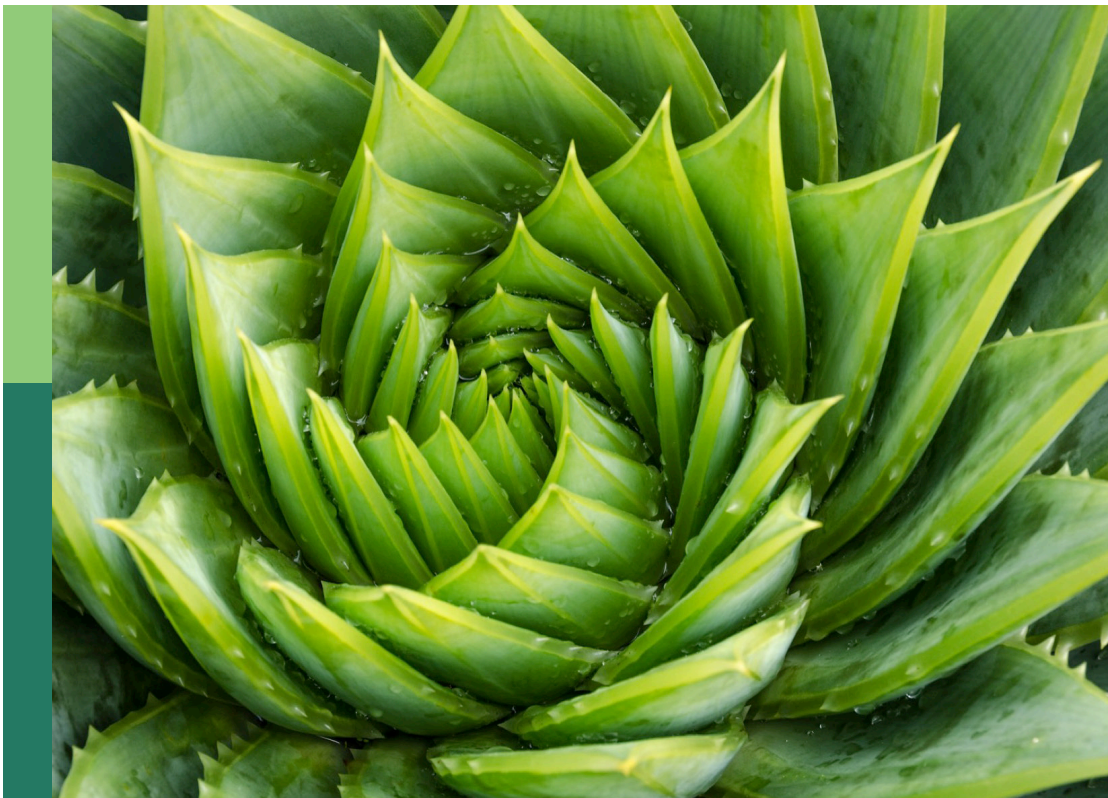
frontiers | Research Topics

The plant cell wall: Advances and current perspectives

Edited by
Wagner Rodrigo De Souza, Igor Cesarino and Rowan Andrew Craig Mitchell

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-2934-8
DOI 10.3389/978-2-8325-2934-8

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

The plant cell wall: Advances and current perspectives

Topic editors

Wagner Rodrigo De Souza — Federal University of ABC, Brazil

Igor Cesarino — University of São Paulo, Brazil

Rowan Andrew Craig Mitchell — Rothamsted Research, United Kingdom

Citation

De Souza, W. R., Cesarino, I., Mitchell, R. A. C., eds. (2023). *The plant cell wall: Advances and current perspectives*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-2934-8

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 22 June 2023
DOI 10.3389/fpls.2023.1235749

OPEN ACCESS

EDITED AND REVIEWED BY
Anna N Stepanova,
North Carolina State University,
United States

*CORRESPONDENCE
Wagner Rodrigo de Souza
wagner.souza@ufabc.edu.br
Rowan A.C. Mitchell
rowan.mitchell@rothamsted.ac.uk
Igor Cesarino
icesarino@usp.br

RECEIVED 06 June 2023
ACCEPTED 13 June 2023
PUBLISHED 22 June 2023

CITATION
de Souza WR, Mitchell RAC and Cesarino I (2023) Editorial: The plant cell wall: advances and current perspectives. *Front. Plant Sci.* 14:1235749. doi: 10.3389/fpls.2023.1235749

Editorial: The plant cell wall: advances and current perspectives

Wagner Rodrigo de Souza[1]*, Rowan A.C. Mitchell[2]* and Igor Cesarino[3,4]*

[1]Center for Natural and Human Sciences, Federal University of ABC, Santo André, Brazil, [2]Plant Sciences, Rothamsted Research, Hertfordshire, United Kingdom, [3]Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil, [4]Synthetic and Systems Biology Center, InovaUSP, São Paulo, Brazil

KEYWORDS

plant cell wall, biomass, cellulose, hemicellulose, lignin, bioeconomy

Editorial on the Research Topic
The plant cell wall: advances and current perspectives

Plant cells are surrounded by the cell wall, a dynamic component that shapes the cell and is key to their function. In developing tissues with an active cell elongation process, plant cells are surrounded by a primary cell wall (PCW), comprised of cellulose, hemicellulose, pectins, structural proteins, and, in grasses, phenolic compounds (Loqué et al., 2015). The PCW is responsible for maintaining cell shape and expanding cell size, bestowing it with mechanical strength, minimizing water loss and protecting it against stresses. A secondary cell wall (SCW) is deposited internally to the PCW once cell elongation ceases to allow specialized cells to perform their function. SCWs are composed of cellulose and hemicellulose impregnated by the phenolic polymer lignin, resulting in a complex and rigid structure that provides physical strength and hydrophobicity to supportive and water-transporting tissues (Meents et al., 2018). Cell walls also constitute the majority of plant biomass and thus play a crucial role in the food and biofuel industries (Burton and Fincher, 2014). Given that PCWs are thought to contain just a few layers of cellulose microfibrils whereas the SCWs contain hundreds, lignified SCWs account for the majority of plant biomass (Zeng et al., 2014). The proportions and chemical composition of the major components of the SCW also vary among cell types and plant species, and their physicochemical properties will ultimately determine biomass digestibility, thereby modulating nutrient release, gut biota, and health (Zhao et al., 2012; Burton and Fincher, 2014). This property also influences the production of biofuels, e.g. biomass fermentation to produce bioethanol. Understanding the distinct aspects of cell wall biology is therefore crucial for improving the production of plant-based biomaterials and developing plants with important characteristics for the food, agricultural, and bioenergy industries.

This Research Topic aimed to collate a wide spectrum of perspectives and advances in plant cell wall research. Thirteen articles were accepted for publication, and they are organized into two sections: 1) Advances in plant cell wall deposition/assembly/biogenesis; and 2) Biotechnological strategies toward optimized plant cell walls for the bioeconomy.

Advances in plant cell wall deposition/assembly/biogenesis

In this Research Topic, important advances in plant cell wall biogenesis were reported. Paterlini et al. performed biochemical characterization of the cell wall of *Arabidopsis thaliana* plasmodesmata, which are membrane-lined pores involved in the symplastic transport of biological molecules between neighbouring cells. Xyloglucans and pectins were shown to account for around 60% of the plasmodesmata cell wall, whereas enzymatic fingerprinting revealed specific polysaccharide signatures: most xyloglucans were fucosylated, homogalacturonans were not extensively methyl-esterified, rhamnogalacturonan I showed limited branching and rhamnogalacturonan II was highly methyl-acetylated. These data open new opportunities for the study of plasmodesmata function. Li et al. found numerous cell wall-related transcripts were differentially regulated when comparing transcriptomes during endosperm cellularization and endosperm differentiation of Arabidopsis wild-type and the *N-terminal acetyltransferase A subunit 15* (*naa15*) mutant where these processes are abnormal. Chen et al. used a rice rolling-leaf mutant, *rlm1-D*, to demonstrate that its phenotype, characterized by rolling leaves, is mainly caused by abnormal secondary cell wall (SCW) deposition. *RLM1* was cloned by a map-based method and found to encode an R2R3 MYB transcription factor that can bind to the promoter of *CINNAMYL ALCOHOL DEHYDROGENASE 2* (*OsCAD2*), a key gene responsible for lignin biosynthesis in rice. An interacting partner of RLM1, MITOGEN-ACTIVATED PROTEIN KINASE 10 (OsMAPK10), was also identified and the authors proposed a MAPK-MYB-*OsCAD2* genetic regulatory network controlling SCW deposition, providing novel insights into the molecular regulatory mechanisms controlling leaf morphology in rice.

By analyzing published pressure–volume curves and measures of succulence in 25 species of the order Caryophyllales, Fradera-Soler et al. showed that elastic adjustment, whereby plants change cell wall elasticity, is uniquely beneficial to succulents for avoiding turgor loss. The authors also studied phylogenetically diverse succulent species to demonstrate several differences in cell wall biochemistry between succulent and non-succulent leaves, pointing to the existence of what they called "succulent glycome". Dabravolski and Isayenkov discuss in a review the roles of cell wall components in salt stress tolerance and the regulatory mechanisms underlying cell wall maintenance under salt stress conditions. Xylan is emerging as the key polymer in SCW for linking polysaccharide and lignin components but is also present in PCW (Terrett and Dupree, 2019; Tryfona et al., 2023). New results suggesting separate xylan synthase complexes for PCW and SCW in Arabidopsis are discussed by Anders et al. In grass PCW and SCW, xylan is decorated with arabinose some of which have hydroxycinnamates attached (Scheller and Ulvskov, 2010; Terrett and Dupree, 2019). There are two articles on AT10, the BAHD acyltransferase enzyme responsible for addition of *p*-coumarate to arabinoxylan. Moller et al. showed that abolition of *OsAT10* by CRISPR/Cas9 almost abolished this linkage in rice. Houston et al. reported that a natural knock-out variant of the *HvAT10* ortholog in barley caused much lower levels of this linkage in barley grain cell walls. The role of *p*-coumarate on arabinoxylan is unclear but the similar ferulate decoration appears to be crucial in cross-linking arabinoxylan chains to each other and to lignin in grass cell walls as discussed in the review article by Chandrakanth et al. which also covers hydroxycinnamate decoration of lignin. Another feature of grass cell walls is the presence of (1,3;1,4)- β-glucan synthesized by CELLULOSE SYNTHASE-LIKE F6 (CSLF6). New evidence on control of transcription of this gene in barley grain is presented in Garcia-Gimenez et al.

Biotechnological strategies toward optimized plant cell walls for the bioeconomy

Engineering cell walls is a key strategy to generate optimized crops with enhanced processability to produce biofuels and other bioproducts in biorefineries (Loqué et al., 2015). When targeting lignin, two strategies are envisaged: reducing lignin content and altering lignin structure/composition. De Meester et al. reported on the engineering of curcumin, a natural metabolite harboring two phenolic rings linked by a labile aliphatic chain, as an alternative monomer incorporating into poplar lignin. By expressing two curcumin biosynthetic genes under the control of a SCW-specific promoter, curcumin was produced and incorporated into the lignified cell walls of poplar. However, different from what has been reported for Arabidopsis (Oyarce et al., 2019), the curcumin-producing transgenic poplars suffered from yield penalties in addition to altered cell wall composition. More importantly, the saccharification efficiency of the transgenic lines was not different from that of the control plants, suggesting that translating this strategy from Arabidopsis to crops will likely demand further optimization. In another strategy to alter lignin composition, Shafiei et al. reported on the down-regulation of the lignin biosynthetic gene *ferulate 5-hydroxylase* (*F5H*) in barley, which resulted in reduced syringyl/guaiacyl (S/G) ratio in the straw. Interestingly, parameters such as lignin content, straw mechanical properties, plant growth habit, grain characteristics, and saccharification efficiency all remained unaffected. These results suggest that altering S/G composition had little effect on plant development and biomass processability in barley. For grass biomass in general, targeting the ferulate responsible for cross-linking lignin to polysaccharide in SCW is seen as a promising approach to enhance saccharification, as is boosting of ester-linked ferulate on lignin in all plant biomass, both achieved by manipulation of *BAHD* genes (Chandrakanth et al.). Bioenergy and biorefining require specific biomass characteristics of a crop variety. Based on a detailed cell wall analyses of above-ground biomass (comprised of stem and leaf material) of 49 representative genotypes of the genus Miscanthus, Iacono et al. identified a number of cell wall related variables important for biomass recalcitrance. Their results emphasize the inter- and intra- specific variation in cell wall characteristics and biomass recalcitrance and

the importance of also considering yield- and organ-related parameters when analyzing cell wall properties and biomass recalcitrance aimed at improving Miscanthus as a biomass crop.

The characterization of the molecular mechanisms underlying the biosynthesis and deposition of the major cell wall components is essential not only for our understanding of how cell walls evolved as a dynamic component playing key roles in plant growth and development but also to allow the rational engineering of plant cell walls for the bioeconomy. Articles in this Research Topic provide novel insights into cell wall biogenesis and exciting biotechnological strategies for the optimization of plant biomass for biorefineries.

Author contributions

WS, IC and RM wrote about every article they each edited. All authors provided feedback on the Editorial. All authors contributed to the article and approved the submitted version.

Funding

IC is indebted to Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for the research grant n° 2021/06142-8 and to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the research fellowship 302626/2022-0. WRS would like to thank FAPESP for the research grant 2019/04878-7.

Acknowledgments

We thank all the authors that have participated in this topic for their important contributions.

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.

Publisher's note

References

Burton, R. A., and Fincher, G. B. (2014). Plant cell wall engineering: applications in biofuel production and improved human health. *Curr. Opin. Biotechnol.* 26, 79–84. doi: 10.1016/j.copbio.2013.10.007

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

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

Oyarce, P., Meester, B. D., Fonseca, F., de Vries, L., Goeminne, G., Pallidis, A., et al. (2019). Introducing curcumin biosynthesis in arabidopsis enhances lignocellulosic biomass processing. *Nat. Plants* 1, 225–237. doi: 10.1038/s41477-018-0350-3

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

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

Tryfona, T., Bourdon, M., Delgado Marques, R., Busse-Wicher, M., Vilaplana, F., Stott, K., et al. (2023). Grass xylan structural variation suggests functional specialization and distinctive interaction with cellulose and lignin. *Plant J.* 113, 1004–1020. doi: 10.1111/tpj.16096

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

Zhao, X., Zhang, L., and Liu, D. (2012). Biomass recalcitrance. part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuels Bioprod. Biorefining-Biofpr* 6, 465–482. doi: 10.1002/bbb.1331

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 31 May 2022
doi: 10.3389/fpls.2022.905111

RLM1, Encoding an R2R3 MYB Transcription Factor, Regulates the Development of Secondary Cell Wall in Rice

Zhenhua Chen[1†], Shouzhen Teng[1†], Di Liu[1†], Yuan Chang[1], Liying Zhang[1], Xuean Cui[1], Jinxia Wu[1], Pengfei Ai[2], Xuehui Sun[1], Tiegang Lu[1*] and Zhiguo Zhang[1*]*

[1]Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing, China, [2]College of Food Science and Biology, Hebei University of Science and Technology, Shijiazhuang, China

OPEN ACCESS

Edited by:
Wagner Rodrigo De Souza,
Federal University of ABC, Brazil

Reviewed by:
Ruibo Hu,
Qingdao Institute of Bioenergy and Bioprocess Technology (CAS), China
Yuanhu Xuan,
Shenyang Agricultural University, China

***Correspondence:**
Xuehui Sun
sunxuehui@caas.cn
Tiegang Lu
lutiegang@caas.cn
Zhiguo Zhang
zhangzhiguo@caas.cn

[†]These authors have contributed equally to this work

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

Received: 26 March 2022
Accepted: 02 May 2022
Published: 31 May 2022

Citation:
Chen Z, Teng S, Liu D, Chang Y, Zhang L, Cui X, Wu J, Ai P, Sun X, Lu T and Zhang Z (2022) RLM1, Encoding an R2R3 MYB Transcription Factor, Regulates the Development of Secondary Cell Wall in Rice. Front. Plant Sci. 13:905111. doi: 10.3389/fpls.2022.905111

Leaf morphology is an important component of rice ideal plant type. To date, many regulatory genes influencing leaf morphology in rice have been cloned, and their underlying molecular regulatory mechanism has been preliminarily clarified. However, the fine regulation relationship of leaf morphogenesis and plant type remains largely elusive. In this study, a rolling-leaf mutant, named *rlm1-D*, was obtained and controlled by a pair of dominant nuclear genes. Cytological observations revealed that the *rlm1* was mainly caused by abnormal deposition of secondary cell walls. Molecular evidence showed ectopic expression of a MYB-type transcription factor LOC_Os05g46610 was responsible for the phenotype of *rlm1-D*. A series of experiments, including the transcription factor-centered technology, DNA-binding assay, and electrophoretic mobility shift assay, verified that RLM1 can bind to the promoter of *OsCAD2*, a key gene responsible for lignin biosynthesis in rice. An interacting partner of RLM1, OsMAPK10, was identified. Multiple biochemical assays confirmed that OsMAPK10 interacted with RLM1. OsMAPK10 positively regulated the lignin content in the leaves and stems of rice. Moreover, OsMAPK10 contributes to RLM1 activation of downstream target genes. In particular, *RLM1* is exclusively expressed in the stems at the mature plant stage. The yield of RLM1 knockdown lines increased by over 11% without other adverse agricultural trait penalties, indicating great practical application value. A MAPK-MYB-OsCAD2 genetic regulatory network controlling SCW was proposed, providing a theoretical significance and practical value for shaping the ideal plant type and improving rice yield.

Keywords: rice, RLM1, R2R3 MYB transcription factor, secondary cell wall, yield

INTRODUCTION

Leaf morphology plays an important role in rice ideal plant type. Leaf morphology improvement focuses on erect leaves and moderately rolled leaves. Moderate rolling can keep rice leaves upright with no drooping, improving the light reception conditions at the base of the canopy at the middle and later stages, increase the light energy use rate, and enhancing root activity

to improve lodging resistance (Zhang et al., 2009; Zou et al., 2011). Some rolling-leaf cultivars, such as Liangyoupeiyou and Miyang, have become popular with growers (Wang et al., 2003). Therefore, the study of leaf morphogenesis has become the focus of breeding for high-yield rice.

Leaf morphogenesis is a complex developmental process that is influenced by many factors. In *Arabidopsis*, the genetic regulatory network of leaf morphogenesis is relatively well known. HD-Zip III, YABBY, and KANADI are involved in the establishment and development of leaf polarity (Emery et al., 2003; Ha et al., 2010). By regulating the expression of HD-Zip III and *ARF3/ARF4*, miRNA165/166 and ta-siRNA, respectively, are also involved in the establishment and development of leaf polarity (Hunter et al., 2006; Merelo et al., 2016). To date, more than 10 regulatory genes controlling leaf morphogenesis have been cloned in rice, and their underlying molecular regulatory mechanisms have been preliminarily studied. The genes involved mainly include *LC2* (Zhao et al., 2010), *OsHB1* (Itoh et al., 2008), *ACL1/ACL2* (Li et al., 2010), *Brd1* (Hong et al., 2002), *CLD1/SRL1* (Li et al., 2017); *Yab1* (Dai et al., 2007), *OsCSLD4/NRL1* (Hu et al., 2010), *OsAGO7* (Shi et al., 2007), *RL14* (Fang et al., 2012), *CFL1* (Wu et al., 2011), *SLL1* (Zhang et al., 2009), *ROC5* (Zou et al., 2011), *ROC8* (Sun et al., 2020), and *PSL1* (Zhang et al., 2021). These genes mainly play a role in regulating the number and size of bulliform cells. *RL14* and *PSL1* positively regulate the development of bulliform cells, while *Brd1*, *Yab1*, *RL14, LC2*, *ROC5, ROC8*, *OSHb1*, and *ACL1/ACL2* negatively regulate the development of these cells. In addition to the size and numbers of bulliform cells that cause leaf rolling in rice, the distribution of mesophyll cells, the formation of sclerenchyma cells, and the development of leaf cuticles can affect leaf rolling. *SLL1* encodes a KANADI family MYB transcription factor. Mutation of *SLL1* leads to a disruption of the programmed death process of abaxial mesophyll cells and inhibits the normal formation of abaxial sclerenchyma. *CFL1* encodes an unknown protein with a WW domain and promotes leaf rolling. CFL1 interacts with HDG1 (HD-ZIP IV) proteins and negatively regulates the development of cuticles by affecting the expression of *BDG* and *FDH* genes related to cuticle development. Totally, leaf morphology underlying molecular regulatory mechanism has been preliminarily clarified. However, the fine regulation relationship of leaf morphogenesis and plant type remains largely elusive.

To better understand the molecular mechanism of leaf morphogenesis and plant type, the rolling-leaf mutant *rlm1* was studied in depth. The *rlm1* is controlled by a pair of dominant nuclear genes. Cytological observations revealed that RLM1 directly affects secondary cell wall (SCW) deposition, indirectly regulates the development of bulliform cells, and promotes leaf rolling. *RLM1* was cloned by a map-based method and found to encode an R2R3 MYB transcription factor. The *rlm1-D* mutant phenotype was caused by the ectopic expression of the R2R3 MYB transcription factor. Combining biochemical, cytological, and genetic strategies, we proposed a new molecular regulatory network based on MAPK-MYB-OsCAD2. This study improves our understanding of the relationship between leaf morphogenesis and plant type and provides genetic resources for cultivating ideal plant-type rice materials.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

The *rlm1-D*, WT, overexpression lines RLM1-OE1 and RLM1-OE2, mutant lines *RLM1-m1* and *RLM1-m2*, and plants used for mapping populations were grown in the greenhouse and at a field station of the Chinese Academy of Agriculture Science. All the plants tested at the seedling stage were grown in a growth chamber with 10 h daytime/14 h night-time photoperiod at 28°C, under a light intensity of 150–200 μmol/m^2/s and under 60–70% relative humidity (RH).

Mapping of RLM1

A mapping population was generated by the crossing of *rlm1-D* heterozygous plants with the *Dular* cultivar (*Indica*). Ten WT leaf samples were collected for preliminary mapping. Nearly 500 plants were used for fine mapping, and the location of *RLM1* locus between Indel2670 and Indel2672 was narrowed. A T-DNA insertion was found in the promoter region of LOC_Os05g46610.

Gene Overexpression and CRISPR-Cas9

RLM1-pCAMBIA1302 vectors containing the full RLM1 ORF were constructed to reproduce the phenotype. Two Cas9 targets were selected *via* CRISPR-gene editing (GE),[1] and RLM1-pP1C.7 was constructed. The full-length OsMAPK10 CDS was cloned into pCAMBIA1302, and two Cas9 targets were selected *via* CRISPR-GE to generate the OsMAPK10-pP1C.7 knockdown plasmids. All the plasmids were then transformed into the *Agrobacterium* AGL1 strain, which were then transformed into *Nipponbare* calli.

Analysis of Expression Patterns by qRT-PCR and GUS Staining

RNA samples were isolated from WT plants at the seedling and heading stages. The cDNAs were synthesized *via* FastKing gDNA Dispelling RT SuperMix (KR118, TIANGEN, Beijing). qRT-PCR was subsequently performed with SYBR qPCR Mix (ZF501, ZOMANBIO, Beijing) according to the manufacturer's instructions. The housekeeping gene *OsActin1 (LOC_Os03g0718100)* was used as an internal control to normalize gene expression.

ProRLM1-1391Z vectors were used to determine the expression pattern, and the 2,000 bp sequence upstream was amplified from the RLM1 codon start site. The recombinant pRLM1-1391Z *Agrobacterium* strain was subsequently transformed into *Nipponbare* calli. At the seedling and heading stage, GUS staining was conducted, with 2 mM X-Gluc applied to different tissues.

[1]http://skl.scau.edu.cn

Subcellular Localization of RLM1

The full-length CDS of RLM1 was cloned, amplified, and inserted into pAN580. The resulting RLM1-pAN580 constructs were transformed into *Nipponbare* protoplasts, and mCherry was selected as a nuclear marker. The signals were detected using a laser scanning microscope (LSM980, Zesis, Germany).

SEM Observations

Stem samples from leaves or the 2nd internode were collected from rice at the heading stage. The tissues were cut to 0.5 cm and fixed in 2.5% glutaraldehyde for 12 h. The samples were prepared according to the methods of Wang et al. (2022) and then imaged with a scanning electron microscope (H-7500, Hitachi, Japan).

Lignin and Cellulose Content Measurements

Samples of leaves and stems from WT, *rlm1-D*, RLM1-OE, *RLM1-m*, *OsMAPK10-m*, and OsMAPK10-OE plants at the heading stage were collected. The lignin and cellulose contents of 2 g samples were quantified by Wuhan Metwell Biotechnology Co., Ltd. (Wuhan, China).

Paraffin Sectioning and Observations

Leaves from *rlm1-D* and WT plants were collected and fixed in Carnoy's solution (ethanol/glacial acetic acid = 3/1) overnight. The samples were then treated according to the methods of Sun et al. (2020) and observed as previously described (Cui et al., 2019).

Phylogenetic Analysis

A BLAST search of *Arabidopsis* and *Zea mays* database was conducted, with the RLM1 protein sequence used as a query sequence. A phylogenetic tree was generated by MEGA XI software; the unrooted tree was constructed according to the neighbor-joining method.

Transcriptome Data Analysis

Total RNA was extracted from WT and *rlm1-D* plants in the 3rd-leaf stage. Seqhealth Technology Co., Ltd. (Wuhan, China) prepared the libraries and performed high-throughput sequencing. The reference genome sequence was downloaded from the NCBI database (PRJDB1747). KEGG enrichment analysis of differentially expressed genes was implemented by KOBAS software (version 2.1.1) with a corrected *p*-value cut-off of 0.05 to judge statistically significant enrichment.

Y2H Assays

The full-length CDSs of RLM1 and OsMAPK10 were inserted into the *EcoRI* and *BamHI* sites of pGBKT7 (bait) and pGADT7 (prey), respectively, by an In-Fusion kit (Clontech, Japan). All the resulting constructs were verified by sequencing before being transformed into yeast strain AH109. The bait and prey plasmids were cotransformed into AH109 cells by an Ex-Yeast Transformation Kit (ZC135; ZOMANBIO, Beijing, China). The yeast colonies were transferred to plates containing SD/-Leu/-Trp and SD/-His/-Trp/-Leu/60 mM 3-AT media and then allowed to grow at 28°C for 3 days.

Pull-Down Assays

Pull-Down assays were performed with GST-RLM1 and His-OsMAPK10. The RLM1 and OsMAPK10 coding sequences were cloned into pGEX6p-1 or pACYCDuet1 and purified with GST-Tag or His-Tag magnetic beads, respectively. GST-RLM1 and His-MAPK10 were subsequently incubated in binding buffer [20 mM sodium dihydrogen phosphate, 300 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM dithiothreitol (DTT; pH 7.4)] for 2 h, after which binding buffer was used to wash the beads 5 times. Then, elution buffer [20 mM sodium dihydrogen phosphate, 300 mM NaCl, 500 mM imidazole (pH 7.4)] was applied to elute the proteins for subsequent western blot assays. Anti-RLM1 and anti-GST antibodies were utilized to detect the pulled down RLM1, and anti-His antibodies were utilized to detect the input.

Split-LUC Complementation Assays

The RLM1 CDS was cloned into pCAMBIA1300-nLUC by an In-Fusion kit (Clontech, Japan), and the OsMAPK10 CDS was cloned into pCAMBIA1300-cLUC. The plasmid was sequenced for verification and then transformed into *Agrobacterium* strain GV3101. Four-week-old *N. benthamiana* plants were infiltrated by the transformed *Agrobacterium* cells and incubated at 23°C under dark conditions for 2 days; pCAMBIA1300-nLUC with pCAMBIA1330-cLUC served as a blank control. The leaves were sprayed with D-luciferin potassium salt (Yeasen, Shanghai, China), and the fluorescence was detected by an IVIS Lumina LT series III instrument (PerkinElmer).

Y1H Assays

The full-length CDS of RLM1 was inserted into the *EcoRI* and *BamHI* sites of pGADT7 by an In-Fusion kit (Clontech). The MYB1AT motif from the *OsCAD2* promoter was inserted (as three tandem repeats) into pHis2 at both the *EcoRI* and *SacI* cloning sites. The plasmids were subsequently transformed into yeast strain Y187 by an Ex-Yeast Transformation Kit. The yeast colonies were transferred to plates containing SD/-Trp/-Leu and SD/-His/-Trp/-Leu media supplemented with 60 mM 3-AT and allowed to grow at 28°C for 3 days.

EMSA Assays

The direct binding of RLM1 to the *OsCAD2* promoter was detected using an EMSA kit (GS009, Beyotime, Shanghai, China), following the manufacturer's protocol, in conjunction with probes. The CDS of *RLM1* was cloned into the *Escherichia coli* expression vector pGEX6p-1 with a GST tag at the *BamHI* and *EcoRI* sites. The resulting plasmid was then transformed into the *E. coli* Rosetta strain and induced with 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Recombinant GST-RLM1 was purified using GST purification magnetic beads (abs9902, Absin, Shanghai China) according to the manufacturer's protocol.

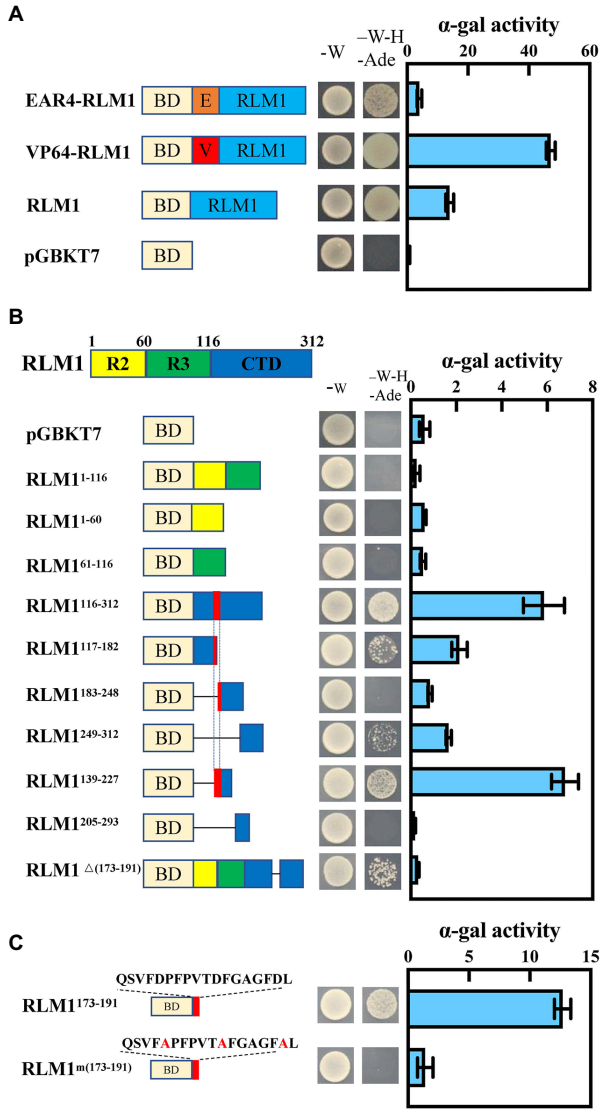


FIGURE 1 | RLM1 acts as a transcriptional activator. **(A)** Transactivation analysis of RLM1, VP64-RLM1, and EAR4-RLM1 *via* a yeast assay. VP64 and EAR4 are an exogenous activation domain and an exogenous repression module, respectively. pGBKT7 alone was used as a negative control. Plate auxotroph and α-galactosidase assay showing the transcriptional activation of each protein. The bars represent the means ± SDs of three independent experiments. α-gal, α-galactosidase; -W, SD/-Trp; -W-H-Ade, SD/-Trp/-His/-Ade. **(B)** Mapping of the transactivation motif of RLM1 *via* a yeast assay. R2, aa 1–60; R3, aa 61–116; CTD, aa 117–312. Plate auxotroph and α-galactosidase assay showing the transcriptional activation of each protein. The bars represent the means ± SDs of three independent experiments. α-gal, α-galactosidase; pGBKT7, GAL4 DNA-binding domain; -W, SD; -W-H-Ade, SD/-Trp/-His/-Ade. **(C)** Transactivation analysis of the C-terminal motif QSVFDPFPVTDFGAGFDL *via* a yeast assay. BD represents the GAL4 DNA-binding domain. Mutations introduced into the C-terminal motif are indicated by the red font. Plate auxotroph and α-galactosidase assay showing the transcriptional activation by each protein. The bars represent the means ± SDs of three independent experiments. α-gal, α-galactosidase; -W, SD/-Trp; -W-H-Ade, SD/-Trp/-His/-Ade.

The biotin-labeled oligonucleotide probes were synthesized and labeled by Shanghai Sangon Company (Shanghai, China).

TF-Centered Y1H Assays

A 7 bp random motif library from yeast strain Y187 was purchased from Nanjing Ruiyuan Biotechnology Co., Ltd. (Nanjing, China). A TF-centered Y1H assay was performed as previously described (Ji et al., 2018). The yeast motif library was incubated overnight, after which an Ex-Yeast Transformation Kit was utilized to generate competent cells. Then, 35 μg of pGADT7-RLM1 was transformed into competent yeast library cells. The yeast was transferred to plates of SD/-His/-Trp/-Leu media supplemented with 60 mM 3-AT and allowed to grow at 28°C for 5 days. Monoclonal colonies were selected for sequencing, and random motif sequences between "GGG" and "CCC" (the *SmaI* site) were screened. The insertion sequences were analyzed using PLACE[2] and PlantCARE[3] to identify whether they were known motifs.

Transactivation Activity Assay in Yeast

VP64 (encoding an exogenous activation domain) or EAR4 (encoding an exogenous repressor domain) was fused to the CDS of RLM1, which was then inserted into pGBKT7, and an Ex-Yeast Transformation Kit was used to transform these vectors into yeast strain AH109; the resulting yeast were then cultured on SD/-Trp media. Then, single colonies were selected and cultured on SD/-Ade-Trp-His media at 28°C for 3 days. The α-gal enzyme activity was assayed according to the Yeast Protocols Handbook (No. PT3024-1, Clontech). To finely map the active domain of OsRLM1, we divided the protein into different parts, as shown in **Figure 1**, and measured the α-gal enzyme activity.

Transactivation Assay in Rice Protoplast

The promoter of *OsCAD2*, which was defined as the region 2 kb upstream from the transcription start site, was cloned and inserted into pGreenII-0800-Luc *via* the In-Fusion strategy to generate the corresponding reporter vector proOsCAD2-0800. The *RLM1* and *OsMAPK10* ORFs were cloned and inserted into pCAMBIA1307 to generate effector constructs. The resulting plasmids were transformed into rice protoplasts and cultured for 16 h. pCAMBIA1307 with pOsCAD2-0800 was used as a control. The LUC/REN ratio was determined by a Dual-Luciferase Reporter Assay System (Vazyme, DL101-01) according to the manufacturer's instructions.

In vitro Phosphorylation Assay

RLM1 phosphorylation by OsMAPK10 was performed with purified GST-RLM1 and His-OsMAPK10. The proteins were incubated in kinase buffer [50 mM hydroxyethyl piperazineethanesulfonic acid (HEPES; pH 7.4), 10 mM $MnCl_2$, 1 mM DTT, 30 μM ATP] at 30°C for 30 min. The reaction was stopped, electrophoresis was performed on an SDS-PAGE gel (6% SDS, 5 μM Phos-tag, 10 μM $MnCl_2$), and RLM1 was detected with an anti-GST antibody.

[2]http://www.dna.affrc.go.jp/PLACE/

[3]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

Statistical Analysis

The experiments were repeated at least three times. The means with standard deviation are shown in the figures. Significant differences were determined by Student's *t* test, and are marked with asterisks (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$).

Primer and Gene Sequences

The primers used are shown in **Supplementary Table S3**.

RESULTS

Identification of *rlm1-D*

To explore the regulatory factors controlling leaf morphogenesis in rice, a rolling-leaf mutant, named *rlm1-D*, was obtained from a rice T-DNA mutant library (Wan et al., 2009). The *rlm1-D* presented dwarfism, rolling leaves, and a dark leaf color at the tillering stage (**Figures 2A–C**). In particular, owing to the excessive rolling of its leaves, the *rlm1-D* mutant did not produce seeds normally. There were three phenotypes of the *rlm1-D* heterozygous offspring: excessive-rolling, half-rolling, and wild-type like (WT-like) phenotypes, with which the corresponding segregation ratio was 1/2/1 (**Supplementary Figure S1C**). This suggests that *rlm1-D* may be controlled by a pair of dominant nuclear genes and that the *rlm1-D* gene may have a dose-dependent effect on leaf rolling.

To observe the cytological changes of *rlm1-D*, paraffin sectioning demonstrated that the size and number of bulliform cells in *rlm1-D* were significantly reduced compared to those in the WT (**Supplementary Figures S1D,E**). Since the *rlm1-D* plants exhibited a rolling-leaf phenotype accompanied by a dwarf-type phenotype, the features of *rlm1-D* were similar to those of the *roc8* and *rl14* mutants (Fang et al., 2012; Sun et al., 2020). Thus, we determined the leaf structure of *rlm1-D* and the WT using scanning electron microscopy (SEM). SEM results revealed that sclerenchyma cell wall (SCW) deposition was abnormal in *rlm1-D* internode tissues compared to WT tissues (**Figure 2D**). Therefore, we measured the contents of lignin and cellulose in *rlm1-D* and WT leaf and stem tissues. The results showed that the lignin contents of the stems and leaves of *rlm1-D* were significantly higher than those of the WT (**Figures 2E,G**). However, the cellulose content in *rlm1-D* was comparable to that in the WT (**Figures 2F,H**). The results of the abovementioned experiments indicated that abnormal SCWs in *rlm1-D* were perhaps the main factor responsible for leaf rolling. The abnormal SCWs in *rlm1-D* perhaps influenced water transport and indirectly regulated the size of bulliform cells, which promoted leaf rolling in *rlm1-D,* similar as *rl14* mutant (Fang et al., 2012).

Map-Based Cloning of *RLM1*

To clone *RLM1*, a genetic mapping population was generated by a cross between *rlm1-D* heterozygous plants and the rice *indica* cultivar *Dular*. *RLM1* was preliminarily mapped to a 25.2–27.72 cM range on chromosome 5. Further fine mapping narrowed *RLM1* to a 100 kb region (**Figure 3A**). Sequence analysis revealed a T-DNA insertion 750 bp upstream of the start codon (ATG) of LOC_Os05g46610 (**Figure 3B**). We speculated that the CaMV35S promoter in the T-DNA region activated LOC_Os05g46610 expression. To confirm this hypothesis, a quantitative reverse transcription PCR (qRT-PCR) experiment was performed to judge the expression levels of LOC_Os05g46570, LOC_Os05g46580, LOC_Os05g46610, LOC_Os05g46620, and LOC_Os05g46630 in *rlm1-D*, *rlm1-D* heterozygous plants and WT plants. The results demonstrated that the expression levels of the abovementioned genes (except for LOC_Os05g46610) in *rlm1-D* were comparable to those in the WT. The LOC_Os05g46610 expression level in *rlm1-D* increased by ~1,000-fold. Additionally, the expression level of LOC_Os05g46610 in the *rlm1-D* heterozygous plant increased to a level that was ~800-fold greater (**Figures 3C–G**). Overall, the map-based cloning and qRT-PCR analysis indicated that LOC_Os05g46610 was a candidate gene of *RLM1*.

To further verify that LOC_Os05g46610 is responsible for the *rlm1-D* phenotype, an overexpression experiment was conducted. *Agrobacterium tumefaciens* strain AGL1 containing the recombinant construct (LOC_Os05g46610-pCambia1302) was used to infect the *Nipponbare* calli. Fifteen lines among the T_1 progeny revealed rolling-leaf and dwarf-type phenotypes (**Figures 4A–C**). qRT-PCR analysis showed that the degree of the leaf rolling among the overexpression lines was positively correlated with the expression level of LOC_Os05g46610 (**Figure 4D**). SEM experiments also demonstrated that SCW deposition was abnormal in RLM1-OE1 and RLM1-OE2 (**Figure 4E**). Lignin and cellulose content measurements also verified the increased lignin content and normal cellulose content in RLM1-OE1 and RLM1-OE2, the contents of which were similar to those of the *rlm1-D* mutant (**Figures 4F–I**). The results of the abovementioned experiments confirmed that ectopic expression of LOC_Os05g46610 was responsible for the rolling-leaf phenotype of *rlm1-D*.

Bioinformatics and Molecular Characteristics of RLM1

BLAST analysis demonstrated that RLM1 encoded an R2R3 MYB transcription factor. R2R3 MYB transcription factors are major transcriptional regulators that mediate a variety of plant biological processes, including seed germination, drought tolerance, stomatal conductance, lateral root development, hormone biosynthesis, anthocyanin accumulation, and cuticle wax biosynthesis (Jiang and Rao, 2020). The function of RLM1, which encodes an R2R3 MYB transcription factor, is unknown. RLM1 contains three exons and two introns, and its predicted ORF was 936 bp and encoded 312 amino acids (aa). Protein motif analysis demonstrated that RLM1 contained an N-terminal domain (R2R3) and a C-terminal domain (CTD). The N-terminus likely mainly interacts with other proteins, and the C-terminus activates or inhibits downstream target genes (Meng et al., 2019).

The amino acid sequence of RLM1 is 55% homologous with that of AtMYB4 from *Arabidopsis* (**Supplementary Figure S2**).

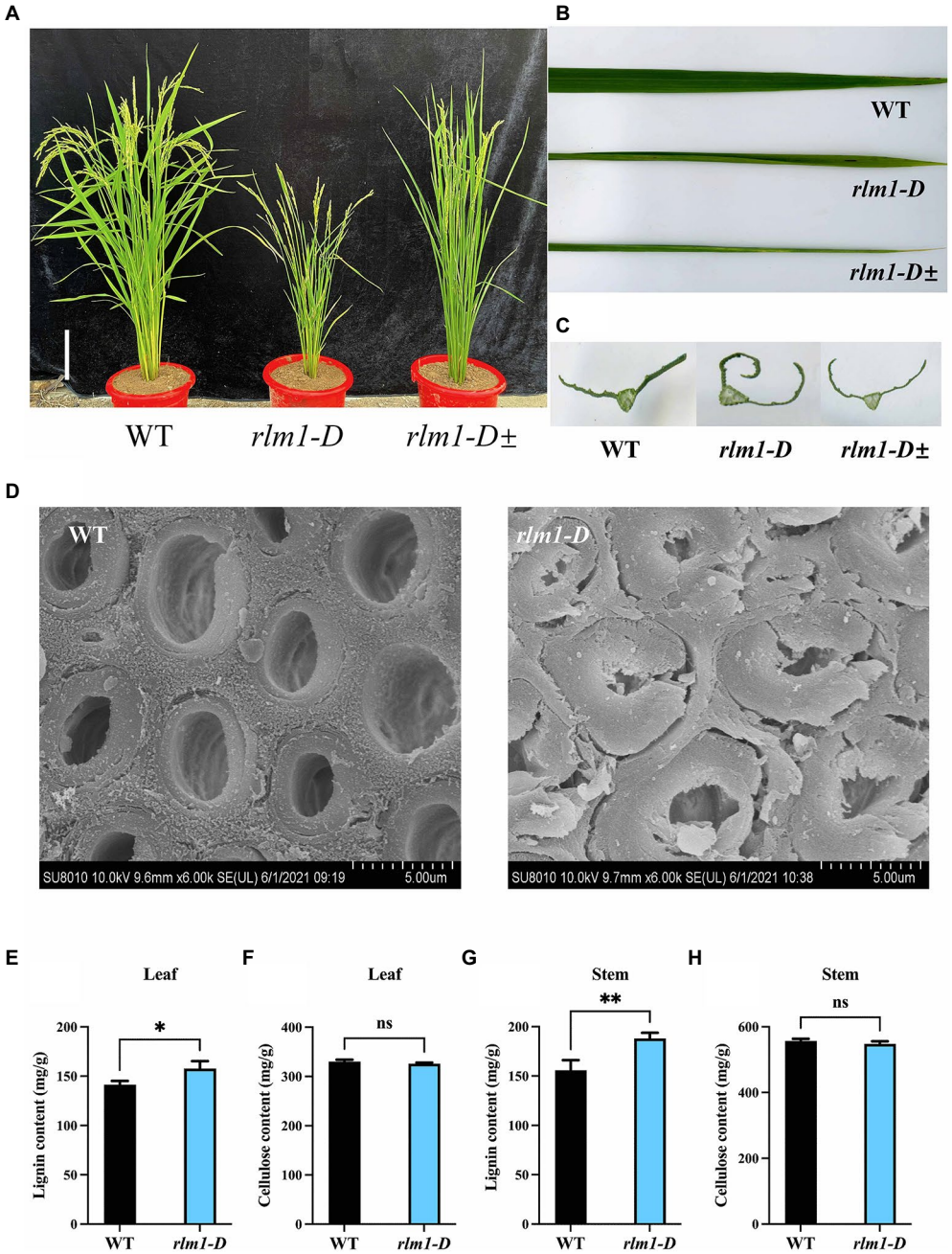


FIGURE 2 | Clarification of the *rlm1-D* phenotype. **(A)** Phenotype comparisons of WT, *rlm1-D*, and *rlm1* heterozygous (*rlm1-D*±) plants at the heading stage. Bar = 5 cm. **(B)** Flag leaf comparison of the plants shown **(A)**. **(C)** Flag leaf cross-section comparisons of the plants shown in **(A)**. **(D)** SEM observations of sclerenchyma cell walls in the internodes from the 3-month-old WT and *rlm1-D* plants. Bar = 5 μm. **(E)** Lignin content measurements in WT and *rlm1-D* leaf tissue. **(F)** Cellulose content measurements in WT and *rlm1-D* leaf tissue. **(G)** Lignin content measurements in the internodes of the 3-month-old WT and *rlm1-D* plants. **(H)** Cellulose content measurements in secondary stem tissues of WT and *rlm1-D* plants. $*p < 0.05$, $**p < 0.01$ (Student's *t*-test).

