Sun, P., Zhao, S., et al. (2020). Genome resequencing reveals demographic history and genetic architecture of seed salinity tolerance in *Populus euphratica*. *J. Exp. Bot.* 71, 4308–4320. doi: 10.1093/jxb/eraa172

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, 1–12. doi: 10.1186/s13068-017-0734-z

Joseph, M. P., Papdi, C., Kozma-Bognár, L., Nagy, I., López-Carbonell, M., Rigó, G., et al. (2014). The Arabidopsis ZINC FINGER PROTEIN3 interferes with abscisic acid and light signaling in seed germination and plant development. *Plant Physiol.* 165, 1203–1220. doi: 10.1104/pp.113.234294

Kalluri, U. C., Difazio, S. P., Brunner, A. M., and Tuskan, G. A. (2007). Genome-wide analysis of Aux/IAA and ARF gene families in *Populus trichocarpa*. *BMC Plant Biol.* 7, 1–14. doi: 10.1186/1471-2229-7-59

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S. Y., 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

Kawasaki, T., Koita, H., Nakatsubo, T., Hasegawa, K., Wakabayashi, K., Takahashi, H., et al. (2006). Cinnamoyl-CoA reductase, a key in lignin biosynthesis, is an effector of small GTPase Rac in defense signaling in rice. *Proc. Natl. Acad. Sci. U. S. A.* 103, 230–235. doi: 10.1073/pnas.0509875103

Keuskamp, D. H., Pollmann, S., Voesenek, L. A. C. J., Peeters, A. J. M., and Pierik, R. (2010). Auxin transport through PIN-FORMED 3 (PIN3) controls shade avoidance and fitness during competition. *Proc. Natl. Acad. Sci. U. S. A.* 107, 22740–22744. doi: 10.1073/pnas.1013457108

Khan, M., Rozhon, W., Bigeard, J., Pflieger, D., Husar, S., Pitzschke, A., et al. (2013). Brassinosteroid-regulated GSK3/Shaggy-like kinases phosphorylate mitogen-activated protein (MAP) kinase kinases, which control stomata development in *Arabidopsis thaliana*. *J. Biol. Chem.* 288, 7519–7527. doi: 10.1074/jbc.M112.384453

Labbe, J., Muchero, W., Czarnecki, O., Wang, J., Wang, X., Bryan, A. C., et al. (2019). Mediation of plant-mycorrhizal interaction by a lectin receptor-like kinase. *Nat. Plants* 5, 676–680. doi: 10.1038/s41477-019-0469-x

Lachenbruch, B., Moore, J. R., and Evans, R. (2011). "Radial Variation in Wood Structure and Function in Woody Plants, and Hypotheses for its Occurence," in *Size and Age Related Changes in Tree Structure and Function*. Eds. F. C. Meinzer, B. Lachenbruch and T. E. Dawson (Dordrecht: Springer), 121–164.

Lacombe, E., Hawkins, S., Van Doorsselaere, J., Piquemal, J., Goffner, D., Poeydomenge, O., et al. (1997). Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: Cloning, expression and phylogenetic relationships. *Plant J.* 11, 429–441. doi: 10.1046/j.1365-313X.1997.11030429.x

Lease, K. A., Wen, J., Li, J., Doke, J. T., Liscum, E., and Walker, J. C. (2001). A mutant Arabidopsis heterotrimeric G-protein β subunit affects leaf, flower, and fruit development. *Plant Cell* 13, 2631–2641. doi: 10.1105/tpc.13.12.2631

Lee, J., Kim, D. H., and Hwang, I. (2014). Specific targeting of proteins to outer envelope membranes of endosymbiotic organelles, chloroplasts, and mitochondria. *Front. Plant Sci.* 5, 173. doi: 10.3389/fpls.2014.00173

Li, G., Lin, F., and Xue, H. W. (2007). Genome-wide analysis of the phospholipase D family in Oryza sativa and functional characterization of PLDB1 in seed germination. *Cell Res.* 17, 881–894. doi: 10.1038/cr.2007.77

Li, Y., Liu, P., Takano, T., and Liu, S. (2016). A chloroplast-localized rubredoxin family protein gene from Puccinellia tenuiflora (PutRUB) increases NaCl and NaHCO3 tolerance by decreasing H2O2 accumulation. *Int. J. Mol. Sci.* 17, 804. doi: 10.3390/ijms17060804

Lira-Navarrete, E., Valero-González, J., Villanueva, R., Martínez-Júlvez, M., Tejero, T., Merino, P., et al. (2011). Structural insights into the mechanism of protein O-fucosylation. *PloS One* 6, e25365. doi: 10.1371/journal.pone.0025365