In *Arabidopsis*, AtMYB4 acts as a transcriptional repressor (Jin et al., 2000). To confirm whether RLM1 has transcriptional inhibitory activity, the constructs (RLM1-pGBKT7, VP64-RLM1-pGBKT7, EAR4-RLM1-pGBKT7) were expressed in yeast, respectively. The yeast colony status on the selected media (SD/−Ade-Trp-His) showed that the activation activity of VP64-RLM1-pGBKT7, RLM1-pGBKT7, pGBKT7, and EAR4-RLM1-pGBKT7 decreased successively (**Figure 1A**). The above results showed that, in contrast to AtMYB4, RLM1 mainly had a transcriptional activation activity.

To further study which region of RLM1 had activation activity, the full-length RLM1 sequence was divided into three regions: R2 ($RLM1^{1-60}$), R3 ($RLM1^{61-116}$), and CTD ($RLM1^{116-312}$). Then, three different constructs (R2-pGBKT7, R3-pGBKT7 and

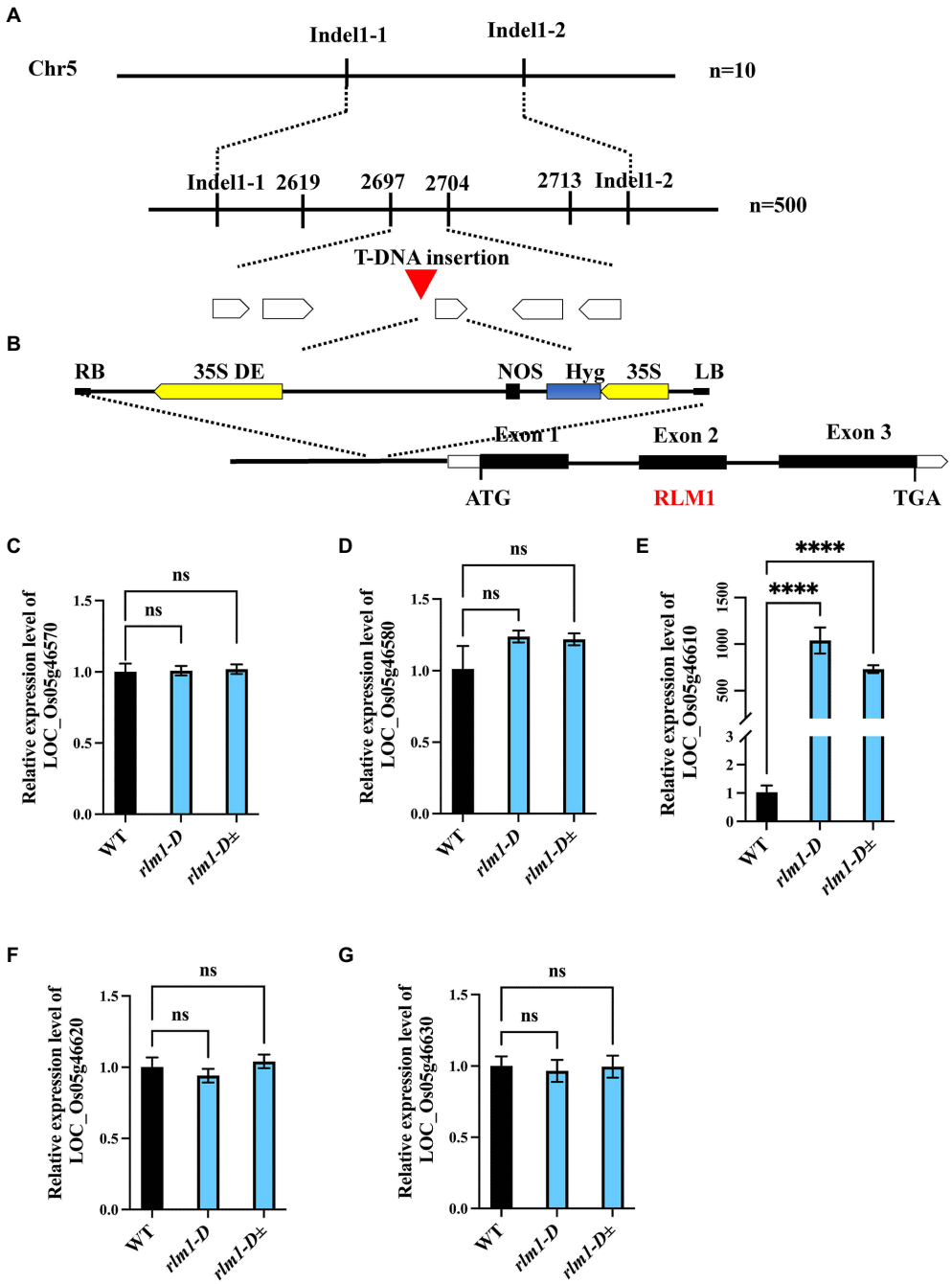


FIGURE 3 | Map-based cloning of RLM1. **(A)** The RLM1 locus was mapped to a 2.5 cM interval on chromosome 5, between Indel markers Indel1-1 and Indel1-2. The location of the RLM1 locus was narrowed to a 10 kb region between insertion–deletion (InDel) markers 2,697 and 2,704 *via* 500 F2 WT-like plants. **(B)** The T-DNA insert site is located upstream of LOC_Os05g46610. The structure of RLM1 is shown. ATG and TGA, start and stop codons, respectively. Boxes, exons; lines between black boxes, introns. **(C–G)** Expression analysis of LOC_Os05g46570, LOC_Os05g46580, LOC_Os05g46570610, LOC_Os05g46620, and LOC_Os05g46630 in WT, *rlm1-D*, and heterozygous (*rlm1-D*±) plants. ****$p < 0.0001$ (Student's *t*-test).

CTD-pGBKT7) were expressed in yeast, respectively. Yeast activation assays revealed that R2-pGBKT7 and R3-pGBKT7 had no activation effect in yeast growing on selective media (SD/-Ade-Trp-His). However, CTD- pGBKT7 had a stronger activation effect on yeast growing on the selective medium (SD/-Ade-Trp-His; **Figure 1B**).

Next, we tested the activation activities of different CTD deletion constructs. Yeast activation experiments revealed that an 18 amino acid (aa) sequence (QSVFDPFPVTDFGAGFDL) was very important for the activation activity of RLM1. Partial mutation or full deletion of these 18 aa had a great impact on the activation activity of RLM1. Moreover, the 18 aa motif

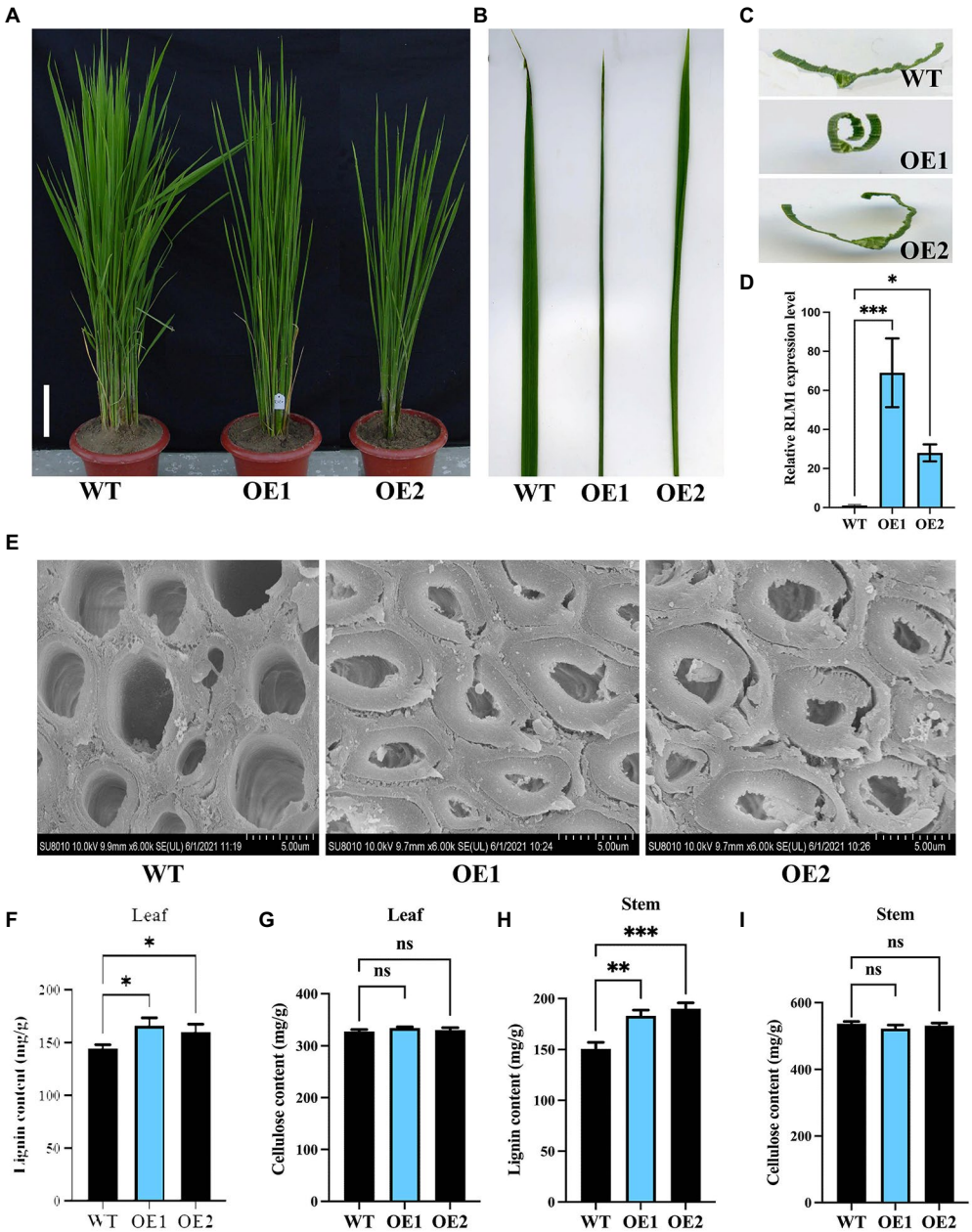


FIGURE 4 | Analysis of RLM1 overexpression in the WT background. **(A)** Comparison of WT and RLM1 overexpression plants (OE1, OE2) at the late-tiller stage. Bar = 5 cm. **(B)** Flag leaf comparisons of the plants shown in **(A)**. **(C)** Flag leaf cross-section comparisons of the plants shown in **(A)**. **(D)**. RLM1 expression level analysis of WT and RLM1-overexpressing plants (OE1, OE2). **(E)** SEM observations of sclerenchyma cell walls in the internodes from the 3-month-old WT, OE1, and OE2 plants. Bar = 5 μm. **(F)** Lignin content measurements in WT, OE1, and OE2 leaf tissues. **(G)** Cellulose content measurements in WT, OE1, and OE2 leaf tissue. **(H)** Lignin content measurements in the second internode tissues from 3-month-old WT, OE1, and OE2 plants. **(I)** Cellulose content measurements in the second internode tissues from 3-month-old WT, OE1, and OE2 plants. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ (Student's *t*-test).

alone had strong transcriptional activation activity, similar to that of the full RLM1-pGBKT7. This 18 aa motif contains some acidic amino acids, and after these acidic amino acids were mutated (Glu/Asp replaced by Ala), the activation activity was no longer detected, indicating that the acidic amino acids of these 18 aa were very important for the activation activity of RLM1 (**Figure 1C**).

Temporal and Spatial Expression Patterns of RLM1

To determine the tissue specificity of *RLM1*, *A. tumefaciens* strain containing proRLM1-1391Z construct were used to infect *Nipponbare* calli. β-Glucuronidase staining of T_1 progeny revealed that RLM1 is mainly expressed in young seedlings. In particular, RLM1 was also specifically expressed in the stem, possibly at

the site of the intermediate meristem at the mature stage, and RLM1 was not expressed in the leaf, panicle, or root tissue at the mature stage (**Supplementary Figures S3A–H**). The qRT-PCR experiment showed that *RLM1* was highly expressed at the young seedling stage and in stem tissue at the mature stage but was not expressed in the other tissues, which is generally consistent with the GUS staining assay results (**Supplementary Figure S3I**).

RLM1 encodes an R2R3-type MYB transcription factor and may localize to the nucleus. To judge whether RLM1 is localized in the nucleus, a RLM1-pAN580 construct was transiently transformed into rice protoplasts. Confocal observations revealed that RLM1 indeed localized to the nucleus and acted as a transcription factor (**Supplementary Figures S4A–D**).

Transcriptome Analysis of RLM1

To verify the regulatory pathways in which RLM1 is involved, RNA samples from 3rd-leaf stage WT and *rlm1-D* seedlings were collected. Then, high-throughput sequencing and analysis were performed by Seqhealth Technology Co., Ltd. (Wuhan, China). We obtained high-quality data (BioProject accession: PRJNA813008), with 97% of the reads mapped to the reference transcriptome. A total of 1,777 genes were differentially expressed in the *rlm1-D* plants, of which 973 genes were upregulated and 804 genes were downregulated, which is consistent with RLM1 acting as a transcriptional activator (at least three repeats, $p \leq 0.01$; **Supplementary Table S1**). Kyoto Encyclopedia of Gene and Genomes (KEGG) analysis demonstrated that the upregulated genes were involved in the phenylpropanoid metabolic pathway and that the downregulated genes played a role in the biosynthesis of secondary metabolites (**Supplementary Figures S5A,B**). We selected 19 upregulated genes involved in the phenylpropanoid metabolic pathway. The gene expression levels were confirmed by qRT-PCR, and the results were generally consistent with the transcriptome data (**Supplementary Figure S6A**).

RLM1 Target Binding Gene Analysis

To identify the target sites that may be directly regulated by RLM1, using transcription factor-centered (TF-centered) yeast one-hybrid (Y1H) technology, we screened the motifs of RLM1-binding target sites. A pHis2 yeast library containing 7 bp random fragments was constructed and screened by using RLM1-AD according to the mentioned methods (Ji et al., 2014). A total of 122 motifs were obtained to perform enrichment analysis,[4] and we found that 20 motifs possibly bind to RLM1 (**Supplementary Table S2**). Selection on yeast screening media (SD/-Trp-Leu-His with 60 mM 3-AT) revealed that all 20 motifs could be recognized by RLM1 (**Supplementary Figure S7**). These motifs are mainly composed of cis-acting elements whose core sequence is "MYB1AT." However, other cis-acting elements, such as *CAATBOX1* and *GATABOX*, were also found.

[4]www.dna.affrc.go.jp/PLACE/?action=newplac

Combined with the transcriptome sequencing results, additional results showed that these motifs were present in the promoter of genes upregulated in the phenylpropanoid metabolic pathway. Because OsCCR1 and OsCAD2 showed higher expression in *rlm1-D* and its promoter contained an MYB1AT motif (**Supplementary Figures S6B,C**) and *OsCAD2* was responsible for rice lignin synthesis (Ookawa et al., 2014). We selected *OsCAD2* gene for further study.

Y1H assays confirmed that RLM1 could recognize and bind to the promoter of *OsCAD2* (**Figures 5A,B**). A DNA-binding assay was performed to detect whether RLM1 can bind to the *OsCAD2* promoter region *in vitro*. The results revealed that RLM1 fused to a glutathione S-transferase (GST) tag could bind to the MYB1AT motif, while the binding was significantly weakened in a mutant MYB1AT motif (TAACCATG to TAATTATA; **Figure 5C**). Next, EMSA experiment was used to judge the RLM1 binding to the promoter of OsCAD2. We incubated GST-RLM1 proteins together with a biotin-labeled probe. The protein-DNA complex showed an obvious band shift, but the mutant probe did not (**Figure 5D**). Thus, these results verified that RLM1 could bind to the *OsCAD2* promoter containing the MYB1AT motif. In addition, a luciferase (LUC) activity assay was used to determine whether RLM1 can activate *OsCAD2*. In rice protoplasts cotransfected with the effector and reporter vectors, the ratio of firefly LUC to Renilla luciferase (REN) of the effector OsCAD2 pro-LUC was fourfold higher than that of the empty vector control (**Figure 5E**).

RLM1 Interaction Partners

To clarify the regulatory network of RLM1, a yeast library screening assay was conducted. First, a self-activation test of RLM1-pGBKT7 was performed. We found that RLM1-pGBKT7 could not grow in selective media (SD/-Trp-His-Leu + 60 mM 3-AT). Then, the yeast screening library assay was carried out with a concentration of 60 mM 3-AT, and multiple clones were obtained by screening and sequencing. OsMAPK10 was repeated at least five times (**Figure 6A**). Through the RiceXPro database, we found that OsMAPK10 had a similar expression pattern as RLM1 (**Supplementary Figures S8A,B**). We subsequently carried out a follow-up study of OsMAPK10.

Three experiments were applied to verify the relationship between OsMAPK10 and RLM1. First, the full-length OsMAPK10 coding sequence was amplified and fused to pGADT7, and the yeast two-hybrid (Y2H) assay showed that OsMAPK10 interacted with RLM1-pGBKT7 in selective media (SD/-Trp-His-Leu + 60 mM 3AT; **Figure 6A**). Second, we used the split-LUC complementation experiment to judge the interaction of both. The pCAMBIA1300-nLUC (RLM1-nLUC), and pCAMBIA1300-cLUC (OsMAPK10-cLUC) were transformed into *Agrobacterium* GV3101, respectively. The different combinations were injected into tobacco, and LUC fluorescence was detected. The results suggested that RLM1-nLUC and MAPK10-cLUC could interact, causing LUC to be expressed and fluoresce (**Figure 6B**). Third, the pull-down assay also confirmed the interaction of OsMAPK10 and RLM1. The GST-RLM1 and GST proteins were incubated with beads bound to His-MAPK10, washed several times after

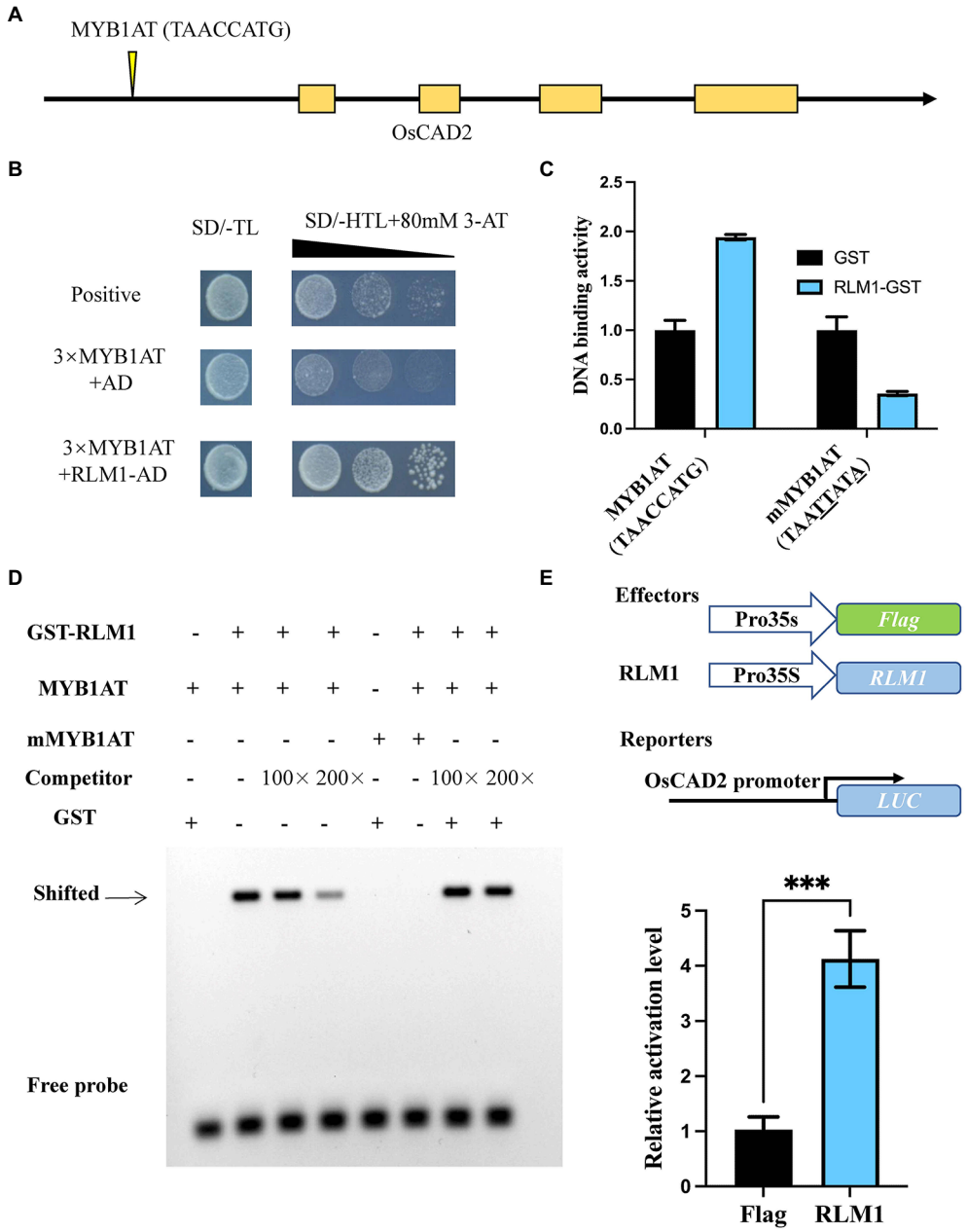


FIGURE 5 | RLM1 binds to the motif of OsCAD2 and activates its expression. **(A)** The MYB1AT motif box is located approximately 1,931–1,938 bp upstream of the *OsCAD2* start codon. **(B)** A Y1H assay was used to test RLM1 binding to the MYB1AT motif. An AD-only vector was used as a control. **(C)** DNA-protein binding assay revealing that RLM1 binds to the MYB1AT motif in the *OsCAD2* promoter. MYB1AT, TAACCATG; mMYB1AT, TAATTATA. **(A)** Schematic diagram of *OsCAD2*. MYB1AT (TAACCATG) indicates the position in the *OsCAD2* promoter. **(B)** Y1H assays show that RLM1 can bind to the MYB1AT motif, compared with the binding of the negative and positive controls. **(C)** DNA-binding activity of each group. GST was used as a negative control. The data are presented as the means ± SDs ($n = 3$). **(D)** EMSA shows that the GST-RLM1 recombinant protein directly binds to the biotin-labeled probe of the *OsCAD2* promoter fragment. The signal in the RLM1 group and the MYB1AT motif in the *OsCAD2* promoter shifted, but no signal shift was found in the RLM1 or mutant MYB1AT group. The shifted signal became increasingly weaker and eventually was no longer detected when unlabeled competitive probes were added. MYB1AT, TAACCATG; mMYB1AT, TAATTATA. **(E)** RLM1 induces proOsCAD2::LUC signaling. LUC activity was detected in rice protoplasts. The expression of REN was used as an internal control. The LUC/REN ratio indicates the expression level of the promoter ($n = 3$). ***$p < 0.001$ (Student's *t*-test).

incubation, denatured, detected *via* western blotting, and hybridized to anti-GST and anti-RLM1 antibodies. We found that GST-RLM1 proteins could be detected, but GST proteins could not be (**Figure 6C**). Taken together, the results of the three above mentioned experiments proved that RLM1 interacts with OsMAPK10.

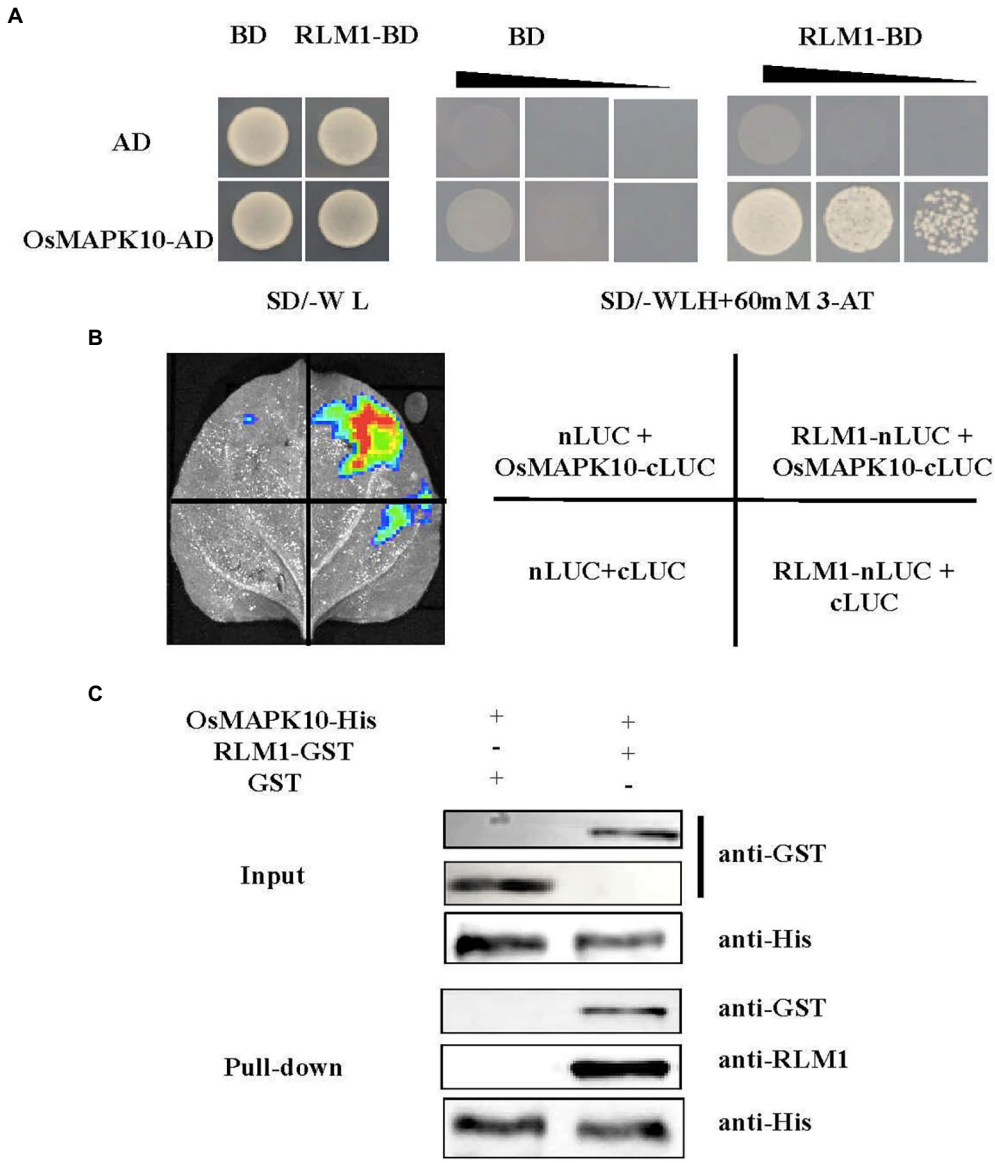


FIGURE 6 | RLM1 physically interacts with OsMAPK10. **(A)** A Y2H interaction assay was performed between RLM1 and OsMAPK10. A pGADT7 vector (AD) was used as the negative control. RLM1 was fused to the pGBKT7 vector (RLM1-BD). OsMAPK10 proteins were fused to the pGADT7 vector (OsMAPK10-AD). **(B)** The full OsMAPK10 protein was fused to the pCAMBIA1300-cLUC vector (OsMAPK10-cLUC). The full RLM1 protein was fused to pCAMBIA1300-nLUC (RLM1-nLUC). Split-LUC complementation assay showing that RLM1 can interact with OsMAPK10 in cells of *N. benthamiana* leaves. The LUC signals were not detected in the corresponding negative controls. **(C)** Pull-Down assays of RLM1 with OsMAPK10 *in vitro*. RLM1-GST was incubated together with OsMAPK10-His beads, and GST proteins were used as controls. Western bolt was detected *via* anti-GST, anti-RLM1, and anti-his antibodies, respectively. Compared with the input and Pull-Down band results, the results showed that GST did not interact with OsMAPK10, but RLM1 interacted with OsMAPK10.

Generation of OsMAPK10 Knockdown Lines and Overexpression Lines

To further study whether *OsMAPK10* participates in leaf morphogenesis, *OsMAPK10* knockout lines were created *via* gene-editing technology. The mutation of knockout line 1 (*OsMAPK10-m1*) occurred in the second exon, and 10 bp was deleted. The mutation of knockout line 2 (*OsMAPK10-m2*) occurred in the second exon, and 1 bp was deleted (**Supplementary Figure S9A**). Both mutations resulted in frameshift mutations (**Supplementary Figure S9B**). qRT-PCR analysis demonstrated that the expression level of *OsMAPK10* reached a lower level than it did in the WT (**Supplementary Figure S9D**). The *OsMAPK10-m1* and *OsMAPK10-m2* lines were considered knockdown mutants. The *OsMAPK10-m1* and *OsMAPK10-m2* lines had a dwarf-type phenotype in the field (**Supplementary Figure S9C**). There were lower lignin contents in the leaves and stems in the *OsMAPK10-m1* and *OsMAPK10-m2* lines compared to the WT

(**Supplementary Figures S9E–H**). Overexpression of *OsMAPK10* driven by CaMV35S resulted in increased lignin contents in the leaves and stems of OE1 and OE2, similar to those of *rlm1-D* (**Supplementary Figures S10A–G**). The knockdown and overexpression experiments of *OsMAPK10* showed that OsMAPK10 may directly participate in SCW deposition.

Relationship Between OsMAPK10 and RLM1

Three experiments were performed to study the relationship between OsMAPK10 and RLM1. MAPK family proteins mainly function by phosphorylating interacting partners. To judge whether OsMAPK10 phosphorylates RLM1, His-OsMAPK10 and GST-RLM1 fusion constructs were transformed into the *E. coli* Rosetta strain (DE3). Both fusion proteins were incubated and hybridized with GST antibody with or without the addition of a Phos-tag. The results demonstrated that when Phos-tag was added to the SDS-PAGE gel, the hybridization band was lagged without the addition of the Phos-tag (**Supplementary Figure S11A**). These results indicated that RLM1 could be phosphorylated by OsMAPK10. A LUC activity assay was applied to detect whether OsMAPK10 can facilitate RLM1 activation of downstream target genes. In rice protoplasts cotransfected with the effector and reporter vectors, the ratio of LUC to REN of the effector pOsCAD2-LUC was threefold higher than that of the empty vector control. While OsMAPK10 and RLM1 were transferred into rice protoplast together, the ratio of LUC to REN of the effector pOsCAD2-LUC was sixfold higher than that of the empty vector control (**Supplementary Figure S11B**). The above experiments preliminarily suggested that RLM1 is phosphorylated by OsMAPK10 and that phosphorylated RLM1 may activate the expression of downstream target genes.

RLM1 Knockdown Lines Have a Potential Effect on Rice Yield

To explore the function of *RLM1*, two *RLM1*-deletion mutant lines (*rlm1-m1* and *rlm1-m2*) were generated using CRISPR editing technology. The *rlm1-m1* line and *rlm1-m2* line contained a 69 bp deletion in the second exon of RLM1 (from 199 to 267 bp) and a 1 bp deletion in the second exon of RLM1 (at 203 bp), respectively (**Figures 7A,B**). The qRT-PCR of *rlm1-m1* and *rlm1-m2* plants showed that the expression level of *RLM1* was reduced significantly in the *rlm1-m1* and *rlm1-m2* plants, which indicated that both plants were knockdown mutants (**Figure 7D**). The leaf morphology phenotypes of the two knockdown lines were similar to that of the WT (**Figure 7C**). Because RLM1 was expressed in the stems only at maturity, SEM observation was applied, which revealed that the SCWs in the *rlm1-m1* and *rlm1-m2* stems were significantly thinner than those in the WT (**Figure 7K**). The result in the *rlm1-m1* and *rlm1-m2* stems was just contrary to the cytological observation of *RLM1* overexpression lines. The lignin content of the *rlm1-m1* and *rlm1-m2* stems also decreased slightly compared to that of the WT (**Figure 7G**). However, the cellulose and lignin contents of the leaves of the *rlm1-m1* and *rlm1-m2* plants did not change significantly (**Figures 7E,F,H**). Moreover, the *rlm1-m1* and *rlm1-m2* had more grains per panicle (**Figure 7I**), and the yield per plant increased by 11% compared with that of the WT (**Figure 7L**), whereas the 1,000 grain weight in *rlm1-m1* and *rlm1-m2* were similar to WT (**Figure 7J**).

DISCUSSION

Rice yield is mainly determined by leaf morphology and plant type. At present, many genes controlling leaf morphogenesis have been identified. However, how to understand the regulation relationship of leaf morphogenesis and plant type remains unclear. In this study, by screening a T-DNA insertion mutant population, we obtained a rolling-leaf mutant, *rlm1-D*, and cloned the *RLM1* gene *via* the map-based cloning method. *RLM1* encodes a typical R2R3 MYB transcription factor, and we proposed a preliminary a regulatory network based on cytological, biochemical, and genetic evidence. First, most of the reported genes, such as *LC2* (Zhao et al., 2010), *OsHB1* (Itoh et al., 2008), *ACL1/ACL2* (Li et al., 2010), *BRD1* (Hong et al., 2002), *CLD1/SRL1* (Li et al., 2017), *YAB1* (Dai et al., 2007), *OsCSLD4/NRL1* (Hu et al., 2010), and *OsAGO7* (Shi et al., 2007), directly regulate the development of bulliform cells and promote leaf rolling. A series of experiments, such as those involving paraffin sectioning, SEM observations, and measurements of cellulose and lignin contents, confirmed that the abnormal SCW disposition in *rlm1-D* plants was the direct cause of the *rlm1-D* rolling-leaf and dwarf-type phenotypes, and the reduction in the number and size of bulliform cells in *rlm1-D* was a secondary effect; these mechanisms were different from that governing the rolling-leaf phenotype. This mechanism controlling leaf morphology in *rlm1-D* was similar to that involving *RL14* (Fang et al., 2012) and *ROC8* (Sun et al., 2020). Second, multiple conclusions supported the *rlm1-D* was caused by the ectopic expression of *RLM1*. The expression level analysis showed that *RLM1* is highly expressed in all parts of *rlm1-D* mutant (**Supplementary Figure S12**). Overexpression of *RLM1* in *Nipponbare* derived from CaMV35S reproduces the rolling-leaf and dwarf-type phenotype, similar to the *rlm1-D* mutant phenotype. There was a dose-dependent effect on RLM1 expression due to heterologous activation of RLM1. Third, the biochemical analysis of RLM1 also confirmed that RLM1 had transcriptional activity, and this activity was determined by the CTD motif. Further experiments involving the deletion of the CTD motif determined that 18 aa was important for RLM1 transcriptional activity. Fourth, transcriptome analysis of *rlm1-D* and WT plants confirmed that RLM1 regulated the phenylpropanoid pathway. qRT-PCR, Y1H assay, and electrophoretic mobility shift assay (EMSA) experiments determined that RLM1 targets the promoter of OsCAD2 responsible for lignin synthesis by binding the MYB1AT motif. Fifth, Expression pattern analysis at the young seedling and mature stages suggested that RLM1 was mainly expressed in young seedlings and in mature stems, not in the other tissues of mature plants. The temporal and spatial expression

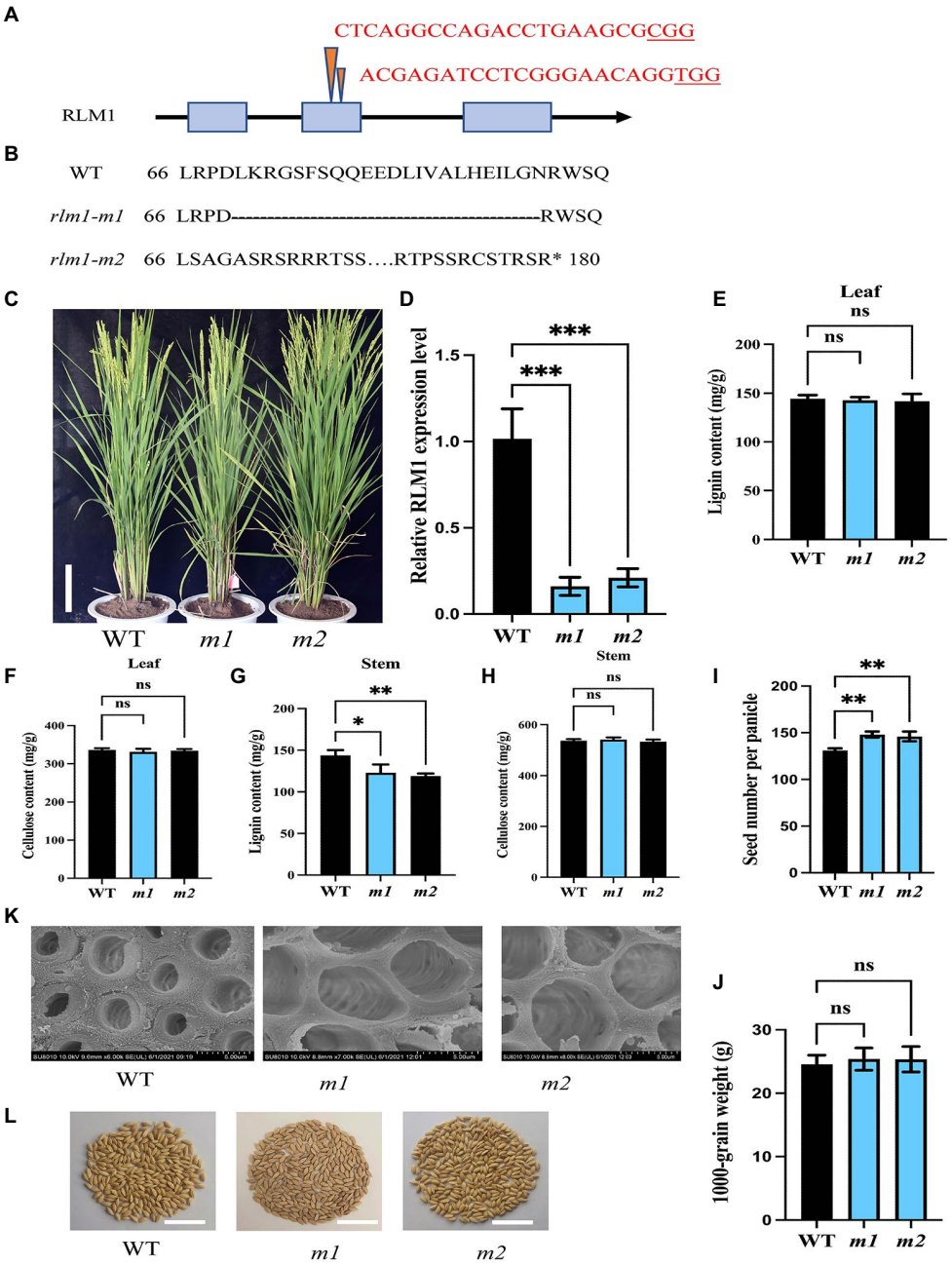


FIGURE 7 | CRISPR-Cas9-induced mutations in RLM1. **(A)** Schematic map of the genomic region of RLM1 and the sgRNA target site. The arrow shows the sgRNA target site of the RLM1 genomic sequence, and the PAM motif (NGG) is shown in red. The blue boxes represent RLM1 exons, and the black lines indicate intron sequences. **(B)** Amino acid alignment surrounding the sgRNA target region showing the predicted peptide sequence of the WT and mutant alleles. The site of the predicted frame shifted sequence is underlined, and new stop codons are indicated by asterisks. **(C)** Phenotypes of the WT and homozygous mutants *rlm1-m1* and indicated *rlm1-m2* at the heading stage (Bar = 5 cm). **(D)** qRT-PCR analysis showing RLM1 expression levels in WT plants and *rlm1-m1* and *rlm1-m2* homozygous mutant plants. *OsActin1* was selected as an internal reference. **(E)** Lignin content measurements in WT, *rlm1-m1,* and *rlm1-m2* leaf tissue. **(F)** Cellulose content measurements in WT, *rlm1-m1,* and *rlm1-m2* leaf tissue. **(G)** Lignin content measurements in the second internodes from 3-month-old WT, *rlm1-m1*, and *rlm1-m2* plants. **(H)** Cellulose content measurements in the second internodes from 3-month-old WT, *rlm1-m1*, and *rlm1-m2* plants. **(I)** Comparison of seed number per panicle among WT, *rlm1-m1*, and *rlm1-m2* plants ($n = 15$). **(K)** SEM observations of sclerenchyma cell walls in the internodes from the 3-month-old WT, *rlm1-m1,* and *rlm1-m2 plants*. Bar = 5 μm. **(J)** Comparison of 1,000-grain weight among WT, *rlm1-m1*, and *rlm1-m2* plants ($n = 15$). **(L)** Seeds per panicle of WT, *rlm1-m1,* and *rlm1-m2*. Bar = 5 cm. The bars represent the SDs of the means. The student's *t*-test was performed to determine the significance: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$ (Student's *t*-test).

specificity of RLM1 determines the role of RLM1 in leaf morphology and plant type. Totally, the above five aspects support that *RLM1* influences SCW deposition by regulating lignin synthesis in leaves and stems in rice.

The components of the cell wall include cellulose, hemicellulose, and lignin, which provide support and defensive ability to plants. Together, the components of the cell wall also play the most important role in terms of biomass energy, degradation, and the transformation of straw. Therefore, the pyramidal hierarchy of SCW disposition regulation was determined in model plant species. NACs are the top-layer transcription factors for SCW deposition, and MYB transcription factors constitute the most important hub (Wang et al., 2021). Many MYBs in rice have been found to be involved in regulating the development of the SCW. *OsMYB103L* mediates cellulose biosynthesis and secondary wall formation in leaves and stems mainly by directly binding the promoters of *CESA4*, *CESA7*, *CESA9*, and *BC1* and regulating their expression (Ye et al., 2015). NAC29/31 directly regulates *OsMYB61*, which in turn activates CESA expression (Ye et al., 2018). However, the function of *RLM1* was different from that previously reported for *OsMYB61* and *OsMYB103L*. We found that *RLM1* regulates SCW deposition by regulating lignin synthesis. Lignin synthesis is specifically initiated in the cells that form SCWs, providing mechanical support for the upright growth of plants, helping to establish long-distance transport channels for water and materials, and strengthening the stress resistance barrier of plants. Notably, AtMYB4 is highly homologous to RLM1. AtMYB4 was shown to regulate the accumulation of the UV protectant compound sinapoyl malate through its ability to repress the transcription of the gene encoding the phenylpropanoid pathway enzyme cinnamate 4-hydroxylase (Jin et al., 2000). RLM1 acts as an activator, and AtMYB4 acts as a repressor. The opposite function of the two depends on 18 aa.

As components of complexes, MAPKs and MYBs participate in different biological development processes in plants. Mitogen-activated protein kinase 6 negatively regulates SCW biosynthesis by modulating MYB46 protein stability in *Arabidopsis* (Im et al., 2021). MYB75 phosphorylation by MPK4 is required for light-induced anthocyanin accumulation in *Arabidopsis* (Li et al., 2016). However, the biological functions of MAPK and MYB complexes remain unknown in rice. OsMAPK10 regulates RLM1 based on the following considerations. Like that of RLM1, the expression of OsMAPK10 occurs predominantly in the stems. Y2H, split-LUC complementation, and pull-down assays confirmed the interaction between OsMAPK10 and RLM1. OsMAPK10 can phosphorylate RLM1. LUC activity assays revealed that OsMAPK10 facilitates the activation of downstream target genes by RLM1. We will study the specific phosphorylation sites in the future. Nevertheless, the knockdown lines of OsMAPK10 displayed a semi dwarf-type phenotype, and knockdown lines of RLM1 showed a normal-leaf phenotype, indicating that OsMAPK10 is perhaps involved in other plant development processes. The above mentioned findings suggest that a OsMAPK10-RLM1-OsCAD2 regulatory network controls leaf morphogenesis and plant type. However, the RLM1 phosphorylation sites modified by OsMAPK10 still need to be identified.

Potential Application of RLM1

Temporal and spatial expression specificity of RLM1 indicates that *RLM1* has potential application in production. In the *RLM1* knockout lines, the morphological development of the leaves is not affected, but the SCW deposition of stems is modified. The aperture around stem SCW became bigger and was conducive to more water and mineral transport. Thus, there is reasonable to increase the productivity in *RLM1* knockout lines. RLM1 perhaps coordinates the balance of C and N, which is very similar to OsMYB61 (Gao et al., 2020). *RLM1* is highly expressed in the stems, and ectopic expression of *RLM1* affects both leaf morphology and plant height, indicating that *RLM1* perhaps participates in the gibberellin (GA) biosynthetic or signaling pathway. GA acts as the key hub between nitrogen and carbon (Gao et al., 2020), and *RLM1* may coordinate the nitrogen and carbon balance in plants.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA813008.

AUTHOR CONTRIBUTIONS

ZC and ST conceived the main study. ZZ and YC obtained the mutant and constructed the genetic population. PA and JW constructed the genetic transformation. XC, XS, and DL constructed the expression analysis and sub-cell localization. LZ constructed the yeast interaction. ZZ, XC, and ZC contributed to the data analysis. ZZ and TL designed and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Key Research and Development Program of China (2020YFA0907603 and 2020YFE0202300), NSFC (32001532), NSFC-CGIAR (31861143006), the Agricultural Science and Technology Innovation Program for TL, XC, and ZZ support, and the Fundamental Research Funds for Central Non-Profit of Institute of Crop Sciences for TL, XC, and ZZ support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.905111/full#supplementary-material

Supplementary Table S1 | Up regulated genes and down regulated genes in WT and *rlm1-D* lists.

Supplementary Table S2 | List of RLM1 binding motifs.

REFERENCES

Cui, X., Zhang, Z., Wang, Y., Wu, J., Han, X., Gu, X., et al. (2019). TWI1 regulates cell-to-cell movement of OSH15 to control leaf cell fate. *New Phytol.* 221, 326–340. doi: 10.1111/nph.15390

Dai, M., Hu, Y., Zhao, Y., Liu, H., and Zhou, D. X. (2007). A WUSCHEL-LIKE HOMEOBOX gene represses a YABBY gene expression required for rice leaf development. *Plant Physiol.* 144, 380–390. doi: 10.1104/pp.107.095737

Emery, J. F., Floyd, S. K., Alvarez, J., Eshed, Y., Hawker, N. P., Izhaki, A., et al. (2003). Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. *Curr. Biol.* 13, 1768–1774. doi: 10.1016/j.cub.2003.09.035

Fang, L., Zhao, F., Cong, Y., Sang, X., Du, Q., Wang, D., et al. (2012). Rolling-leaf14 is a 2OG-Fe (II) oxygenase family protein that modulates rice leaf rolling by affecting secondary cell wall formation in leaves. *Plant Biotechnol. J.* 10, 524–532. doi: 10.1111/j.1467-7652.2012.00679.x

Gao, Y., Xu, Z., Zhang, L., Li, S., Wang, S., Yang, H., et al. (2020). MYB61 is regulated by GRF4 and promotes nitrogen utilization and biomass production in rice. *Nat. Commun.* 11:5219. doi: 10.1038/s41467-020-19019-x

Ha, C. M., Jun, J. H., and Fletcher, J. C. (2010). Control of Arabidopsis leaf morphogenesis through regulation of the YABBY and KNOX families of transcription factors. *Genetics* 186, 197–206. doi: 10.1534/genetics.110.118703

Hong, Z., Ueguchi-Tanaka, M., Shimizu-Sato, S., Inukai, Y., Fujioka, S., Shimada, Y., et al. (2002). Loss-of-function of a rice brassinosteroid biosynthetic enzyme, C-6 oxidase, prevents the organized arrangement and polar elongation of cells in the leaves and stem. *Plant J.* 32, 495–508. doi: 10.1046/j.1365-313x.2002.01438.x

Hu, J., Zhu, L., Zeng, D., Gao, Z., Guo, L., Fang, Y., et al. (2010). Identification and characterization of NARROW AND ROLLED LEAF 1, a novel gene regulating LEAF morphology and plant architecture in rice. *Plant Mol. Biol.* 73, 283–292. doi: 10.1007/s11103-010-9614-7

Hunter, C., Willmann, M. R., Wu, G., Yoshikawa, M., De La Luz Gutiérrez-Nava, M., and Poethig, S. R. (2006). Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. *Development* 133, 2973–2981. doi: 10.1242/dev.02491

Im, J. H., Ko, J. H., Kim, W. C., Crain, B., Keathley, D., and Han, K. H. (2021). Mitogen-activated protein kinase 6 negatively regulates secondary wall biosynthesis by modulating MYB46 protein stability in Arabidopsis thaliana. *PLoS Genet.* 17:e1009510. doi: 10.1371/journal.pgen.1009510

Itoh, J., Hibara, K., Sato, Y., and Nagato, Y. (2008). Developmental role and auxin responsiveness of class III homeodomain leucine zipper gene family members in rice. *Plant Physiol.* 147, 1960–1975. doi: 10.1104/pp.108.118679

Ji, X., Wang, L., Nie, X., He, L., Zang, D., Liu, Y., et al. (2014). A novel method to identify the DNA motifs recognized by a defined transcription factor. *Plant Mol. Biol.* 86, 367–380. doi: 10.1007/s11103-014-0234-5

Ji, X., Wang, L., Zang, D., and Wang, Y. (2018). "Transcription factor-centered yeast one-hybrid assay," in *Two-Hybrid Systems: Methods and Protocols*. ed. L. Oñate-Sánchez (New York, NY: Springer New York).

Jiang, C. K., and Rao, G. Y. (2020). Insights into the diversification and evolution of R2R3-MYB transcription factors in plants. *Plant Physiol.* 183, 637–655. doi: 10.1104/pp.19.01082

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

Li, L., Shi, Z. Y., Li, L., Shen, G. Z., Wang, X. Q., An, L. S., et al. (2010). Overexpression of ACL1 (abaxially curled leaf 1) increased bulliform cells and induced abaxial curling of leaf blades in rice. *Mol. Plant* 3, 807–817. doi: 10.1093/mp/ssq022

Li, S., Wang, W., Gao, J., Yin, K., Wang, R., Wang, C., et al. (2016). MYB75 phosphorylation by MPK4 is required for light-induced anthocyanin accumulation in Arabidopsis. *Plant Cell* 28, 2866–2883. doi: 10.1105/tpc.16.00130

Li, W. Q., Zhang, M. J., Gan, P. F., Qiao, L., Yang, S. Q., Miao, H., et al. (2017). CLD1/SRL1 modulates leaf rolling by affecting cell wall formation, epidermis integrity and water homeostasis in rice. *Plant J.* 92, 904–923. doi: 10.1111/tpj.13728

Meng, Y., Wang, Z., Wang, Y., Wang, C., Zhu, B., Liu, H., et al. (2019). The MYB activator WHITE PETAL1 associates with MtTT8 and MtWD40-1 to regulate carotenoid-derived flower pigmentation in Medicago truncatula. *Plant Cell* 31, 2751–2767. doi: 10.1105/tpc.19.00480

Merelo, P., Ram, H., Pia Caggiano, M., Ohno, C., Ott, F., Straub, D., et al. (2016). Regulation of MIR165/166 by class II and class III homeodomain leucine zipper proteins establishes leaf polarity. *Proc. Natl. Acad. Sci. U. S. A.* 113, 11973–11978. doi: 10.1073/pnas.1516110113

Ookawa, T., Inoue, K., Matsuoka, M., Ebitani, T., Takarada, T., Yamamoto, T., et al. (2014). Increased lodging resistance in long-culm, low-lignin gh2 rice for improved feed and bioenergy production. *Sci. Rep.* 4:6567. doi: 10.1038/srep06567

Shi, Z., Wang, J., Wan, X., Shen, G., Wang, X., and Zhang, J. (2007). Over-expression of rice OsAGO7 gene induces upward curling of the leaf blade that enhanced erect-leaf habit. *Planta* 226, 99–108. doi: 10.1007/s00425-006-0472-0

Sun, J., Cui, X., Teng, S., Kunnong, Z., Wang, Y., Chen, Z., et al. (2020). HD-ZIP IV gene Roc8 regulates the size of bulliform cells and lignin content in rice. *Plant Biotechnol. J.* 18, 2559–2572. doi: 10.1111/pbi.13435

Wan, S., Wu, J., Zhang, Z., Sun, X., Lv, Y., Gao, C., et al. (2009). Activation tagging, an efficient tool for functional analysis of the rice genome. *Plant Mol. Biol.* 69, 69–80. doi: 10.1007/s11103-008-9406-5

Wang, Y., Sun, J., Deng, C., Teng, S., Chen, G., Chen, Z., et al. (2022). Plasma membrane-localized SEM1 protein mediates sugar movement to sink rice tissues. *Plant J.* 109, 523–540. doi: 10.1111/tpj.15573

Wang, Y., Yu, W., Ran, L., Chen, Z., Wang, C., Dou, Y., et al. (2021). DELLA-NAC interactions mediate GA signaling to promote secondary Cell Wall formation in cotton stem. *Front. Plant Sci.* 12:655127. doi: 10.3389/fpls.2021.655127

Wang, R., Zhang, Y., Qian, L., and Yu, J. (2003). Photooxidation characteristics of super-hybrid rice "Liangyoupeijiu" and its parents. *Ying Yong Sheng Tai Xue Bao* 14, 1309–1312.

Wu, R., Li, S., He, S., Wassmann, F., Yu, C., Qin, G., et al. (2011). CFL1, a WW domain protein, regulates cuticle development by modulating the function of HDG1, a class IV homeodomain transcription factor, in rice and Arabidopsis. *Plant Cell* 23, 3392–3411. doi: 10.1105/tpc.111.088625

Ye, Y., Liu, B., Zhao, M., Wu, K., Cheng, W., Chen, X., et al. (2015). CEF1/OsMYB103L is involved in GA-mediated regulation of secondary wall biosynthesis in rice. *Plant Mol. Biol.* 89, 385–401. doi: 10.1007/s11103-015-0376-0

Ye, Y., Wu, K., Chen, J., Liu, Q., Wu, Y., Liu, B., et al. (2018). OsSND2, a NAC family transcription factor, is involved in secondary cell wall biosynthesis through regulating MYBs expression in rice. *Rice* 11:36. doi: 10.1186/s12284-018-0228-z

Zhang, G., Hou, X., Wang, L., Xu, J., Chen, J., Fu, X., et al. (2021). PHOTOSENSITIVE LEAF ROLLING 1 encodes a polygalacturonase that modifies cell wall structure and drought tolerance in rice. *New Phytol.* 229, 890–901. doi: 10.1111/nph.16899

Zhang, G. H., Xu, Q., Zhu, X. D., Qian, Q., and Xue, H. W. (2009). SHALLOT-LIKE1 is a KANADI transcription factor that modulates rice leaf rolling by regulating leaf abaxial cell development. *Plant Cell* 21, 719–735. doi: 10.1105/tpc.108.061457

Zhao, S. Q., Hu, J., Guo, L. B., Qian, Q., and Xue, H. W. (2010). Rice leaf inclination2, a VIN3-like protein, regulates leaf angle through modulating cell division of the collar. *Cell Res.* 20, 935–947. doi: 10.1038/cr.2010.109

Zou, L. P., Sun, X. H., Zhang, Z. G., Liu, P., Wu, J. X., Tian, C. J., et al. (2011). Leaf rolling controlled by the homeodomain leucine zipper class IV gene Roc5 in rice. *Plant Physiol.* 156, 1589–1602. doi: 10.1104/pp.111.176016

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 | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 July 2022
doi: 10.3389/fpls.2022.943349

Engineering Curcumin Biosynthesis in Poplar Affects Lignification and Biomass Yield

*Barbara De Meester[1,2], Paula Oyarce[1,2], Ruben Vanholme[1,2], Rebecca Van Acker[1,2], Yukiko Tsuji[3,4], Thijs Vangeel[5], Sander Van den Bosch[5], Jan Van Doorsselaere[6], Bert Sels[5], John Ralph[3,4] and Wout Boerjan[1,2]**

[1] Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, [2] VIB Center for Plant Systems Biology, Ghent, Belgium, [3] Department of Biochemistry, University of Wisconsin, Madison, WI, United States, [4] US Department of Energy, Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, Madison, WI, United States, [5] Center for Sustainable Catalysis and Engineering, KU Leuven, Leuven, Belgium, [6] VIVES, Roeselare, Belgium

OPEN ACCESS

Edited by:
Wagner Rodrigo De Souza, Federal University of ABC, Brazil

Reviewed by:
Yunjun Zhao, Center for Excellence in Molecular Plant Sciences (CAS), China
An-Shan Hsiao, University of Exeter, United Kingdom

***Correspondence:**
Wout Boerjan
woboe@psb.ugent.be

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

Received: *13 May 2022*
Accepted: *14 June 2022*
Published: *04 July 2022*

Citation:
De Meester B, Oyarce P, Vanholme R, Van Acker R, Tsuji Y, Vangeel T, Van den Bosch S, Van Doorsselaere J, Sels B, Ralph J and Boerjan W (2022) Engineering Curcumin Biosynthesis in Poplar Affects Lignification and Biomass Yield. Front. Plant Sci. 13:943349. doi: 10.3389/fpls.2022.943349

Lignocellulosic biomass is recalcitrant toward deconstruction into simple sugars mainly due to the presence of lignin. By engineering plants to partially replace traditional lignin monomers with alternative ones, lignin degradability and extractability can be enhanced. Previously, the alternative monomer curcumin has been successfully produced and incorporated into lignified cell walls of Arabidopsis by the heterologous expression of *DIKETIDE-CoA SYNTHASE* (*DCS*) and *CURCUMIN SYNTHASE2* (*CURS2*). The resulting transgenic plants did not suffer from yield penalties and had an increased saccharification yield after alkaline pretreatment. Here, we translated this strategy into the bio-energy crop poplar. Via the heterologous expression of *DCS* and *CURS2* under the control of the secondary cell wall *CELLULOSE SYNTHASE A8-B* promoter (*ProCesA8-B*), curcumin was also produced and incorporated into the lignified cell walls of poplar. *ProCesA8-B:DCS_CURS2* transgenic poplars, however, suffered from shoot-tip necrosis and yield penalties. Compared to that of the wild-type (WT), the wood of transgenic poplars had 21% less cellulose, 28% more matrix polysaccharides, 23% more lignin and a significantly altered lignin composition. More specifically, *ProCesA8-B:DCS_CURS2* lignin had a reduced syringyl/guaiacyl unit (S/G) ratio, an increased frequency of *p*-hydroxyphenyl (H) units, a decreased frequency of *p*-hydroxybenzoates and a higher fraction of phenylcoumaran units. Without, or with alkaline or hot water pretreatment, the saccharification efficiency of the transgenic lines was equal to that of the WT. These differences in (growth) phenotype illustrate that translational research in crops is essential to assess the value of an engineering strategy for applications. Further fine-tuning of this research strategy (e.g., by using more specific promoters or by translating this strategy to other crops such as maize) might lead to transgenic bio-energy crops with cell walls more amenable to deconstruction without settling in yield.

Keywords: lignin, lignin engineering, alternative lignin monomers, poplar, curcumin, translational research

INTRODUCTION

Lignocellulosic biomass, which is largely composed of plant cell walls, is a promising renewable feedstock for the production of biofuels and bio-based materials (Vanholme et al., 2013a). The polysaccharides present in the plant cell wall, cellulose and hemicelluloses, can be hydrolyzed into fermentable primary sugars by a process called saccharification. These monomers can be further processed into bulk chemicals, such as bioethanol, levulinic acid, and 2,5-furandicarboxylic acid (Bozell et al., 2007; Corma et al., 2007; Vanholme et al., 2013a; Isikgor and Becer, 2015; de Vries et al., 2021b). However, the saccharification process is hindered by the presence of lignin. This complex aromatic heteropolymer is crucial to plant growth and development; it provides rigidity to plant cells, provides imperviousness to water conducting cells, and plays a role in defense to pathogens. The lignin polymer is mainly derived from the monolignols coniferyl and sinapyl alcohol and low levels of *p*-coumaryl alcohol. After polymerization in the cell wall, the monolignols produce guaiacyl (G), syringyl (S), and *p*-hydroxyphenyl (H) units, respectively (Vanholme et al., 2019). To make the polysaccharides accessible to the hydrolytic enzymes, however, lignocellulosic biomass needs to be pretreated with heat and/or (physico)chemical methods. These pretreatments are costly and energy-demanding. There is therefore interest in engineering lignin of biomass crops to facilitate cell wall deconstruction (Sederoff et al., 1999; Boerjan et al., 2003; Ralph et al., 2004b, 2019; Chen and Dixon, 2007; Vanholme et al., 2012; Mottiar et al., 2016; Chanoca et al., 2019; Halpin, 2019; Mahon and Mansfield, 2019; De Meester et al., 2020).

Studies of different naturally occurring or transgenic and mutant plants have led to the discovery that plants can tolerate large shifts in lignin composition, often without visible effects on plant development and morphology (Meyer et al., 1998; Marita et al., 1999; Franke et al., 2000; Stewart et al., 2009; Eudes et al., 2017; de Vries et al., 2018). This notion has steered research toward the biosynthesis and incorporation of alternative monomers into the lignin (Grabber et al., 2008, 2010, 2019; Vanholme et al., 2012; Smith et al., 2015, 2022; Mottiar et al., 2016; del Río et al., 2022). Enhancing the incorporation of atypical monomers that are rare, or even absent, in the lignin of wild-type (WT) plants could permit a more efficient conversion of lignocellulosic biomass for the industry while maintaining the biological role of lignin. Over the past years, several studies have evaluated the incorporation of alternative monomers into the lignin polymer of several plant species. For example, overexpression of the gene encoding the enzyme FERULOYL-CoA MONOLIGNOL TRANSFERASE in poplar has led to higher amounts of monolignol ferulates incorporated into the lignin, which considerably increased the saccharification efficiency and chemical pulping (Wilkerson et al., 2014; Kim et al., 2017; Zhou et al., 2017). Similarly, (enhanced) *p*-coumaroylation of Arabidopsis and poplar lignins was achieved by the expression of *p-COUMAROYL-CoA MONOLIGNOL TRANSFERASES* from *Brachypodium distachyon* or *Oryza sativa*, resulting in an increased saccharification efficiency without compromising the plant phenotype (Smith et al., 2015; Sibout et al., 2016; Lapierre et al., 2021). Expression of a bacterial *HYDROXYCINNAMOYL-COA HYDRATASE-LYASE* (*HCHL*) in Arabidopsis resulted in the incorporation of *p*-hydroxybenzaldehyde and *p*-hydroxybenzoate into the lignin (Eudes et al., 2012). The resulting HCHL-engineered plants had a lignin polymer with a reduced molecular weight; biomass from these plants was more easily saccharified after pretreatment. The combinatorial downregulation of *COMT* and overexpression of *F5H* in Arabidopsis led to more than 30% of easily cleavable benzodioxane linkages in the lignin due to an increased incorporation of 5-hydroxyconiferyl alcohol (Vanholme et al., 2010; Weng et al., 2010). Disrupting *FLAVONE SYNTHASE II* (*FNSII*) in rice resulted in an altered cell wall, incorporating the intermediate flavanone naringenin, instead of tricin, into the lignin polymer (Lam et al., 2017). The *fnsII* mutant plants showed a normal growth and an enhanced saccharification efficiency. The incorporation of naringenin into poplar lignin was achieved by expressing an apple *CHALCONE SYNTHASE* 3 (Mahon et al., 2022). Similarly to the engineered rice plants, biomass from the transgenic trees had increased saccharification yields from plants that grew normally. Finally, heterologous expression of a gene encoding a bacterial 3-DEHYDROSHIKIMATE DEHYDRATASE in hybrid poplar resulted in reduced lignin amounts, the incorporation of monolignol-3,4-dihydroxybenzoate conjugates in the lignin, and improved saccharification without affecting growth (Unda et al., 2022).

We previously engineered Arabidopsis to incorporate the alternative monomer curcumin into the lignin polymer (Oyarce et al., 2019). Curcumin, a metabolite naturally occurring in *Curcuma longa* (turmeric), has two phenolic rings linked by a seven carbon chain containing two conjugated carbonyl functionalities (**Figure 1A**; Chen and Tan, 1998). The conjugated 9-carbonyl function at the *para*-position of the two phenolic hydroxyl functions eases the cleavage of 8-*O*-4 inter-unit linkages under alkaline conditions (Oyarce et al., 2019). More specifically, the required hydrolysis temperature of the 8-*O*-4 linkage can be lowered from ∼170°C (for traditional lignin β-ether units) to ∼65°C (Criss et al., 1998; Imai et al., 2007; Tsuji et al., 2015; Mnich et al., 2017). In addition, the β-diketone functionality in the aliphatic chain of curcumin can be cleaved under alkaline conditions at temperatures as low as 30°C (Hauser et al., 1948; Pearson and Mayerle, 1951; Tønnesen and Karlsen, 1985; Price and Buescher, 1997; Wang et al., 1997; Tomren et al., 2007). In transgenic Arabidopsis, curcumin biosynthesis was achieved by the expression of the genes encoding the enzymes DIKETIDE-CoA SYNTHASE (DCS) and CURCUMIN SYNTHASE 2 (CURS2), that catalyze the conversion of one malonyl-CoA and two feruloyl-CoA molecules into curcumin (**Figure 1A**; Katsuyama et al., 2009; Oyarce et al., 2019). Using the secondary cell wall *CELLULOSE SYNTHASE A4* promoter (*ProCesA4*), the expression of the biosynthetic genes was restricted to lignified cells. The *ProCesA4:DCS_CURS2* Arabidopsis lines were successful at incorporating curcumin into the lignified cell wall without affecting the plant's biomass yield. Moreover, the transgenic Arabidopsis lines had an increased

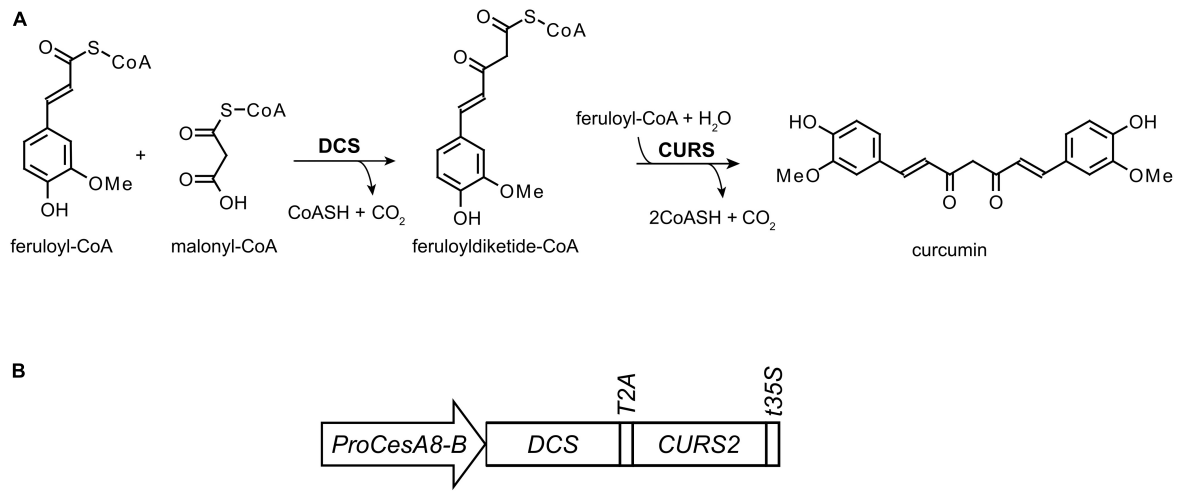


FIGURE 1 | Curcumin biosynthesis. **(A)** In turmeric, DIKETIDE-CoA SYNTHASE (DCS) and CURCUMIN SYNTHASE (CURS) convert two molecules of feruloyl-CoA and one molecule of malonyl-CoA into curcumin, which is characterized by two phenolic ring structures linked via a 7-carbon unsaturated aliphatic chain. **(B)** For the biosynthesis of curcumin in poplar, the coding sequences of *DCS* and *CURS2* genes were linked together by the *T2A* sequence, and put under control of the *CELLULOSE SYNTHASE 8-B* promoter (*ProCesA8-B*) and terminated with a *35S* terminator sequence (*t35S*).

saccharification efficiency after alkaline pretreatment. In this work, we investigated the translation of this strategy into the bio-energy crop poplar.

MATERIALS AND METHODS

Plant Material and Growth Conditions

The creation of the vector *pDONR221-DCS_CURS2* is described in Oyarce et al. (2019). The *CELLULOSE SYNTHASE 8-B* promoter (*ProCesA8-B*) from *P. trichocarpa* (Song et al., 2010) was synthesized by Genscript (Piscataway, NJ, United States) in a pUC57-Kan vector. Sequence information can be found in **Supplementary Table 1**. The expression vector *ProCesA8-B:DCS_CURS2* was created by an LR multisite II reaction with donor vectors *pUC57-ProCesA8B* and *pDONR221-DCS_CURS2* and destination vector *pALS7m24GW*. After confirmation by PCR, the expression vector was introduced into *Agrobacterium tumefaciens* strain C58C1 PMP90 by electroporation and positive colonies harboring the construct were selected by PCR. Agrobacterium-mediated transformation of *Populus tremula* × *P. alba* 717-B4 was performed according to Brasileiro et al. (1992).

In total, 28 independent transgenic lines (one individual plant for each line) and WT controls (eight plants) were transferred from *in vitro* culture to soil in pots of 5.5 cm diameter, placed in a tray filled with water and covered with a cage liner (Tecniplast APET disposable cage liner for cage body 1291H) for acclimatization. After 2 weeks, one side of the cage liner was lifted and kept accordingly for 1 day, after which the other side was also lifted, to gradually reduce humidity. After 2 weeks of growing in the greenhouse (16 h light/8 h dark photoperiod at ± 21°C), the plants were transferred to 10-L pots filled with a Saniflor® 691 commercial soil (Van Israel). After growing for 87 days in the greenhouse (first growth cycle), the *ProCesA8-B:DCS_CURS2* and WT poplars were harvested for further analysis. After harvesting, new shoots developed from the stool. After growing for another 4 months in the greenhouse (second growth cycle), the *ProCesA8-B:DCS_CURS2* and WT poplars were harvested for bright-field microscopy.

Harvesting of Plant Material

After 87 days of growth in soil in the greenhouse (first growth cycle), the WTs reached heights of approximately 130 cm. At this time point, all poplars were harvested by cutting the stem 10 cm above the base leaving 2–3 axillary buds to allow development of new shoots. The lower 10–15 cm part was debarked and stored in ethanol until future use for fluorescence microscopy. Subsequently, the top 20 cm of the stem was removed. The 6-cm stem part below the removed 20 cm top was debarked, frozen in liquid nitrogen, cut in about 30 pieces and stored at –70°C until further use for phenolic profiling. The leftover stem piece was debarked, air-dried and ground in a ball mill for catalytic hydrogenolysis, cell wall analysis and saccharification.

After 4 months of growth in soil in the greenhouse (second growth cycle), the WTs reached heights of approximately 2 m. At this time point, the poplars were harvested by cutting the main stem 10 cm above the base. The lower 10- to 15-cm part was debarked, stored in tap water and imaged by bright-field microscopy the same day.

Phenolic Profiling

Soluble phenolic compounds were extracted from approximately 100 mg of stem material (about 6 to 10 pieces of stem from the first growth cycle; see above) with 1 mL methanol at 70°C for 15 min under 1,000 rpm shaking. After centrifugation at room temperature and maximum speed, 800 μL of the supernatant was dried under vacuum and the pellet was resuspended in 100 μL

of cyclohexane and 100 μL of 1/1 DMSO/water (v/v). The tubes were vortexed and centrifuged at 14,000 rpm (20,000 x g) for 10 min, after which a 15-μL aliquot of the aqueous phase was injected on a Ultra-High-Performance Liquid Chromatography (UHPLC) system (Waters Acquity UPLC®) equipped with a BEH C18 column (2.1 × 150 mm, 1.7 μM, Waters) and hyphenated to a time-of-flight mass spectrometer (TOF MS, Synapt Q-Tof, Waters Corporation, Milford, Massachusetts, United States), using gradient elution. Gradient elution information, negative-ion mode mass spectrometry setting, chromatogram integration and alignment via Progenesis QI software (Waters) were performed as previously described (Eloy et al., 2017). Peak abundances were normalized to the dry weight (mg) of the pellet remaining after methanol extraction and drying on a SpeedVac. Principal Component Analysis (PCA) was performed in R software via the "prcomp" command.

Microscopy

For fluorescence microscopy, the stem piece (see above) from the first growth cycle was incubated for three days in 70% ethanol. Slices of 15 μm thick were made using a Reichert-Jung 2040 Autocut Microtome (Leica). Next, the slices were incubated for (i) 2 h in tap water or (ii) 2 h in acetone followed by 2 h in tap water to wash away the soluble compounds. The fluorescence of curcumin and lignin was visualized using the Zeiss LSM 780 microscope with an iLCI Plan-Neofluar 25x/0.8 Imm Korr DIC M27 objective (0.6 zoom). The fluorescence signal of curcumin was obtained at an excitation wavelength of 488 nm and emission wavelength between 490 and 578 nm. For lignin visualization, an excitation wavelength of 350 nm and emission wavelength between 407 and 479 nm was used.

For bright-field microscopy, the stem piece (see above) from the second growth cycle was cut into slices of 15 μm thick using a Reichert-Jung 2040 Autocut Microtome (Leica). Next, the slices were incubated for 2 h in acetone followed by 2 h in tap water. The sections were imaged with an Olympus BX51 microscope (Olympus) with an Olympus PlanC N 10x (0.25 NA) objective.

Catalytic Hydrogenolysis

Ground wood powder (see above) from the first growth cycle was used to prepare Cell Wall Residue (CWR) by sequentially washing the wood powder for 30 min each with milliQ water at 98°C, ethanol at 76°C, chloroform at 59°C and acetone at 54°C. Three hundred mg of CWR was incubated with 40 mg (5 wt% Pd) Pd/C catalyst and 40 mL methanol in a 100-mL Parr batch reactor at 250°C for 3 h with an initial H_2-pressure of 20 bar (at room temperature). After this reaction, the reactor was cooled and washed with acetone to collect all products. The reaction mixture was filtered and the filtrate was evaporated using a rotary evaporator. A part (roughly two thirds) of this product mixture was further separated into carbohydrate products and lignin products via liquid/liquid extraction: 2.5 mL of water was added and subsequently threefold extracted with 1 mL of dichloromethane (DCM). The DCM phases were combined and the DCM was evaporated using a N_2 flow, followed by oven drying at 80°C overnight, yielding a brown, viscous lignin oil. The resulting lignin oil samples were transferred quantitatively in 10-mL crimp cap vials, extracted in 150 μL methanol/150 μL milliQ water and centrifuged for 15 min at 19,757 × *g*. The supernatant (100 μL) was analyzed via UHPLC–MS following the parameters described by Vanholme et al. (2013b) using a gradient of two buffers [buffer A (99/1/0.1 H_2O/ACN/formic acid pH3), buffer B (99/1/0.1 ACN/H_2O/formic acid pH3)], but with the following modification: between 0 to 30 min from 95% A and 5% B to 50% A and 50% B, between 30 to 40 min from 50% A and 50% B to 0% A and 100% B. Chromatograms were visualized in Masslynx V4.1 software (Waters) and peaks were integrated using the default parameters (using automatic noise measurement and 2x smoothing).

Cell Wall Analysis

To determine the cellulose, matrix polysaccharide and lignin amount, ground wood powder (see above) from the first growth cycle was used for preparing CWR by sequentially washing for 30 min each with milliQ water at 98°C, ethanol at 76°C, chloroform at 59°C and acetone at 54°C. The remaining CWR was dried under vacuum and was determined gravimetrically (expressed as mass percentage of dry weight). To determine the crystalline cellulose level, the Updegraff method was used on 10 mg of CWR essentially as described by Updegraff (1969) and modified according to De Meester et al. (2020). The mass loss upon trifluoroacetic acid digestion was used to determine the matrix polysaccharide content (including mainly hemicelluloses, but also pectins and amorphous cellulose). Lignin content was determined by the acetyl bromide method on 5 mg of CWR essentially as described by Dence (1992) and modified according to Van Acker et al. (2013).

Lignin composition was determined by two-dimensional Heteronuclear Single-Quantum Coherence Nuclear Magnetic Resonance (2D HSQC NMR). Two hundred and fifty mg of ground wood powder (see above) from the first growth cycle was extracted three times with distilled water, three times with 80% aqueous ethanol, and once with acetone. The extracted samples were ball-milled using a Fritsch Planetary micro mill Pulverisette 7 vibrating at 600 rpm with zirconium dioxide (ZrO_2) vessels containing ZrO_2 ball bearings (10 mm x 10). One cycle of the ball-milling condition consists of 5 min milling and 5 min cooling cycle, and cycle numbers were dependent on each amount of sample. Samples were digested (72 h x 2) with Cellulysin® Cellulase, *Trichoderma viridae* (Calbiochem), at 35°C in acetate buffer (pH 5.0) in order to obtain the Enzyme Lignin (EL). ELs containing small amounts of residual polysaccharides were dissolved into DMSO-d_6/pyridine-d_5 (4:1) and subjected to NMR using a Bruker Biospin NEO 700-MHz spectrometer fitted with a cryogenically-cooled QCI $^{1}H/^{31}P/^{13}C/^{15}N$ gradient cryoprobe with inverse geometry (proton coil closest to the sample). 2D-^{1}H–^{13}C HSQC spectra were acquired using Bruker's pulse program (hsqcetgpsip2.2). Bruker's Topspin 3.2 (Mac) software was used to process spectra. The central DMSO peak was used as internal reference (δ_C: 39.50, δ_H: 2.49 ppm).

Saccharification Assays

Saccharification assays were performed on 10 mg of ground wood powder (see above) from the first growth cycle as described by

Van Acker et al. (2016). The activity of the 10× diluted enzyme mix was 0.14 Filter Paper Units/mL. For the alkali pretreatment, the stem material was treated with 1 mL 0.25% (v/v) NaOH at 90°C for 3 h while shaking at 750 rpm. For the hot water pretreatment, the stem material was incubated for 3 h with water at 90°C. The cellulose-to-glucose conversion was calculated based on the amount of glucose released upon saccharification and the original cellulose content that was measured for each sample.

RESULTS

ProCesA8-B:DCS_CURS2 Poplars Produce Curcumin and Curcumin-Derived Metabolites

To produce curcumin in the lignified tissues of poplar (*Populus tremula x P. alba*), the curcumin biosynthetic genes *DCS* and *CURS2,* linked together by a sequence coding for the self-cleaving T2A peptide, were expressed under the control of the *P. trichocarpa* secondary cell wall *CesA8-B 8-B* promoter (**Figure 1B**). This promoter was previously shown to be successful for lignin engineering in poplar (Wilkerson et al., 2014) and confers high expression in developing xylem cells (Joshi et al., 2004; Suzuki et al., 2006; Song et al., 2010). In total, 28 independent *ProCesA8-B:DCS_CURS2* poplar lines were obtained after Agrobacterium-mediated transformation and selection. After transfer to soil, all 28 transgenic lines and their WT controls were grown for 87 days in the greenhouse, after which they were harvested for further analysis.

To investigate whether the expression of *DCS* and *CURS2* resulted in the production of the envisioned compounds, phenolic metabolites were extracted from 87-day-old poplar xylem, and subsequently analyzed via UHPLC-MS. Upon a targeted search, five "curcuminoids" were found to be produced in all 28 transgenic poplars, whereas these compounds were absent in the WT (**Table 1** and **Supplementary Figure 1**). In addition to free curcumin (**1**), different coupling products between curcumin and coniferyl alcohol were identified: two isomers of curcumin(4-*O*-8)G (**2**–**3**), curcumin(5–8)G (**4**) and curcumin(8-8)G (**5**), reflecting the three main interunit bonds that result from radical coupling of curcumin with coniferyl alcohol.

In addition to curcumin-derived compounds, *ProCesA8-B:DCS_CURS2* poplars also had an increased abundance of derivatives of the phenylpentanoid intermediate (**Table 1** and **Supplementary Figure 1**). More specifically, six phenylpentanoid monomers, defined as compounds in which the phenylpentanoid structure was free or linked either via a 4-*O*-ether or γ-*O*-ester to a non-phenolic moiety, were found: dihydroferuloyl-β-keto acid was characterized as a free compound (**6**), or coupled to glycerol (**7**), malate (**8**), hexose (**9**), or a subunit of unknown identity (**10**). Additionally, tetrahydroferuloyl-β-keto acid coupled to hexose was characterized (**11**). Three phenylpentanoid-containing dimers, that are defined as compounds in which a phenylpentanoid moiety was linked via 4-*O*-8, 8-8, 8-5 or a cyclobutane structure to a phenylpropanoid or a second phenylpentanoid moiety, were also characterized: dihydroferuloyl-β-keto acid coupled with a 4-*O*-8 bond or a cyclobutane structure to coniferyl alcohol (**12**–**13**), and an 8-5 coupling product of dihydroferuloyl-β-keto acid and coniferyl alcohol with a hexose (**14**).

Curcumin Is Incorporated Into the Secondary Cell Wall of *ProCesA8-B:DCS_CURS2* Lines

Curcumin needs to be incorporated into the lignin polymer before it can render the lignin polymer more susceptible to alkaline pretreatments. Curcumin is fluorescent at an excitation wavelength of 488 nm, a spectral property that was successfully used to visualize the localization of curcumin in the cell wall region of inflorescence stem cross-sections of *ProCesA4:DCS_CURS2* Arabidopsis (Oyarce et al., 2019). To investigate the localization of curcumin in 87-day-old *ProCesA8-B:DCS_CURS2* poplars, cross-sections of the stem were examined via fluorescence microscopy. The curcumin-specific fluorescence signal was present in the cell walls of *ProCesA8-B:DCS_CURS2* xylem cells, but was absent in those of WT cells (**Figure 2**). The curcumin-specific fluorescence signal remained after washing the cross sections in acetone to remove soluble metabolites (**Supplementary Figure 2**), establishing that curcumin was tightly linked to the secondary cell wall. Additionally, lignin autofluorescence was used to visualize the architecture of the vessels in WT and *ProCesA8-B:DCS_CURS2* stems. The *ProCesA8-B:DCS_CURS2* xylem was similar to that of the WT and displayed no vascular defects (**Figure 2**).

Because the NMR signals from curcumin and G-lignin units overlap (Oyarce et al., 2019), this technique could not be used to confirm that curcumin was incorporated into the lignified cell wall of the (Arabidopsis and poplar) transgenic lines. Therefore, similar as for *ProCesA4:DCS_CURS2* Arabidopsis, catalytic hydrogenolysis was performed on extract-free cell wall material derived from the dried, debarked stems of 87-day-old *ProCesA8-B:DCS_CURS2* poplars to confirm curcumin incorporation into their cell wall. During catalytic hydrogenolysis, cell wall material is processed with a redox catalyst under reductive conditions, resulting in the cleavage of alkyl aryl ether bonds in the lignin, the removal of secondary (benzylic) alcohols, and the reduction of aliphatic double bonds (Galkin and Samec, 2016; Renders et al., 2019). However, aromatic structures and phenolic moieties are not hydrogenated and the cell wall polysaccharides remain largely intact (Van den Bosch et al., 2015). Here, Pd/C was used as a catalyst, retaining the largest fraction of primary alcohols (Li et al., 2018; Abu-Omar et al., 2021). Next, the lignin oil fraction was analyzed via UHPLC-MS. Based on principal component analysis of all detected peaks, *ProCesA8-B:DCS_CURS2* samples separated from WT samples (explaining 25% of the variation in PC1; **Supplementary Figure 3A**). Subsequently, we performed a targeted search for peaks that were derived from curcumin (coupling products) based on the structurally characterized products found in *ProCesA4:DCS_CURS2* Arabidopsis lignin oils (i.e., tetrahydrocurcumin, hexahydrocurcumin,

TABLE 1 | Targeted analysis of phenolic metabolites in *ProCesA8-B:DCS_CURS2* poplar stems.

No.	Compound name	*m/z*	R.T. (min)	WT	*ProCesA8-B:DCS_CURS2*	Ratio
				mean ± SD	mean ± SD	
Curcuminoids						
1	Curcumin (enol)[1]	367.120	26.55	n.d.	2307 ± 1308**	∞
2	Curcumin(4-*O*-8)G 1[1]	563.193	23.86	n.d.	4846 ± 1899**	∞
3	Curcumin(4-*O*-8)G 2[1]	563.193	27.60	n.d.	1328 ± 991**	∞
4	Curcumin(5-8)G[2]	545.184	27.63	n.d.	909 ± 732**	∞
5	Curcumin(8-8)G[2]	545.183	25.17	n.d.	1290 ± 666**	∞
Phenylpentanoid monomers						
6	Dihydroferuloyl-β-keto acid[1]	237.076	6.22	208 ± 118	16451 ± 23292**	79
7	Dihydroferuloyl-β-keto acid + glycerol[2]	311.112	4.80	n.d.	1005 ± 817**	∞
8	Dihydroferuloyl-β-keto acid + malate[2]	353.087	6.93	n.d.	626 ± 460**	∞
9	Dihydroferuloyl-β-keto acid + hexose[1]	399.130	3.09	1990 ± 1229	103329 ± 70530**	52
10	Dihydroferuloyl-β-keto acid + 302 Da[1]	539.177	10.26	n.d.	31468 ± 19507**	∞
11	Tetrahydroferuloyl-β-keto acid + hexose[2]	401.145	4.13	n.d.	4177 ± 4544**	∞
Phenylpentanoid-containing dimers						
12	Dihydroferuloyl-β-keto acid(4-*O*-8)G[1]	433.142	7.78	216 ± 139	12467 ± 10109**	58
13	Dihydroferuloyl-β-keto acid coniferyl alcohol cyclobutane dimer[2]	417.157	12.47	282 ± 155	3536 ± 2048**	13
14	[Dihydroferuloyl-β-keto acid(8-5)G or G(8-5)dihydroferuloyl-β-keto acid] + hexose[2]	577.192	8.30	n.d.	606 ± 526**	∞

*When compared to WT, the abundance of curcuminoids and phenylpentanoids is increased in the transgenic lines (Student's t-test; **P < 0.01; WT, n = 8 biologically independent replicates; ProCesA8-B:DCS_CURS2, n = 28 biologically independent lines). Peak area (mean) ± standard deviation (SD) are expressed in counts. R.T., retention time; G, guaiacyl unit. Remark: signals below 100 counts are considered as not detected (n.d.). [1]Tentatively structurally characterized based on MS/MS spectral analysis for ProCesA8-B:DCS_CURS2 and based on m/z and R.T. for WT. [2]Annotated based on m/z and R.T. similarity with structurally characterized compounds in Oyarce et al. (2019) for both ProCesA8-B:DCS_CURS2 and WT.*

deoxyhexahydrocurcumin and deoxyoctahydrocurcumin(5-8)G; Oyarce et al. (2019). Two peaks were found for which the intensity was at least tenfold higher in *ProCesA8-B:DCS_CURS2* compared to WT (**Supplementary Figures 3B,C**). These two peaks matched with hexahydrocurcumin (an incomplete reduction product of curcumin) and with deoxyoctahydrocurcumin(5-8)G, hereby confirming the ability of curcumin to couple with traditional monomers in the cell wall of poplar (**Supplementary Figures 3B,C**).

In conclusion, fluorescence microscopy and catalytic hydrogenolytic analysis showed that curcumin incorporated into the secondary cell wall of *ProCesA8-B:DCS_CURS2* poplars.

Engineering the Curcumin Pathway Into Poplar Affects Plant Growth

To determine the influence of the introduction of the curcumin biosynthetic pathway on the development of poplar, the height of WT and *ProCesA8-B:DCS_CURS2* poplars was monitored over their 87-day growth period (**Figure 3A**). After growing for 57 days in soil, no significant differences in height were observed between WT and *ProCesA8-B:DCS_CURS2* poplars. However, from day 60 onward, the growth rate of the transgenic poplars gradually declined until most of them ceased to grow (**Supplementary Table 2**). During this growth period, a lot of variation was observed in the height of *ProCesA8-B:DCS_CURS2* poplars (**Figures 3A,B** and **Supplementary Table 2**). At the time of harvest on day 87 (first growth cycle), three out of 28 transgenic poplars had shoot tips (several tens of centimeters, depending on the plant) that started to desiccate (**Figure 3C**), while already developing many new secondary shoots from the basal part of the stem and the roots (**Figure 3D**). After growing for 87 days in the greenhouse, the root system of the transgenic poplars was underdeveloped when compared to the WT (**Figures 3F,G**).

After pruning, the poplars were grown again for a second growth cycle of 4 months. In this case, we observed that,

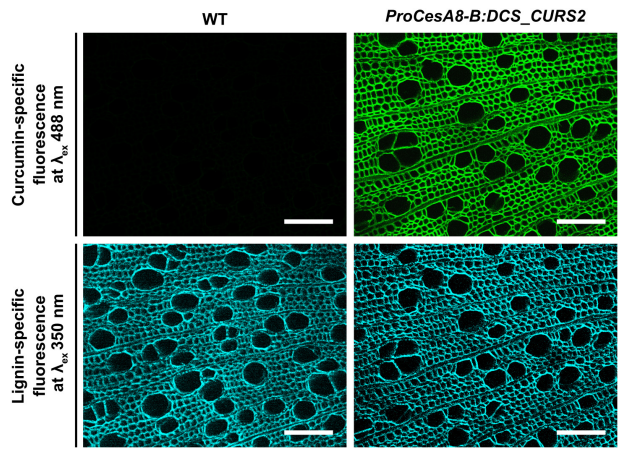


FIGURE 2 | Fluorescence microscopy on transverse stem sections of *ProCesA8-B:DCS_CURS2* poplars. The curcumin-specific fluorescence signal (excitation wavelength of 488 nm) was observed in the cell wall of *ProCesA8-B:DCS_CURS2* poplars, but not in that of WT poplars. The lignin-specific fluorescence (excitation wavelength of 350 nm) was observed in both the WT and *ProCesA8-B:DCS_CURS2* lines. Scale bars: 100 μm.