Lu, W., Xiao, L., Quan, M., Wang, Q., El-Kassaby, Y. A., Du, Q., et al. (2020). Linkage-linkage disequilibrium dissection of the epigenetic quantitative trait loci (epiQTLs) underlying growth and wood properties in *Populus*. *New Phytol.* 225, 1218–1233. doi: 10.1111/nph.16220

Ma, T., Wang, J., Zhou, G., Yue, Z., Hu, Q., Chen, Y., et al. (2013). Genomic insights into salt adaptation in a desert poplar. *Nat. Commun.* 4, 2797. doi: 10.1038/ncomms3797

Major, I. T., and Peter Constabel, C. (2008). Functional analysis of the kunitz trypsin inhibitor family in poplar reveals biochemical diversity and multiplicity in defense against herbivores. *Plant Physiol.* 146, 888–903. doi: 10.1104/pp.107.106229

Markakis, M. N., Boron, A. K., Van Loock, B., Saini, K., Cirera, S., Verbelen, J. P., et al. (2013). Characterization of a small auxin-up RNA (SAUR)-like gene involved in *Arabidopsis thaliana* development. *PloS One* 8, 1–13. doi: 10.1371/journal.pone.0082596

Máthé, C., Garda, T., Freytag, C., and Hamvas, M. M. (2019). The role of serine-threonine protein phosphatase pp2a in plant oxidative stress signaling—facts and hypotheses. *Int. J. Mol. Sci.* 20, 3028. doi: 10.3390/ijms20123028

McKown, A. D., Guy, R. D., Quamme, L., Klápště, J., La Mantia, J., Constabel, C. P., et al. (2014a). Association genetics, geography and ecophysiology link stomatal patterning in *Populus trichocarpa* with carbon gain and disease resistance trade-offs. *Mol. Ecol.* 23, 5771–5790. doi: 10.1111/mec.12969

McKown, A. D., Klápště, J., Guy, R. D., Geraldes, A., Porth, I., Hannemann, J., et al. (2014b). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytol.* 203, 535–553. doi: 10.1111/nph.12815

Mckown, A. D., Guy, R. D., Klapste, J., Geraldes, A., Friedmann, M., Cronk, Q. C. B., et al. (2014c). Geographical and environmental gradients shape phenotypic trait variation and genetic structure in *Populus trichocarpa*. *New Phytol.* 201, 1263–1276. doi: 10.1111/nph.12601

McKown, A. D., Klápště, J., Guy, R. D., Corea, O. R. A., Fritsche, S., Ehlting, J., et al. (2019). A role for SPEECHLESS in the integration of leaf stomatal patterning with the growth vs disease trade-off in poplar. *New Phytol.* 223, 1888–1903. doi: 10.1111/nph.15911

McSteen, P., and Zhao, Y. (2008). Plant Hormones and Signaling: Common Themes and New Developments. *Dev. Cell* 14, 467–473. doi: 10.1016/j.devcel.2008.03.013

Muchero, W., Guo, J., DiFazio, S. P., Chen, J., 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

Neale, D. B., and Kremer, A. (2011). Forest tree genomics: Growing resources and applications. *Nat. Rev. Genet.* 12, 111–122. doi: 10.1038/nrg2931

Neale, D. B., and Savolainen, O. (2004). Association genetics of complex traits in conifers. *New Phytol.* 9, 325–330. doi: 10.1111/j.1469-8137.2010.03593.x

Nolan, T. M., Vukašinović, N., Liu, D., Russinova, E., and Yin, Y. (2020). Brassinosteroids: Multidimensional Regulators of Plant Growth, Development, and Stress Responses. *Plant Cell* 32, 295–318. doi: 10.1105/tpc.19.00335

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

Orfila, C., Sørensen, S. O., Harholt, J., Geshi, N., Crombie, H., Truong, H. N., et al. (2005). QUASIMODO1 is expressed in vascular tissue of *Arabidopsis thaliana* inflorescence stems, and affects homogalacturonan and xylan biosynthesis. *Planta* 222, 613–622. doi: 10.1007/s00425-005-0008-z

Pandey, P., Irulappan, V., Bagavathiannan, M. V., and Senthil-Kumar, M. (2017). Impact of combined abiotic and biotic stresses on plant growth and avenues for crop improvement by exploiting physio-morphological traits. *Front. Plant Sci.* 8, 537. doi: 10.3389/fpls.2017.00537

Pendergrass, S. A., Brown-Gentry, K., Dudek, S., Frase, A., Torstenson, E. S., Goodloe, R., et al. (2013). Phenome-Wide Association Study (PheWAS) for Detection of Pleiotropy within the Population Architecture using Genomics and Epidemiology (PAGE) Network. *PloS Genet.* 9, e1003087. doi: 10.1371/journal.pgen.1003087