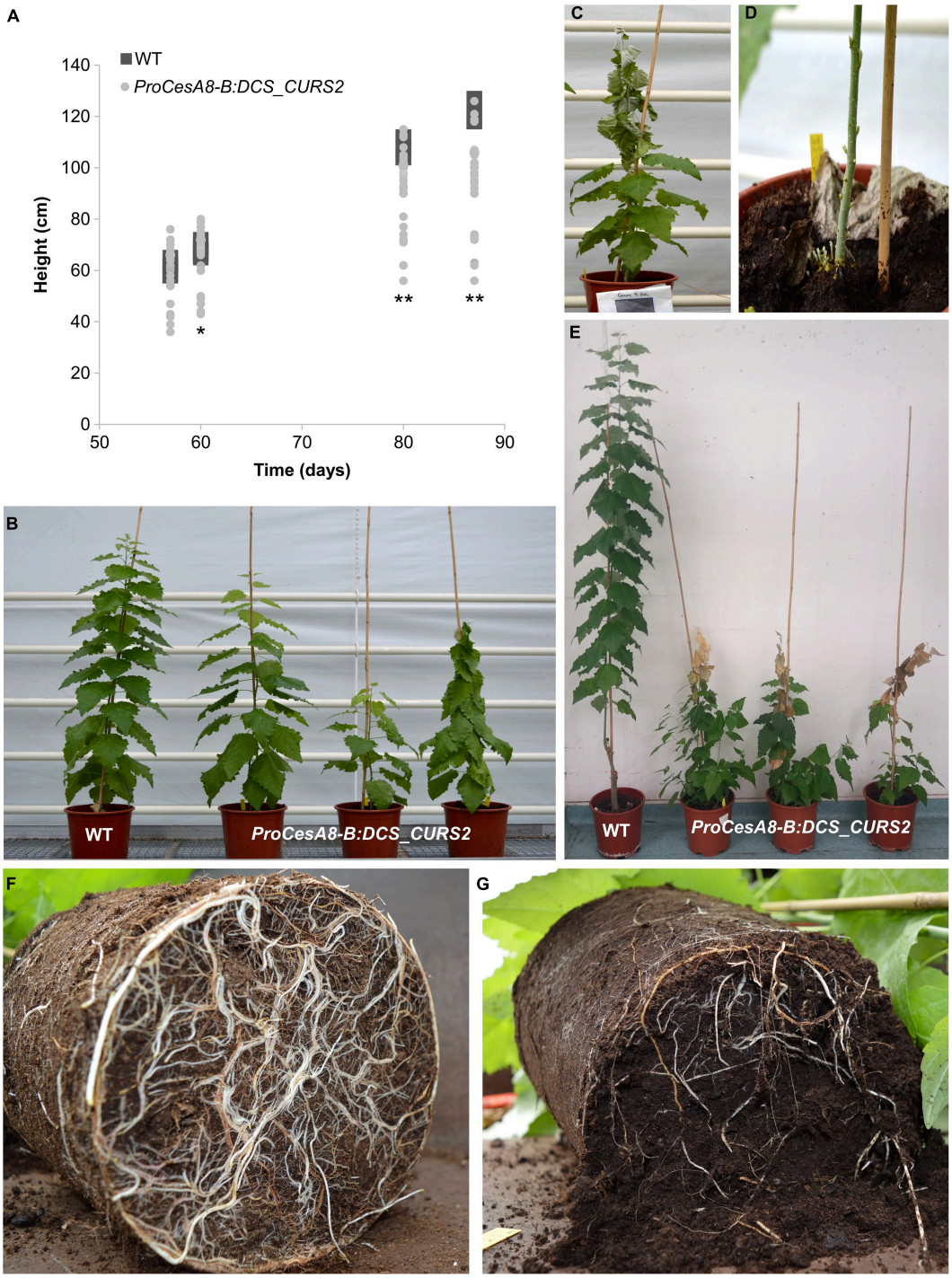


FIGURE 3 | *ProCesA8-B:DCS_CURS2* poplars display growth perturbations. **(A)** Growth curves of WT and *ProCesA8-B:DCS_CURS2* poplars over a period of 87 days. Individual values are represented by squares (WT) or dots (*ProCesA8-B:DCS_CURS2*). Differences in growth between the WT and the transgenic lines were assessed with Student's *t*-test (*$0.01 < P < 0.05$; **$P < 0.01$; WT, $n = 8$ biologically independent replicates; *ProCesA8-B:DCS_CURS2*, $n = 28$ biologically independent lines). **(B)** During the 87-day growth period, a lot of variation was observed in the severity of the growth perturbation of *ProCesA8-B:DCS_CURS2* poplars. **(C,D)** After ceasing growth, the shoot tip of some transgenic poplars started to senesce **(C)**, after which many new secondary shoots emerged from the roots and basal part of the stem **(D)**. **(E)** Image taken after a second 4-month growth period (after pruning the poplars after their first 87-day-long growth period); the WT reached a height of approximately 2 m, at which point the tip of the main stem of the *ProCesA8-B:DCS_CURS2* poplars had become necrotic at a height of approximately 1 m. However, a lot of secondary shoots sprouted from the roots and the basal part of the main stem. **(F,G)** The root system of *ProCesA8-B:DCS_CURS2* poplars **(G)** was underdeveloped compared to that of the WT **(F)** after growing in the greenhouse for 87 days. Scale bars: 20 cm in **(B,C,E)** and 10 cm in **(D,F,G)**.

after ceasing growth, all 28 *ProCesA8-B:DCS_CURS2* lines had shoot tips that desiccated and became necrotic. As in the first growth cycle, this shoot-tip necrosis was accompanied by the formation of secondary shoots leading to a "bushy" phenotype (**Figure 3E**). To investigate whether the shoot-tip necrosis of 4-month-old *ProCesA8-B:DCS_CURS2* poplars was correlated with vascular collapse, stem cross sections were visualized using bright-field microscopy (**Supplementary Figure 4**). The cross-sections were washed with acetone to remove soluble metabolites. Interestingly, the walls of secondary-thickened cells of *ProCesA8-B:DCS_CURS2* transgenic plants showed a yellow coloration, indicative of the presence of curcumin (or a curcumin-derived compound), whereas those of the WT were gray. Both WT and transgenic lines showed round, open vessels.

ProCesA8-B:DCS_CURS2 Poplars Have Altered Cell Wall Composition

ProCesA8-B:DCS_CURS2 poplars displayed yield penalties and produced and incorporated curcumin into their cell wall. To study additional effects on the cell wall, the CWR, the cellulose, matrix polysaccharide and lignin contents and the lignin composition of the basal part of the dried, debarked stem of 87-day-old WT and *ProCesA8-B:DCS_CURS2* poplars were determined (**Table 2**). Interestingly, compared to the WT, transgenic poplars had ∼10% less CWR, and thus relatively more soluble compounds. The crystalline cellulose content per CWR was decreased from 40.4% in the WT to 32.1% in the *ProCesA8-B:DCS_CURS2* poplars, while the amount of matrix polysaccharides (including hemicelluloses, pectins and amorphous cellulose) in the CWRs was increased from 40.8% in the WT to 52.4% in the *ProCesA8-B:DCS_CURS2* lines. The acetyl bromide lignin content per CWR was increased from 15.8% in the WT to 19.4% in *ProCesA8-B:DCS_CURS2*.

In order to gain further insights into the structural changes caused by curcumin biosynthesis and/or cross-coupling into the lignin polymer, 2D-NMR was performed on lignins from dried, debarked stems of 87-day-old WT and *ProCesA8-B:DCS_CURS2* lines. By analyzing the aromatic and aliphatic regions of the 2D ^{1}H–^{13}C correlation HSQC spectra, it was possible to visualize differences in lignin monomeric composition and interunit linkages (**Figure 4** and **Table 2**; Mansfield et al., 2012). The fraction of S units was decreased from 66.1% in the WT to 52.2% in the transgenic lines, whereas the fraction of G units was increased from 33.8% in WT to 45.8% in *ProCesA8-B:DCS_CURS2*. Consequently, the S/G ratio (calculated by dividing the proportion of S units by the proportion of G units) was reduced from 1.96 in the WT to 1.16 in *ProCesA8-B:DCS_CURS2*. Notably, the NMR resonance signals originating from the phenolic ring structures of curcumin are not distinguishable from those originating from G units (Oyarce et al., 2019), thus (part of) the increase in G units measured in the transgenic lines could potentially be attributed to the incorporation of curcumin into the lignin polymer. The frequency of H units was increased from 0.2% in the WT to 1.9% in the *ProCesA8-B:DCS_CURS2* lines, whereas the frequency of *p*-hydroxybenzoates in the lignin polymer was reduced from 5.9% in the WT to 0.7% in *ProCesA8-B:DCS_CURS2*. The interunit linkage types were deduced from the oxygenated aliphatic region of the HSQC. In WT, the relative fraction of β-aryl ether (8-*O*-4) linkages and resinol (8-8) linkages was 89.1 and 9.1%, respectively. These fractions did not significantly differ in the *ProCesA8-B:DCS_CURS2* lines. On the other hand, WT contained 1.8% phenylcoumaran (8-5) linkages and this fraction was significantly increased to 3.8% in *ProCesA8-B:DCS_CURS2* lines.

TABLE 2 | Cell wall characteristics.

	WT	*ProCesA8-B:DCS_CURS2*
CWR (% dry weight)	90.2 ± 2.8	79.5 ± 4.5**
Cellulose (% CWR)	40.4 ± 6.2	32.1 ± 8.0*
Matrix polysaccharides (% CWR)	40.8 ± 3.5	52.4 ± 8.8**
Acetyl bromide lignin (% CWR)	15.8 ± 1.2	19.4 ± 1.0**
NMR-derived aromatic units		
% H	0.2 ± 0.1	1.9 ± 0.8**
% S	66.1 ± 1.2	52.2 ± 5.0**
% G	33.8 ± 1.2	45.8 ± 4.6**
S/G	1.96 ± 0.10	1.16 ± 0.23**
% PB	5.9 ± 0.9	0.7 ± 0.9**
NMR-derived interunit linkages		
% β-Aryl ether (8-*O*-4; A)	89.1 ± 1.9	87.2 ± 2.1
% Phenylcoumaran (8-5; B)	1.8 ± 1.3	3.8 ± 1.3**
% Resinol (8-8; C)	9.1 ± 0.7	9.0 ± 1.0

*The cell wall residue (CWR) expressed as mass percentage of dry weight was determined gravimetrically after a sequential extraction. Crystalline cellulose content was determined by the Updegraff method and the mass loss during TFA extraction was used as an estimate of the amount of matrix polysaccharides. Lignin content was determined via the acetyl bromide (AcBr) method and expressed as mass percentage of CWR. Lignin composition was determined via 2D HSQC NMR. Differences between the WT and the transgenic lines were assessed with Student's t-test (*0.01 < P < 0.05; **P < 0.01; WT, n = 8 biologically independent replicates; ProCesA8-B:DCS_CURS2, n = 28 biologically independent lines). H, p-hydroxyphenyl; S, syringyl; G, guaiacyl; PB, p-hydroxybenzoate (see also **Figure 4**).*

To exclude the possibility that the observed shifts in cell wall composition were the consequence of the developmental delay in *ProCesA8-B:DCS_CURS2* lines, the cell wall characteristics of 87-day-old WT and four transgenic poplars with comparable heights (ranging from 118 to 127 cm and no shoot-tip necrosis) at the time of harvest were determined. These selected transgenic poplars still showed the previously described changes in cell wall composition (**Supplementary Table 3**).

Wood of *ProCesA8-B:DCS_CURS2* Lines Has a Saccharification Potential That Is Equal to That of the WT

Curcumin-containing lignin in Arabidopsis is more easily cleaved under alkaline conditions, leading to an increase in saccharification potential (Oyarce et al., 2019). To test the impact of the structural changes in the lignin polymer of *ProCesA8-B:DCS_CURS2* stems, saccharification assays were performed on CWRs of 87-day-old WT and transgenic poplars after an alkaline pretreatment, with untreated and hot-water pretreated samples for comparison (**Figure 5**).

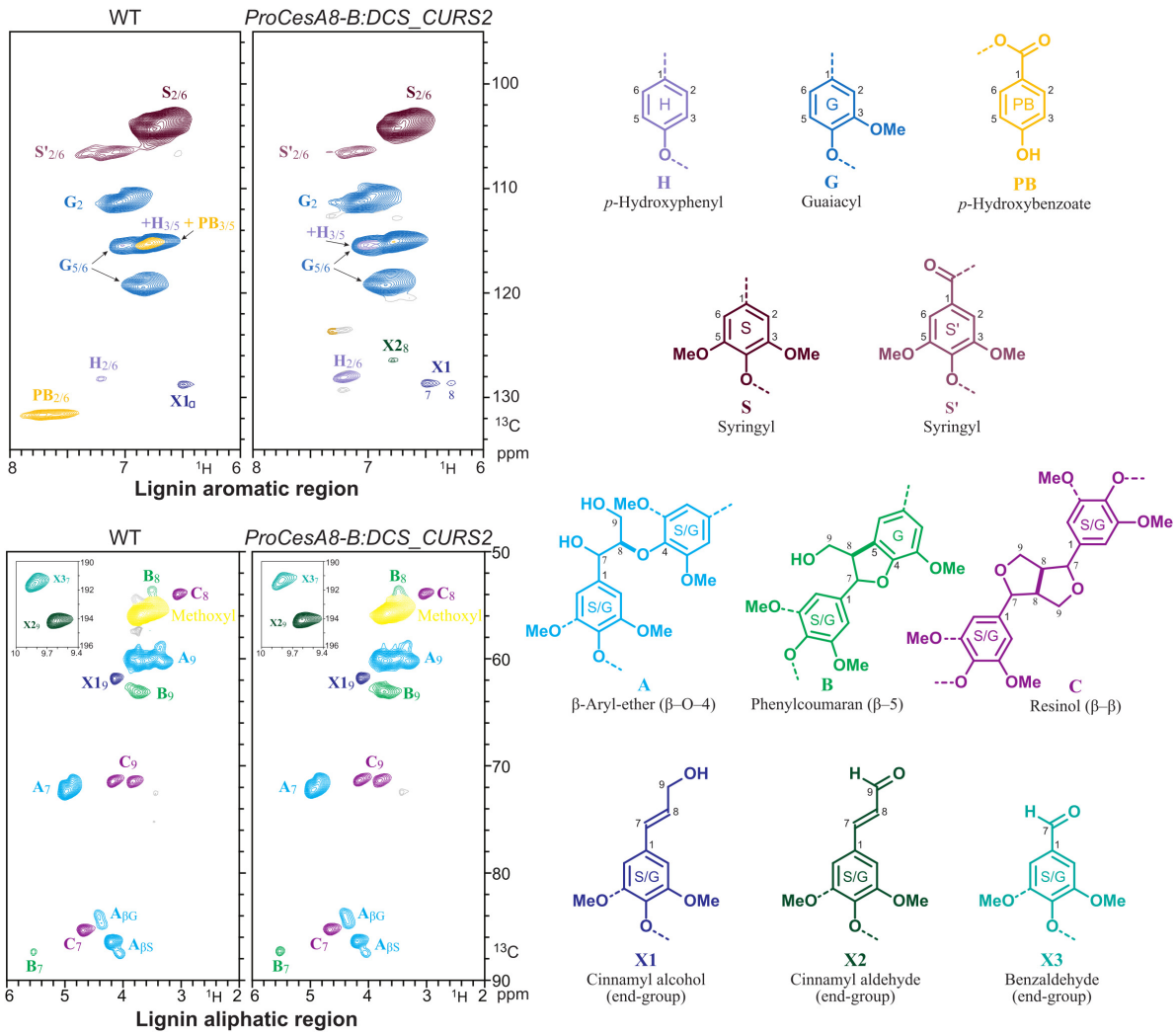


FIGURE 4 | Structural characterization of lignin via NMR. HSQC spectra of the aromatic and oxygenated aliphatic regions of whole cell walls from stems of WT and *ProCesA8-B:DCS_CURS2* poplars. Integrated values for each monomeric unit H, G, S and *p*-hydroxybenzoate and the α-C/H correlation peaks from the major lignin interunit structures A–C are provided in **Table 2**. The colors of the substructures shown match those of the corresponding signals in the HSQC spectra (where they are resolved). Representative figure for WT, *n* = 8 biologically independent replicates and *ProCesA8-B:DCS_CURS2*, *n* = 28 biologically independent lines.

For all tested pretreatments, the cellulose-to-glucose conversion after 48 h of saccharification of the transgenic poplars was equal to that of the WT (**Figure 5**). The cellulose-to-glucose conversion without pretreatment was about 26%, after alkaline pretreatment about 57% and after hot water pretreatment about 24% both for WT and *ProCesA8-B:DCS_CURS2* samples.

DISCUSSION

Translation of Fundamental Knowledge From Arabidopsis Into Poplar

Arabidopsis has been widely used as a model system for basic plant research and is a valuable model to study lignification (Meinke et al., 1998; Humphreys and Chapple, 2002; Koornneef and Meinke, 2010). The lignin biosynthetic pathways in Arabidopsis and the bioenergy crop poplar are similar, although poplar also incorporates additional units, like coniferyl and sinapyl *p*-hydroxybenzoates, into its lignin polymer (Lu et al., 2004, 2015; Morreel et al., 2004; Vanholme et al., 2019; Zhao et al., 2021; de Vries et al., 2022). Several lignin engineering strategies gave similar results in Arabidopsis and poplar. For example, greenhouse-grown *CCR*- or *CSE*-deficient Arabidopsis and poplar had lower amounts of lignin, higher saccharification efficiencies and yield penalties that depend on the level of residual target enzyme activity (Van Acker et al., 2013, 2014; Vanholme et al., 2013b; De Meester et al., 2020; de Vries et al., 2021a), whereas greenhouse-grown *4CL1*-deficient Arabidopsis and poplar had reduced lignin amounts and normal growth (Van Acker et al., 2013; Zhou et al., 2015). Introducing the curcumin biosynthetic pathway into Arabidopsis resulted in transgenic plants that produce and incorporate curcumin into their lignified

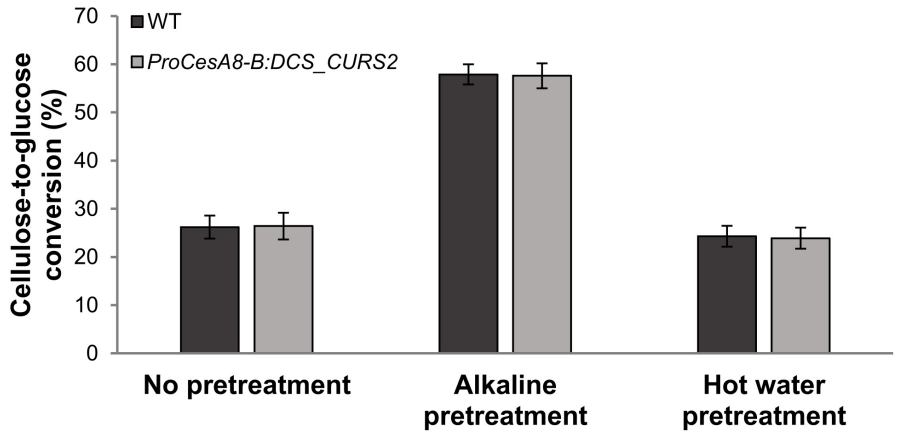


FIGURE 5 | Saccharification potential of stem biomass from WT and *ProCesA8-B:DCS_CURS2* poplars. Cellulose-to-glucose conversion efficiency after 48 h of saccharification. Samples were saccharified using no pretreatment, alkaline pretreatment (62.5 mM NaOH, 3 h, 90°C) or hot water pretreatment (100% H_2O, 3 h, 90°C). No significant differences between the WT and the transgenic lines were observed at the 0.01 significance level (Student's *t*-test; WT, *n* = 8 biologically independent replicates; *ProCesA8-B:DCS_CURS2*, *n* = 28 biologically independent lines). Error bars indicate standard deviation.

cell walls while growing normally and having an up to 24% increased saccharification efficiency after alkaline pretreatment (Oyarce et al., 2019). To assess its value for applications, we evaluated this lignin engineering strategy in the bioenergy crop poplar. Introducing the curcumin biosynthetic pathway into poplar resulted in transgenic plants that also produce and incorporate curcumin into their lignified cell walls, but that exhibit altered growth phenotypes and no improvements in saccharification efficiency.

ProCesA8-B:DCS_CURS2 Poplars Produce and Incorporate Curcumin Into Their Lignified Cell Walls

In *ProCesA4:DCS_CURS2* Arabidopsis, phenylpentanoid- and curcumin-derived compounds were produced (Oyarce et al., 2019). These compounds, including phenylpentanoid monomers and dimers, free curcumin, and curcumin coupled to coniferyl alcohol, were also found here in *ProCesA8-B:DCS_CURS2* poplar. These findings suggest that, just like in Arabidopsis, (i) it is possible to engineer curcumin biosynthesis into poplar solely via the expression of *DCS* and *CURS2* under a secondary cell wall promoter, and (ii) curcumin is also capable of coupling with traditional monomers via radical-radical interactions during lignification in poplar.

As with *ProCesA4:DCS_CURS2* Arabidopsis (Oyarce et al., 2019), *ProCesA8-B:DCS_CURS2* poplar stem-cross sections showed a yellow coloration of their secondary-thickened cell walls under bright-field microscopy. As curcumin is a yellow pigment, this yellow color is indicative of the presence of curcumin (or a curcumin-derived compound) in the transgenic lines (Oyarce et al., 2019). Moreover, *ProCesA8-B:DCS_CURS2* poplar stem-cross sections also showed a strong fluorescence in their secondary-thickened cell walls under fluorescence microscopy (at an excitation wavelength of 488 nm). Because the fluorescence was similar to that seen after *in vitro* polymerization of curcumin in WT cross sections and as seen in *ProCesA4:DCS_CURS2* Arabidopsis (Oyarce et al., 2019), this confirmed the presence of curcumin in the secondary-thickened cell walls of the transgenic poplar lines. Finally, just like in *ProCesA4:DCS_CURS2* Arabidopsis (Oyarce et al., 2019), catalytic hydrogenolysis showed curcumin's capability of translocating to and incorporating into the secondary cell wall of transgenic poplars.

ProCesA8-B:DCS_CURS2 Poplars Display a Significantly Altered Growth Phenotype

In contrast to *ProCesA4:DCS_CURS2* Arabidopsis that grew normally (Oyarce et al., 2019), *ProCesA8-B:DCS_CURS2* poplars suffered from yield penalties. At heights between 90 and 100 cm, the transgenic lines ceased growth, followed by desiccation and eventually necrosis of their shoot tips and the outgrowth of many new secondary shoots from the lower parts of the plant. The shoot-tip necrosis phenotype observed in *ProCesA8-B:DCS_CURS2* poplars clearly differed from the growth perturbations previously reported for lignin-modified plants (Chen and Dixon, 2007; Shadle et al., 2007; Coleman et al., 2008; Voelker et al., 2010; Ralph et al., 2012; Bonawitz and Chapple, 2013; Van Acker et al., 2013, 2014; Vanholme et al., 2013b). The latter, also coined lignin modification induced dwarfism (LMID), is typically associated with biomass and seed yield penalties. However, to the best of our knowledge, no shoot tip necrosis followed by the outgrowth of secondary branches was ever reported in lignin-modified plants. Additionally, the LMID observed in lignin-modified plants might be due to the loss of vessel cell wall integrity, leading to vascular collapse, which impedes water and nutrient transport from the roots to the aerial part of the plant (De Meester et al., 2018, 2021; Muro-Villanueva et al., 2019; Cao et al., 2020). Vascular collapse is often observed in growth-impeded lignin-engineered plants

including poplars (Leplé et al., 2007; Coleman et al., 2008; Voelker et al., 2011; Cao et al., 2020; De Meester et al., 2021). However, *ProCesA8-B:DCS_CURS2* poplars deposited a higher fraction of lignin into their secondary cell walls and, similarly to *ProCesA4:DCS_CURS2* Arabidopsis (Oyarce et al., 2019), did not suffer from collapsed vessels indicating that the origin of their observed growth perturbations is not related to vascular collapse. Nonetheless, due to their underdeveloped root system, water (and nutrient) uptake might be limited in the transgenic poplars, leading to the observed growth phenotypes.

Although, to the best of our knowledge, not observed previously in lignin-modified plants, shoot tip necrosis followed by the development of secondary shoots is commonly observed in *in vitro* cultured plants and can be caused by a number of factors, including nutrient deficiency (calcium and boron), the concentration of cytokinins and nutrients in the medium, aeration, gelling agents, pH of the growth medium, and subculture period (Bairu et al., 2009; Srivastava and Joshi, 2013). Of all these causes, calcium and boron deficiencies are the most common causes of shoot tip necrosis in different plant species, including avocado, potato, chestnut, pear and, as here, poplar. Calcium deficiency in plants results in poor root growth (George et al., 2008), which is also observed in the *ProCesA8-B:DCS_CURS2* poplars. Furthermore, calcium starvation was shown to lead to a strong inhibition of shoot growth in poplar (Lautner et al., 2007). Also boron deficiency leads to inhibition of growth of shoot meristems as it is a structural constituent of cell walls (Hu and Brown, 1994; Apostol and Zwiazek, 2004). As curcumin can form a complex with boron (Hayes and Metcalfe, 1962) and was shown to regulate calcium-related processes and mobilization in human cell cultures (Kliem et al., 2012; Mayadevi et al., 2012), it is possible that the curcumin present in *ProCesA8-B:DCS_CURS2* poplars makes calcium and boron unavailable, causing deficiencies followed by shoot tip necrosis. Due to their underdeveloped root system, the uptake of calcium and boron from the soil in *ProCesA8-B:DCS_CURS2* poplars could additionally be compromised.

The severe reduction in the amount of *p*-hydroxybenzoates in *ProCesA8-B:DCS_CURS2* lignin (see also below) might also lead/contribute to the decreased health of the transgenic trees. To date, the biological function of lignin acylation remains unknown (de Vries et al., 2022). However, *phbmt1* poplars with nearly depleted *p*-hydroxybenzoates in stem lignin also display compromised early-stage growth in soil, although no shoot tip necrosis was observed (Zhao et al., 2021). When these poplars were grown for more than 3 months in soil, their growth appeared normal.

Introducing the Curcumin Biosynthesis Pathway Into Poplar Significantly Alters Cell Wall Composition but Does Not Affect Saccharification Efficiency

Introducing the curcumin biosynthetic pathway into Arabidopsis and poplar redirects (part of) the feruloyl-CoA pool from the biosynthesis of the traditional monolignols toward curcumin biosynthesis. Nevertheless, the total lignin amount and H/G/S monomeric composition of *ProCesA4:DCS_CURS2* Arabidopsis remained unaltered, meaning that either the level of curcumin incorporation was minimal, or a feedback mechanism is in place to keep the total lignin amount at WT levels (Oyarce et al., 2019). In contrast to *ProCesA4:DCS_CURS2* Arabidopsis, *ProCesA8-B:DCS_CURS2* poplars had increased amounts of lignin, a decreased S/G ratio and an increased percentage of H units in the lignin when compared to the WT. These changes in lignin structure are most likely not a consequence of their aberrant growth, as they also occur in *ProCesA8-B:DCS_CURS2* poplars with a height comparable to that of the WT at the time of harvest. Hence, the structural changes to the lignin polymer were the consequence of the introduction of the *DCS* and *CURS2* transgenes leading to changes in the flux through the phenylpropanoid pathway and/or the possible activation of stress response pathways (i.e., the production of defense/stress lignin). Abiotic stresses typically induce the biosynthesis of H-rich lignin in both angiosperms and gymnosperms (Cesarino, 2019). Ozone-stressed poplar leaves were also shown to have higher lignin amounts, a lower S/G and a higher amount of H units incorporated into their lignin polymers (Cabane et al., 2004). A logical consequence of the relative increase in G and H units in the lignin of the transgenic lines is the observed relative increase in lignin 8-5 carbon-carbon linkages. Moreover, the relative decrease in S units in the lignin of the transgenic lines is in line with their relative decrease in *p*-hydroxybenzoates as, in poplar as well as in palms, it has been observed that *p*-hydroxybenzoates are almost exclusively bound to S units in the lignin of xylem fibers (Stewart et al., 2009; Lu et al., 2015; Regner et al., 2018; Goacher et al., 2021).

Although the cellulose amount of *ProCesA4:DCS_CURS2* Arabidopsis remained unaltered when compared to the WT control (Oyarce et al., 2019), that of *ProCesA8-B:DCS_CURS2* wood was significantly decreased. *ProCesA8-B:DCS_CURS2* wood also had increased amounts of matrix polysaccharides, which are mainly composed of hemicelluloses, when compared to the WT. As hemicelluloses might also covalently bind to lignin in trees (Terrett and Dupree, 2019), the observed changes in the cell wall composition of the transgenic lines might be the consequence of the activation of stress response pathways intended to produce lignin-carbohydrate complexes that strengthen the cell wall, as has been hypothesized for the ferulate-mediated cross-linking of polysaccharides with both polysaccharides and lignin in grasses (Monties, 1989; Bolwell, 1993; Ralph et al., 2004a; Ralph, 2010).

The cellulose-to-glucose conversion efficiency after saccharification using alkaline pretreatment was increased by up to 24% in *ProCesA4:DCS_CURS2* Arabidopsis (Oyarce et al., 2019), whereas that of *ProCesA8-B:DCS_CURS2* poplar was equal to that of the WT. The latter probably means that the effect of the incorporation of curcumin is nullified by other structural changes that occur in the secondary cell wall of the transgenic lines such as (i) their increased amount of lignin, which is the major limiting factor in the enzymatic hydrolysis of cell wall polysaccharides into simple sugars (Chen and Dixon, 2007; Van Acker et al., 2013; Vanholme et al., 2013b), (ii) the

increased amount of 8-5 structures in their lignin that, unlike 8-*O*-4 ether linkages, resist harsh alkaline or acidic pretreatment conditions (Sarkanen and Ludwig, 1971), (iii) their increased amount of matrix polysaccharides and thus lignin-hemicellulose complexes that inhibit saccharification by enhancing cell wall recalcitrance (Yang and Wyman, 2004), and (iv) potentially, the reduced amount of *p*-hydroxybenzoates in their lignin; enhanced *p*-coumaroylation of poplar lignin leads to improved saccharification yields after alkaline pretreatment (Lapierre et al., 2021), suggesting that reduced *p*-hydroxybenzoylation of poplar lignin might have the opposite effect.

Conclusion and Future Perspectives

Introducing *DCS* and *CURS*2 under the control of a secondary cell wall promoter results in the production of curcumin and subsequent incorporation of this alternative monomer into the secondary cell walls of Arabidopsis and poplar. However, in contrast to Arabidopsis, this lignin engineering strategy results in increased lignification, significantly altered lignin composition, growth perturbations and shoot tip necrosis when applied in poplar. These differences in (growth) phenotype display the importance of translational research in crops early during scientific development.

To generate transgenic poplars with cell walls more amenable to deconstruction without settling in yield, the proposed strategy might still be valuable if it can be fine-tuned to avoid the stress responses and growth defects. For example, using a different promoter to drive the expression of the curcumin biosynthesis genes (with a different spatial and/or temporal expression pattern) might lead to transgenic trees with more desirable characteristics. Indeed, here, the secondary cell-wall specific *CesA8-B* promoter was used that mainly confers expression in the xylem, but also has activity in phloem fibers and shoot tips (Joshi et al., 2004; Suzuki et al., 2006). As phloem fibers are mostly involved in providing support, biosynthesis of curcumin in these cells probably does not affect growth. However, the expression of the curcumin biosynthesis genes in the shoot tip, albeit to low levels, might contribute to the observed shoot tip necrosis. By restricting *DCS* and *CURS2* expression to the xylem, this adverse phenotype could potentially be avoided. To achieve this, a xylem-specific promoter is required to drive the curcumin biosynthesis genes. Examples include *ProDX15*, that confers expression in developing xylem cells (Jeon et al., 2016), *ProMX3*, that confers expression in mature xylem tissue (Nguyen et al., 2016), and *ProSNBE*, that confers expression in vessels (and sometimes also ray cells; De Meester et al., 2021). Alternatively, translation of this engineering strategy to other crops (e.g., maize) might lead to successful production and incorporation of curcumin into the lignified cell wall without affecting yield (just like in Arabidopsis).

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

BD, RuV, and WB designed the experiments. BD, PO, ReV, YT, TV, and JV performed the experiments. BD, PO, RuV, ReV, YT, TV, and JR collected and analyzed data. BD wrote the article with contributions from all authors. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from Ghent University (Multidisciplinary Research Partnership "Biotechnology for a Sustainable Economy" Grant 01MR0510W) and from the Agency for Innovation by Science and Technology (IWT) through the IWT-SBO project BIOLEUM (grant 130039). BD was indebted to IWT for a predoctoral fellowship and Research Foundation Flanders (grant G020618N) and the Energy Transition Fund (ETF) projects ADV_BIO and AD-LIBIO for a postdoctoral fellowship, PO was funded by the National Commission for Scientific and Technological Research (Chile) for a predoctoral fellowship and by SBO-FISH through the ARBOREF project (grant 140894) for a postdoctoral fellowship, RuV was indebted to Research Foundation Flanders for a postdoctoral fellowship (grant 12B0415N). YT and JR were partially funded by Stanford University's Global Climate and Energy Program (GCEP), and the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE- SC0018409).

ACKNOWLEDGMENTS

We thank Annick Bleys for critically reading the manuscript, Eef Parthoens from the VIB BioImaging Core for fluorescence microscopy assistance and Sandrien Desmet and Geert Goeminne from the VIB Metabolomics Core for assistance with the metabolomics experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.943349/full#supplementary-material

REFERENCES

Abu-Omar, M. M., Barta, K., Beckham, G. T., Luterbacher, J. S., Ralph, J., Rinaldi, R., et al. (2021). Guidelines for performing lignin-first biorefining. *Energy Environ. Sci.* 14:262.

Apostol, K. G., and Zwiazek, J. J. (2004). Boron and water uptake in jack pine (*Pinus banksiana*) seedlings. *Environ. Exp. Bot.* 51, 145–153. doi: 10.1016/j.envexpbot.2003.09.002

Bairu, M. W., Stirk, W. A., and Van Staden, J. (2009). Factors contributing to in vitro shoot-tip necrosis and their physiological interactions.

Plant Cell Tissue Organ Cult. 98, 239–248. doi: 10.1007/s11240-009-9560-8

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi: 10.1146/annurev.arplant.54.031902.134938

Bolwell, G. P. (1993). Dynamic aspects of the plant extracellular matrix. *Int. Rev. Cytol* 146, 261–324.

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

Bozell, J. J., Holladay, J. E., Johnson, D. B., and White, J. F. (2007). *Top Value added Chemicals from Biomass. Volume II: Results of Screening for Potential Candidates from Biorefinery Lignin*. Golden, CO: NREL. .

Brasileiro, A. C. M., Tourneur, C., Leple, J.-C., Combes, V., and Jouanin, L. (1992). Expression of the mutant *Arabidopsis thaliana* acetolactate synthase gene confers chlorsulfuron resistance to transgenic poplar plants. *Transgenic Res.* 1, 133–141.

Cabane, M., Pireaux, J. C., Leger, E., Weber, E., Dizengremel, P., Pollet, B., et al. (2004). Condensed lignins are synthesized in poplar leaves exposed to ozone. *Plant Physiol.* 134, 586–594. doi: 10.1104/pp.103.031765

Cao, S., Huang, C., Luo, L., Zheng, S., Zhong, Y., Sun, J., et al. (2020). Cell-specific suppression of 4-coumarate-CoA ligase gene reveals differential effect of lignin on cell physiological function in Populus. *Front. Plant Sci.* 11:589729. doi: 10.3389/fpls.2020.589729

Cesarino, I. (2019). Structural features and regulation of lignin deposited upon biotic and abiotic stresses. *Curr. Opin. Biotechnol.* 56, 209–214. doi: 10.1016/j.copbio.2018.12.012

Chanoca, A., De Vries, L., and Boerjan, W. (2019). Lignin engineering in forest trees. *Front. Plant Sci.* 10:912. doi: 10.3389/fpls.2019.00912

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, Y.-R., and Tan, T.-H. (1998). Inhibition of the c-Jun N-terminal kinase (JNK) signaling pathway by curcumin. *Oncogene* 17, 173–178. doi: 10.1038/sj.onc.1201941

Coleman, H. D., Park, J.-Y., Nair, R., Chapple, C., and Mansfield, S. D. (2008). RNAi-mediated suppression of p-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *Proc. Natl. Acad. Sci. U.S.A* 105, 4501–4506. doi: 10.1073/pnas.0706537105

Corma, A., Iborra, S., and Velty, A. (2007). Chemical routes for the transformation of biomass into chemicals. *Chem. Rev.* 107, 2411–2502. doi: 10.1021/cr050989d

Criss, D. L., Fisher, T. H., and Schultz, T. P. (1998). Alkaline hydrolysis of nonphenolic α-carbonyl β-0-4 lignin dimers substituted on the leaving phenoxide ring: comparison with benzylic hydroxyl analogues. *Holzforschung* 52, 57–60. doi: 10.1515/hfsg.1998.52.1.57

De Meester, B., De Vries, L., Özparpucu, M., Gierlinger, N., Corneillie, S., Pallidis, A., et al. (2018). Vessel-specific reintroduction of CINNAMOYL-COA REDUCTASE1 (CCR1) in dwarfed ccr1 mutants restores vessel and xylary fiber integrity and increases biomass. *Plant Physiol.* 176, 611–633. doi: 10.1104/pp.17.01462

De Meester, B., Madariaga Calderón, B., De Vries, L., Pollier, J., Goeminne, G., Van Doorsselaere, J., et al. (2020). Tailoring poplar lignin without yield penalty by combining a null and haploinsufficient CINNAMOYL-CoA REDUCTASE2 allele. *Nat. Commun.* 11:5020. doi: 10.1038/s41467-020-18822-w

De Meester, B., Vanholme, R., De Vries, L., Wouters, M., Van Doorsselaere, J., and Boerjan, W. (2021). Vessel-and ray-specific monolignol biosynthesis as an approach to engineer fiber-hypolignification and enhanced saccharification in poplar. *Plant J.* 108, 752–765. doi: 10.1111/tpj.15468

de Vries, L., Brouckaert, M., Chanoca, A., Kim, H., Regner, M. R., Timokhin, V. I., et al. (2021a). CRISPR-Cas9 editing of CAFFEOYL SHIKIMATE ESTERASE 1 and 2 shows their importance and partial redundancy in lignification in *Populus tremula* × P. alba. *Plant Biotechnol. J.* 19, 2221–2234. doi: 10.1111/pbi.13651

de Vries, L., Guevara-Rozo, S., Cho, M., Liu, L.-Y., Renneckar, S., and Mansfield, S. D. (2021b). Tailoring renewable materials via plant biotechnology. *Biotechnol. Biofuels* 14:167. doi: 10.1186/s13068-021-02010-z

de Vries, L., Mackay, H. A., Smith, R. A., Mottiar, Y., Karlen, S. D., Unda, F., et al. (2022). pHBMT1, a BAHD-family monolignol acyltransferase, mediates lignin acylation in poplar. *Plant Physiol.* 188:1014–1027. doi: 10.1093/plphys/kiab546

de Vries, L., Vanholme, R., Van Acker, R., De Meester, B., Sundin, L., and Boerjan, W. (2018). Stacking of a low-lignin trait with an increased guaiacyl and 5-hydroxyguaiacyl unit trait leads to additive and synergistic effects on saccharification efficiency in *Arabidopsis thaliana*. *Biotechnol. Biofuels* 11:257. doi: 10.1186/s13068-018-1257-y

del Río, J. C., Rencoret, J., Gutiérrez, A., Kim, H., and Ralph, J. (2022). (in press) Unconventional lignin monomers – extension of the lignin paradigm. *Adv. Bot. Res.* doi: 10.1016/bs.abr.2022.02.001Get

Dence, C. W. (1992). "The determination of lignin," in *Methods in Lignin Chemistry*, eds S. Y. Lin and C. W. Dence (Berlin: Springer-Verlag). 33–61.

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., Dutta, T., Deng, K., Jacquet, N., Sinha, A., Benites, V. T., et al. (2017). SbCOMT (Bmr12) is involved in the biosynthesis of tricin-lignin in sorghum. *PLoS One* 12:e0178160. doi: 10.1371/journal.pone.0178160

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, 609–620. doi: 10.1111/j.1467-7652.2012.00692.x

Franke, R., Mcmichael, C. M., Meyer, K., Shirley, A. M., Cusumano, J. C., and Chapple, C. (2000). Modified lignin in tobacco and poplar plants over-expressing the *Arabidopsis* gene encoding ferulate 5-hydroxylase. *Plant J.* 22, 223–234. doi: 10.1046/j.1365-313x.2000.00727.x

Galkin, M. V., and Samec, J. S. M. (2016). Lignin valorization through catalytic lignocellulose fractionation: a fundamental platform for the future biorefinery. *ChemSusChem* 9, 1544–1558. doi: 10.1002/cssc.201600237

George, E. F., Hall, M. A., and De Klerk, G.-J. (2008). *Plant Propagation by Tissue Culture*, 3rd Edn. Dordrecht: Springer.

Goacher, R. E., Mottiar, Y., and Mansfield, S. D. (2021). ToF-SIMS imaging reveals that p-hydroxybenzoate groups specifically decorate the lignin of fibres in the xylem of poplar and willow. *Holzforschung* 75, 452–462. doi: 10.1515/hf-2020-0130

Grabber, J. H., Davidson, C., Tobimatsu, Y., Kim, H., Lu, F., Zhu, Y., et al. (2019). Structural features of alternative lignin monomers associated with improved digestibility of artificially lignified maize cell walls. *Plant Sci.* 287:110070. doi: 10.1016/j.plantsci.2019.02.004

Grabber, J. H., Hatfield, R. D., Lu, F., and Ralph, J. (2008). Coniferyl ferulate incorporation into lignin enhances the alkaline delignification and enzymatic degradation of maize cell walls. *Biomacromolecules* 9, 2510–2516. doi: 10.1021/Bm800528f

Grabber, J. H., Schatz, P. F., Kim, H., Lu, F., and Ralph, J. (2010). Identifying new lignin bioengineering targets: 1. monolignol-substitute impacts on lignin formation and cell wall fermentability. *BMC Plant Biol.* 10:114. doi: 10.1186/1471-2229-10-114

Halpin, C. (2019). Lignin engineering to improve saccharification and digestibility in grasses. *Curr. Opin. Biotechnol.* 56, 223–229. doi: 10.1016/j.copbio.2019.02.013

Hauser, C. R., Swamer, F. W., and Ringler, B. I. (1948). Alkaline cleavage of unsymmetrical β-diketones. Ring opening of acylcylohexanones to form ε-acyl caproic acids. *J. Am. Chem. Soc.* 70, 4023–4026. doi: 10.1021/ja01192a016

Hayes, M. R., and Metcalfe, J. (1962). The boron-curcumin complex in the determination of trace amounts of boron. *Analyst* 87, 956–969. doi: 10.1039/AN9628700956

Hu, H., and Brown, P. H. (1994). Localization of boron in cell walls of squash and tobacco and its association with pectin (evidence for a structural role of boron in the cell wall). *Plant Physiol.* 105, 681–689. doi: 10.1104/pp.105.2.681

Humphreys, J. M., and Chapple, C. (2002). Rewriting the lignin roadmap. *Curr. Opin. Plant Biol* 5, 224–229. doi: 10.1016/s1369-5266(02)00257-1

Imai, A., Yokoyama, T., Matsumoto, Y., and Meshitsukat, G. (2007). Significant lability of guaiacylglycerol β-phenacyl ether under alkaline conditions. *J. Agric. Food Chem.* 55, 9043–9046. doi: 10.1021/jf071147d

Isikgor, F. H., and Becer, C. R. (2015). Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chem.* 6:4497.

Jeon, H. W., Cho, J. S., Park, E. J., Han, K. H., Choi, Y. I., and Ko, J. H. (2016). Developing xylem-preferential expression of PdGA20ox1, a gibberellin 20-oxidase 1 from *Pinus densiflora*, improves woody biomass production in a hybrid poplar. *Plant Biotechnol. J.* 14, 1161–1170. doi: 10.1111/pbi.12484

Joshi, C. P., Bhandari, S., Ranjan, P., Kalluri, U. C., Liang, X., Fujino, T., et al. (2004). Genomics of cellulose biosynthesis in poplars. *New Phytol.* 164, 53–61. doi: 10.1111/j.1469-8137.2004.01155.x

Katsuyama, Y., Kita, T., and Horinouchi, S. (2009). Identification and characterization of multiple curcumin synthases from the herb *Curcuma longa*. *FEBS Lett.* 583, 2799–2803. doi: 10.1016/j.febslet.2009.07.029

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* 10:101. doi: 10.1186/s13068-017-0784-2

Kliem, C., Merling, A., Giaisi, M., Köhler, R., Krammer, P. H., and Li-Weber, M. (2012). Curcumin suppresses T cell activation by blocking Ca2+ mobilization and nuclear factor of activated T cells (NFAT) activation. *J. Biol. Chem.* 287, 10200–10209. doi: 10.1074/jbc.M111.318733

Koornneef, M., and Meinke, D. (2010). The development of *Arabidopsis* as a model plant. *Plant J.* 61, 909–921. doi: 10.1111/j.1365-313X.2009.04086.x

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

Lapierre, C., Sibout, R., Laurans, F., Lesage-Descauses, M.-C., Déjardin, A., and Pilate, G. (2021). p-Coumaroylation of poplar lignins impacts lignin structure and improves wood saccharification. *Plant Physiol.* 187, 1374–1386. doi: 10.1093/plphys/kiab359

Lautner, S., Ehlting, B., Windeisen, E., Rennenberg, H., Matyssek, R., and Fromm, J. (2007). Calcium nutrition has a significant influence on wood formation in poplar. *New Phytol.* 173, 743–752. doi: 10.1111/j.1469-8137.2007.01972.x

Leplé, J.-C., Dauwe, R., Morreel, K., Storme, V., Lapierre, C., Pollet, B., et al. (2007). Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell* 19, 3669–3691. doi: 10.1105/tpc.107.054148

Li, Y., Shuai, L., Kim, H., Motagamwala, A. H., Mobley, J. K., Yue, F., et al. (2018). An "ideal lignin" facilitates full biomass utilization. *Sci. Adv.* 4:eaau2968. doi: 10.1126/sciadv.aau2968

Lu, F., Karlen, S. D., Regner, M., Kim, H., Ralph, S. A., Sun, R.-C., et al. (2015). Naturally *p*-hydroxybenzoylated lignins in palms. *BioEnergy Res.* 8, 934–952. doi: 10.1007/s12155-015-9583-4

Lu, F., Ralph, J., Morreel, K., Messens, E., and Boerjan, W. (2004). Preparation and relevance of a cross-coupling product between sinapyl alcohol and sinapyl *p*-hydroxybenzoate. *Org. Biomol. Chem.* 2, 2888–2890. doi: 10.1039/B411428K

Mahon, E. L., De Vries, L., Jang, S.-K., Middar, S., Kim, H., Unda, F., et al. (2022). Exogenous chalcone synthase expression in developing poplar xylem incorporates naringenin into lignins. *Plant Physiol.* 188, 984–996. doi: 10.1093/plphys/kiab499

Mahon, E. L., and Mansfield, S. D. (2019). Tailor-made trees: engineering lignin for ease of processing and tomorrow's bioeconomy. *Curr. Opin. Biotechnol.* 56, 147–155. doi: 10.1016/j.copbio.2018.10.014

Mansfield, S. D., Kim, H., Lu, F., and Ralph, J. (2012). Whole plant cell wall characterization using solution-state 2D NMR. *Nat. Protoc.* 7, 1579–1589. doi: 10.1038/nprot.2012.064

Marita, J. M., Ralph, J., Hatfield, R. D., and Chapple, C. (1999). NMR characterization of lignins in *Arabidopsis* altered in the activity of ferulate 5-hydroxylase. *Proc. Natl. Acad. Sci. U.S.A* 96, 12328–12332. doi: 10.1073/pnas.96.22.12328

Mayadevi, M., Sherin, D. R., Keerthi, V. S., Rajasekharan, K. N., and Omkumar, R. V. (2012). Curcumin is an inhibitor of calcium/calmodulin dependent protein kinase II. *Bioorg. Med. Chem.* 20, 6040–6047. doi: 10.1016/j.bmc.2012.08.029

Meinke, D. W., Cherry, J. M., Dean, C., Rounsley, S. D., and Koornneef, M. (1998). Arabidopsis thaliana: a model plant for genome analysis. *Science* 282, 679–682. doi: 10.1126/science.282.5389.662

Meyer, K., Shirley, A. M., Cusumano, J. C., Bell-Lelong, D. A., and Chapple, C. (1998). Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A* 95, 6619–6623. doi: 10.1073/pnas.95.12.6619

Mnich, E., Vanholme, R., Oyarce, P., Liu, S., Lu, F., Goeminne, G., et al. (2017). Degradation of lignin β−aryl ether units in Arabidopsis thaliana expressing LigD. LigF and LigG from *Sphingomonas paucimobilis* SYK-6. *Plant Biotechnol. J.* 15, 581–593.

Monties, B. L. (1989). "Lignins," in *Methods in Plant Biochemistry*, ed. J. Harborne (London: Academic Press). 113–157.

Morreel, K., Ralph, J., Kim, H., Lu, F., Goeminne, G., Ralph, S., et al. (2004). Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. *Plant Physiol.* 136, 3537–3549. doi: 10.1104/pp.104.049304

Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., and Mansfield, S. D. (2016). Designer lignins: harnessing the plasticity of lignification. *Curr. Opin. Biotechnol.* 37, 190–200. doi: 10.1016/j.copbio.2015.10.009

Muro-Villanueva, F., Mao, X., and Chapple, C. (2019). Linking phenylpropanoid metabolism, lignin deposition, and plant growth inhibition. *Curr. Opin. Biotechnol.* 56, 202–208. doi: 10.1016/j.copbio.2018.12.008

Nguyen, V. P., Cho, J. S., Choi, Y. I., Lee, S. W., Han, K. H., and Ko, J. H. (2016). Evaluation of a novel promoter from *Populus trichocarpa* for mature xylem tissue specific gene delivery. *Plant Physiol. Biochem.* 104, 226–233. doi: 10.1016/j.plaphy.2016.03.033

Oyarce, P., De Meester, B., Fonseca, F., De Vries, L., Goeminne, G., Pallidis, A., et al. (2019). Introducing curcumin biosynthesis in *Arabidopsis* enhances lignocellulosic biomass processing. *Nat. Plants* 5, 225–237. doi: 10.1038/s41477-018-0350-3

Pearson, R. G., and Mayerle, E. A. (1951). Mechanism of the hydrolytic cleavage of carbon—carbon bonds. I. alkaline hydrolysis of β-diketones. *J. Am. Chem. Soc.* 73, 926–930. doi: 10.1021/ja01147a012

Price, L. C., and Buescher, R. W. (1997). Kinetics of alkaline degradation of the food pigments curcumin and curcuminoids. *J. Food Sci.* 62, 267–269. doi: 10.1111/j.1365-2621.1997.tb03982.x

Ralph, J. (2010). Hydroxycinnamates in lignification. *Phytochem. Rev.* 9, 65–83. doi: 10.1007/s11101-009-9141-9

Ralph, J., Akiyama, T., Coleman, H. D., and Mansfield, S. D. (2012). Effects on lignin structure of coumarate 3-hydroxylase downregulation in poplar. *BioEnergy Res.* 5, 1009–1019. doi: 10.1007/s12155-012-9218-y

Ralph, J., Bunzel, M., Marita, J. M., Hatfield, R. D., Lu, F., Kim, H., et al. (2004a). Peroxidase-dependent cross-linking reactions of p-hydroxycinnamates in plant cell walls. *Phytochem. Rev.* 3, 79–96. doi: 10.1023/B:PHYT.0000047811.13837.fb

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

Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P. F., et al. (2004b). Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochem. Rev.* 3, 29–60. doi: 10.1023/B:PHYT.0000047809.65444.a4

Regner, M., Bartuce, A., Padmakshan, D., Ralph, J., and Karlen, S. D. (2018). Reductive cleavage method for quantitation of monolignols and low-abundance monolignol conjugates. *ChemSusChem* 11, 1600–1605. doi: 10.1002/cssc.201800617

Renders, T., Van Den Bossche, G., Vangeel, T., Van Aelst, K., and Sels, B. (2019). Reductive catalytic fractionation: state of the art of the lignin-first biorefinery. *Curr. Opin. Biotechnol.* 56, 193–201. doi: 10.1016/j.copbio.2018.12.005

Sarkanen, K. V., and Ludwig, C. H. (1971). *Lignins, Occurrence, Formation, Structure and Reactions*. New York: Wiley-Interscience.

Sederoff, R. R., Mackay, J. J., Ralph, J., and Hatfield, R. D. (1999). Unexpected variation in lignin. *Curr. Opin. Plant Biol.* 2, 145–152.

Shadle, G., Chen, F., Reddy, M. S. S., Jackson, L., Nakashima, J., and Dixon, R. A. (2007). Down-regulation of hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry* 68, 1521–1529. doi: 10.1016/J.Phytochem.2007.03.022

Sibout, R., Le Bris, P., Legée, F., Cézard, L., Renault, H., and Lapierre, C. (2016). Structural redesigning arabidopsis lignins into alkali-soluble lignins through the expression of p-coumaroyl-CoA:monolignol transferase PMT. *Plant Physiol.* 170, 1358–1366. doi: 10.1104/pp.15.01877

Smith, R. A., Gonzales-Vigil, E., Karlen, S. D., Park, J.-Y., Lu, F., Wilkerson, C. G., et al. (2015). Engineering monolignol p-coumarate conjugates into poplar and *Arabidopsis* lignins. *Plant Physiol.* 169, 2992–3001. doi: 10.1104/pp.15.00815

Smith, R. A., Lu, F., Muro Villanueva, F., Cusumano, J. C., Chapple, C., and Ralph, J. (2022). Manipulation of lignin monomer composition combined with the

introduction of monolignol conjugate biosynthesis leads to synergistic changes in lignin structure. *Plant Cell Physiol.* 63, 744–754. doi: 10.1093/pcp/pcac031

Song, D., Shen, J., and Li, L. (2010). Characterization of cellulose synthase complexes in Populus xylem differentiation. *New phytol.* 187, 777–790. doi: 10.1111/j.1469-8137.2010.03315.x

Srivastava, A., and Joshi, A. G. (2013). Control of shoot tip necrosis in shoot cultures of *Portulaca grandiflora* hook. *Not. Sci. Biol.* 5, 45–49. doi: 10.15835/nsb519009

Stewart, J. J., Akiyama, T., Chapple, C., Ralph, J., and Mansfield, S. D. (2009). The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol.* 150, 621–635. doi: 10.1104/pp.109.137059

Suzuki, S., Li, L., Sun, Y. H., and Chiang, V. L. (2006). The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*. *Plant Physiol.* 142, 1233–1245. doi: 10.1104/pp.106.086678

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

Tomren, M. A., Másson, M., Loftsson, T., and Tønnesen, H. H. (2007). Studies on curcumin and curcuminoids: XXXI. symmetric and asymmetric curcuminoids: stability, activity and complexation with cyclodextrin. *Int. J. Pharm.* 338, 27–34. doi: 10.1016/j.ijpharm.2007.01.013

Tønnesen, H. H., and Karlsen, J. (1985). Studies on curcumin and curcuminoids. V. Alkaline degradation of curcumin. *Z. Lebensm. Unters. Forsch.* 180, 132–134. doi: 10.1007/bf01042637

Tsuji, Y., Vanholme, R., Tobimatsu, Y., Ishikawa, Y., Foster, C. E., Kamimura, N., et al. (2015). Introduction of chemically labile substructures into *Arabidopsis* lignin through the use of LigD, the Cα-dehydrogenase from Sphingobium sp. strain SYK-6. *Plant Biotechnol. J.* 13, 821–832. doi: 10.1111/pbi.12316

Unda, F., Mottiar, Y., Mahon, E. L., Karlen, S. D., Kim, K. H., Loqué, D., et al. (2022). A new approach to zip-lignin? 3,4-dihydroxybenzoate is compatible with lignification. *New Phytol.* [Online], (in press), doi: 10.1111/nph.18136

Updegraff, D. M. (1969). Semimicro determination of cellulose in biological materials. *Anal. Biochem.* 32, 420–424. doi: 10.1016/S0003-2697(69)80009-6

Van Acker, R., Leplé, J.-C., Aerts, D., Storme, V., Goeminne, G., Ivens, B., et al. (2014). Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. *Proc. Natl. Acad. Sci. U.S.A* 111, 845–850. doi: 10.1073/pnas.1321673111

Van Acker, R., Vanholme, R., Piens, K., and Boerjan, W. (2016). Saccharification protocol for small-scale lignocellulosic biomass samples to test processing of cellulose into glucose. *Bio-Protocol* 6:e1701.

Van Acker, R., Vanholme, R., Storme, V., Mortimer, J. C., Dupree, P., and Boerjan, W. (2013). Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnol. Biofuels* 6:46. doi: 10.1186/1754-6834-6-46

Van den Bosch, S., Schutyser, W., Koelewijn, S. F., Renders, T., Courtin, C. M., and Sels, B. F. (2015). Tuning the lignin oil OH-content with Ru and Pd catalysts during lignin hydrogenolysis on birch wood. *Chem. Commun.* 51, 13158–13161. doi: 10.1039/c5cc04025f

Vanholme, B., Desmet, T., Ronsse, F., Rabaey, K., Van Breusegem, F., De Mey, M., et al. (2013a). Towards a carbon-negative sustainable bio-based economy. *Front. Plant Sci.* 4:174. doi: 10.3389/fpls.2013.00174

Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., et al. (2013b). 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., 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

Vanholme, R., Morreel, K., Darrah, C., Oyarce, P., Grabber, J. H., Ralph, J., et al. (2012). Metabolic engineering of novel lignin in biomass crops. *New Phytol.* 196, 978–1000. doi: 10.1111/j.1469-8137.2012.04337.x

Vanholme, R., Ralph, J., Akiyama, T., Lu, F., Pazo, J. R., Kim, H., et al. (2010). Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in *Arabidopsis*. *Plant J.* 64, 885–897. doi: 10.1111/j.1365-313X.2010.04353.x

Voelker, S. L., Lachenbruch, B., Meinzer, F. C., Jourdes, M., Ki, C., Patten, A. M., et al. (2010). Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiol.* 154, 874–886. doi: 10.1104/pp.110.159269

Voelker, S. L., Lachenbruch, B., Meinzer, F. C., and Strauss, S. H. (2011). Reduced wood stiffness and strength, and altered stem form, in young antisense 4CL transgenic poplars with reduced lignin contents. *New Phytol.* 189, 1096–1109. doi: 10.1111/j.1469-8137.2010.03572.x

Wang, Y.-J., Pan, M.-H., Cheng, A.-L., Lin, L.-I., Ho, Y.-S., Hsieh, C.-Y., et al. (1997). Stability of curcumin in buffer solutions and characterization of its degradation products. *J. Pharm. Biomed. Anal.* 15, 1867–1876. doi: 10.1016/S0731-7085(96)02024-9

Weng, J.-K., Mo, H., and Chapple, C. (2010). Over-expression of F5H in COMT-deficient *Arabidopsis* leads to enrichment of an unusual lignin and disruption of pollen wall formation. *Plant J.* 64, 898–911. doi: 10.1111/j.1365-313X.2010.04391.x

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

Yang, B., and Wyman, C. E. (2004). Effect of xylan and lignin removal by batch and flowthrough pretreatment on the enzymatic digestibility of corn stover cellulose. *Biotechnol. Bioeng.* 86, 88–95. doi: 10.1002/bit.20043

Zhao, Y., Yu, X., Lam, P.-Y., Zhang, K., Tobimatsu, Y., and Liu, C.-J. (2021). Monolignol acyltransferase for lignin p-hydroxybenzoylation in Populus. *Nat. Plants* 7, 1288–1300. doi: 10.1038/s41477-021-00975-1

Zhou, S., Runge, T., Karlen, S. D., Ralph, J., Gonzales-Vigil, E., and Mansfield, S. D. (2017). Chemical pulping advantages of zip-lignin hybrid poplar. *ChemSusChem* 10, 3565–3573. doi: 10.1002/cssc.201701317

Zhou, X., Jacobs, T. B., Xue, L. J., Harding, S. A., and Tsai, C. J. (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate:CoA ligase specificity and redundancy. *New Phytol.* 208, 298–301. doi: 10.1111/nph.13470

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 22 July 2022
DOI 10.3389/fpls.2022.926300

OPEN ACCESS

EDITED BY
Igor Cesarino,
University of São Paulo, Brazil

REVIEWED BY
Aimin Wu,
South China Agricultural University, China
Rachel Renae Schendel,
University of Kentucky,
United States

*CORRESPONDENCE
Leonardo D. Gomez
leonardo.gomez@york.ac.uk

†In memory of our dear friend and colleague Simon J. McQueen-Mason

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

RECEIVED 22 April 2022
ACCEPTED 05 July 2022
PUBLISHED 22 July 2022

CITATION
Möller SR, Lancefield CS, Oates NC, Simister R, Dowle A, Gomez LD and McQueen-Mason SJ (2022) CRISPR/Cas9 suppression of OsAT10, a rice BAHD acyltransferase, reduces p-coumaric acid incorporation into arabinoxylan without increasing saccharification. *Front. Plant Sci.* 13:926300. doi: 10.3389/fpls.2022.926300

CRISPR/Cas9 suppression of OsAT10, a rice BAHD acyltransferase, reduces p-coumaric acid incorporation into arabinoxylan without increasing saccharification

Svenning R. Möller[1], Christopher S. Lancefield[2], Nicola C. Oates[1], Rachael Simister[1], Adam Dowle[3], Leonardo D. Gomez[1]*and Simon J. McQueen-Mason[1†]

[1]CNAP, Biology Department, University of York, York, United Kingdom, [2]School of Chemistry and Biomedical Science Research Complex, University of St. Andrews, St. Andrews, United Kingdom, [3]Biology Department, Bioscience Technology Facility, University of York, York, United Kingdom

Ester-linked hydroxycinnamic acids ferulic acid (FA) and para-coumaric acid (*p*-CA) play important roles in crosslinking within cell wall arabinoxylans (AX) and between AX and lignin in grass cell walls. The addition of hydroxycinnamates to AX, is mediated by the Mitchell clade of BAHD acyl-coenzyme A-utilizing transferases. Overexpression of OsAT10 (a Mitchell clade BAHD acyl transferase) in rice, has previously been shown to increase *p*-CA content in AX in leaves and stems, leading to increased cell wall digestibility, potentially associated with a concomitant decrease in FA content. To investigate the physiological role of OsAT10 we established CRISPR/Cas9 rice knock-out mutants devoid of OsAT10. Our analysis of hydroxycinnamic acid content in wild type plants revealed that AX associated *p*-CA is found almost exclusively in rice husks, with very little found in other tissues. Mutant plants were essentially devoid of ester-linked *p*-CA associated with AX, indicating that OsAT10 represents the major enzyme responsible for the addition of *p*-CA to arabinoxylan in rice plants. We found no change in the digestibility of rice husk lacking AX-associated *p*-CA, suggesting that the changes in digestibility seen in OsAT10 overexpressing plants were solely due to compensatory decreases in AX-associated FA.

KEYWORDS

p-coumaric acid, ferulic acid, hydroxycinnamic acid, BAHD acyl transferase, husk, rice, arabinoxylan, cell wall

Introduction

In plants, cell walls play a multitude of roles from providing support and integrity to cells and tissues to providing an important interface between the plant and its environment. Strength and integrity are key aspects of cell wall functions, but these have to be accompanied by dynamic flexibility to allow growth and development. Cellulose

microfibrils composed of paracrystalline assemblages of β1-4, glucans provide the structural framework of the cell wall. Cellulose microfibrils are encased in a matrix of hemicellulose and lignin, and the interactions between these polymers underpins the integrity of the cell wall. The major hemicellulose of graminaceous monocots, such as rice, is a complex arabinoxylan (AX; Scheller and Ulvskov, 2010). AX is made up of a β1-4 xylan backbone, decorated with a range of monosaccharide side chains, predominated by arabinose, but including xylose, galactose and glucuronic acid. Many of the xylosyl backbone residues of AX that lack sugar side chains may be acetylated, and the patterning of xylan decorations impacts the conformation of the hemicellulose and its ability to associate with cellulose (Simmons et al., 2016; Grantham et al., 2017). One of the key features of AX is its ability to form covalent associations between neighboring chains and with the lignin network (Mnich et al., 2020). These linkages primarily occur through ferulic acid (FA) residues, which are present as esters on a subset of the arabinosyl side chains of AX. FA moieties form crosslinks between one another through free radical chemistry to form intra-AX crosslinks (Ralph et al., 1994). The FA sidechains of AX can also form covalent linkages with lignin (Ralph et al., 1992, 1995). Both of these associations play important roles in determining the integrity and digestibility of cereal cell walls (de Oliveira et al., 2015). In addition to FA, the arabinosyl side chains of AX can be ester linked to another hydroxycinnamic acid (HCA), *p*-coumaric acid (*p*-CA), which is not thought to crosslink directly with lignin or neighboring AX chains (Mueller-Harvey et al., 1986; Ishii et al., 1990; Ralph, 2010).

FA and *p*-CA are both also present as esters in the lignin network. The addition of FA and *p*-CA to AX and lignin are believed to be mediated by enzymes belonging to the so-called core Mitchell clade BAHD acyl-CoA transferases (MCBAT), using HCA-CoA as an acyl-donor during biosynthesis (Mitchell et al., 2007; Piston et al., 2010; de Souza et al., 2018, 2019). Using a bioinformatics approach, Mitchell et al. (2007) identified 10 putative MCBAT genes in the rice genome, known as *OsAT1-10*.

Experimental evidence regarding the role of specific MCBATs in different plant species has been produced over the last 10 years since Piston et al. Piston et al. (2010) first demonstrated that simultaneous RNAi knockdown of multiple MCBATs, reduced total FA content in rice. Buanafina et al. (2016) showed that RNAi suppression of BdAT1 (*OsAT1* homolog) in *Brachypodium* resulted in a reduction of total FA content. Furthermore, a T-DNA knockout of *OsAT7* was shown to reduce FA content in rice sheaths (Bartley et al., 2013).

A number of studies have implicated MCBATs in the incorporation of HCAs into lignin. For example, overexpression of *OsAT5* in rice stems led to increased lignin associated ferulic acid (Bartley et al., 2013; Karlen et al., 2016). *OsAT4* has been shown to transfer *p*-CA to lignin *in vitro* (Withers et al., 2012). *OsAT3* homologs in maize (*ZmPMTs*) transfer *p*-CA, FA and to a lesser degree caffeic acid to sinapyl alcohol *in vitro*. However, RNAi knockdown of *ZmPMT* only resulted in a reduction of *p*-CA in cell walls (Marita et al., 2014). In addition, *Brachypodium* knockouts of the *OsAT3* homolog *BdPMT*, completely eliminated sinapyl alcohol bound *p*-CA, while overexpression produced changes to the lignin content and S:G lignin ratio (Petrik et al., 2014). Overexpression of *BdPMT2*, the *OsAT8* homolog, in Arabidopsis resulted in an increase in lignin-associated *p*-CA, similar to *BdPMT1* (Sibout et al., 2016). *BdPMT1* overexpression, in turn, transferred *p*-CA to guaiacyl alcohol, highlighting the broader range of specificities for this enzyme. Taken together, these studies indicate that OsAT3 and 4 (and possibly OsAT8) participate in the addition of *p*-CA to lignin, and OsAT5 in the addition of FA to lignin.

RNAi gene silencing of an *OsAT9* orthologue shows reduced FA-Ara in *Setaria viridis*, and to a lesser degree in *Brachypodium*, indicating it is likely responsible for the addition of FA to arabinosyl side chains in AX. *OsAT9* is also the closest relative to *OsAT10* phylogenetically. These gene silencing studies also demonstrated how changing the content of AX-linked FA has a profound influence on the cell wall and applied plant traits such as digestibility (de Souza et al., 2018, 2019; Li et al., 2018).

OsAT10 overexpression in rice increased incorporation of *p*-CA to AX, and this was accompanied by a decrease in FA on AX, and a concomitant increase in saccharification in stems and leaves (Bartley et al., 2013). A similar effect was seen in switchgrass (Li et al., 2018), suggesting that OsAT10 is responsible for the incorporation of *p*-CA to AX. However, RNAi plants with simultaneous knock downs of *OsAT6* through *10* showed no reduction in *p*-CA in stems and leaves, despite a significant reduction in *OsAT10* transcript (Piston et al., 2010). However, as RNAi can easily reduce the expression of multiple homologous genes, results may have resulted from off-target effect. In addition, RNAi does not fully remove a transcript allowing one to identify leading to potentially ambiguous results. BAHDs may have multiple overlapping specificities (D'Auria, 2006) and overexpression thus may not accurately describe the main target of the enzyme. Using the CRISPR/Cas9 system, targeted cuts can be made to DNA, generating insertions or deletion of base pairs which result in frameshift mutations completely removing the translated protein (Shan et al., 2013). The CRISPR/Cas9 guide only requires 20 base pair to determine the cut and can be designed to minimise potential off-target effects in homologs (Young et al., 2019), unlike RNAi approaches, which typically make use of a much longer hairpin RNA for silencing (Watson et al., 2005).

Mota et al. (2021) recently used *RNAi* gene suppression to reduce the activity of SvBAHD05 (*OsAT1* homolog), a MCBAT in *S. viridis*. Their results showed that suppressing this gene led to a significant reduction in *p*-CA-AX in these plants and this resulted in a modest increase in the digestibility of the affected tissues with commercial cellulases, suggesting that *p*-CA-AX may be a determinant of cell wall recalcitrance to digestion. It has been suggested that *p*-CA can act as a catalyst for lignin formation by radical transfer to sinapyl alcohol (Takahama and Oniki, 1994; Hatfield et al., 2008), but Mota et al. (Currie and Perry, 2007) reported no apparent changes in lignin in their plants with reduced *p*-CA-AX content.

Several plant cell wall components have been suggested to interact with, or bind to, silica (Leroux et al., 2013; He et al., 2015; Guerriero et al., 2018). Several authors have suggested that AX-HCA complexes may act as nucleation sites for silica deposition (He et al., 2013; Soukup et al., 2017). Husks of cereal grain are notably rich in both *p*-CA and silica (Currie and Perry, 2007; Ikram et al., 2017; Gao et al., 2018), suggesting a potential association between these compounds. Lignin content and *p*-CA are also known to increase under silicon deficiency (Goto et al., 2003; Suzuki et al., 2012). These associations between *p*-CA, lignin and silica in the husk underline a possible role in stress defence and grain protection (Boz, 2015).

Whereas most other studies on the Mitchell clade have either utilized RNAi or overexpression, here, we report on the use of CRISPR/Cas9 gene editing to knock out *OsAT10* in rice allowing for full determination of its biological role. Our studies reveal that OsAT10 is responsible for most of the *p*-CA on AX in rice plants. Our studies show that most of the *p*-CA on AX is found in rice husks and that the complete loss of the *p*-CA on AX had no detectable effect on cell wall digestibility or silica content.

Materials and methods

Vector construction

Initially, the sgRNA expression cassette was cloned into vector pJIT163-2NLSCas9 (Addgene: Plasmid #53064; Shan et al., 2013), while simultaneously adding a ubiquitin promoter to the *Cas9* gene. The sgRNA expression cassette was made from two fragments with the guide being added as a primer overhang. The primers used for cloning are listed in Supplementary Table S1. A U3 promoter fragment with guide overhang was amplified from UAP1-U3, a sgRNA scaffold fragment with guide overhang was amplified from pICSL90010 and a ubiquitin promoter fragment was amplified from UAP1-OsUbi. pJIT163-2NLSCas9 was opened with Bam-HI and used in a NEBbuilder (New England Biolabs #E2621S) cloning reaction together with the above fragments. The hygromycin cassette was obtained from GoldenGate cloning using vectors pICH47732, pICSL12009, pICL80036 and pICH41421 (Engler et al., 2014; Vazquez-Vilar et al., 2016).

The final *Agrobacterium* expression vectors were cloned by combining the CRISPR/Cas9 expression cassette with the hygromycin expression cassette into the pAGM4723 (Weber et al., 2011) which had been digested with OliI and KpnI, using NEBbuilder. Apart from pJIT163-2NLSCas9 remaining vectors were obtained from MoClo (Weber et al., 2011) and its expansions from (Engler et al., 2014; Vazquez-Vilar et al., 2016).

Rice transformation

Agrobacterium mediated transformation was performed according to (Nishimura et al., 2006). The CRISPR/Cas9 binary vectors were introduced into Japonica rice Nipponbare cultivar. Hygromycin was used as a selection marker and resistant calli were transferred to regeneration media to obtain mature rice plants.

CRISPR/Cas screening

An approximately 600 bp PCR fragment *of OsAT10* (Os06g39390) was obtained from genomic DNA from all the *Osat10* mutants with primers p161 & p162 (see Supplementary Table S1). The fragments were sequenced at GATC-Eurofins with a nested primer p181. Plants with frameshift mutations were propagated further (Figure 1B). Plants were grown in a greenhouse at (30\u00B0C with added lights) T2 plants water was supplemented with 1 ml of Plant Magic Plus Bio-Silicon per litre of watering water. Leaf blades were harvested after 10 weeks and after 16 weeks watering was stopped and fully mature plants were left to senesce as natural, after which the inner dry stems and rice husks were collected. Plant material was freeze dried and powdered using a ball mill. Five T2 or T3 plants of each mutants were analysed as biological replicates in the following experiments. Total plant height was measured every second week.

Preparation of AIR

Alcohol insoluble residue (AIR) was prepared by adding 2 ml of water to 50 mg of powdered plant material in a 2 ml Eppendorf tube resuspending the powder in the water and shaking by inverting the tube for 30 s before the plant material was pelleted by centrifuging at 12,000RCF for 1 min, the water was removed and replaced with 2 ml of 96% ethanol, followed by two further ethanol washes before the residue was allowed to dry under ambient conditions.

Silica analysis

Plant material was freeze dried and powdered using a ball mill. Powdered plant material was pressed into disks with hydraulic press and measured using X-ray fluorescence according to (Reidinger et al., 2012).

Measurement of p-CA and FA by alkaline hydrolysis

Alkaline extraction of p-CA and FA was performed according to (Fry, 2000). Approximately 10 mg of dry plant material was ball milled in a 2 ml tube using metal beads and a ball mill grinder. One milliliter of 1 M NaOH was added and the samples were incubated shaking at 140 rpm at 30°C

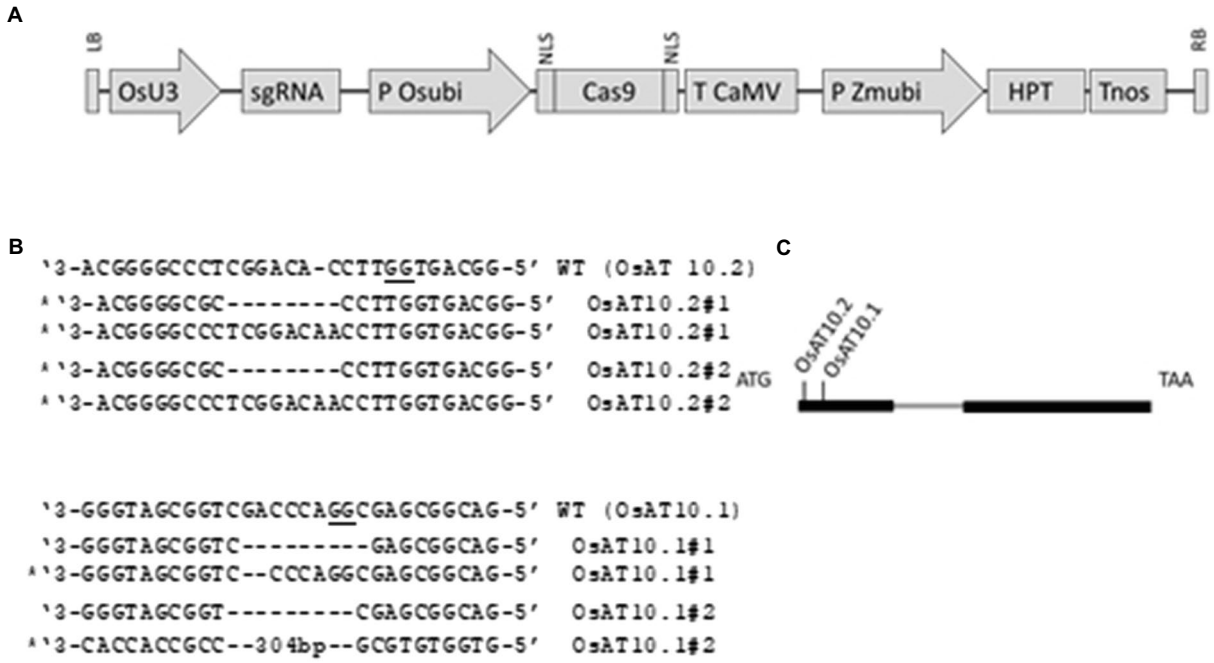


FIGURE 1
Constructs used and targets for CRISPR/Cas within OsAT10. **(A)** Schematic presentation of the T-DNA structure in the vectors utilized in this study. The expression of the sgRNA scaffold is driven by the rice U3 promoter (OsU3), expression of Cas9 by the rice ubiquitin promoter (P Osubi) and expression of hygromycin selection marker (HPT) by Maize ubiquitin promoter. Abbreviations: NLS, nuclear localization signal; Tnos, Nopaline terminator; T CaMV, Cauliflower Mosaic Virus (CaMV) 35S terminator; LB, Left border; RB, right border. **(B)** Indels present in the original lines, KO lines selected are marked and PAM is underlined. **(C)** Schematic of the OsAT10 gene structure and target site. The intron is marked with gray and the two exons are shown with thick black lines. The location of the two crisprs are marked.

overnight. The reaction was neutralized by addition of 100 µl 99% TFA and phase extracted twice using butanol. The butanol was evaporated and the residue resuspended in 200 µl MeOH. Samples were analyzed by LC/MS. Separation was carried out on an Acquity I class LC system (Waters United Kingdom, Elstree) using a BEH C18-1.7u 2.1x100mm (Waters), the mobile phases were A) 0.1% acetic acid in H_2O, and B) 0.1% acetic acid in acetonitrile. The gradient started at 20% B, and increased to 100% B over 3.25 min, where it remained for 0.15 min. The gradient returned to 20% B over 0.05 min, and re-equilibrated over 0.55 min. The total duration of the program was 4 min. The flow rate was 0.5 ml min^{-1}, and the injection volume 3 µl. The detector used was a TSQ Endura™ triple quadrupole mass spectrometer (Thermo Fisher Scientific, Altrincham, United Kingdom) with a HESI source, operated in negative MRM mode. Commercially purchased standards of *p*-CA and FA (Sigma) were used to construct standard curves in the range from 12 to 400 µM. The transitions were m/z 163.05 to 119.11, collision energy 15.3 V (*p*-CA) and m/z 193.03 to 134.06, collision energy 16.8 V (FA), with a 0.25 s cycle time, 1.5 mTorr collision gas pressure, ion transfer tube temperature of 333°C, vaporizer temperature of 317° C and a spray voltage of 3.5 kV. Q1 and Q2 resolutions were 1.2 Da. Data processing was performed using Thermo Xcalibur 4.0.27.10 QualBrowser and QuanBrowser software.

Measurement of p-CA-arabinose and FA-arabinose by mild acidolysis

Measurement of *p*-CA-arabinose (*p*-CA-ara) and FA-arabinose (FA-ara) was performed essentially according to (Lapierre et al., 2018). One millilitre of acidolysis reagent (dioxane, methanol and aq 2 M HCl 60:30:10, v/v/v) was added to 5–10 mg of AIR in a screwcap Eppendorf vial. The vials were incubated for 3 h at 80°C in an Eppendorf tube shaking at 1200 rpm. After cooling, 2 ml of water and 20 µl of 0.1 µg/µl *o*-coumaric acid in methanol was added. *o*-coumaric acid was used to correct for loss during the phase extraction, although this was minor. The diluted mixture was extracted with EtOAc. EtOAc was evaporated and the pellet dissolved in 150 µl of MeOH, which was analysed by HPLC using UV detection. 10 µl of sample was injected onto an XBridge BEH Shield RP18 Column, 130 Å, 3.5 µm, 1 mm × 100 mm (Waters Part Number: 186003149) using 1.23 ml/min flowrate and a binary gradient made up of 0.1% acetic acid and methanol. 0–60 min going from 5 to 65% MeOH, followed by a rinse 60–65 min from 65 to 90% methanol and then 65–70 min back to 5% methanol or 0–30 min going from 5 to 70% MeOH, followed by a rinse 30–35 min from 70 to 90% methanol and then 35–40 min back to 5% methanol. The eluent was monitored using a photodiode array detector and HCA and HCA-conjugates were determined by absorption at 320 nm. Data processing was performed using

Waters Empower Pro software. As an external standard we used a series of five dilution of *p*-CA and FA spanning 0.12 µg/µl and 0.14 µg/µl to 0.0075 µg/µl and 0.0088 µg/µl, respectively. Putative *p*-CA-ara and FA-ara peaks were extracted, evaporated, and resuspended in 100 µl MeOH to subjected to identification using MALDI-TOF/TOF. Acquisition was performed using a Bruker Ultraflex III equipped with a Nd:YAG smart beam laser with positive ionisation in reflectron mode post spotting samples 1:1 with 20 mg/ml 2,5-dihydroxybenzoic acid (DHB) matrix. MS^1 precursor ions unique to the samples and absent in a DHB blank, were manually selected for MS/MS fragmentation performed in LIFT mode without the introduction of a collision gas. Bruker flex Analysis software (version 3.3) was used to perform spectral processing and peak list generation. MS2 product ion spectra were manually inspected and annotated for analyte identification.

Monosaccharide analysis

Non-cellulosic monosaccharide analysis was performed using high-performance anion exchange chromatography (HPAEC; Carbopac PA-10; Dionex). Water extractable arabinoxylan was fractionated from approximately 40 mg of dry powdered plant material with 1 ml of water overnight. The extract was filtered through a 0.45 µm polytetrafluoroethylene (PTFE) filters and evaporated to obtain a pellet. Three milligram of AIR or the pellet from water extractable arabinoxylans were hydrolysed with 0.5 ml 2 M trifluoroacetic acid (TFA) for 4 h at 100°C, cooled to room temperature and evaporated completely. The pellet was resuspended in 200 µl of pure water, filtered through a 0.45 µm PTFE filter and separated by HPAEC as described in Jones et al. (2003). The monosaccharides were quantified by using an external calibration containing seven monosaccharide standards at 0.3, 0.6, and 1 µM (arabinose, fucose, galactose, glucose, mannose, rhamnose, and xylose), that were subjected to acid hydrolysis in parallel with the samples.

Saccharification assays

Powdered AIR was loaded into 96-well plates, using a custom-made robotic platform (Labman Automation, Stokesley, North Yorkshire, United Kingdom), and saccharification assays were performed according to Gomez et al. (2010) after water pretreatment. Enzymatic hydrolysis was carried out using an enzyme cocktail with a 4:1 ratio of Celluclast and Novozyme 188.

Nuclear magnetic resonance material preparation

Rice plant samples were prepared based on previously published methods (Kim and Ralph, 2010; Mansfield et al., 2012). Dried samples (*ca.* 600 mg) were finely milled using a Planetary Fritsch Pulverisette 7 Micro Mill spinning at 600 rpm with nylon vessels (45 ml) containing ZrO_2 ball bearings (14 mm × 2, 10 mm × 9). The total grinding time was 1 h 55 min consisting of 15 min intervals with 5 min interval breaks to avoid excessive heating. To produce lignin enriched fractions ball milled samples were treated with Cellic® CTec2 enzyme cocktail (Novozymes, Denmark, activity 120 FPU/ml) which contains a mixture of plant cell wall degrading enzymes. The milled samples (380 mg) were suspended in 50 mM sodium acetate buffer (10 ml, pH 5.5) and the Cellic® CTec2 enzyme cocktail was added (126 µl, 40 FPU/g substrate). The mixture was incubated on a rotary shaker at 50°C for 5 days. The insoluble residue was collected by centrifugation (4,000 rpm, 2 min), washed twice with water and then freeze dried to give fine powders. Yield from WT rice husk was 151 mg (40%) and from the *Osat10* mutant 141 mg (37%). Lignin enriched fractions were transferred directly to a 5 mm Nuclear Magnetic Resonance (NMR) tube (75 mg) and roughly distributed up the sides of the tube. Pre-mixed DMSO-d_6/pyridine-d_5 (4,1) was directly added into the NMR tube and vigorously shaken for a few seconds to mix the contents before being placed in a sonicator bath for 1 h.

NMR acquisition and processing

NMR spectra were acquired on a 600 MHz Bruker Avance III HD 600 MHz spectrometer with TCI cryoprobe. 2D ^1H–^{13}C HSQC spectra were acquired using a standard Bruker pulse program (hsqcetgpsisp2.2). The NMR spectra were acquired from 12 to −2 ppm in F2 (1H) using 2048 data points for an acquisition time (AQ) of 122 ms, an interscan delay (D1) of 1 s, 47–137 ppm in F1 (^{13}C) using 144 increments of 16 scans. The spectrum was processed using squared cosine bell in both dimensions and LPfc linear prediction (18 coefficients) in F1. Interactive integrations of contours in 2D HSQC plots were carried out using Bruker's TopSpin 3.6 (Windows) software, as was all data processing. Semi-quantitative analysis of the lignin components was accomplished by integration and comparison of the S2/6 and G2 resonances in the aromatic region and the signals for each structural unit in the oxygenated alkyl region. Abundances of the lignin linkages are expressed as the number of linkages per 100 C_9 units. Relative amounts of *p*-CA and FA are expressed as a simple percentage relative to the lignin aromatics. Assignments of NMR spectra are made by comparison to previously published data (Kim et al., 2008; Mansfield et al., 2012; Lam et al., 2017, 2019; Takeda et al., 2019).

Results

Generating CRISPR/Cas knockout mutants

We designed two RNA guide sequences (OsAT10.1 and OsAT10.2; Figure 1) using sgRNA2.0 scorer (Chari et al., 2015, 2017) while taking into account the potential off targets and sites

with restriction enzymes using CRISPRdirect (Kawahara et al., 2013; Naito et al., 2015). Multiple T1 mutants for both guides were obtained. From these two transformations with biallelic mutations we selected two as independent lines for each guide RNA. By having two knockouts made using different guide RNAs we greatly reduced the possibility of the observed phenotypes arising from off-target effects. We identified homozygous T2 plants with frame-shift mutations preventing the protein translation by sequencing PCR products generated across the target sites. The mutant plants did not show any obvious visible phenotype. However, when we measured height in third-generation mutants grown under greenhouse conditions they grew slightly slower than the corresponding wild type plants with $p < 0.05$ *via* Student's *t*-test Figure 2.

Osat10 mutants lack AX-associated p-CA, which is mostly found in the husks of wild type plants

We initially measured total *p*-CA and FA content in rice stems, leaves and husks, using alkaline hydrolysis followed by analysis by LC/MS in 10–16 weeks old T2 rice mutants and wild type plants (Figure 3A). This procedure releases all exclusively ester-linked HCAs from both AX and lignin. In stems there were no discernible differences between wild type and the *Osat10* knockout mutants in terms of total ester-linked HCA content. However, there was a significant reduction in ester-linked *p*-CA content in the mutant husks and a smaller reduction was apparent in mutant leaves compared to wild type. This is compatible with published expression data for *OsAT10* at RiceXPro (Sato et al., 2011, 2013; Supplementary Figure S4) that suggests *OsAT10* is almost exclusively expressed in husks. While leaf expression appears nearly absent this is agreed by data presented by Piston et al. (2010) and RNAseq data from the Rice Genome Annotation Project (Ouyang et al., 2007).

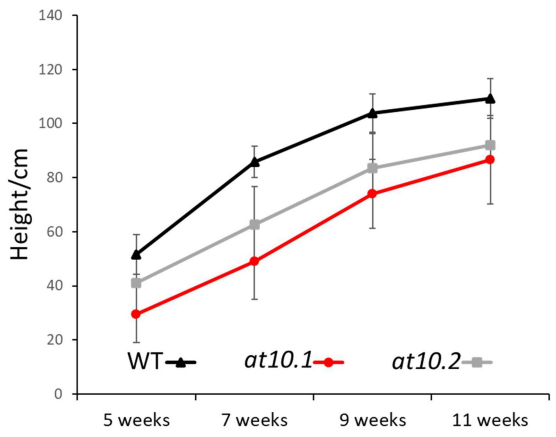


FIGURE 2
Growth of rice plants studied. Total height was measured at different stages of growth. Wild Type (Black), OsAT10.1 (Gray) and OsAT10.1 (Red). Error bars represent standard deviation.

Because ester-linked *p*-CA can be associated with both AX and lignin we used the method of Lapierre et al. (2018) to extract and measure HCAs and their arabinose conjugates in order to distinguish the AX associated HCAs as their arabinose conjugates from those associated with lignin (Figure 3B). This method provides a measure of the relative amount of *p*-CA and FA linked to arabinose, as well as *p*-CA and FA. Although, this amount only contains a partial contribution from arabinose bound *p-CA*. Using HPLC methods, we were able to identify the *p*-CA and FA peaks using standards, as well as the peaks of methylated *p*-CA and FA by running the standards through the acidolysis process. While we did not have the 5-O-*p*-coumaroyl- and 5-O-feruloyl-L-arabinofuranoside standards, the two peaks at 32.9 or 20.1 min and 34.3 or 20.6 min have similar UV–vis spectra to the *p*-CA and FA, respectively. To confirm the identification of these peaks, we subjected them to the mild alkaline hydrolysis treatment and re-analysed them, yielding peaks corresponding to *p*-CA and FA, respectively. Further confirmation of the peak identities was obtained using MALDI-TOF/TOF, confirming their identity as *p*-CA-Ara and FA-Ara derivatives (Supplementary Figure S1).

In wild type husks, we found that the ratio of *p*-CA-ara to FA-ara was about 6:1, making *p*-CA likely the major HCA found in AX in rice husks. In contrast, the *Osat10* knockout mutants were almost completely devoid of *p*-CA-ara (Figures 3B,C). This was corroborated by analysis of multiple rice tissues revealing that the mutants lack *p*-CA-ara in all major tissues. Our analyses reveal that *p*-CA-ara is a minor component in wild type rice leaves, almost absent in wild type stems, and essentially absent in all tissues of the *Osat10* mutants (Figures 3B,C). The *Osat10* mutants exhibit a compensatory increase in FA content of AX in rice husk mutants with roughly twice the amount seen in wild type husks (Figure 3B). There is, however, a small peak remaining at the *p*-CA-ara peak in the *Osat10* mutant husks.

It has been suggested that *p*-CA might aid in the nucleation of lignin by performing radical transfer to sinapyl alcohol (Takahama and Oniki, 1994). To determine if there were any related effects on lignin in rice husks of the mutant plants, we initially used gel-state 2D-HSQC NMR to examine the cell wall composition. Gel-state 2D-NMR was performed according to (Kim and Ralph, 2010; Mansfield et al., 2012) to measure the aromatic units. We observed a 40% decrease in *p*-CA in the knockout mutant, in good agreement with the changes observed by alkaline hydrolysis (Figure 4). To explore if these changes were associated with the lignin or AX fractions, we treated ball milled rice husk with a biomass degrading enzyme cocktail (Cellic CTec2) to create a lignin enriched residue, depleted in cellulose and hemicelluloses which was also subjected to 2D-HSQC NMR analysis. In contrast to the non-digested husks, no change in *p*-CA was apparent in the lignin enriched residue from the *Osat10* mutants (Supplementary Figure S3). These results confirm the almost complete loss *p*-CA in the AX fraction of husk cell walls in the mutant plants, with no associated discernible changes in lignin composition.

10.3389/fpls.2022.926300

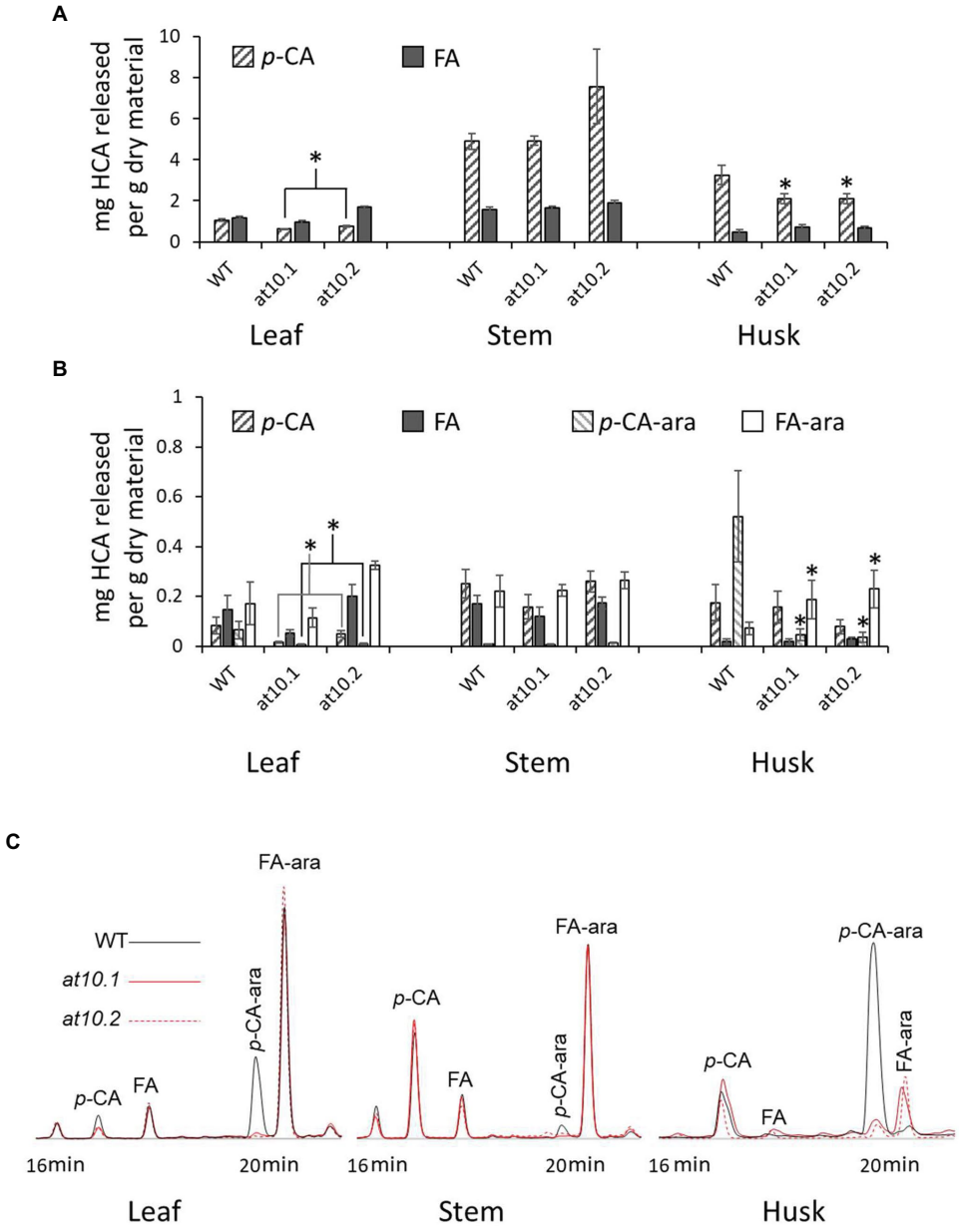


FIGURE 3
Hydroxycinnamic acids contents and identification in *Osat10*. **(A)** Determination of total ester-linked ferulic acid (Grey bars) and *p*-coumaric acid (Hatched bars) content of stems, leaves and husks as obtained from alkaline hydrolysis per gram of dry material. **(B)** HPLC-UV determination of *p*-CA (Hatched bars), FA (Grey bars), *p*-CA-ara (Hatched bars, lighter) and FA-ara (White bars) released by mild acidolysis of rice leaves, stems and husks from *OsAT10.1, OsAT10.2* and WT. Error bars represent standard deviation. *Significant difference at $p<0.05$ *via* Student's t test. **(C)** HPLC chromatograms from leaves, stems and husks, showing UV absorption at 320nm. Wild type is black OsAT10.1 is red and OsAT10.2 is dotted. Compounds: *p*-CA-ara; methyl 5-O-E-*p*-coumaroyl-L-arabinofuranoside, FA-ara; methyl 5-O-E-feruloyl-L-arabinofuranoside, FA; Ferulic acid, *p*-CA; *p*-Coumaric Acid.