Porter, H. F., and O'Reilly, P. F. (2017). Multivariate simulation framework reveals performance of multi-trait GWAS methods. *Sci. Rep.* 7, 1–12. doi: 10.1038/srep38837

Porth, I., Klápště, J., Skyba, O., Friedmann, M. C., Hannemann, J., Ehlting, J., et al. (2013a). 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

Porth, I., Ranjan, P., Klapšte, J., Guy, R. D., Tuskan, G. A., Ehlting, J., et al. (2013b). 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

Porth, I., Klápště, J., McKown, A. D., La Mantia, J., Guy, R. D., Ingvarsson, P. K., et al. (2015). Evolutionary Quantitative Genomics of *Populus trichocarpa*. *PloS One* 10, e0142864. doi: 10.1371/journal.pone.0142864

Preston, K. A., Cornwell, W. K., and DeNoyer, J. L. (2006). Wood density and vessel traits as distinct correlates of ecological strategy in 51 California coast range angiosperms. *New Phytol.* 170, 807–818. doi: 10.1111/j.1469-8137.2006.01712.x

Quan, M., Wang, Q., Phangthavong, S., Yang, X., Song, Y., Du, Q., et al. (2016). Association studies in *Populus tomentosa* reveal the genetic interactions of PTO-MIR156c and its targets in wood formation. *Front. Plant Sci.* 7, 1159. doi: 10.3389/fpls.2016.01159

Quan, M., Du, Q., Xiao, L., Lu, W., Wang, L., Xie, J., et al. (2019). Genetic architecture underlying the lignin biosynthesis pathway involves noncoding RNAs and transcription factors for growth and wood properties in *Populus*. *Plant Biotechnol. J.* 17, 302–315. doi: 10.1111/pbi.12978

Rymaszewski, W., Vile, D., Bediee, A., Dauzat, M., Luchaire, N., Kamrowska, D., et al. (2017). Stress-related gene expression reflects morphophysiological responses to water deficit. *Plant Physiol.* 174, 1913–1930. doi: 10.1104/pp.17.00318

Sade, N., Vinocur, B. J., Diber, A., Shatil, A., Ronen, G., Nissan, H., et al. (2009). Improving plant stress tolerance and yield production: Is the tonoplast aquaporin SlTIP2;2 a key to isohydric to anisohydric conversion? *New Phytol.* 181, 651–661. doi: 10.1111/j.1469-8137.2008.02689.x

Schindelin, J., Rueden, C. T., Hiner, M. C., and Eliceiri, K. W. (2015). The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev.* 82, 518–529. doi: 10.1002/mrd.22489

Shim, H., Chasman, D.II, Smith, J. D., Mora, S., Ridker, P. M., Nickerson, D. A., et al. (2015). A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. *PloS One* 10, 1–20. doi: 10.1371/journal.pone.0120758

Slavov, G. T., Difazio, S. P., Martin, J., Schackwitz, W., Muchero, W., Rodgers-Melnick, E., et al. (2012). Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytol.* 196, 713–725. doi: 10.1111/j.1469-8137.2012.04258.x

Sperry, J. S. (2000). Hydraulic constraints on plant gas exchange. *Agric. For. Meteorol.* 104, 13–23. doi: 10.1016/S0168-1923(00)00144-1

Sperry, J. S. (2003). Evolution of Water Transport and Xylem Structure. *Int. J. Plant Sci.* 164, S115–S127. doi: 10.1086/368398

Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9440–9445. doi: 10.1073/pnas.1530509100

Street, N. R., and Ingvarsson, P. K. (2011). Association genetics of complex traits in plants. *New Phytol.* 189, 909–922. doi: 10.1111/j.1469-8137.2010.03593.x

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.* 108, 6300–6305. doi: 10.1073/pnas.1009252108

Tang, F., Wei, H., Zhao, S., Wang, L., Zheng, H., and Lu, M. (2016). Identification of microRNAs involved in regeneration of the secondary vascular system in populus tomentosa carr. *Front. Plant Sci.* 7, 724. doi: 10.3389/fpls.2016.00724

Tschaplinski, T. J., Abraham, P. E., Jawdy, S. S., Gunter, L. E., Martin, M. Z., Engle, N. L., et al. (2019). The nature of the progression of drought stress drives differential metabolomic responses in *Populus deltoides*. *Ann. Bot.* 124, 617–626. doi: 10.1093/aob/mcz002. In press.