Loss of AX p-Coumaroylation of AX is not accompanied by alterations in the matrix monosaccharide composition, silica content, or cell wall digestibility in rice husk

We sought to establish if the loss of *p*-CA on AX had any impact on the composition or enzyme digestibility of rice husk cell walls. To analyze the monosaccharide composition of cell wall polymers, we treated AIR obtained from rice husks stems and leaves with TFA to hydrolyze non-crystalline polysaccharides, and analyzed the resulting monosaccharides by HPAEC. We observed no significant changes in the monosaccharide composition in the *Osat10* mutants compared to those from wild type (Figure 5). We also investigated whether the loss of *p*-CA in husk cell walls, led to any alterations in the

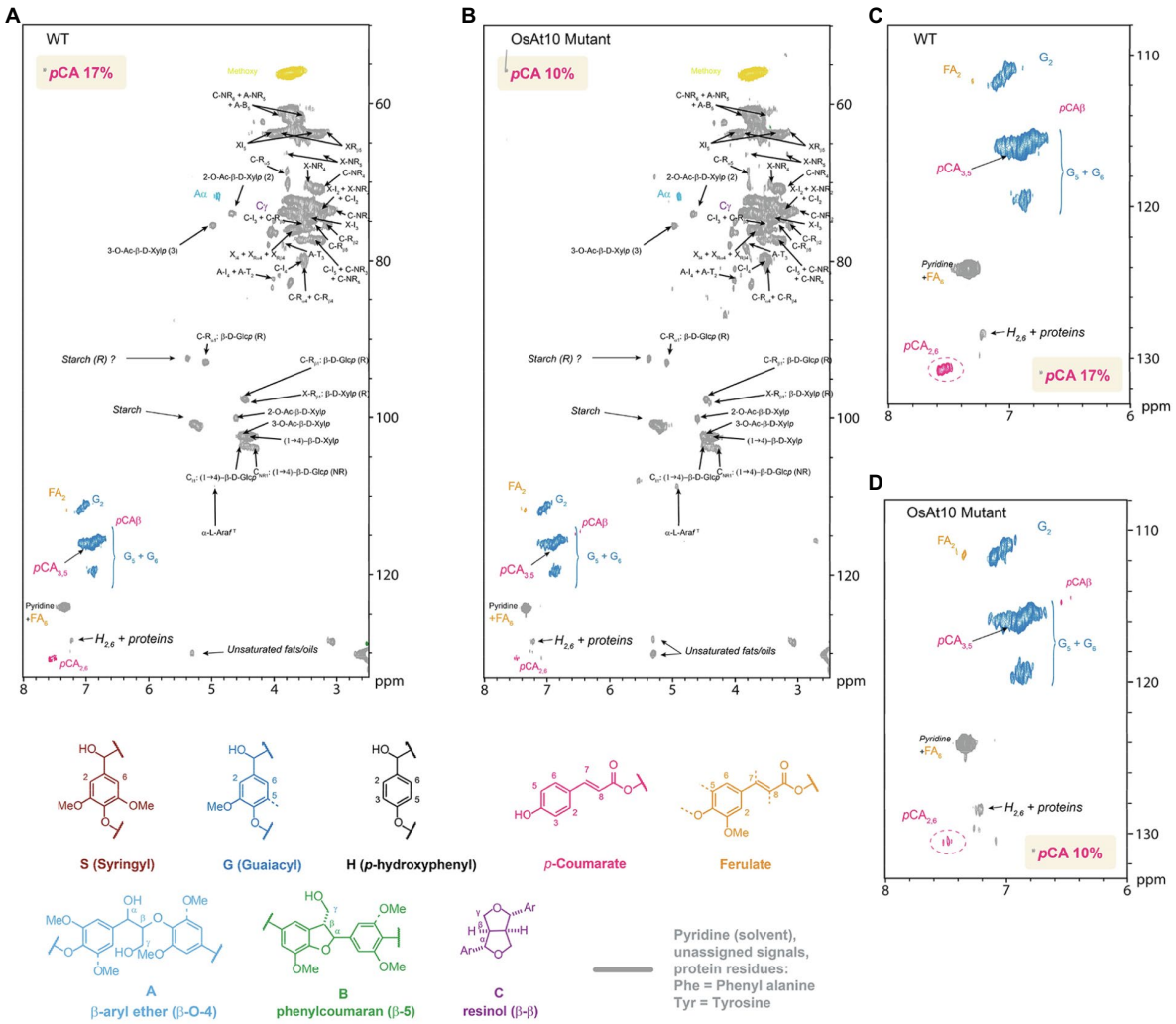


FIGURE 4
Whole cell gel-state 2D-NMR heteronuclear single-quantum coherence (HSQC) analysis of rice husk cell walls from **(A–C)** the WT control and **(B–D)** OsAt10 knockout mutant. The analytical data are from volume integrals of correlation peaks representing reasonably well-resolved C/H pairs in similar environments; thus, they are from G_2 and p-$CA_{2/6}$ in the aromatic region, with corrections applied for units that have two C/H pairs per unit. The p-$CA_{2/6}$ abundances are reported on a per 100 C9 unit basis relative to G=100 C9 units. The small amounts of S units present in the samples were too low to be accurately quantified in these cases.

extractability of wall polysaccharides, by changing the chemical properties of AX. To tests this, we subjected rice husks and leaves to hot water extraction and quantified the sugars extracted after a TFA treatment similar to the one applied to AIR samples (Supplementary Figure S2). No significant differences in water extractability of polysaccharides were observed between mutant and wild type. To measure total lignin content, we treated AIR from leaves, husks and stems with acetyl bromide in glacial acetic acid in order to solubilize the lignin, after which lignin content was determined by absorption at 280 nm (Figure 6). As with wall monosaccharide composition, no significant difference was observed between the wild type and both mutants.

Bartley et al. (2013) showed that overexpression of *OsAT10* increases *p*-CA levels in rice stem AX, with a concomitant decrease in FA and that this led to increased enzymatic digestibility on the straw. We investigated whether the lack of *p*-CA substitutions on AX in rice husks in the *Osat10* mutants resulted in any changes in enzymatic digestibility. These experiments used an automated platform to perform enzymatic saccharification assays on dried biomass powder using a commercial cellulase cocktail (Gomez et al., 2010). We did not observe any significant differences between the digestibility of mutant and wild type husks, leaves or stems in these experiments (Figure 7A).

The husks (lemma) of rice and other cereal seeds play a protective role and are notable for having high contents of silica and *p*-CA (Hunt et al., 1984; Gao et al., 2018), and connections between aromatic cell wall components and silica have been suggested by some authors (He et al., 2015; Zhang et al., 2015;

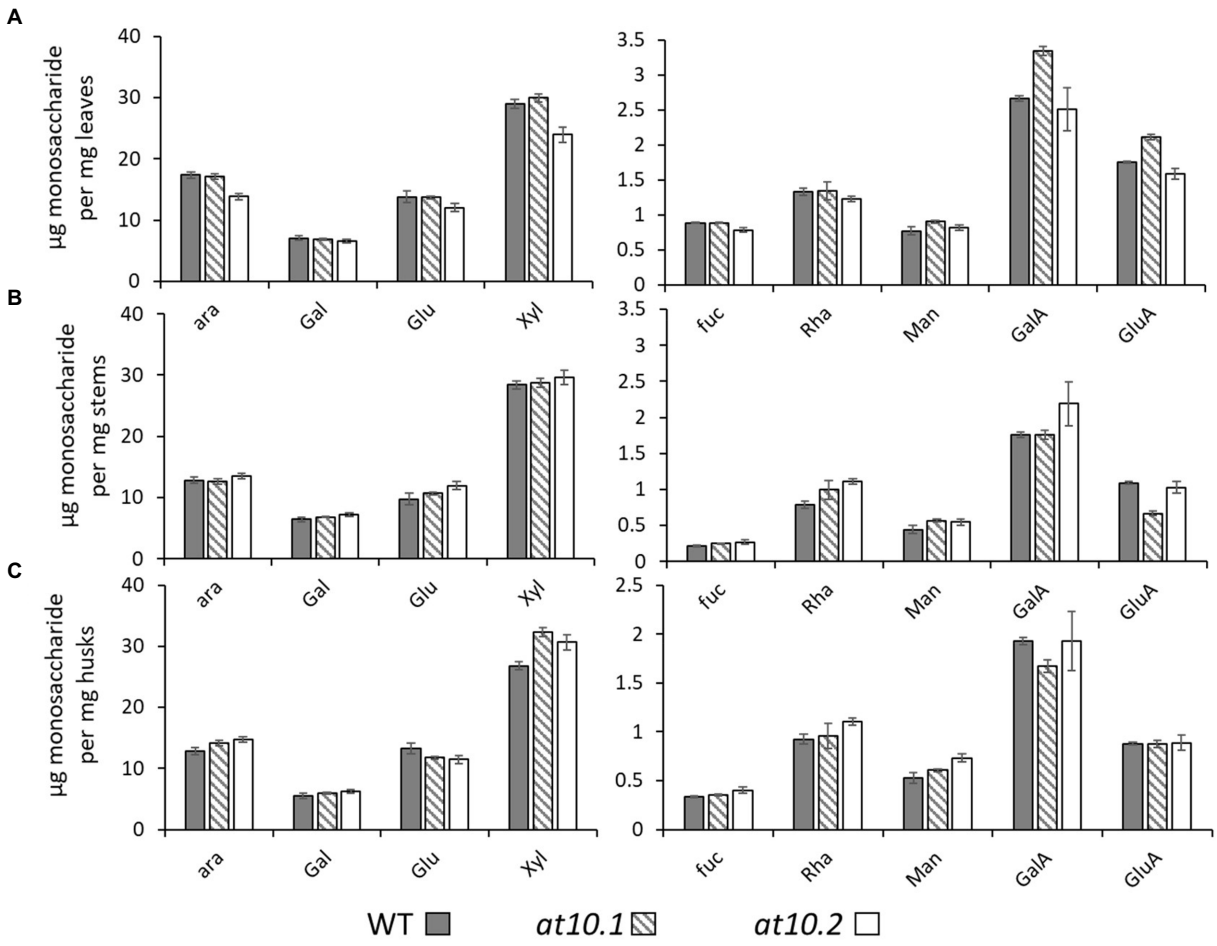


FIGURE 5
Monomeric sugar analysis. µg of Monomeric sugars pr. mg dry material released by TFA. Material was taken from leaves **(A)**, stems **(B)** and husks **(C)**. WT is shown as Gray bars, OsAT10.1 Hatched bars and OsAT10.2 White bars. Error bars represent standard deviation. Student's t-test was run between WT and both mutants to check for significance.

Soukup et al., 2017). We measured the levels of Si in rice husks from mutant and wild type plants, but found no significant difference between the two (Figure 7B).

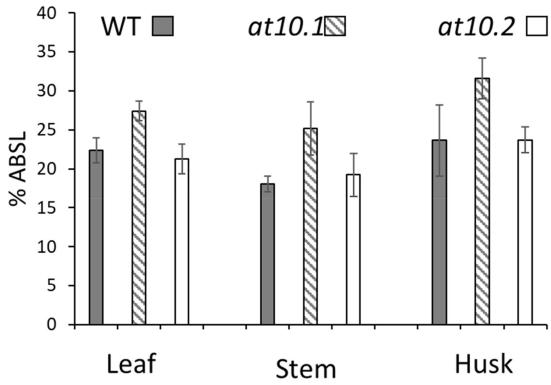


FIGURE 6
Acetyl bromide soluble lignin analysis. The lignin content of WT (Gray bars), Osat10.1 (Hatched bars), Osat10.2 (White bars) as mass % acetyl bromide soluble lignin. The bars indicate the standard errors. Student's t-test was run between WT and both mutants to check for significance.

Discussion

MCBATs are responsible for the incorporation of HCAs into plant cell walls, including *p*-CA and FA into lignin and into AX. The roles of specific MCBAT genes are starting to be defined, but off-target effects in RNAi studies, and reliance on the analysis of total *p*-CA and FA content (rather than lignin- and AX-associated HCAs) has led to ambiguities regarding specific roles. Previous studies implicated OsAT10 as being responsible for *p*-CA incorporation into AX (Bartley et al., 2013; Li et al., 2018). However, these studies involved overexpression of the OsAT10 gene resulting in increased AX- *p*-CA content in rice and switchgrass, but did not demonstrate the role of the gene in normal plants. We used a gene knockout approach to study the role of OsAT10 in rice, and

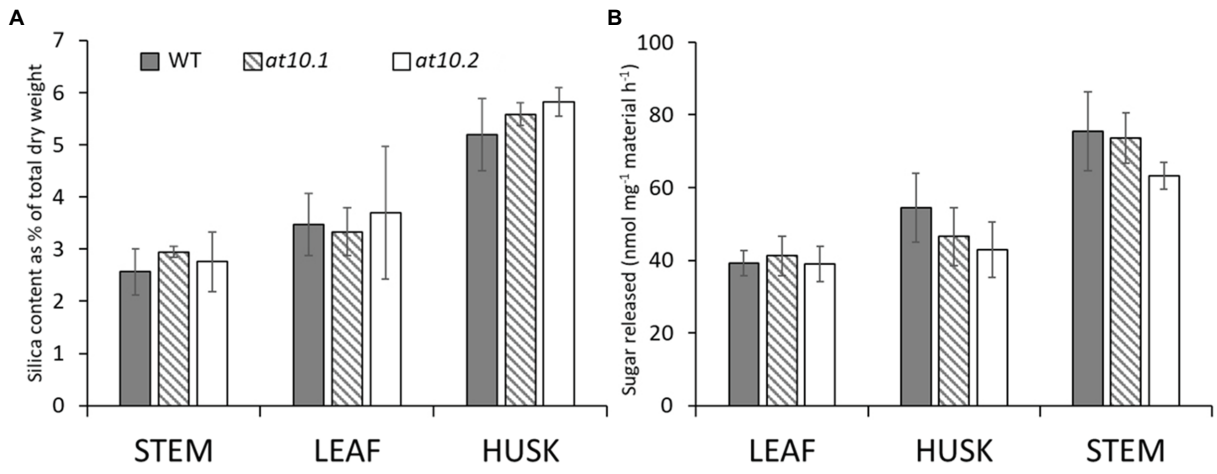


FIGURE 7
Silica and saccharification analysis. **(A)** Analysis of silica content measured by X-ray fluorescence and **(B)** saccharification analysis of leaves, husks and stems from wild type (grey bars) and OsAT10 mutant OsAT10.1 (Hatched bars) OsAT10.2 (Empty bars) The ground material for saccharification analysis was digested with a commercial cellulase for 8h at 50°C following a mild water pre-treatment. Student's t-test was run between WT and both mutants to check for significance.

examined the levels of *p*-CA and FA in both lignin and AX to allow unambiguous assignment of the role of OsAT10 in rice.

The observation that lowering AX bound FA in grass cell walls decrease biomass recalcitrance (de Souza et al., 2018, 2019) has increased the interest to understand the role of each family member of the BAHD acyltransferase clade of genes. In rice, overexpression and suppression have shed light upon the role of some members of this gene family. Indeed, the roles of *OsAT5*, *OsAT7, OsAT4,* and *OsAT10* have been studied and their roles partially established. Bartley et al. (Bartley et al., 2013) studied the role of *OsAT10* using a promoter tag overexpression system and showed that the overexpression of this gene increased the incorporation of *p*-CA to AX, reduced the amount of FA linked to AX and decreased the recalcitrance of the biomass. In the present work we used a CRISPR/Cas9 approach to generate lack-of-function genes for *OsAT10* allowing for detailed analysis if the gene function without the possible side effects of gene overexpression.

OsAT10 is likely singularly responsible for the addition of p-CA to GAX in rice

Our experiments allow an unambiguous appraisal of the role of the OsAT10 in rice plants. In our mutants, *p*-CA-ara appears to be almost completely absent in all tissues examined, indicating that this gene is responsible for most, or all, of the addition of *p*-CA to AX in Nipponbare rice. By quantifying the total *p*-CA and the *p*-CA esterified to arabinose, we show that OsAT10 is not responsible for the addition of *p*-CA to rice lignin. Analysis of the total *p*-CA in rice husks indicates a decrease within the same range that we find using a 2D NMR approach. However, once the arabinoxylan has been removed by Cellic Ctec2 treatment to create a lignin enriched fraction we do not see changes in lignin associated *p*-CA between wild type and mutants. This further demonstrates *OsAT10* function is likely to exclusively transfer *p*-CA onto arabinose for incorporation into AX.

The synthesis of AX occurs in the Golgi, and a cytosolic localisation of BADH activity would be required for the acylated product to be transported into the Golgi (Myton and Fry, 1994; Fry, 2004). UDP-arabinofuranose is synthesised in the cytosol and has therefore been proposed as the substrate for both OsAT10 and OsAT9 (Buanafina, 2009; Rautengarten et al., 2011; Seifert, 2018), requiring a glycosyltransferase to incorporate the product onto AX in the Golgi. The lack of subcellular location signal peptide in *OsAT10* indicating a cytosolic localization, is coherent with the proposed biosynthetic pathway and de Souza et al. (de Souza et al., 2018) claims the *in-situ* localization green fluorescence protein fusion with the wheat ortholog is cytoplasmic. The BAHD enzymes have been found to be cytoplasmic (D'Auria, 2006). The role of OsAT10 MCBAT is underpinned by having close homologs only in the commelinids and the presence of AX conjugated *p*-CA only in grasses (Li et al., 2018). Although data presented by (Piston et al., 2010) indicates that *OsAT10* is significantly expressed in stems, the Rice Xpro database (Sato et al., 2011, 2013) suggests that *OsAT10* is mainly expressed in rice husks. In the present work, we found minimal amounts of *p*-CA-ara in stems of wild type plants. Interestingly, we observed an increase in FA-ara in our *Osat10* mutant husks. (de Souza et al., 2018, 2019) reported an increase in *p*-CA-ara upon down regulation of *OsAT9* homologs in *Setaria* that add FA to AX. It seems possible that, *OsAT10* and *OsAT9* compete for the available UDP sugar substrate and/or accessible arabinose on AX in stems, and when either *OsAt10* or *OsAT9* are downregulated, one activity takes prevalence over the other.

Loss of p-CA-AX does not affect cell wall digestibility

Osat10 mutants do not show any change in saccharification with a commercial cellulase cocktail. This supports the hypothesis that the increased cell wall digestibility seen in OsAT10 overexpressor lines (Bartley et al., 2013; Li et al., 2018) was caused by a reduction in the amount of FA in AX, a side effect of the gene overexpression. The loss of *p*-CA-ara in the OsAT10 mutants also had no impact on the digestibility of the biomass in leaves, stems or husks. Our results contrast with those recently published by Mota et al. (Mota et al., 2021), which showed that gene suppression of *SvBAHD05* in *S. viridis*, resulted in decreased levels of *p*-CA-AX in most tissues, and that this was accompanied by increased biomass digestibility. While species differences might account for this apparent contradiction, we note that the gene suppression experiments in *S. viridis* only resulted in partial suppression of the expression of *SvBAHD05* and this was accompanied by only partial reduction in the amount of *p*-CA-AX in the tissues, compared to our gene knockout experiments with almost complete loss of *p*-CA-AX. A significant reduction in *p*-CA-AX was also seen in the rice xylosyltransferase mutant *xax1* (Chiniquy et al., 2012) indicative of how other enzymes may influence *p*-CA-AX content. In addition, some small reductions in the amount of FA-AX are also evident in the work published by Mota et al. (Mota et al., 2021), which may be due to off-target effects on other MCBATs in *S. viridis,* or to promiscuous activity of the encoded enzyme. It seems possible that the small increases in biomass digestibility seen in these experiments in *S. viridis*, may be the result of the slight reduction in FA-AX that they observed rather than due to the reduction in *p*-CA-AX. In contrast, our work in rice resulted in complete loss of function of the *OsAT10* gene, resulting in almost complete loss of *p*-CA-AX, and no significant change in biomass digestibility. We may observe a small decrease in the growth of the mutant plants. We are not sure if this is indeed due to the plant compensating for the loss of OsAT10 or a result of the plant experiencing a fitness loss from being transformed with Cas9.

OsAt10 knock out mutants show no changes in lignin composition and silica content

Our results show that *p*-CA is a very minor component of AX in most rice tissues, with the exception of rice husks. Rice husks provide a protective layer to the seed and these tissues are notably rich in *p*-CA-ara and silica. The loss of *p*-CA-AX in the *Osat10* mutant rice husk had no impact on silica content and lignin composition, indicating no notable interactions between these wall components. (Hatfield et al., 2008) studied the possible role of *p*-CA as an electron donor in lignin polymerisation but found no effect of free *p*-CA *in vitro*. This lack of effect could be due to either the low AX conjugated *p*-CA concentration or to the low mobility of *p*-CA within the cell wall to contribute as an electron donor. However, other compounds, such as coniferyl alcohol, have been shown to enhance sinapyl alcohol oxidation (Takahama and Oniki, 1997). Our measurements also show that while FA is the predominant AX modification in most rice tissues, *p*-CA is the major HCA in AX in husks. This is in agreement with observations in barley glumes, which can be considered anatomically similar to rice husks, and also have *p*-CA as the predominant HCA in AX (Lapierre et al., 2018).

The silica content and the silica cell morphology (Data not shown) in stem and husks was not changed by the loss of *p*-CA-ara. Several other cell wall compounds have been implicated in silica deposition including, e.g., callose (Guerriero et al., 2018), homogalacturonan (Leroux et al., 2013) and hemicellulose (He et al., 2015). Although it is possible that the total amount of silica is mostly driven by the uptake and not dependant on a binding partner, a recent study has proposed a specialized protein (Kumar et al., 2021) as being responsible for silica precipitation in sorghum silica cells.

Although we did not examine the other potential physiological roles for AX associated *p*-CA, there is abundant evidence that they are a key component of abiotic and biotic stress responses. They have been implicated as fungal pathogens and insect deterrents (Santiago et al., 2007, 2008; Buanafina, 2009; Lanoue et al., 2010) and as UV-protectant and free radical quenchers (Ruhland et al., 2005; Stelzner et al., 2019). The higher content of AX conjugate *p*-CA in husks, pollen, and fruits, relative to other tissues supports their protective role in plant organs that play an essential role in reproduction.

In summary, our work shows that OsAT10 is responsible for AX coumaroylation in rice and may be the only gene conferring this activity in aerial tissues. Our results further show that AX coumaroylation has little or no impact on cell wall digestibility or on lignin or silicon content. This study helps clarify the roles of members of the MCBATs in cell wall biology.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

SM designed experiments and genetic constructs and analyzed the experimental data. CL performed the NMR analysis of the samples. NO performed the HPLC analysis of free and conjugated HCA. RS performed saccharification and monosaccharide analysis. AD identified HCAs by MS. LG and SM-M conceived and coordinated the project, experiments, and data analysis. SM, LG, and SM-M wrote the manuscript with contributions from all authors. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by funding from the BBSRC through grants BB/P022499/1 and BB/N013689/1. The York Centre of Excellence in Mass was supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/K039660/1 and EP/M028127/1). CL thanks the Leverhulme Trust Early Career Fellowship (ECF-2018-480) and the University of St Andrews.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.926300/full#supplementary-material

References

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 Saccharification1[C][W][OA]. *Plant Physiol.* 161, 1615–1633. doi: 10.1104/pp.112.208694

Boz, H. (2015). P-Coumaric acid in cereals: presence, antioxidant and antimicrobial effects. *Int. J. Food Sci. Technol.* 50, 2323–2328. doi: 10.1111/ijfs.12898

Buanafina, M. M. (2009). Feruloylation in grasses: current and future perspectives. *Mol. Plant* 2, 861–872. doi: 10.1093/mp/ssp067

Buanafina, M. M. O., Fescemyer, H. W., Sharma, M., and Shearer, E. A. (2016). Functional testing of a PF02458 homologue of putative rice arabinoxylan feruloyl transferase genes in *Brachypodium distachyon*. *Planta* 243, 659–674. doi: 10.1007/s00425-015-2430-1

Chari, R., Mali, P., Moosburner, M., and Church, G. M. (2015). Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nat. Methods* 12, 823–826. doi: 10.1038/nmeth.3473

Chari, R., Yeo, N. C., Chavez, A., and Church, G. M. (2017). sgRNA scorer 2.0—a species independent model to predict CRISPR/Cas9 activity. *ACS Synth. Biol.* 6, 902–904. doi: 10.1021/acssynbio.6b00343

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

Currie, H. A., and Perry, C. C. (2007). Silica in plants: biological, biochemical and chemical studies. *Ann. Bot.* 100, 1383–1389. doi: 10.1093/aob/mcm247

D'Auria, J. C. (2006). Acyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol.* 9, 331–340. doi: 10.1016/j.pbi.2006.03.016

de Oliveira, D. M., Finger-Teixeira, A., Rodrigues Mota, T., Salvador, V. H., Moreira-Vilar, F. C., Correa Molinari, H. B., et al. (2015). Ferulic acid: a key component in grass lignocellulose recalcitrance to hydrolysis. *Plant Biotechnol. J.* 13, 1224–1232. doi: 10.1111/pbi.12292

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

de Souza, W. R., Pacheco, T. F., Duarte, K. E., Sampaio, B. L., de Oliveira Molinari, P. A., Martins, P. K., et al. (2019). Silencing of a BAHD acyltransferase in sugarcane increases biomass digestibility. *Biotechnol. Biofuels* 12:111. doi: 10.1186/s13068-019-1450-7

Engler, C., Youles, M., Gruetzner, R., Ehnert, T. M., Werner, S., Jones, J. D. G., et al. (2014). A Golden Gate modular cloning toolbox for plants. *ACS Synth. Biol.* 3, 839–843. doi: 10.1021/sb4001504

Fry, S. (2000). *The Growing Plant Cell Wall: Chemical And Metabolic Analysis*. London: Longman, 3.

Fry, S. C. (2004). Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. *New Phytol.* 161, 641–675. doi: 10.1111/j.1469-8137.2004.00980.x

Gao, Y., Guo, X., Liu, Y., Fang, Z., Zhang, M., Zhang, R., et al. (2018). A full utilization of rice husk to evaluate phytochemical bioactivities and prepare cellulose nanocrystals. *Sci. Rep.* 8:10482. doi: 10.1038/s41598-018-27635-3

Gomez, L. D., Whitehead, C., Barakate, A., Halpin, C., and McQueen-Mason, S. J. (2010). Automated saccharification assay for determination of digestibility in plant materials. *Biotechnol. Biofuels* 3:23. doi: 10.1186/1754-6834-3-23

Goto, M., Ehara, H., Karita, S., Takabe, K., Ogawa, N., Yamada, Y., et al. (2003). Protective effect of silicon on phenolic biosynthesis and ultraviolet spectral stress in rice crop. *Plant Sci.* 164, 349–356. doi: 10.1016/S0168-9452(02)00419-3

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

Guerriero, G., Stokes, I., and Exley, C. (2018). Is callose required for silicification in plants? *Biol. Lett.* 14:20180338. doi: 10.1098/rsbl.2018.0338

Hatfield, R., Ralph, J., and Grabber, J. H. (2008). A potential role for sinapyl p-coumarate as a radical transfer mechanism in grass lignin formation. *Planta* 228:919. doi: 10.1007/s00425-008-0791-4

He, C., Ma, J., and Wang, L. (2015). A hemicellulose-bound form of silicon with potential to improve the mechanical properties and regeneration of the cell wall of rice. *New Phytol.* 206, 1051–1062. doi: 10.1111/nph.13282

He, C., Wang, L., Liu, J., Liu, X., Li, X., Ma, J., et al. (2013). Evidence for 'silicon' within the cell walls of suspension-cultured rice cells. *New Phytol.* 200, 700–709. doi: 10.1111/nph.12401

Hunt, L. P., Dismukes, J. P., Amick, J. A., Schei, A., and Larsen, K. (1984). Rice hulls as a raw material for producing silicon. *J. Electrochem. Soc.* 131, 1683–1686. doi: 10.1149/1.2115937

Ikram, S., Huang, L. Y., Zhang, H., Wang, J., and Yin, M. (2017). Composition and nutrient value proposition of brewers spent grain. *J. Food Sci.* 82, 2232–2242. doi: 10.1111/1750-3841.13794

Ishii, T., Hiroi, T., and Thomas, J. R. (1990). Feruloylated xyloglucan and p-coumaroyl arabinoxylan oligosaccharides from bamboo shoot cell-walls. *Phytochemistry* 29, 1999–2003. doi: 10.1016/0031-9422(90)85055-K

Jones, L., Milne, J. L., Ashford, D., and McQueen-Mason, S. J. (2003). Cell wall arabinan is essential for guard cell function. *Proc. Natl. Acad. Sci.* 100, 11783–11788. doi: 10.1073/pnas.1832434100

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

Kawahara, Y., de la Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., et al. (2013). Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6:4. doi: 10.1186/1939-8433-6-4

Kim, H., and Ralph, J. (2010). Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d6/pyridine-d5. *Org. Biomol. Chem.* 8, 576–591. doi: 10.1039/B916070A

Kim, H., Ralph, J., and Akiyama, T. (2008). Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d6. *Bioenergy Res.* 1, 56–66. doi: 10.1007/s12155-008-9004-z

Kumar, S., Adiram-Filiba, N., Blum, S., Sanchez-Lopez, J. A., Tzfadia, O., Omid, A., et al. (2021). Siliplant1 protein precipitates silica in sorghum silica cells. *J. Exp. Bot.* 72:6672. doi: 10.1093/jxb/erab333

Lam, P. Y., Tobimatsu, Y., Matsumoto, N., Suzuki, S., Lan, W., Takeda, Y., et al. (2019). OsCAldOMT1 is a bifunctional O-methyltransferase involved in the biosynthesis of tricin-lignins in rice cell walls. *Sci. Rep.* 9:11597. doi: 10.1038/s41598-019-47957-0

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

Lanoue, A., Burlat, V., Henkes, G. J., Koch, I., Schurr, U., and Röse, U. S. R. (2010). De novo biosynthesis of defense root exudates in response to Fusarium attack in barley. *New Phytol.* 185, 577–588. doi: 10.1111/j.1469-8137.2009.03066.x

Lapierre, C., Voxeur, A., Karlen, S. D., Helm, R. F., and Ralph, J. (2018). Evaluation of Feruloylated and p-Coumaroylated Arabinosyl units in grass Arabinoxylans by Acidolysis in Dioxane/methanol. *J. Agric. Food Chem.* 66, 5418–5424. doi: 10.1021/acs.jafc.8b01618

Leroux, O., Leroux, F., Mastroberti, A. A., Santos-Silva, F., van Loo, D., Bagniewska-Zadworna, A., et al. (2013). Heterogeneity of silica and glycan-epitope distribution in epidermal idioblast cell walls in *Adiantum raddianum* laminae. *Planta* 237, 1453–1464. doi: 10.1007/s00425-013-1856-6

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

Mansfield, S. D., Kim, H., Lu, F., and Ralph, J. (2012). Whole plant cell wall characterization using solution-state 2D NMR. *Nat. Protoc.* 7, 1579–1589. doi: 10.1038/nprot.2012.064

Marita, J. M., Hatfield, R. D., Rancour, D. M., and Frost, K. E. (2014). Identification and suppression of the p-coumaroyl CoA:hydroxycinnamyl alcohol transferase in *Zea mays* L. *Plant J.* 78, 850–864. doi: 10.1111/tpj.12510

Mitchell, R. A. C., Dupree, P., and Shewry, P. R. (2007). A novel bioinformatics approach identifies candidate genes for the synthesis and Feruloylation of Arabinoxylan. *Plant Physiol.* 144, 43–53. doi: 10.1104/pp.106.094995

Mnich, E., Bjarnholt, N., Eudes, A., Harholt, J., Holland, C., Jørgensen, B., et al. (2020). Phenolic cross-links: building and de-constructing the plant cell wall. *Nat. Prod. Rep.* 37, 919–961. doi: 10.1039/C9NP00028C

Mota, T. R., Souza, W. R., Oliveira, D. M., Martins, P. K., Sampaio, B. L., Vinecky, F., et al. (2021). Suppression of a BAHD acyltransferase decreases p-coumaroyl on arabinoxylan and improves biomass digestibility in the model grass *Setaria viridis*. *Plant J* 105, 136–150. doi: 10.1111/tpj.15046

Mueller-Harvey, I., Hartley, R. D., Harris, P. J., and Curzon, E. H. (1986). Linkage of p-coumaroyl and feruloyl groups to cell-wall polysaccharides of barley straw. *Carbohydr. Res.* 148, 71–85. doi: 10.1016/0008-6215(86)80038-6

Myton, K. E., and Fry, S. C. (1994). Intraprotoplasmic feruloylation of arabinoxylans in *Festuca arundinacea* cell cultures. *Planta* 193, 326–330.

Naito, Y., Hino, K., Bono, H., and Ui-Tei, K. (2015). CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* 31, 1120–1123. doi: 10.1093/bioinformatics/btu743

Nishimura, A., Aichi, I., and Matsuoka, M. (2006). A protocol for agrobacterium-mediated transformation in rice. *Nat. Protoc.* 1, 2796–2802. doi: 10.1038/nprot.2006.469

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., et al. (2007). The TIGR Rice genome annotation resource: improvements and new features. *Nucleic Acids Res.* 35, D883–D887. doi: 10.1093/nar/gkl976

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

Piston, F., Uauy, C., Fu, L., Langston, J., Labavitch, J., and Dubcovsky, J. (2010). Down-regulation of four putative arabinoxylan feruloyl transferase genes from family PF02458 reduces ester-linked ferulate content in rice cell walls. *Planta* 231, 677–691. doi: 10.1007/s00425-009-1077-1

Ralph, J. (2010). Hydroxycinnamates in lignification. *Phytochem. Rev.* 9, 65–83. doi: 10.1007/s11101-009-9141-9

Ralph, J., Grabber, J. H., and Hatfield, R. D. (1995). Lignin-ferulate cross-links in grasses: active incorporation of ferulate polysaccharide esters into ryegrass lignins. *Carbohydr. Res.* 275, 167–178. doi: 10.1016/0008-6215(95)00237-N

Ralph, J., Helm, R. F., Quideau, S., and Hatfield, R. D. (1992). Lignin–feruloyl ester cross-links in grasses. Part 1. Incorporation of feruloyl esters into coniferyl alcohol dehydrogenation polymers. *J. Chem. Soc., Perkin Trans. 1*, 2961–2969. doi: 10.1039/P19920002961

Ralph, J., Quideau, S., Grabber, J. H., and Hatfield, R. D. (1994). Identification and synthesis of new ferulic acid dehydrodimers present in grass cell walls. *J. Chem. Soc., Perkin Trans. 1*, 3485–3498. doi: 10.1039/p19940003485

Rautengarten, C., Ebert, B., Herter, T., Petzold, C. J., Ishii, T., Mukhopadhyay, A., et al. (2011). The interconversion of UDP-Arabinopyranose and UDP-Arabinofuranose is indispensable for plant development in Arabidopsis. *Plant Cell* 23, 1373–1390. doi: 10.1105/tpc.111.083931

Reidinger, S., Ramsey, M. H., and Hartley, S. E. (2012). Rapid and accurate analyses of silicon and phosphorus in plants using a portable X-ray fluorescence spectrometer. *New Phytol.* 195, 699–706. doi: 10.1111/j.1469-8137.2012.04179.x

Ruhland, C. T., Xiong, F. S., Clark, W. D., and Day, T. A. (2005). The influence of ultraviolet-B radiation on growth, hydroxycinnamic acids and flavonoids of *Deschampsia antarctica* during springtime ozone depletion in Antarctica. *Photochem. Photobiol.* 81, 1086–1093. doi: 10.1562/2004-09-18-RA-321

Santiago, R., Reid, L. M., Arnason, J. T., Zhu, X. Y., Martinez, N., and Malvar, R. A. (2007). Phenolics in maize genotypes differing in susceptibility to Gibberella stalk rot (Fusarium graminearum Schwabe). *J. Agric. Food Chem.* 55, 5186–5193. doi: 10.1021/jf070641e

Santiago, R., Sandoya, G., Butrón, A., Barros, J., and Malvar, R. A. (2008). Changes in phenolic concentrations during recurrent selection for resistance to the Mediterranean corn borer (Sesamia nonagrioides Lef.). *J. Agric. Food Chem.* 56, 8017–8022. doi: 10.1021/jf800922j

Sato, Y., Takehisa, H., Kamatsuki, K., Minami, H., Namiki, N., Ikawa, H., et al. (2013). RiceXPro version 3.0: expanding the informatics resource for rice transcriptome. *Nucleic Acids Res.* 41, D1206–D1213. doi: 10.1093/nar/gks1125

Sato, Y., Antonio, B. A., Namiki, N., Takehisa, H., Minami, H., Kamatsuki, K., et al. (2011). RiceXPro: a platform for monitoring gene expression in japonica rice grown under natural field conditions. *Nucleic Acids Res.* 39, D1141–D1148. doi: 10.1093/nar/gkq1085

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

Seifert, G. J. (2018). Mad moves of the building blocks – nucleotide sugars find unexpected paths into cell walls. *J. ExBot.* 69, 905–907. doi: 10.1093/jxb/ery026

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 686–688. doi: 10.1038/nbt.2650

Sibout, R., le Bris, P., Legée, F., Cézard, L., Renault, H., and Lapierre, C. (2016). Structural Redesigning Arabidopsis Lignins into Alkali-Soluble Lignins through the Expression of p-Coumaroyl-CoA:Monolignol Transferase PMT. *Plant Physiol.* 170, 1358–1366. doi: 10.1104/pp.15.01877

Simmons, T. J., Mortimer, J. C., Bernardinelli, O. D., Pöppler, A. C., Brown, S. P., deAzevedo, E. R., et al. (2016). Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* 7:13902. doi: 10.1038/ncomms13902

Soukup, M., Martinka, M., Bosnić, D., Čaplovičová, M., Elbaum, R., and Lux, A. (2017). Formation of silica aggregates in sorghum root endodermis is predetermined by cell wall architecture and development. *Ann. Bot.* 120, 739–753. doi: 10.1093/aob/mcx060

Stelzner, J., Roemhild, R., Garibay-Hernández, A., Harbaum-Piayda, B., Mock, H. P., and Bilger, W. (2019). Hydroxycinnamic acids in sunflower leaves serve as UV-A screening pigments. *Photochem. Photobiol. Sci.* 18, 1649–1659. doi: 10.1039/C8PP00440D

Suzuki, S., Ma, J. F., Yamamoto, N., Hattori, T., Sakamoto, M., and Umezawa, T. (2012). Silicon deficiency promotes lignin accumulation in rice. *Plant Biotechnol (Tokyo)* 29, 391–394. doi: 10.5511/plantbiotechnology.12.0416a

Takahama, U., and Oniki, T. (1994). Effects of ascorbate on the oxidation of derivatives of Hydroxycinnamic acid and the mechanism of oxidation of Sinapic acid by Cell Wall-bound peroxidases. *Plant Cell Physiol.* 35, 593–600. doi: 10.1093/oxfordjournals.pcp.a078634

Takahama, U., and Oniki, T. (1997). Enhancement of peroxidase-dependent oxidation of Sinapyl alcohol by an Apoplastic component, 4-Coumaric acid Ester isolated from epicotyls of *Vigna angularis* L. *Plant Cell Physiol.* 38, 456–462. doi: 10.1093/oxfordjournals.pcp.a029189

Takeda, Y., Tobimatsu, Y., Yamamura, M., Takano, T., Sakamoto, M., and Umezawa, T. (2019). Comparative evaluations of lignocellulose reactivity and usability in transgenic rice plants with altered lignin composition. *J. Wood Sci.* 65:6. doi: 10.1186/s10086-019-1784-6

Vazquez-Vilar, M., Bernabé-Orts, J. M., Fernandez-del-Carmen, A., Ziarsolo, P., Blanca, J., Granell, A., et al. (2016). A modular toolbox for gRNA–Cas9 genome engineering in plants based on the GoldenBraid standard. *Plant Methods* 12:10. doi: 10.1186/s13007-016-0101-2

Watson, J. M., Fusaro, A. F., Wang, M. B., and Waterhouse, P. M. (2005). RNA silencing platforms in plants. *FEBS Lett.* 579, 5982–5987. doi: 10.1016/j.febslet.2005.08.014

Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS One* 6:e16765. doi: 10.1371/journal.pone.0016765

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

Young, J., Zastrow-Hayes, G., Deschamps, S., Svitashev, S., Zaremba, M., Acharya, A., et al. (2019). CRISPR-Cas9 editing in maize: systematic evaluation of off-target activity and its relevance in crop improvement. *Sci. Rep.* 9:6729. doi: 10.1038/s41598-019-43141-6

Zhang, J., Zou, W., Li, Y., Feng, Y., Zhang, H., Wu, Z., et al. (2015). Silica distinctively affects cell wall features and lignocellulosic saccharification with large enhancement on biomass production in rice. *Plant Sci.* 239, 84–91. doi: 10.1016/j.plantsci.2015.07.014

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 September 2022
DOI 10.3389/fpls.2022.883139

OPEN ACCESS

EDITED BY
Rowan Andrew Craig Mitchell,
Rothamsted Research,
United Kingdom

REVIEWED BY
Paula Del Carmen Fernandez,
Instituto Nacional de Tecnología Agropecuaria, Argentina
Shin Taketa,
Okayama University, Japan

*CORRESPONDENCE
Kelly Houston
Kelly.Houston@hutton.ac.uk

†PRESENT ADDRESS
Guillermo Garcia-Gimenez,
Department of Animal, Plant and Soil Sciences, School of Life Sciences, La Trobe Institute for Agriculture and Food, La Trobe University, Bundoora, VIC, Australia

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

RECEIVED 24 February 2022
ACCEPTED 17 August 2022
PUBLISHED 08 September 2022

CITATION
Garcia-Gimenez G, Schreiber M, Dimitroff G, Little A, Singh R, Fincher GB, Burton RA, Waugh R, Tucker MR and Houston K (2022) Identification of candidate MYB transcription factors that influence *CslF6* expression in barley grain. *Front. Plant Sci.* 13:883139. doi: 10.3389/fpls.2022.883139

Identification of candidate MYB transcription factors that influence *CslF6* expression in barley grain

Guillermo Garcia-Gimenez[1†], Miriam Schreiber[2], George Dimitroff[3], Alan Little[3], Rohan Singh[3], Geoffrey B. Fincher[3], Rachel A. Burton[3], Robbie Waugh[1,2], Matthew R. Tucker[3] and Kelly Houston[1*]

[1]The James Hutton Institute, Dundee, United Kingdom, [2]Plant Sciences Division, College of Life Sciences, University of Dundee, Dundee, United Kingdom, [3]School of Agriculture, Food and Wine, The University of Adelaide, Adelaide, SA, Australia

(1,3;1,4)-β-Glucan is a non-cellulosic polysaccharide required for correct barley grain fill and plant development, with industrial relevance in the brewing and the functional food sector. Barley grains contain higher levels of (1,3;1,4)-β-glucan compared to other small grain cereals and this influences their end use, having undesirable effects on brewing and distilling and beneficial effects linked to human health. *HvCslF6* is the main gene contributing to (1,3;1,4)-β-glucan biosynthesis in the grain. Here, the transcriptional regulation of *HvCslF6* was investigated using an *in-silico* analysis of transcription factor binding sites (TFBS) in its putative promoter, and functional characterization in a barley protoplast transient expression system. Based on TFBS predictions, TF classes AP2/ERF, MYB, and basic helix-loop-helix (bHLH) were over-represented within a 1,000 bp proximal *HvCslF6* promoter region. Dual luciferase assays based on multiple *HvCslF6* deletion constructs revealed the promoter fragment driving *HvCslF6* expression. Highest *HvCslF6* promoter activity was narrowed down to a 51 bp region located −331 bp to −382 bp upstream of the start codon. We combined this with TFBS predictions to identify two MYB TFs: *HvMYB61* and *HvMYB46/83* as putative activators of *HvCslF6* expression. Gene network analyses assigned *HvMYB61* to the same co-expression module as *HvCslF6* and other primary cellulose synthases (*HvCesA1*, *HvCesA2*, and *HvCesA6*), whereas *HvMYB46/83* was assigned to a different module. Based on RNA-seq expression during grain development, *HvMYB61* was cloned and tested in the protoplast system. The transient over-expression of *HvMYB61* in barley protoplasts suggested a positive regulatory effect on *HvCslF6* expression.

KEYWORDS

cell wall, (1,3;1,4)-β-glucan, barley, transcription factors, MYB

Introduction

The primary cell walls of certain members of the Poaceae, including barley, are enriched with (1,3;1,4)-β-glucan, a distinctive non-cellulosic polysaccharide that predominantly accumulates during cell expansion (Carpita et al., 2001). In contrast, secondary walls of the Poaceae are mainly characterized by the presence of heteroxylans as the major non-cellulosic component, and tissue-specific lignin accumulation (Burton and Fincher, 2014). This differs from eudicots where xyloglucan and pectins are the most abundant non-cellulosic polysaccharides found in primary cell walls and secondary walls are characterized by the presence of heteroxylans, predominantly 4-O-methylglucuronoxylans, heteromannans and lignin (Harris, 2006).

The discovery of *Cellulose synthase-like* (*Csl*) genes required for (1,3;1,4)-β-glucan biosynthesis, including members of the *CslF*, *H* and *J* families (Burton et al., 2006; Doblin et al., 2009; Little et al., 2018), provided the basis for detailed biochemical and molecular studies of (1,3;1,4)-β-glucan structure, assembly and function. These genetic discoveries complemented initial biochemical and physicochemical studies of (1,3;1,4)-β-glucan carried out predominantly in barley (Woodward et al., 1983), wheat (Bacic and Stone, 1980), and oats (Aaman and Graham, 1987), which reported remarkable differences in relative polysaccharide abundance and structure across cereal species (Bacic and Stone, 1981; Burton and Fincher, 2012). Previous studies identified polymorphisms in the *HvCslF6* gene, a key gene involved in grain (1,3;1,4)-β-glucan biosynthesis (Cory et al., 2012; Taketa et al., 2012; Hu et al., 2014), variations in *CslF* transcript abundance (Burton et al., 2008; Wong et al., 2015), *HvCslF6* subcellular location (Wilson et al., 2015) and amino acids contributing to *HvCslF6* protein structure and (1,3;1,4)-β-glucan solubility (Jobling, 2015; Dimitroff et al., 2016). Despite this, no clear relationship has been established between barley cultivar-specific polymorphisms in the *HvCslF6* upstream region, or the introns of this gene, and the wide range in grain (1,3;1,4)-β-glucan levels observed in barley (Houston et al., 2014; Garcia-Gimenez et al., 2019).

Little is known about regulatory mechanisms affecting primary cell wall formation in grasses, and specifically the biosynthesis of (1,3;1,4)-β-glucan in developing barley grain. In a co-expression analysis of transcribed genes during the cellularization of developing barley endosperm, Zhang et al. (2016a) identified several classes of candidate TF genes that might be involved in cell wall synthesis, but at that stage direct interactions between TF and specific cell wall genes were not investigated. In the model grass *Brachypodium distachyon* a trihelix transcription factor (THX1) was validated as a positive regulator of *BdCslF6* binding to a GT-rich cis-element in the second intron of this gene (Fan et al., 2018). In rice, Zhao et al. (2019) carried out a comprehensive study of transcriptional regulators within the context of cell wall composition. They identified several genes, including members of the MYB gene family, that changed the transcript levels of known cell wall synthesis genes in a transient expression system. One gene, *OsMYB61a*, was of interest due to its ability within this system to regulate the expression of numerous cell wall genes and the observation that knock out lines had altered cell wall composition (Zhao et al., 2019).

Regulation of cell wall composition has also been studied in *Arabidopsis thaliana*, particularly in the context of secondary cell walls. This led to the identification of key transcription factors (i.e., secondary wall NAC; SWN) (Taylor-Teeples et al., 2015) that act as regulatory switches enhancing the expression of downstream MYB master regulators, MYB46 and MYB83 (McCarthy et al., 2009). These MYBs interact with a wide range of cell wall-associated transcription factors (Kim et al., 2014) and networks affecting polysaccharide biosynthesis. In rice and maize, functional orthologs of these *Arabidopsis* secondary cell wall MYB master regulators were identified, suggesting a conserved regulatory mechanism across vascular plants (Zhong et al., 2011; Rao and Dixon, 2018).

Here, our work aimed to explore the transcriptional regulation of *HvCslF6* by identifying putative cis-elements in the 5′ promoter region of *HvCslF6* and the transcription factor(s) binding to them. Such data would provide another avenue to investigate the regulation of *HvCslF6* gene expression and how this affects (1,3;1,4)-β-glucan abundance.

First, the ability of a 3,000 bp *HvCslF6* upstream region to drive expression was tested by GFP tagging in transgenic barley lines. To identify regulatory motifs in the *HvCslF6* promoter, a series of 5′ deletion constructs generated from the 3,000 bp *HvCslF6* upstream region were fused to the dual luciferase reporter system. A barley protoplast-based transient expression system was adapted (Rao, 2007; Yoo et al., 2007) and used to test activity of the *HvCslF6* promoter deletion constructs, permitting the analysis of six deletion constructs simultaneously. Our results from the *HvCslF6* promoter characterization were combined with *in silico* predictions for transcription factor binding sites (TFBS), gene network and co-expression analyses, and previous findings from other cereal species, which suggested several candidate transcription factors that may regulate *HvCslF6* expression.

Results

A 3,000 bp *HvCslF6* putative promoter region is sufficient to drive expression in grain and vegetative tissues

HvCslF6 is expressed across a range of tissues including grain and young/developing tissues such as the coleoptile, root and leaves (Burton et al., 2008). To study the transcriptional regulation of this gene *in planta*, transgenic barley lines carrying

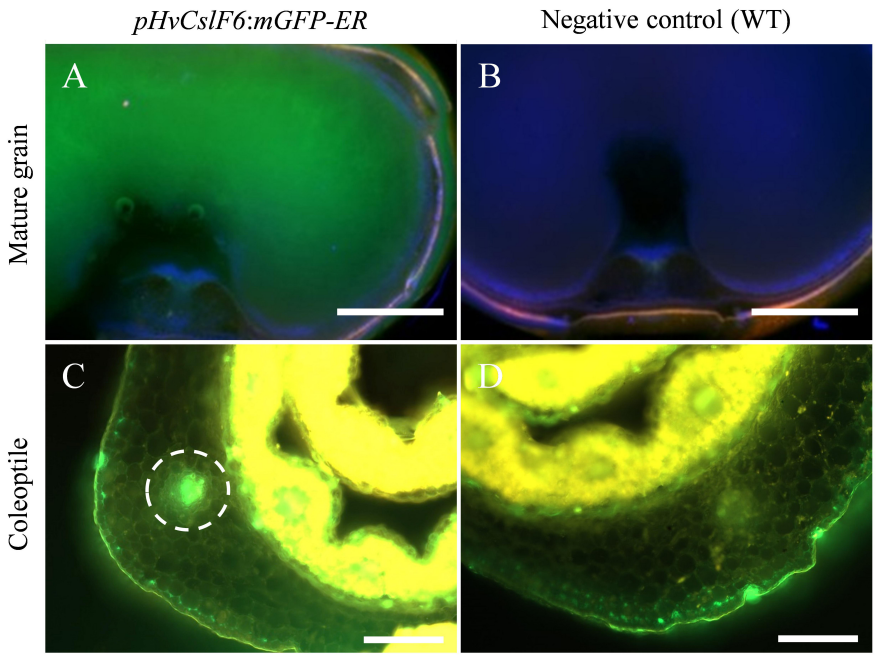


FIGURE 1
Fluorescent images of *pHvCslF6:mGFP-ER* (−3,000 bp) barley transverse sections in mature grain and coleoptile tissues (cv. Golden Promise). **(A)** Mature grain from line transformed with *pHvCslF6:mGFP-ER*, **(B)** negative control of wild type grain section, **(C)** coleoptile from line transformed with *pHvCslF6:mGFP-ER*, **(D)** negative control of wild type coleoptile section. The positive control is mGFP-ER driven by the 35S CaMV promoter. The dashed circle indicates the location of a vascular bundle. Images were collected on a Zeiss M2 AxioImager equipped with DIC optics and an Apotome.2 (Zeiss, Germany). GFP was excited at 488 nm and emission collected from 499 to 530 nm, in green. Scale bars are 1 mm.

endoplasmic reticulum-targeted GFP (mGFP-ER) driven by a 3,000 bp 5′ sequence upstream of *HvCslF6* (*pHvCslF6*) were generated and screened in grain and vegetative tissues. Transverse sections of transgenic mature grain showed strong mGFP-ER expression in the starchy endosperm (**Figure 1A**) compared to the negative control (empty vector) (**Figure 1B**). Similarly, increased mGFP-ER expression around vascular bundles was detected in transverse coleoptile sections of transgenic lines (**Figure 1C**) compared to the negative control (**Figure 1D**). Therefore, we confirmed that the 3,000 bp 5′ sequence upstream of *HvCslF6* is sufficient to drive expression of this GFP gene in both grain and vegetative tissues.

Analysis of transcription factor binding sites in the *HvCslF6* putative promoter region reveals overrepresentation of MYB/SANT, AP2/ERF, and basic helix-loop-helix-related motifs

As the 3,000 bp 5′ promotor region was sufficient to drive the temporal and spatial expression of *HvCslF6* in developing and young tissues, we explored whether known transcription factor binding motifs were either present or enriched within this sequence. First, three TFBS prediction software packages (TRANSFAC® v2014, JASPAR v2020 and PlantPAN v3.0) were compared to identify which produced the most detailed motif predictions with the least number of redundant results. TFBS prediction outputs were compared across the three programs, revealing differences in frequency (number of times a particular motif was predicted within the 3,000 bp *pHvCslF6*) and annotation (profile summary of each motif comprising TF class and origin species) of predicted TFBSs (**Supplementary Figure 1** and **Supplementary Table 1**). It should be noted that a high rate of false positives in predicted TFBSs has been reported in most of these prediction tools (Shahmuradov et al., 2005). Therefore, choosing a comprehensive and experimentally validated database of TFBSs was key to obtain reliable *in silico* predictions. TRANSFAC® was discarded due to the relatively low number of predicted motifs that were also identified using JASPAR and PlantPAN. A comparative motif analysis between PlantPAN and JASPAR identified seven classes of TFBSs with similar relative frequencies (allowing a ± 10% variation in motif prediction results across both programs). These were: MYB/SANT (12.5% in JASPAR and 14.0% in PlantPAN, respectively), DNA-binding with one zinc finger DOF (13.3% and 8.9%), APETALA2/Ethylene response factors AP2/ERF (9.3% and 8.4%), basic leucine zipper bZIP (4.7% and 6.5%), WRKY (5.4% and 3.2%), auxin response factors ARF (B3

domain; 3.1% and 5.2%), and NAC (4.6% and 3.6%). Notably, 52% of PlantPAN predicted motifs could not be classified because no information/description was available regarding their hit sequence.

Based on these results, JASPAR was chosen for subsequent *in silico* analyses of the proximal promoter region of *HvCslF6* (1,000 bp upstream of the start codon) and 355 non-redundant motifs were identified. JASPAR TFBS scores are based on position frequency matrices (PFMs), defined as occurrences of each nucleotide at each position in a set of observed TF-DNA interactions (Fornes et al., 2020). We selected JASPAR for further analyses due to the higher number of non-redundant predicted TFBS and motif-associated information compared to TRANSFAC and PlantPAN. In addition, JASPAR is a regularly maintained open-access, manually curated and experimentally defined database for plant-specific TFBSs (Castro-Mondragon et al., 2022).

These 355 predicted motifs were filtered in accordance with an arbitrary prediction score threshold of ≥10 (maximum score was 15.9) to minimize the detection of false positives/redundant motifs in identical promoter locations (unfiltered TFBS predictions are shown in **Supplementary Figure 2**). By applying this filter, 107 unique predicted motifs with prediction scores between 0.7 and 15.9 were identified and grouped into eight classes of TFBSs including: MYB/SANT, AP2/ERF, basic helix-loop-helix (bHLH), ARF, DOF, Homeodomain-leucine zipper (HD-ZIP), Zinc Finger (ZF; SBP and SWIM type), NAC/NAM and a mixed class of TFBSs (low-represented putative TF classes with high prediction scores) (**Figure 2A**). Predicted motifs grouped into AP2/ERF, MYB/SANT and bHLH TF classes were the most abundant, hence over-represented within the *HvCslF6* proximal promoter region (−1,000 bp) when compared to the total of 107 non-redundant motifs identified. JASPAR prediction scores for the three most abundant TF classes were similar to each other (11.6, MYB/SANT; 11.4, bHLH and 11.7, AP2/ERF; Shown on top of each column in orange, **Figure 2A**).

Comparative analysis of *CslF6* transcription factor binding sites in other cereal species and in genes co-expressed with *HvCslF6* identifies common motifs for certain transcription factor families

To understand if there was conservation of binding site motifs across grass species, we carried out *in-silico* analysis of the putative *CslF6* promoter sequences in wheat (*Triticum aestivum*), *Brachypodium distachyon* and rice (*Oryza sativa*), and compared the TFBS to those identified in barley (filtered by $a \geq 10$ prediction accuracy score cut-off, as previously applied to *HvCslF6* full length promoter analysis). This revealed 59 common motifs across all species (**Figure 2B** and **Supplementary Table 2**) corresponding to: MYBs/SANTs, ARFs, AP2/ERFs, AT-Hook DNA-binding factors, bHLHs, DREBs, and HD-ZIPs TF classes. The second largest group of common motifs (26) was shared between barley, wheat and *Brachypodium CslF6* upstream sequences and corresponded to putative binding sites for MYB-related TFs, (predominantly MYBs/SANTs: MYB55/61, MYB59, MYB111 and MYB113; GARP/G2-LIKE: PHL11, PHL12; MYB-related: UIF1), and other less abundant bHLH and DOF-related motifs. The rice *CslF6* upstream sequence contained the highest number of unique TFBS predictions (84) compared to other cereals (13, barley; 13, wheat; and 14, *Brachypodium*). The frequency, location, and description of predicted TFBS across all species is shown in **Supplementary Table 2**.

Next, the analysis of TFBSs was expanded to include three more members of the *Cellulose synthase* gene superfamily, namely *HvCesA1*, *HvCesA2* and *HvCesA6*, which are involved in cellulose biosynthesis in the primary cell wall (Burn et al., 2002; Burton et al., 2006) and known to be co-expressed with *HvCslF6* (Burton et al., 2004; Burton and Fincher, 2009). Common predicted motifs across *HvCesA1*, *HvCesA2*, *HvCesA6*, and *HvCslF6* corresponded to MYB-related, AP2/ERF, HD-ZIP, DOF, ZF-HD, Homeo Domain (HD), and MADS box factors. The largest set of common motifs was found across *HvCesA2* and *HvCesA6* putative promoters (77), followed by *HvCesA2* *HvCesA6*, and *HvCslF6* (60) (**Figure 2C**). As a comparison three *HvCesA* genes (*HvCesA4*, *HvCesA7*, and *HvCesA8*) which are not co-expressed with *HvCslF6* were also analyzed for TFBSs. This analysis identified one common predicted TFBS, a SQUAMOSA promoter-binding protein-like (SPL) motif across all putative promoters (SPL TFs are mainly involved in plant growth and development; Tripathi et al., 2018) and abundant gene-specific, not shared, motif subsets. The largest set of TFBSs was found in the *HvCslF6* promoter (57) containing MYBs, AP2/ERFs HD-ZIP among other TF classes. (**Supplementary Figure 3** and **Supplementary Table 3**).

Identification of the essential *HvCslF6* promoter region for transcription

To identify the *HvCslF6* promoter region necessary for transcriptional activation, seven progressively deleted promoter constructs were generated from the 3,000 bp region upstream of the *HvCslF6* start codon (−3,000, −1,846, −1,357, −858, −607, −382, and −199 bp) and tested in a barley leaf-derived protoplast system (previously optimized using GFP/YFP-expressing constructs; **Supplementary Figure 4**). The region of the promoter that produced the highest relative luciferase assay activity was a −382 bp region upstream of the *HvCslF6* start codon (**Figure 3A**). Additionally, deletion constructs containing a −607 bp or −199 bp *HvCslF6* promoter region showed a

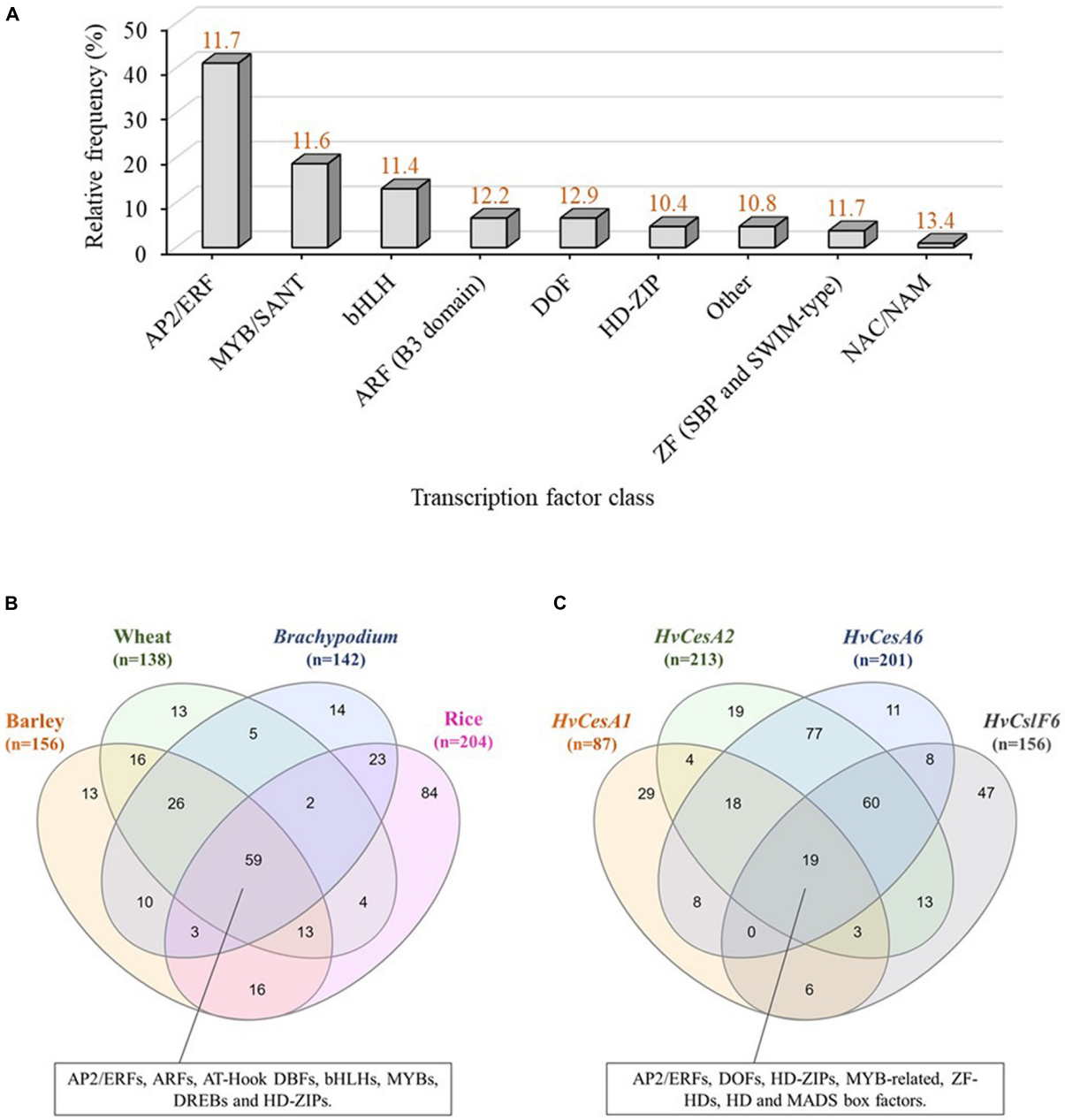


FIGURE 2
In-silico analysis of transcription factor binding sites (TFBSs) in the *CslF6* promoter. **(A)** TFBS predictions identified within a 1,000 bp upstream region of *HvCslF6* start codon using JASPAR (**Fornes et al., 2020**). TFBS predictions were grouped into TF classes and expressed as relative frequency % of TFBSs corresponding to a particular TF family from the total TFBSs predicted (n = 107, after applying $a \geq 10$ matrix score cut-off to maximize TFBS prediction accuracy). Other TF class indicate a mixed class of low-represented TFBSs (frequency within *HvCslF6* full length promoter $\leq 2\%$). Average prediction scores for each TF class are shown in orange on top of each bar and were inferred from position frequency matrices (PFMs) and TF flexible models (TFFMs) in JASPAR. **(B)** Venn diagram of TFBS predicted within putative *CslF6* promoter regions (−1,000 bp) across barley (*Hordeum vulgare*), wheat (*Triticum aestivum*), *Brachypodium distachyon* and rice (*Oryza sativa*). **(C)** Venn diagram of TFBSs predicted within *HvCesA1*, *HvCesA2*, *HvCesA6*, and *HvCslF6* putative promoter regions (−1,000 bp) with identical filtering as described above. Venn diagrams were created using InteractiVenn (**Heberle et al., 2015**).

significant decrease in promoter activity based on the relative luciferase assay compared to the −382 bp region. Therefore, we had delimited a 183 bp sequence within the proximal *HvCslF6* promoter between the two smallest deletion constructs (from −382 to −199 bp) that showed a relative increase in luciferase activity and could be responsible for upregulation of *HvCslF6* expression. *HvCslF6* promoter activity was comparatively low from −607 to −3,000 bp upstream of the *HvCslF6* CDS,

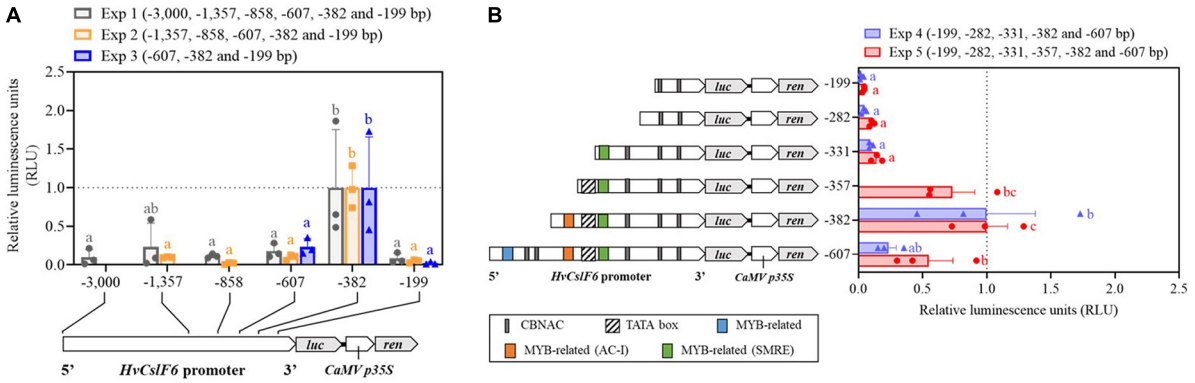


FIGURE 3
Characterization of the *HvCslF6* promoter using dual luciferase assays to identify region(s) response for increase in expression. Promoter activity in panels **(A,B)** is expressed as relative luciferase activity (luciferase/renilla ratio) and normalized to the −382 construct (with highest promoter activity, dashed gray line) on each experiment, respectively. Three independent protoplasts transfections (biological replicates) were performed per deletion construct and averaged to calculate mean reporter activity 24 h after transfection for each experiment. Negative controls (n = 3 empty vector lacking *HvCslF6* promoter) were performed in parallel and luciferase background activity subtracted from tested constructs (n = 3 dual luciferase assays per construct, as technical replicates). **(A)** *HvCslF6* promoter activity analyzed in six deletion constructs (from −3,000 bp full length promoter to −199 bp) using dual luciferase assays across three independent experiments (Exp1, Exp2, and Exp3). Letters above each construct indicate significant differences within each experiment determined by one-way ANOVA (p-value 0.041, Exp1; p-value < 0.01, Exp2 and p-value 0.027, Exp3) followed by Tukey's multi-comparison test. Error bars represent standard error (n = 3). **(B)** *HvCslF6* promoter activity of 5′ *HvCslF6* promoter deletion constructs from −607 bp to −199 bp in barley protoplasts. Gray bar indicates CBNAC calmodulin-binding NAC motif, blue bars indicate MYB-related motifs, yellow bars denote MYB-related AC-I motifs, green bars indicate secondary wall MYB-responsive element, (SMRE) and the diagonally stripped bar denotes predicted location of the TATA box consensus sequence. Different letters above each construct (i.e., a,b) indicate significant differences determined by one-way ANOVA (p-value < 0.01) followed by Tukey's test. Error bars represent standard error (n = 3).

indicating that repressors of *HvCslF6* expression could be binding to this −382 to −607 bp region. Furthermore, no significant differences in promoter activity were observed across constructs carrying the longest *HvCslF6* promoter fragments (−3,000, −1,846, and −1,357 bp; **Supplementary Figure 5**).

We retrieved *HvCslF6* proximal promoter regions (1,000 bp upstream from start codon) from cv. Morex and cv. Golden Promise and carried out a pairwise sequence alignment. Both promoters were identical based on the barley pangenome sequences available at GrainGenes[1] therefore subsequent *in silico* prediction of TF binding sites could also be extended to cv. Golden Promise.

To further delimit putative TF binding sites within the region that showed increased luciferase activity, three additional deletion constructs were generated. Constructs containing −282 bp, −331 bp, and −357 bp of the *HvCslF6* promoter, were tested alongside three previously tested deletion constructs (−199 bp, −382 bp, and −607 bp) to normalize dual luciferase measurements and compare promoter activity across deletion constructs (**Figure 3B**). A significant increase in promoter activity (p-Value 0.029, Tukey's test) was detected in the −357 bp deletion construct. *In-silico* prediction of this promoter fragment indicates that it includes a putative TATA box element (5′–TATAAA–3′). This likely explains the increase in *HvCslF6* promoter activity with the inclusion of this region of the promoter, and the low/zero levels of promoter activity in the first three constructs which based on our *in-silico* prediction lack a TATA box. Highest *HvCslF6* promoter activity was narrowed down to a 51 bp region located −331 bp to −382 bp upstream of the start codon, although no significant differences (p-Value 0.321, Tukey's test) were determined between the −357 bp and −382 bp deletion constructs. Moreover, PlantPAN and JASPAR *in-silico* motif analysis of the additional 25 bp included in the −382 bp fragment compared to the −357 bp fragment predicted a single MYB-related motif, AC-I, enriched in AC bases. Results also indicated low *HvCslF6* promoter activity from −199 bp to −331 bp based on three deletion constructs analyzed (−199 bp, −282 bp, and −331 bp). This region contained two (three in the case of −331 bp construct) predicted calmodulin-binding NAC elements and a MYB-related motif, described as a secondary wall MYB-responsive element (SMRE) based on JASPAR analysis (**Figure 3B**).

MYB transcription factor binding sites are located within a promoter region showing increased luciferase activity

Based on previous studies which suggest a key role for NAC and MYB TF families in SCW polysaccharide biosynthesis (Taylor-Teeples et al., 2015) and the involvement of MYB

1 https://wheat.pw.usda.gov/blast/

TABLE 1 Description of putative binding sites found in the *HvCslF6* promoter based on dual luciferase results of deletion constructs with differential promoter activity.

Putative motif/s in p*HvCslF6* (5′–3′)	*Cis*-acting element/s (5′–3′)	Description	Candidate TF	References
ACCTAC (−364 bp)	MYB consensus: ACC(A/T)A(A/C)	AC-I element, bound by R2R3-MYB proteins	HvMYB61	Romano et al., 2012; Prouse and Campbell, 2013; Zhao et al., 2019
CGTTGGT (−299, −235, −364, −386, −478 bp) (3′–ACCAACG–5′)	MYB consensus /SMRE: ACC(A/T)A(A/C) (T/C)	Secondary wall MYB-responsive element (SMRE)	HvMYB46/83	Kim et al., 2012; Zhong and Ye, 2012
	MYB consensus /M46RE: (T/C)ACC(A/T)A(A/C) (T/C)	MYB46-responsive cis-regulatory element (M46RE)		Ko et al., 2014
GGTAGGTAGGT (−478 bp) (5′–GGTAGGT–3′) (3′–ACCTACC–5′)	R2R3-MYB/MYB3: GGTAGGT(A/G) MYB consensus /SMRE: ACC(A/T)A(A/C) (T/C)	MYB3; subgroup S4 AC-I/ Secondary wall MYB-responsive element (SMRE)	R2R3-MYB/s	Dubos et al., 2010; Ko et al., 2014

TFs in grass cell wall synthesis including hemicelluloses (Zhao et al., 2019), we mapped two putative MYB-related *cis*-elements (among other predicted motifs from our *in-silico* analysis) onto *HvCslF6* promoter regions that conveyed increased promoter activity (**Figure 3B** and **Table 1**). A six-nucleotide, AC-I element (5′–ACCTAC–3′) was predicted from −364 to −358 bp upstream from the *HvCslF6* start codon. Prouse and Campbell (2013), demonstrated that the R2R3-MYB *AtMYB61* can bind to this AC-I element, therefore we were interested in the barley ortholog of this gene as a potential regulator of *HvCslF6* expression. A putative binding site in construct −331 (5′–CGTTGGT–3′/3′–ACCAACG–5′) corresponds to another MYB consensus motif, also described as a secondary wall MYB-responsive element (SMRE) or a MYB46-responsive cis-regulatory element, M46RE (Zhong and Ye, 2012). This *cis*-element was detected in reverse orientation (3′–5′) and in *Arabidopsis* the same motif can be bound by AtMYB46 and AtMYB83 (Kim et al., 2012; Ko et al., 2014). Additionally, a secondary wall MYB-responsive element (SMRE) motif was predicted from −299 to −292 bp, downstream of the TATA box in the *HvCslF6* promoter (see construct −331, **Figure 3B**). However, this had limited impact on promoter activity, likely due to the absence of the TATA box.

To assess protein sequence similarities in MYB46, MYB83, and MYB61 orthologs between barley (**Table 2**), *Arabidopsis* (Romano et al., 2012) and rice (Zhong et al., 2011), we constructed an unrooted phylogenetic tree of R2R3-type MYB transcription factors which includes those that could putatively bind to the −382 bp *HvCslF6* promoter region (**Supplementary Figure 6**). The candidate MYB transcription factors that correspond to the predicted binding sites belong to the plant specific R2R3-MYB subfamily, which is characterized by the presence of two highly conserved MYB DNA-binding domains (Pfam PF00249; Bateman et al., 2017) and constitutes the largest MYB subgroup. HvMYB61 was grouped together with their orthologs from *Arabidopsis* (AtMYB61; Prouse and Campbell, 2013) and rice (OsMYB61a and OsMYB61b; Zhao et al., 2019). The best fitting barley ortholog for AtMYB46 and AtMYB83 was a single gene named HvMYB46/83, following previous studies in rice (OsMYB46/83; Rao and Dixon, 2018; Zhao et al., 2019) and grouped with its orthologs (**Supplementary Figure 6**).

TABLE 2 Description of candidate transcription factors retrieved from GrainGenes (https://wheat.pw.usda.gov/GG3/) based on cv. MorexV3 (Mascher et al., 2021).

Candidate TF	Barley gene ID	Position	5 DPA FPKM	15 DPA FPKM
HvMYB46/83	HORVU.MOREX.r3.5HG0447760	chr5H:141063373..141065705	0.03	0.00
HvMYB61	HORVU.MOREX.r3.1HG0018590	chr1H:61436701..61439073	4.49	0.89

Expression data (fragments per kilobase of exon per million, FPKM across three biological replicates) in developing barley grain at 5 and 15 days post anthesis (DPA), were retrieved from barleyGenes RNA-Seq Database (Available from: https://ics.hutton.ac.uk/barleyGenes/index.html).

Gene network analysis of candidate transcription factors

Using gene expression data from 808 individual samples (**Supplementary Table 4**; Milne et al., 2021), we constructed a co-expression network to identify genes closely connected to *HvCslF6*. The samples covered a wide range of biotic and abiotic stresses, tissue types and cultivars, resulting in a robust dataset. Low expressed genes were removed (see section "Materials and methods") and a weighted gene correlation network analysis (WGCNA) performed (Langfelder and Horvath, 2008). We identified 33 modules, with a color assigned to each module, plus one module (gray) which contains the unallocated genes. *HvCslF6*, along with the primary cell wall cellulose synthases (*HvCesA1*, *HvCesA2*, *HvCesA6*) were assigned to the yellow module (**Supplementary Figure 7**). *HvMYB61* (one of the candidate *HvCslF6* regulators) and *HvTHX1*, the barley ortholog of *BdTHX1*, which binds to the intronic region of *BdCslF6* and acts as a positive regulator of expression of

this gene (Fan et al., 2018), were all members of the yellow module. Visualization of the yellow network in Gephi showed a gene network comprising 992 genes and 54,090 gene connections (**Supplementary Figure 8**). To better investigate the connections, an intermediate stringency threshold of 0.188 (the correlation ranged from 0.15 to 0.335) was applied to define a subnetwork of the top 50 genes connected to *HvCslF6*.

This subnetwork based around *HvCslF6* still included *HvMYB61* (BART1_0-p01380 and the three cellulose synthases; *HvCesA1, HvCesA2, HvCesA6*, BART1_0-p60277, BART1_0-p40943, BART1_0-p44934) (**Figure 4**). All are either known to be involved in primary cell wall synthesis in barley, or in the case of *HvMYB61* the ortholog of *OsMYB61* has been shown to influence to cell wall synthesis (Zhao et al., 2019). Additionally, this module contained other proteins that have been linked to cell wall processes (**Supplementary Table 5**). *COBRA-like protein 3* (BART1_0-p39134) located on chromosome 5H was shown to be co-expressed with *HvCesA1*, *HvCesA2*, *and HvCesA6* (Houston et al., 2015). The other candidate regulator, *HvMYB46/83* did not pass the filtering threshold when removing genes with low levels of expression or were assigned to the black instead of yellow module. Hence, this candidate regulator is unlikely to be a global regulator of *HvCslF6*.

Then we carried out a correlation analysis using transcript expression data from the eoRNA database (Milne et al., 2021), which contains 22 publicly available datasets and >800 RNA-seq datasets from a wide range of tissues, development stages and experimental treatments. We focused on two subsets of genes. The first included our candidate TF gene *HvMYB61,* a selection of genes from the network analysis and others known to be co-expressed with *HvCslF6* (i.e., *HvCesA1, HvCesA2,* and *HvCesA6*). The second subset was comprised of secondary cell wall genes not co-expressed with *HvCslF6* (i.e., *HvCesA4, HvCesA7, HvCesA8*), and included other members of the *HvCslF/H/J* gene family, and *HvTHX1*, the barley ortholog of *BdTHX1*, a positive regulator of *BdCslF6* (**Figure 5A** and **Supplementary Figure 9**). We hypothesized that if *HvMYB61* shows comparable or greater correlation with *HvCslF6* compared to *HvCesA1, HvCesA2, HvCesA6*, this would provide further confidence in a relationship between these candidates. However, if strong positive correlation is observed between the expression of *HvMYB61*, our candidate gene, and the group considered to be secondary cell wall *HvCesA* genes (*HvCesA4, HvCesA7*, and *HvCesA8*), which show no correlation with *HvCslF6*, it might suggest that our candidate gene is unlikely to be involved in the broader regulation *HvCslF6* expression.

For the transcripts of genes considered to be primary cell wall related *CesA* genes and *HvCslF6,* we observed Pearson correlation coefficients of between r = 0.9 to −0.1, while the secondary cell wall *CesA* genes had r = 0.4 to −0.2 (**Supplementary Figure 9**). The range of values is not only due to several genes being compared but also multiple transcripts per gene. For several transcripts there was a high degree of correlation between the expression profile for *HvMYB61* and *HvCslF6* (Person correlation coefficient r = 0.7 to 0.1) and *HvTHX1* and *HvCslF6* (Person correlation coefficient r = 0.6 to −0.1). Therefore, this analysis from an extensive range of samples and conditions provides additional evidence supporting the similar expression dynamics of *HvMYB61* and *HvCslF6*.

To further investigate the expression profile of candidate MYB transcription factors we surveyed several smaller transcript datasets from barley that are independent of those described above. An RNA-seq dataset from cv. Morex across 16 barley tissues showed that *HvMYB61* is expressed abundantly across the entire barley plant (**Figure 5B**). *HvMYB61* expression was detected in 10-day seedling, shoot (10 cm), tiller (3rd internode), rachis, developing grain at 5 days post anthesis (DPA) and embryo; these are tissues where (1,3;1,4)-β-glucan is present. Similarly, *HvTHX1* is also expressed in multiple tissues including developing grain (5 and 15 DPA). Expanding the RNA-seq expression data across barley grain development, an independent expression dataset based on six different spring barley cultivars (cvs. Hopper, Sloop, Extract, Taphouse, Alabama, and Pewter; Aubert et al., 2018; Matros et al., 2021) was used to retrieve the expression profiles of the candidate genes from early to mid-late grain development. *HvMYB61* was mainly expressed during early grain development, from 7 to 11 DPA, whereas *HvMYB46/83* expression remained almost undetectable from 7 to 20 DPA. In the RNA-seq dataset (developing grain) Pearson correlation coefficients with *HvCslF6* were: 0.95 for *HvMYB61* and 0.98 for *HvTHX1* (showing abundant expression from 7 to 20 DPA; **Supplementary Table 6**). Our results showed that the positive correlation between *HvMYB61* and *HvCslF6* across different tissues and datasets, in the context of HvTHX1 (ortholog of BdTHX1, known activator of BdCslF6; Fan et al., 2018) and other *HvCesA*s (co-expressed and not co-expressed), are consistent with a regulatory effect of HvMYB61.

Candidate transcription factor over-expression in barley protoplasts impacts *HvCslF6* expression

Based on expression profiles, our luciferase assay results and data from gene network analyses, we focused subsequent assays on the *HvMYB61* transcription factor, since evidence suggested this gene may influence *HvCslF6* expression. The candidate regulator *HvMYB46/83* was ruled out based on low levels of expression from gene co-expression analyses. The candidate transcription factor *HvMYB61* was cloned and transfected into barley protoplasts plus an empty vector control, respectively. After 24 h, protoplasts were harvested and *HvCslF6* relative transcript abundance was measured. We observed a significant increase in *HvCslF6* expression in the presence of *HvMYB61,*

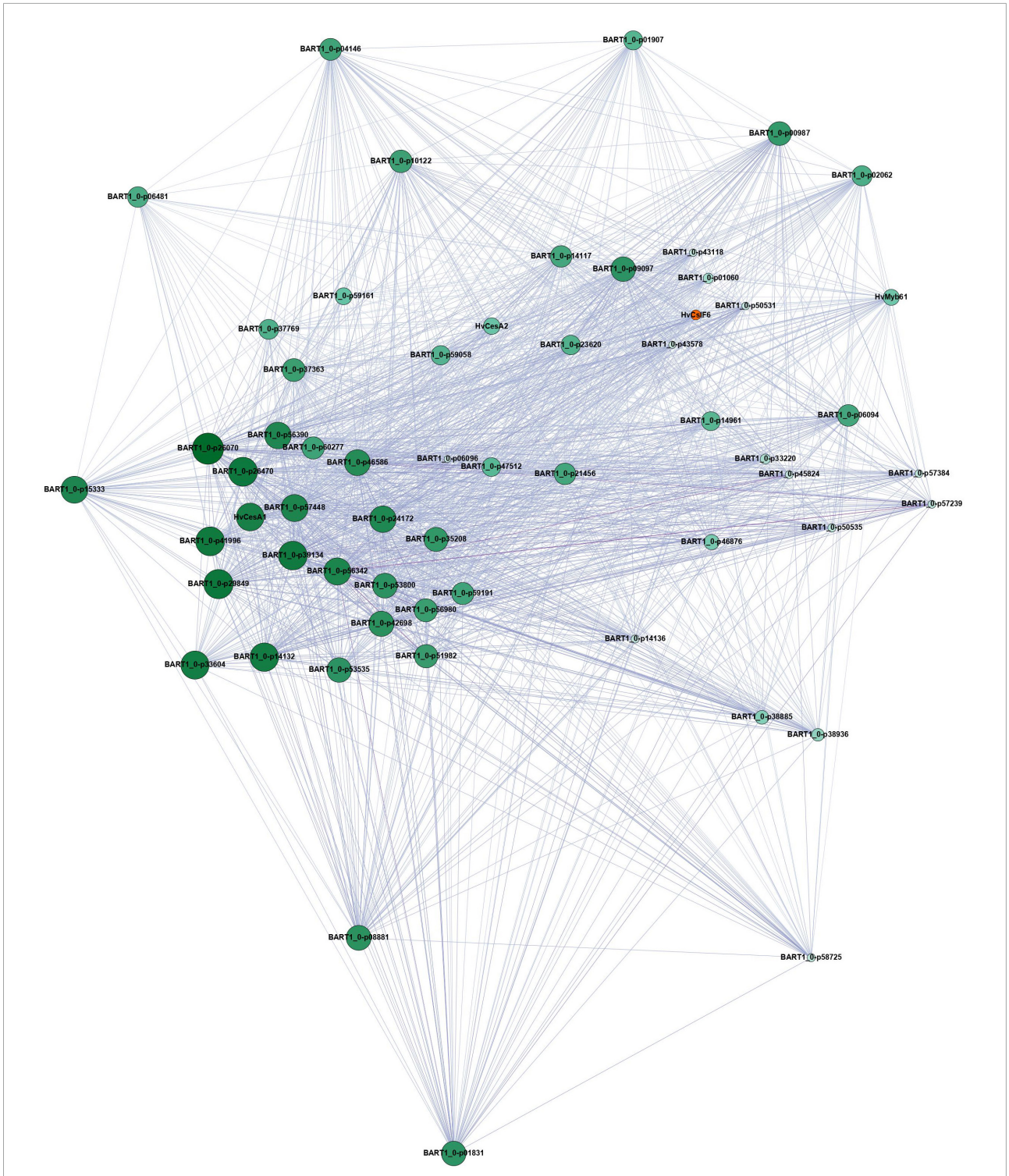


FIGURE 4
Filtered yellow module network based on genes connected to *HvCslF6* (*n* = 50), visualized using Gephi. This subnetwork was obtained from the publicly available EoRNA database (Milne et al., 2021). Filtering parameters are described in the corresponding methods section. Node size and color correspond to gene degree centrality. Big and dark green nodes have a high number of linked nodes, while small and light green nodes only have a small number of linked nodes. The edges are colored by weight. A dark purple corresponds to a high weight and therefore correlation, while a light purple color corresponds to a lower weight and less strong correlation between the nodes. *HvCslF6* is highlighted in orange. Gene names correspond to the description in **Supplementary Table 5**. Genes not listed in the table do not contain protein annotation and are potential non-coding genes.

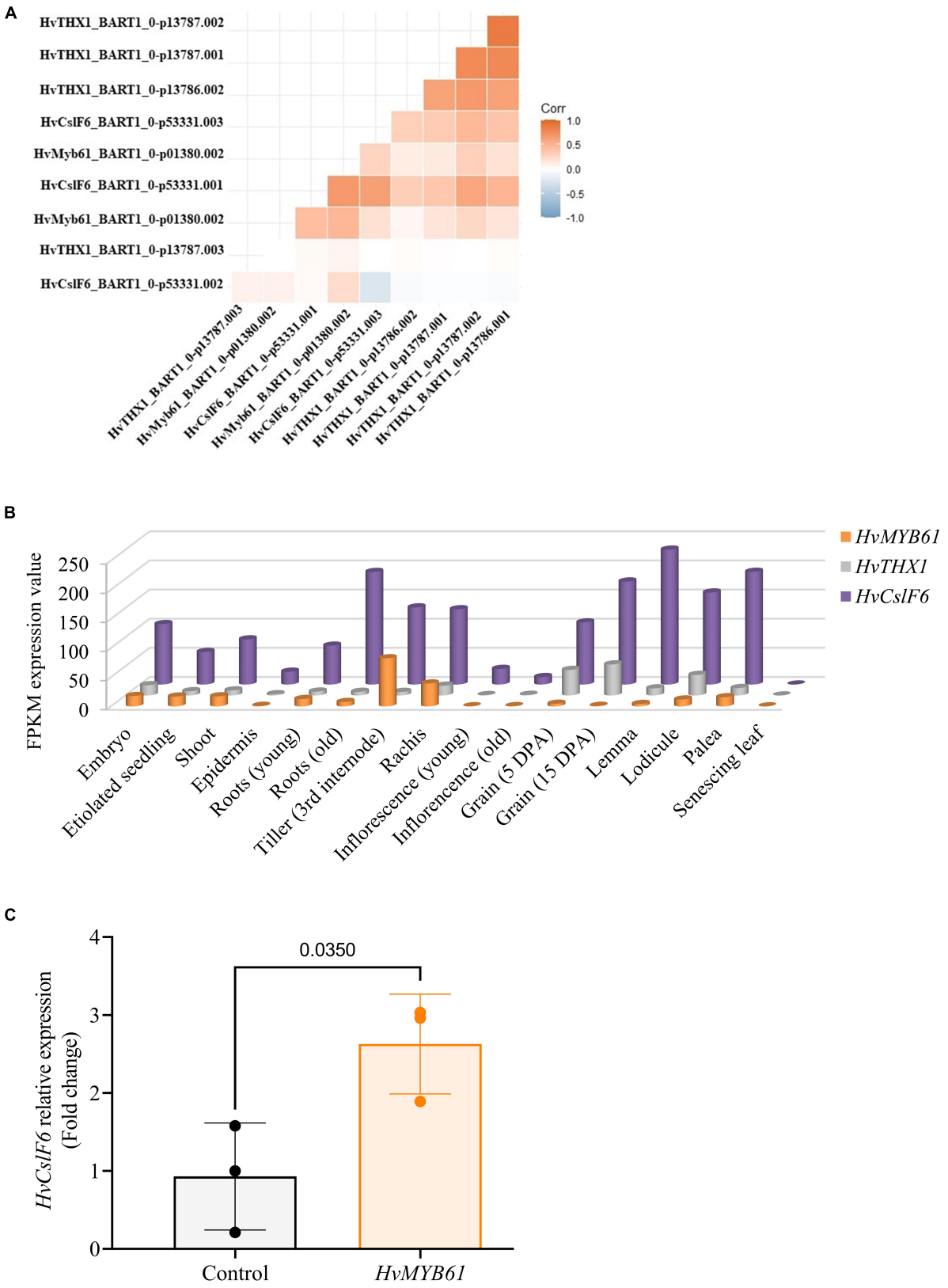


FIGURE 5
Expression profiles of candidate upregulators of *HvCslF6* expression in several datasets. **(A)** Correlation between TPM values of selected candidate genes from the EoRNA database ordered based on hierarchical clustering. **(B)** Expression levels of *HvMYB61* candidate transcription
(Continued)

FIGURE 5 (Continued)
factor, *HvTHX1* and *HvCslF6* across 16 barley tissues obtained from barleyGenes RNA-seq database. (Available from: https://ics.hutton.ac.uk/barleyGenes/, The James Hutton Institute), expressed as FPKM (fragments per kilobase of exon per million fragments mapped) across three biological replicates per tissue. **(C)** Relative expression of *HvCslF6* in barley protoplasts transfected with an empty vector and over-expressing *HvMYB61*, respectively, after 24 h. Internal control genes: α-*tubulin*, *GAPDH* and *HSP70* were used for normalization of relative expression ($2^{-\Delta\Delta CT}$ method; Vandesompele et al., 2002). Error bars represent standard deviation associated with three independent experiments (transfection assays $n = 3$). qRT-PCR reactions were performed in triplicate. *P*-value was calculated based on Student's *t*-test (two-tailed).

representing a fold change of 2.63 ± 0.64 in *HvCslF6* expression compared to protoplasts transfected with an equal amount of empty vector (*p*-Value 0.035, Student's *t*-test; **Figure 5C**). Taken together, our results from the functional characterization of the *HvCslF6* promoter, TF identification and gene co-expression analyses, indicate that the *HvMYB61* transcription factor may be a positive regulator of *HvCslF6* expression in barley.