Tuteja, N., and Mahajan, S. (2007). Calcium signaling network in plants: An overview. *Plant Signal. Behav.* 2, 79–85. doi: 10.4161/psb.2.2.4176

Vaid, N., Macovei, A., and Tuteja, N. (2013). Knights in action: Lectin receptor-like kinases in plant development and stress responses. *Mol. Plant* 6, 1405–1418. doi: 10.1093/mp/sst033

Verger, S., Chabout, S., Gineau, E., and Mouille, G. (2016). Cell adhesion in plants is under the control of putative O-fucosyltransferases. *J. Cell Sci.* 129, e1.2–e1.2. doi: 10.1242/jcs.195537

Verma, A., Lucas, A., Verma, S. S., Zhang, Y., Josyula, N., Khan, A., et al. (2018). PheWAS and Beyond: The Landscape of Associations with Medical Diagnoses and Clinical Measures across 38,662 Individuals from Geisinger. *Am. J. Hum. Genet.* 102, 592–608. doi: 10.1016/j.ajhg.2018.02.017

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

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M.II, Brown, M. A., et al. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* 101, 5–22. doi: 10.1016/j.ajhg.2017.06.005

Wang, Y., Sun, T., Li, T., Wang, M., Yang, G., and He, G. (2016). A CBL-interacting protein kinase TaCIPK2 confers drought tolerance in transgenic tobacco plants through regulating the stomatal movement. *PloS One* 11, e167962. doi: 10.1371/journal.pone.0167962

Wegrzyn, J. L., Eckert, A. J., Choi, M., Lee, J. M., Stanton, B. J., Sykes, R., et al. (2010). Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. *New Phytol.* 188, 515–532. doi: 10.1111/j.1469-8137.2010.03415.x

Weichselbaum, D., Zagrovic, B., and Polyansky, A. A. (2017). Fuento: functional enrichment for bioinformatics. *Bioinformatics* 33, 2604–2606. doi: 10.1093/bioinformatics/btx179

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

Weighill, D., Jones, P., Bleker, C., Ranjan, P., Shah, M., Zhao, N., et al. (2019). Multi-Phenotype Association Decomposition: Unraveling Complex Gene-Phenotype Relationships. *Front. Genet.* 10, 417. doi: 10.3389/fgene.2019.00417

Wu, H. C., Bulgakov, V. P., and Jinn, T. L. (2018). Pectin methylesterases: Cell wall remodeling proteins are required for plant response to heat stress. *Front. Plant Sci.* 871, 1612. doi: 10.3389/fpls.2018.01612

Xi, W., Song, D., Sun, J., Shen, J., and Li, L. (2017). Formation of wood secondary cell wall may involve two type cellulose synthase complexes in *Populus*. *Plant Mol. Biol.* 93, 419–429. doi: 10.1007/s11103-016-0570-8

Xie, M., Zhang, J., Tschaplinski, T. J., Tuskan, G. A., 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

Yang, Y., Labbé, J., Muchero, W., Yang, X., Jawdy, S. S., Kennedy, M., et al. (2016). Genome-wide analysis of lectin receptor-like kinases in *Populus trichocarpa*. *BMC Genomics* 17, 1–16. doi: 10.1186/s12864-016-3026-2

Yee, D., and Goring, D. R. (2009). The diversity of plant U-box E3 ubiquitin ligases: From upstream activators to downstream target substrates. *J. Exp. Bot.* 60, 1109–1121. doi: 10.1093/jxb/ern369

Yu, J., Holland, J. B., McMullen, M. D., and Buckler, E. S. (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics*. 178, 539–551. doi: 10.1534/genetics.107.074245

Zeng, H., Xu, L., Singh, A., Wang, H., Du, L., and Poovaiah, B. W. (2015). Involvement of calmodulin and calmodulin-like proteins in plant responses to abiotic stresses. *Front. Plant Sci.* 6, 600. doi: 10.3389/fpls.2015.00600

Zentella, R., Zhang, Z.-L., Park, M., Thomas, S. G., Endo, A., Murase, K., et al. (2007). Global Analysis of DELLA Direct Targets in Early Gibberellin Signaling in Arabidopsis. *Plant Cell* 19, 3037–3057. doi: 10.1105/tpc.107.054999

Zentella, R., Sui, N., Barnhill, B., Hsieh, W. P., Hu, J., Shabanowitz, J., et al. (2017). The Arabidopsis O-fucosyltransferase SPINDLY activates nuclear growth repressor della. *Nat. Chem. Biol.* 13, 479–485. doi: 10.1038/nchembio.2320

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

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–826. doi: 10.1038/ng.2310

Zhou, X., and Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome- wide association studies. *Nat. Genet.* 11, 407–411. doi: 10.1038/nmeth.2848

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.

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