Discussion

The transcriptional regulation of primary cell wall formation and in particular (1,3;1,4)-β-glucan biosynthesis, which is present across grasses including many economically important crop species, remains largely unresolved. The interest in the identification of regulators able to fine-tune *HvCslF6* expression and grain (1,3;1,4)-β-glucan content ultimately arises from the beneficial health effects associated with (1,3;1,4)-β-glucan-rich diets (Bai et al., 2019) and its importance for efficient brewing and distilling (Gupta et al., 2010). Several transcription factor families (predominantly NACs and MYBs) are known to regulate polysaccharide biosynthesis in secondary cell walls, some of them with conserved functions between grasses and *Arabidopsis* (McCarthy et al., 2009; Zhou et al., 2009; Ko et al., 2014).

We delimited a region of the functional promoter of *HvCslF6* which contained a binding site motif for *MYB61*. We observed that when this gene was overexpressed in the presence of *HvCslF6* it led to an increase in *HvCslF6* expression relative to control samples which lacked *HvMYB61*. This reflects the results of Zhao et al. (2019) who observed that *OsCslF6* and other cell wall related genes were upregulated in rice protoplasts overexpressing *OsMYB61*. The authors also reported a 31% decrease in leaf (1,3;1,4)-β-glucan content was observed in *OsMYB61a* knockout mutants compared to WT. Barley and rice have distinct patterns of (1,3;1,4)-β-glucan distribution throughout the plant, with barley having relatively high levels of this polysaccharide in the grain compared to vegetative tissue (Burton and Fincher, 2014); the opposite is true of rice with extremely low levels of (1,3;1,4)-β-glucan present in the grain (Vega-Sanchez et al., 2012). Phenotypic analyses of *AtMYB61* loss- and gain-of-function mutants showed similar results to *OsMYB61* RNAi lines (Hirano et al., 2013), confirming a role for this gene in secondary cell wall formation in *Arabidopsis*, in addition to regulating other traits (Penfield et al., 2001; Newman et al., 2004; Romano et al., 2012; Romero-Romero et al., 2018). At present, evidence for a direct interaction between HvMYB61 and the putative binding site in the promoter of *HvCslF6* is lacking, as we did not carry out a yeast-1-hybrid experiment. Although Zhao et al. (2019) did not observe any evidence of direct binding between OsMYB61 and *OsCslF6*, it is important to consider that the levels of (1,3;1,4)-β-glucan are typically 5-fold lower in rice than barley, and hence the expression or function of regulatory genes might be different between these two species.

Unlike rice, which contains two MYB61 genes (OsMYB61a and OsMYB61b; Zhao et al., 2019), *HvMYB61* was only present as a single gene on barley chromosome 1H at 47.8 cM. Notably, this co-located with an association peak for grain (1,3;1,4)-β-glucan content (48.4 cM) which was identified in a previous genome wide association scan using elite barley germplasm (Houston et al., 2014). A QTL for malt (1,3;1,4)-β-glucan was detected in the same genomic region using a cv. Steptoe × Morex population (Han et al., 1995). Other genes described in this region of chromosome 1H are *HvCslF9*, a putative grain (1,3;1,4)-β-glucan synthase based on sequence similarity to other *Csl* genes, and *HvGlbI*, a (1,3;1,4)-β-glucan endohydrolase. *HvGlbI* has been shown to hydrolyse both malt (1,3;1,4)-β-glucan and (1,3;1,4)-β-glucan from germinating grains (Slakeski and Fincher, 1992; Betts et al., 2017) and therefore is a plausible candidate for contributing to variation underlying the QTL's mentioned above (Han et al., 1995, Houston et al., 2014). However, because knockout mutants for *HvCslF9* exhibit similar (1,3;1,4)-β-glucan content to wild-type barley grain, it seems likely that this gene plays a minor role, if any, in determining mature grain (1,3;1,4)-β-glucan content (Garcia-Gimenez et al., 2020). Hence, while *HvGlbI* and *HvCslF9*'s contribution to grain (1,3;1,4)-β-glucan content is at least already partially understood,*HvMYB61* appears to be an enticing candidate for further investigations of variation in *HvCslF6* expression during grain development.

We also identified another R2R3-MYB binding site located from −299 to −292 bp in the *HvCslF6* promoter. This was functionally characterized as a secondary cell wall MYB-responsive element, SMRE which can be bound by MYB46 and MYB83 in *Arabidopsis thaliana*, regulating secondary wall biosynthesis (Kim et al., 2012; Zhong and Ye, 2012). This SMRE motif was also described as a MYB46-responsive *cis*-regulatory

element, M46RE (Ko et al., 2014). However, *HvMYB4/83* mRNA levels were barely detectable in most barley tissues, inconsistent with a major role in controlling (1,3;1,4)-β-glucan content. In contrast, *HvMYB61* expression was detected across several tissues where (1,3;1,4)-β-glucan is present, (and *HvCslF6* is expressed) including early stages of grain development when (1,3;1,4)-β-glucan biosynthesis is occurring. No significant differences in promoter activity were observed for constructs carrying the longer *HvCslF6* promoter fragments (−3,000, −1,846 and −1,357 bp; **Supplementary Figure 5**), with these fragments generating comparatively low levels of promoter activity. One of these fragments encompassed the section of the *HvCslF6* promoter (from −1,560 to −1,567 bp of the *HvCslF6* start codon) which Wong et al. (2015) had identified an 8 bp insertion in cv. TR251. However, this deletion was absent from all other cultivars in Wong et al. (2015) and the present study.

We observed further support for a relationship between *HvMYB61* and *HvCslF6* from a comprehensive gene network analysis of *HvCslF6* using 807 gene expression sets from different developmental stages and treatments. This analysis assigned *HvMYB61* to the same co-expression module as *HvCslF6*, the primary cellulose synthases (*HvCesA1*, *HvCesA2*, and *HvCes6*), and *HvTHX1*, the ortholog of a known enhancer of *BdCslF6* expression (Fan et al., 2018), consistent with our hypothesis that *HvMYB61* could act as positive regulator of *HvCslF6* expression. Subsequent analysis of *HvMYB61* by transient over-expression in a barley protoplast system led to a significant increase in *HvCslF6* expression in the presence of *HvMYB61* compared to protoplasts transfected with an equal amount of empty vector across three independent experiments. In line with the study from Zhao et al. (2019) in rice, our results suggest that in cereals, *MYB61* may have a conserved role in regulating the expression of grass-specific cell wall biosynthetic enzymes including *HvCslF6*.

The focus of the current study was to delimit putative cis-elements and transcription factors controlling the expression of *HvCslF6*, and to use these findings to further refine our knowledge of what determines (1,3;1,4)-β-glucan content and ultimately provide diagnostics to enable selection for variation in this trait. Based on our findings, we speculate that selected MYB transcription factors may act as positive regulators of (1,3;1,4)-β-glucan accumulation in barley primary cell walls. This relationship relates predominantly to *HvMYB61* and the proximal promoter region of *HvCslF6* and is supported by co-expression in a range of tissues including the young endosperm. However, we anticipate that the regulation of *HvCslF6* expression could be tissue- and genotype-specific. We previously demonstrated that *HvCslF6* expression is genotype-dependent in a relatively genetically narrow selection of barley cultivars (Garcia-Gimenez et al., 2019). Additionally, the temporal and spatial variation of the expression of this gene has been well characterized previously (Burton et al., 2008). Therefore, expanding this work to include a wider set of germplasm, tissues and generating expression profiles for additional genes such as *HvMYB61* would provide additional insight into the relationship between these genes.

The decrease in *HvCslF6* promoter activity observed in the −607 bp construct compared to the −382 bp construct requires further investigation. Recent studies also describe the emerging role of MYBs as repressors, negatively affecting secondary cell wall biosynthesis (Zhao and Bartley, 2014; Rao and Dixon, 2018) and cold acclimation (Zhang et al., 2016b), among other traits (Zhou et al., 2017; Ma and Constabel, 2019). It is also possible that repressive cis-elements are present in the *HvCslF6* promoter. The dual luciferase system allowed us to screen a more distal promoter region from −382 to −607 bp, which showed comparatively low levels of luciferase activity. This suggests that other potential regulatory factors may bind to the *HvCslF6* promoter in the protoplast system, inhibiting the expression of this gene. TFBS predictions in this 225 bp region correspond to several TF families and include additional MYB-related elements. Additionally, it will be intriguing to assess the function of *THX1* barley orthologs, and any potential additive effect with *HvMYB61*.

Materials and methods

p*HvCslF6*:mGFP-ER transgenic lines

Barley transgenic lines (cv. Golden Promise) carrying a 3,000 bp upstream region from *HvCslF6* start codon fused to a endoplasmic reticulum-targeted GFP (p*HvCslF6*:mGFP-ER) construct were generated via *Agrobacterium*-mediated transformation at The University of Adelaide following the protocol outlined in Burton et al. (2011). Transgenic lines (T_2) were screened using a Zeiss Axioskop 2 Plus microscope (Carl Zeiss Microscopy GmbH, Jena, Germany) equipped with an external UV light source (HBO 100). mGFP was excited at 488 nm and emission collected at 500–530 nm. For coleoptile screening, lines were grown in a petri dish with filter paper for 3–4 days in the dark and screened 5 days after gemination. Transverse section images were captured with an AxioCam 512. Subsequent to collection, images were processed with Zen Software v6.0, utilizing global adjustment tools only.

Transcription factor binding sites prediction in barley and other species

A 3,000 bp region upstream of the *HvCslF6* start codon (cv. Morex) was retrieved from the barley genome explorer (Mascher et al., 2021) and used for TFBS prediction comparison using TRANSFAC® v2014 (Matys et al., 2006), JASPAR v2020 (Fornes et al., 2020) and PlantPAN v3.0 (Chow et al., 2019).

Motif over-representation analysis was carried out using *HvCslF6* TFBS predictions from JASPAR and filtered based on ≥ 10.0 prediction score (maximum score = 15) to minimize false positive motif predictions. Upstream *CslF6* sequences from wheat (*Triticum aestivum*), *Brachypodium distachyon* and rice (*Oryza sativa*) were retrieved from Ensembl Plants[2]. TFBS comparisons across species were carried out within a 1,000 bp *CslF6* putative promoter region using JASPAR (Fornes et al., 2020) and predicted motifs filtered based on ≥ 10.0 prediction score. The same filter was also applied for the TFBS analysis of *HvCesA* upstream regions (1,000 bp). Venn diagrams were created using InteractiVenn (Heberle et al., 2015).

Plant material for protoplast assays

Barley plants (cv. Golden Promise) were grown in a growth chamber with a 16-h light/8-h dark photoperiod at the Functional Genomics (FUNGEN) facility, at The James Hutton Institute, United Kingdom. After 3 weeks, 0.25 g of primary leaves were harvested for protoplast assays. Leaves were cut lengthwise and crosswise (∼20 mm × 10 mm) and peeled (epidermis removal from the abaxial side) in sterile conditions using a sharp razor, scalpel and tweezers under a Leica MZ6 stereo microscope.

Isolation of leaf-derived protoplasts and PEG-mediated transfection

The barley protoplast isolation protocol was mainly based on the protocol for preparation of *Arabidopsis* mesophyll protoplasts (Yoo et al., 2007) with modifications described in **Supplementary Method 1**. Briefly, barley leaves without the epidermis were transferred into 5 mL of the enzyme solution [2% w/v Cellulase Onozuka R-10 (Duchefa, Haarlem, Netherlands), 0.1% w/v Macerozyme R-10 (Duchefa, Haarlem, Netherlands), 0.1% w/v BSA (Sigma-Aldrich, St. Louis, MO, United States), 0.55 M Mannitol (Sigma-Aldrich, St. Louis, MO, United States), pH 5.7]. Leaves were incubated in a 6-well cell culture plate (Sigma-Aldrich, St. Louis, MO, United States) in the dark for 2 h at 28°C. Released protoplasts were mixed with one volume of W5 solution [154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 2 mM MES (Sigma-Aldrich, St. Louis, MO, United States, all), pH 5.7], filtered through a 70 μM gauze (Sigma-Aldrich, St. Louis, MO, United States), washed with 2.5 mL of W5 solution and centrifuged for 3 min at 70 × g, RT (Centrifuge 5810 R, Sigma-Aldrich, St. Louis, MO, United States). Protoplasts were gently re-suspended in 5 mL W5 solution and kept on ice for 30 min to allow cell sedimentation by gravity. Protoplasts were re-suspended to a final concentration of approximately 2×10^5 protoplasts/mL in MMG solution (0.6 M Mannitol, 4 mM MES, 15 mM $MgCl_2$, pH 5.7). Fluorescein diacetate (FDA) was used to determine cell viability under a using a UV-light Zeiss Axioskop 2 Plus microscope and Canon EOS D100 digital camera.

For each transfection assay, 100 μL of protoplasts were mixed with 7 μg of construct DNA and 110 μL of freshly prepared 40% PEG-solution [0.4 M mannitol, 40% w/v PEG-4000 (Sigma-Aldrich, St. Louis, MO, United States), 0.1 M $CaCl_2$ (Sigma-Aldrich, St. Louis, MO, United States), pH 5.7]. After a 10 min incubation in the dark, protoplasts were washed twice with 1.5 mL of W5 solution and centrifuged at 200 × g for 5 min (Centrifuge 5415 D, Sigma-Aldrich, St. Louis, MO, United States). Protoplasts were re-suspended in 80 μL of W5 solution and incubated in the dark for 24 h, at room temperature. Three independent transfection assays were performed per construct and averaged as biological replicates.

HvCslF6 promoter deletion series

The *HvCslF6* (HORVU.MOREX.r3.7HG0698110) promoter deletion constructs were generated by PCR amplification of a 3,000 bp region upstream *HvCslF6* start codon as follows: 1 μL cv. Golden Promise gDNA, 1.25 μL each primer (forward and reverse listed in **Supplementary Table 7**) at 10 mM, 5 μL 5X Phusion HF Buffer, 0.2 μL Phusion HF DNA polymerase (Thermo Fisher Scientific, Waltham, MA, United States) and 11.3 μL sdH_2O in a total volume of 20 μL. Phusion PCR reaction conditions: 98°C for 30 s, 35 cycles (98°C for 10 s, 55°C for 30 s, 72°C for 30 s/kb), 72°C for 5 min and 4°C hold. Promoter fragments were ligated by NotI restriction enzyme cloning into pGreenII 0800-LUC vector which contains *luciferase* and *renilla* reporter genes, enabled to be used in Dual Luciferase Reporter Assays (Promega, Madison, WI, United States). A total of seven PCR amplicons corresponding to *HvCslF6* promoter fragments were NotI digested and purified with the QIAquick Gel Extraction Kit (Qiagen GmbH, Hilden, Germany). pGreenII-0800 vector was linearized in a similar digestion reaction using NotI with the addition of 1 μL TSAP (Promega, Madison, WI, United States) and purified as described for the *pHvCslF6* fragments. Following this, T4 DNA ligations (Promega, Madison, WI, United States) were set up for each construct and 1 μL of the ligation reaction was used for *E. coli* transformation. Plasmids were purified with QIAprep Spin Miniprep Kit (Qiagen GmbH, Hilden, Germany) and confirmed by Sanger sequencing using M13F primer at the Genome Technology facility, The James Hutton Institute.

2 http://plants.ensembl.org

Dual luciferase assays

The Dual-Luciferase® Reporter Assay System (Promega, Madison, WI, United States) was used to determine the promoter activity of *HvCslF6* deletion constructs in barley protoplasts, 24 h after transfection. Luminescence results were obtained using a Varioskan LUX Multimode Microplate Reader (Thermo Fisher Scientific, Waltham, MA, United States) following the dual luciferase kit instructions. Protoplasts were lysed with 80 μL of a 1X passive lysis buffer and incubated for 15 min at room temperature. In parallel, 100 μL of LAR II (Luciferase Assay Reagent II) were pre-dispensed in a Corning® 96-well opaque flat bottom microplate (Thermo Fisher Scientific, Waltham, MA, United States). For all assays, 20 μL of lysed protoplasts were transferred to the 96-well plate and incubated for 1 min, room temperature. After Firefly luciferase (Fluc) activity was measured (3 s per sample with a 1.5 s delay between wells), 100 μL of Stop and Glo Reagent were added enabling Renilla luciferase (Rluc) detection (1.5 s with a 0.5 s delay between wells). For each construct tested, three technical replicates were performed per transfection assay. A negative control, which included protoplasts transformed with empty vector and dual luciferase reagents, was used in each batch of transfections. Differences in dual luciferase reporter assays across *HvCslF6* promoter deletion constructs were determined by one-way ANOVA followed by Tukey's test using GraphPad Prism 8.4.2 Software (CA, United States). Pairwise comparisons were determined by two-tailed Student's *t*-tests.

R2R3-MYB phylogeny construction

The protein sequences of cv. Morex v3 genome[3] were used to identify R2R3 MYB family members. The sequences were uploaded to a webserver[4] for automatic identification of MYB gene family members (Pucker, 2022) using default parameters. Sequences which were not classified as R2R3 MYB were removed for the phylogenetic analysis. Protein sequences of selected *Arabidopsis* and rice MYB proteins (MYB61, including OsMYB61a and OsMYB61b and MYB46/83) were added and the sequences aligned using ClustalW in MEGA (Kumar et al., 2018). Unreliable positions in the alignment were removed using BMGE (Block Mapping and Gathering with Entropy; Criscuolo and Gribaldo, 2010) v2.0. Model selection for amino acid substitution was done in MEGA[5] resulting in LG model plus gamma distribution with invariant sites (LG+G+I) as the best choice. The unrooted phylogenetic tree was build using Maximum Likelihood with 100 bootstrap replications to estimate bootstrap support. Only bootstrap values above 0.5 are displayed.

Gene network construction and analysis

Gene expression values from 808 individual samples (**Supplementary Table 4**) were obtained and are available from Milne et al. (2021). In short, the reads were downloaded from the sequence read archive[6] and mapped against BaRTv1 (Rapazote-Flores et al., 2019) using Salmon (Patro et al., 2017). Expression of the isoforms were added up to obtain gene expression. For network construction low expressed genes were removed by filtering for a TPM (transcript per million) of above 5 in at least a third of the samples. This condensed the number of expressed genes from 60,444 to 17,678. Weighted gene correlation network analysis (WGCNA; Langfelder and Horvath, 2008) was used for network construction. Euclidean clustering beforehand highlighted one sample from E15.MicrosporeEmbryogenic (SRR6433018) as outlier. Therefore, this sample was removed, leaving 807 samples for further analysis. The soft-thresholding power was set to 10 to achieve scale-free topology. Network analysis was carried out using the blockwise module method (Langfelder and Horvath, 2008), with a power of 10, TOMType = unsigned, corType = bicor, networkType = signed hybrid, minModuleSize = 30, maxPOutliers = 0.05, mergeCutHeight = 0.15, deepSplit = 3. Visualization of the *HvCslF6* network was performed in Gephi[7]. Size and coloring of the nodes was determined by degree centrality (the number or proportion of other nodes linked to a specific node). Edge coloring was based on weight. The subnetwork was extracted by filtering for nodes connected to *HvCslF6* (BART1_0-p53331). For this we applied an edge weight which is a threshold, whereby connections between gene represent by values greater or equal to 0.188 are correlated.

Cloning of candidate transcription factor

A nested PCR was used to amplify *HvMYB61* (HORVU.MOREX.r3.1HG0018590) containing: 1 μL of cv. Golden Promise protoplast cDNA (1:10 diluted in sdH_2O), 1.25 μL of each primer (forward and reverse listed on **Supplementary Table 7**) at 10 mM, 5 μL of 5X Phusion HF Buffer, 0.2 μL of Phusion HF DNA polymerase (Thermo Fisher Scientific, Waltham, MA, United States) and 11.3 μL of sdH_2O in a total volume of 20 μL. Phusion PCR reaction

3 https://plants.ensembl.org/Hordeum_vulgare/Info/Index
4 http://pbb.bot.nat.tu-bs.de/MYB_annotator
5 https://www.megasoftware.net/
6 https://www.ncbi.nlm.nih.gov/sra
7 https://gephi.org/

conditions are described above. A 1 μL aliquot of the PCR mix was used as a template to attach Gateway® attB sites, 8x His-tag and thrombin cleavage site at the C-terminus in a similar Phusion HF PCR reaction and gel purified. *HvMYB61* was cloned into pDONR207^{TM}, linearized with *EcoRI* by a BP reaction containing: 3 μL of the purified attB-PCR product (34 ng/μL), 1 μL of linearized pDONR207^{TM} (50 ng/μL), 2 μL 5X BP Clonase^{TM} Reaction Buffer (Thermo Fisher Scientific, Waltham, MA, United States) and 4 μL of TE buffer, pH 8.0 in a total volume of 10 μL. The BP reaction was incubated at RT overnight and stopped by adding 2 μL of Proteinase K and incubated for 10 min at 37°C. 1 μL of the BP reaction was used for *E. coli* transformation. Positive colonies were transferred to 4 mL of LB medium with gentamycin (100 μg/μL) and incubated overnight, 37°C and shaking at 230 rpm. Plasmids were purified using the QIAprep Spin Miniprep Kit (Qiagen GmbH, Hilden, Germany) and confirmed by Sanger sequencing. *HvMYB61* was transferred to pBract214m-HSPT plant expression vector by LR reaction containing: 2 μL of EcoRI linearized pDONR207^{TM}-HvMYB61 (40 ng/μL), 5 μL of pBract214m-HSPT (30 ng/μL), 1 μL TE Buffer, pH 8.0 and 2 μL LR Clonase^{TM} II enzyme mix (Thermo Fisher Scientific, Waltham, MA, United States) in a total volume of 10 μL. The reaction was incubated at 25°C for 3 h and stopped by adding 1 μL of the Proteinase K (37°C for 10 min). After *E. coli* transformation, positive colonies were grown on LB medium with kanamycin (100 μg/μL) for plasmid purification. pBract214m-HSPT-HvMYB61 construct sequence was confirmed as described above using the ZmUbi forward primer (**Supplementary Table 7**).

RNA isolation, cDNA synthesis, and quantitative real-time PCR on protoplasts

Total mRNA was extracted from ∼2 × 10^5 protoplasts (100 μL) using TRIzol® Reagent (Thermo Fisher Scientific, Waltham, MA, United States) according to the manufacturer's instructions. RNA concentration and purity were measured using NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA, United States). Total mRNA to cDNA conversion was performed using cDNA EcoDry^{TM} Premix (Takara, Kyoto, Japan). For each reaction, 100 ng of total mRNA was added to the lyophilized master mix following the manufacturer's instructions. Quantitative real-time PCR (qRT-PCR) was performed in a StepOne Real-Time PCR machine (Thermo Fisher Scientific, Waltham, MA, United States) using PowerUp SYBR Green Master Mix (Thermo Fisher Scientific, Waltham, MA, United States) to determine *HvCslF6* relative transcript abundance. Three independent experiments were performed, each of them included: HvMYB61 PEG-mediated transfection, RNA extraction, cDNA synthesis and qPCRs, respectively. Three replicate qRT-PCR reactions were performed for each cDNA sample (technical replicates) including a negative control (sdH_2O as template). Each qRT-PCR reaction contained: 2 μL protoplast cDNA (1:2 dilution), 5 μL SYBR Green, 1 μL each forward and reverse primer at 4 mM and 1 μL of sdH_2O in a total volume of 10 μL. Primer sequences for *HvCslF6* and qRT-PCR conditions were used as in Burton et al. (2008) and described in **Supplementary Table 7**. Relative *HvCslF6* gene expression was normalized to α-*Tubulin*, *Gapdh* and *Hsp70* housekeeping genes and calculated using the $2^{-\Delta\Delta CT}$ method (Vandesompele et al., 2002) for multiple control genes. Statistical differences in *HvCslF6* expression were determined by Student's *t*-test compared to WT protoplasts (empty vector) using GraphPad Prism 9.1 Software (CA, United States).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

GG-G, GD, KH, MT, RB, and RW conceived the work. GD initiated the cloning of *HvCslF6* promoter deletion series. GG-G performed the laboratory work, TFBS, and data analyses. KH carried out co-expression analyses. MS performed the gene network construction. RS carried out the barley transformations. GG-G, MS, KH, and MT drafted the manuscript. GD, GF, RB, and RW reviewed the manuscript. All authors have read and approved the manuscript.

Funding

This work was supported by the Australian Research Council, the BBSRC, and Rural and Environment Science and Analytical Services Division of the Scottish Government.

Acknowledgments

We thank Abdellah Barakate for providing pZmUBI:bcoGFP-ER and pBract214m-HSPT constructs. We also thank the Genome Technology team at the James Hutton Institute (United Kingdom) for Sanger sequencing and collaborators at The University of Adelaide (Australia) for barley transformations.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.883139/full#supplementary-material

References

Aaman, P., and Graham, H. (1987). Analysis of total and insoluble mixed-linked (1,3; 1,4)-β-D-glucans in barley and oats. *J. Agric. Food Chem.* 35, 704–709. doi: 10.1021/jf00077a016

Aubert, M. K., Coventry, S., Shirley, N. J., Betts, N. S., Würschum, T., Burton, R. A., et al. (2018). Differences in hydrolytic enzyme activity accompany natural variation in mature aleurone morphology in barley (*Hordeum vulgare* L.). *Sci. Rep.* 8:11025. doi: 10.1038/s41598-018-29068-29064

Bacic, A., and Stone, B. (1980). A (1→3)- and (1→4)-linked β-D-glucan in the endosperm cell-walls of wheat. *Carbohydr. Res.* 82, 372–377. doi: 10.1016/S0008-6215(00)85713-85714

Bacic, A., and Stone, B. A. (1981). Chemistry and organization of aleurone cell wall components from wheat and barley. *Funct. Plant Biol.* 8, 475–495. doi: 10.1071/PP9810475

Bai, J., Ren, Y., Li, Y., Fan, M., Qian, H., Wang, L., et al. (2019). Physiological functionalities and mechanisms of β-glucans. *Trends Food Sci. Technol.* 88, 57–66. doi: 10.1016/j.tifs.2019.03.023

Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., Alpi, E., Antunes, R., et al. (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 45, D158–D169. doi: 10.1093/nar/gkw1099

Betts, N. S., Wilkinson, L. G., Khor, S. F., Shirley, N. J., Lok, F., Skadhauge, B., et al. (2017). Morphology, carbohydrate distribution, gene expression, and enzymatic activities related to cell wall hydrolysis in four barley varieties during simulated malting. *Front. Plant Sci.* 8:1872. doi: 10.3389/fpls.2017.01872

Burn, J. E., Hocart, C. H., Birch, R. J., Cork, A. C., and Williamson, R. E. (2002). Functional analysis of the Cellulose synthase genes *CesA1, CesA2, CesA3*, and *Arabidopsis*. *Plant Physiol.* 129, 797–807. doi: 10.1104/pp.010931

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)-β-d-glucans and alters their fine structure. *Plant Biotechnol. J.* 9, 117–135. doi: 10.1111/j.1467-7652.2010.00532.x

Burton, R. A., and Fincher, G. B. (2009). (1,3;1,4)-β-D-Glucans in cell walls of the poaceae, lower plants, and fungi: a tale of two linkages. *Mol. Plant* 2, 873–882. doi: 10.1093/mp/ssp063

Burton, R. A., and Fincher, G. B. (2012). Current challenges in cell wall biology in the cereals and grasses. *Front. Plant Sci.* 3:130. doi: 10.3389/fpls.2012.00130

Burton, R. A., and Fincher, G. B. (2014). Evolution and development of cell walls in cereal grains. *Front. Plant Sci.* 5:456. doi: 10.3389/fpls.2014.00456

Burton, R. A., Jobling, S. A., Harvey, A. J., Shirley, N. J., Mather, D. E., Bacic, A., et al. (2008). The genetics and transcriptional profiles of the *Cellulose synthase-like, HvCslF* gene family in barley. *Plant Physiol.* 146, 1821–1833. doi: 10.1104/pp.107.114694

Burton, R. A., Shirley, N. J., King, B. J., Harvey, A. J., and Fincher, G. B. (2004). The *CesA* gene family of barley. quantitative analysis of transcripts reveals two groups of co-expressed genes. *Plant Physiol.* 134, 224–236. doi: 10.1104/pp.103.032904

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

Carpita, N., Tierney, M., and Campbell, M. (2001). "Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics," in *Plant Cell Walls*, eds N. C. Carpita, M. Campbell, and M. Tierney (Dordrecht: Springer).

Castro-Mondragon, J. A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., et al. (2022). JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 50, D165–D173. doi: 10.1093/nar/gkab1113

Chow, C. N., Lee, T. Y., Hung, Y. C., Li, G. Z., Tseng, K. C., Liu, Y. H., et al. (2019). PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from chip-seq experiments in plants. *Nucleic Acids Res.* 47, D1155–D1163. doi: 10.1093/nar/gky1081

Cory, A. T., Båga, M., Anyia, A., Rossnagel, B. G., and Chibbar, R. N. (2012). Genetic markers for *CslF6* gene associated with (1,3;1,4)-β-glucan concentration in barley grain. *J. Cereal Sci.* 56, 332–339. doi: 10.1016/j.jcs.2012.02.003

Criscuolo, A., and Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* 10:210. doi: 10.1186/1471-2148-10-210

Dimitroff, G., Little, A., Lahnstein, J., Schwerdt, J. G., Srivastava, V., Bulone, V., et al. (2016). (1,3;1,4)-β-Glucan biosynthesis by the CSLF6 enzyme: position and flexibility of catalytic residues influence product fine structure. *Biochemistry* 55, 2054–2061. doi: 10.1021/acs.biochem.5b01384

Doblin, M. S., Pettolino, F. A., Wilson, S. M., Campbell, R., Burton, R. A., Fincher, G. B., et al. (2009). A barley *Cellulose synthase-like CslH* gene mediates (1,3;1,4)-β-D-glucan synthesis in transgenic *Arabidopsis*. *Proc. Natl. Acad. Sci. U S A.* 106, 5996–6001. doi: 10.1073/pnas.0902019106

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

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

Fornes, O., Castro-Mondragon, J. A., Khan, A., Van Der Lee, R., Zhang, X., Richmond, P. A., et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 48, D87–D92. doi: 10.1093/nar/gkz1001

Garcia-Gimenez, G., Russell, J., Aubert, M. K., Fincher, G. B., Burton, R. A., Waugh, R., et al. (2019). Barley grain (1,3;1,4)-β-glucan content: effects of transcript and sequence variation in genes encoding the corresponding synthase and endohydrolase enzymes. *Sci. Rep.* 9:17250. doi: 10.1038/s41598-019-53798-53798

Garcia-Gimenez, G., Barakate, A., Smith, P., Stephens, J., Khor, S. F., Doblin, M. S., et al. (2020). Targeted mutation of barley (1,3;1,4)−β−glucan synthases reveals complex relationships between the storage and cell wall polysaccharide content. *Plant J.* 104, 1009–1022. doi: 10.1111/tpj.14977

Gupta, M., Abu-Ghannam, N., and Gallaghar, E. (2010). Barley for brewing: characteristic changes during malting, brewing and applications of its by-products. *Compr. Rev. Food Sci. Food Saf.* 9, 318–328. doi: 10.1111/j.1541-4337.2010.00112.x

Han, F., Ullrich, S. E., Chirat, S., Menteur, S., Jestin, L., Sarrafi, A., et al. (1995). Mapping of β-glucan content and β-glucanase activity loci in barley grain and malt. *Theor. Appl. Genet.* 91, 921–927. doi: 10.1007/BF00223901

Harris, P. J. (2006). Primary and secondary plant cell walls: a comparative overview. *New Zeal. J. For. Sci.* 36, 36–53. doi: 10.1111/j.1365-2621.2006.01470.x

Heberle, H., Meirelles, V. G., da Silva, F. R., Telles, G. P., and Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 16:169. doi: 10.1186/s12859-015-0611-613

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

Houston, K., Burton, R. A., Sznajder, B., Rafalski, A. J., Dhugga, K. S., Mather, D. E., et al. (2015). A genome-wide association study for culm cellulose content in barley reveals candidate genes co-expressed with members of the *Cellulose synthase A* gene family. *PLoS One* 10:e0130890. doi: 10.1371/journal.pone.0130890

Houston, K., Russell, J., Schreiber, M., Halpin, C., Oakey, H., Washington, J. M., et al. (2014). A genome wide association scan for (1,3;1,4)-β-glucan content in the grain of contemporary 2-row Spring and Winter barleys. *BMC Genomics* 15:907. doi: 10.1186/1471-2164-15-907

Hu, G., Burton, C., Hong, Z., and Jackson, E. (2014). A mutation of the *Cellulose synthase-like, CslF6* gene in barley (*Hordeum vulgare* L.) partially affects the β-glucan content in grains. *J. Cereal Sci.* 59, 189–195. doi: 10.1016/j.jcs.2013.12.009

Jobling, S. A. (2015). Membrane pore architecture of the CSLF6 protein controls (1-3,1-4)-β-glucan structure. *Sci. Adv.* 1:e1500069. doi: 10.1126/sciadv.1500069

Kim, W. C., Ko, J. H., and Han, K. H. (2012). Identification of a cis-acting regulatory motif recognized by MYB46, a master transcriptional regulator of secondary wall biosynthesis. *Plant Mol. Biol.* 78, 489–501. doi: 10.1007/s11103-012-9880-9887

Kim, W. C., Reca, I. B., Kim, Y. S., Park, S., Thomashow, M. F., Keegstra, K., et al. (2014). Transcription factors that directly regulate the expression of CslA9 encoding mannan synthase in *Arabidopsis thaliana*. *Plant Mol. Biol.* 84, 577–587. doi: 10.1007/s11103-013-0154-159

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

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559

Little, A., Schwerdt, J. G., Shirley, N. J., Khor, S. F., Neumann, K., O'Donovan, L. A., et al. (2018). Revised phylogeny of the *Cellulose Synthase* gene superfamily: insights into cell wall evolution. *Plant Physiol.* 177, 1124–1141. doi: 10.1104/pp.17.01718

Ma, D., and Constabel, C. P. (2019). MYB repressors as regulators of phenylpropanoid metabolism in plants. *Trends Plant Sci.* 24, 275–289. doi: 10.1016/j.tplants.2018.12.003

Mascher, M., Wicker, T., Jenkins, J., Plott, C., Lux, T., Koh, C. S., et al. (2021). Long-read sequence assembly: a technical evaluation in barley. *Plant Cell* 33, 1888–1906. doi: 10.1093/plcell/koab077

Matros, A., Houston, K., Tucker, M. R., Schreiber, M., Berger, B., Aubert, M. K., et al. (2021). Genome-wide association study reveals the genetic complexity of fructan accumulation patterns in barley grain. *J. Exp. Bot.* 72, 2383–2402. doi: 10.1093/jxb/erab002

Matys, V., Kel-Margoulis, O. V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., et al. (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* 34, D108–D110. doi: 10.1093/nar/gkj143

McCarthy, R. L., Zhong, R., and Ye, Z. H. (2009). MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell Physiol.* 50, 1950–1964. doi: 10.1093/pcp/pcp139

Milne, L., Bayer, M., Rapazote-Flores, P., Mayer, C. D., Waugh, R., and Simpson, C. G. (2021). EORNA, a barley gene and transcript abundance database. *Sci. Data* 8:90. doi: 10.1038/s41597-021-00872-874

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

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197

Penfield, S., Meissner, R. C., Shoue, D. A., Carpita, N. C., and Bevan, M. W. (2001). MYB61 is required for mucilage deposition and extrusion in the *Arabidopsis* seed coat. *Plant Cell* 13, 2777–2791. doi: 10.1105/tpc.13.12.2777

Prouse, M. B., and Campbell, M. M. (2013). Interactions between the R2R3-MYB transcription factor, AtMYB61, and target DNA binding sites. *PLoS One* 8:e65132. doi: 10.1371/journal.pone.0065132

Pucker, B. (2022). Automatic identification and annotation of MYB gene family members in plants. *BMC Genomics* 23:220. doi: 10.1186/s12864-022-08452-5

Rao, A. L. N. (2007). Preparation and inoculation of mesophyll protoplasts from monocotyledenous and dicotyledenous hosts. *Curr. Protoc. Microbiol.* Chapter 16:Unit 16D.2. doi: 10.1002/9780471729259.mc16d02s4.

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

Rapazote-Flores, P., Bayer, M., Milne, L., Mayer, C. D., Fuller, J., Guo, W., et al. (2019). BaRTv1.0: an improved barley reference transcript dataset to determine accurate changes in the barley transcriptome using RNA-seq. *BMC Genomics* 20:968. doi: 10.1186/s12864-019-6243-6247

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

Romero-Romero, J. L., Inostroza-Blancheteaud, C., Orellanaa, D., Aquea, F., Reyes-Díaz, M., Gil, M. P., et al. (2018). Stomata regulation by tissue-specific expression of the *Citrus sinensis* MYB61 transcription factor improves water-use efficiency in *Arabidopsis*. *Plant Physiol. Biochem.* 130, 54–60. doi: 10.1016/j.plaphy.2018.06.034

Shahmuradov, I. A., Solovyev, V. V., and Gammerman, A. J. (2005). Plant promoter prediction with confidence estimation. *Nucleic Acids Res.* 33, 1069–1076. doi: 10.1093/nar/gki247

Slakeski, N., and Fincher, G. B. (1992). Developmental regulation of (1,3;1,4)-β-glucanase gene expression in barley. *Plant Physiol.* 99, 1226–1231. doi: 10.1104/pp.99.3.1226

Taketa, S., Yuo, T., Tonooka, T., Tsumuraya, Y., Inagaki, Y., Haruyama, N., et al. (2012). Functional characterization of barley betaglucanless mutants demonstrates a unique role for *CslF6* in (1,3;1,4)-β-D-glucan biosynthesis. *J. Exp. Bot.* 63, 381–392. doi: 10.1093/jxb/err285

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

Tripathi, R. K., Bregitzer, P., and Singh, J. (2018). Genome-wide analysis of the SPL/miR156 module and its interaction with the AP2/miR172 unit in barley. *Sci. Rep.* 8:7085. doi: 10.1038/s41598-018-25349-0

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3:research0034.1. doi: 10.1186/gb-2002-3-7-research0034

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

Wilson, S. M., Ho, Y. Y., Lampugnani, E. R., Van de Meene, A. M. L., Bain, M. P., Bacic, A., et al. (2015). Determining the subcellular location of synthesis and assembly of the cell wall polysaccharide (1,3; 1,4)-β-D-glucan in grasses. *Plant Cell* 27, 754–771. doi: 10.1105/tpc.114.135970

Wong, S. C., Shirley, N. J., Little, A., Khoo, K. H. P., Schwerdt, J., Fincher, G. B., et al. (2015). Differential expression of the *HvCslF6* gene late in grain development may explain quantitative differences in (1,3;1,4)-β-glucan concentration in barley. *Mol. Breed.* 35:20. doi: 10.1007/s11032-015-0208-6

Woodward, J. R., Fincher, G. B., and Stone, B. A. (1983). Water-soluble (1→3), (1→4)-β-D-glucans from barley (*Hordeum vulgare*) endosperm. II. fine structure. *Carbohydr. Polym.* 3, 207–225. doi: 10.1016/0144-8617(83)90019-X

Yoo, S.-D., Cho, Y.-H., and Sheen, J. (2007). *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2, 1565–1572. doi: 10.1038/nprot.2007.199

Zhang, R., Tucker, M. R., Burton, R. A., Shirley, N. J., Little, A., Morris, J., et al. (2016a). The dynamics of transcript abundance during cellularization of

developing barley endosperm. *Plant Physiol.* 170, 1549–1565. doi: 10.1104/pp.15.01690

Zhang, Z., Hu, X., Zhang, Y., Miao, Z., Xie, C., Meng, X., et al. (2016b). Opposing control by transcription factors MYB61 and MYB3 increases freezing tolerance by relieving c-repeat binding factor suppression. *Plant Physiol.* 172, 1306–1323. doi: 10.1104/pp.16.00051

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

Zhao, K., Lin, F., Romero-Gamboa, S. P., Saha, P., Goh, H. J., An, G., et al. (2019). Rice genome-scale network integration reveals transcriptional regulators of grass cell wall synthesis. *Front. Plant Sci.* 10:1275. doi: 10.3389/fpls.2019.01275

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

Zhong, R., and Ye, Z. H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.* 53, 368–380. doi: 10.1093/pcp/pcr185

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

Zhou, M., Zhang, K., Sun, Z., Yan, M., Chen, C., Zhang, X., et al. (2017). LNK1 and LNK2 corepressors interact with the MYB3 transcription factor in phenylpropanoid biosynthesis. *Plant Physiol.* 174, 1348–1358. doi: 10.1104/pp.17.00160

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 03 October 2022
DOI 10.3389/fpls.2022.998664

OPEN ACCESS

EDITED BY
Rowan Andrew Craig Mitchell,
Rothamsted Research,
United Kingdom

REVIEWED BY
Duarte D. Figueiredo,
Max Planck Institute of Molecular Plant Physiology, Germany
Vikash Kumar Yadav,
Goa University, India

*CORRESPONDENCE
Jie Zhao
jzhao@whu.edu.cn

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

RECEIVED 20 July 2022
ACCEPTED 16 September 2022
PUBLISHED 03 October 2022

CITATION
Li C, Hu F, Chen H and Zhao J (2022) Transcriptome characteristics during cell wall formation of endosperm cellularization and embryo differentiation in *Arabidopsis*. *Front. Plant Sci.* 13:998664. doi: 10.3389/fpls.2022.998664

Transcriptome characteristics during cell wall formation of endosperm cellularization and embryo differentiation in *Arabidopsis*

Chengcheng Li, Fan Hu, Hongyu Chen and Jie Zhao*

State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan, China

Embryonic and endosperm development are important biological events during *Arabidopsis* seed development, and are controlled by dynamic changes in a range of gene expression. Nevertheless, the regulatory mechanisms of endosperm cellularization and embryo differentiation remain unclear. Here, we characterized the early embryo and endosperm development of the *naa15* mutant that had abnormal embryo differentiation and incomplete endosperm cellularization compared to WT of *Arabidopsis*, and comparatively investigated the changes of gene expressions in WT seeds at 3, 4, and 5 days after pollination (3W, 4W, and 5W) and the white homozygous aborted *naa15* seeds at 5, 6, and 7 DAP (5M, 6M, and 7M) from *naa15-1*/+ siliques using RNA sequencing and qPCR assays. The transcriptome analyses showed that there were 2040 and 3630 differentially expressed genes (DEGs) in 4W (at endosperm cellularization initiation stage and heart embryo stage) vs 3W (at syncytium stage and globular embryo stage), and 5W (at end of endosperm cellularization stage and torpedo embryo stage) vs 4W, respectively. The KEGG and GO analyses showed that lipid metabolic processes and transmembrane transport related to cell wall biogenesis, cell division and differentiation, the plant hormone signaling pathway, photosynthesis, and transcription regulator activity were evidently enriched in WT and *naa15*. The heatmap and qPCR analyses showed that auxin response genes (*ARFs*), auxin transport genes (*PINs*) cytokinin synthesis genes (*LOGs*), cytokinin dehydrogenase genes (*CKXs*), cytokinin receptor, transcription factors (*MYB*, *bHLH*, *MADS-box*, and *ERF*) were significantly downregulated in *naa15* compared to WT. A series of cell wall genes annotated to xyloglucan endotransglycosylase/hydrolase, pectin methyl esterase, and pectin methyl esterase inhibitor were also identified in these DEGs. Moreover, using an immunofluorescent assay, the features of cell walls displayed that cellulose fluorescence signals in the embryo and endosperm of *naa15* were significantly decreased, and the signals of low- and high- methyl esterification of pectin were also obviously decreased in the endosperm of *naa15*. In summary, we identified a large number of DEGs and investigated the features of cell walls during endosperm cellularization and embryonic differentiation, which provided important information on transcription and gene expression to reveal their regulatory mechanisms.

KEYWORDS

Arabidopsis, endosperm cellularization, embryo differentiation, cell wall, transcriptomics, gene expression

Introduction

Seed development in angiosperms is triggered by double fertilization, in which sperm cells reach the embryo sac, and then two sperm cells fuse with egg cell and central cell resulting in the formation of diploid embryo and triploid endosperm, respectively (Dresselhaus et al., 2016; Wang et al., 2022). The intact seed contains the embryo, endosperm, and seed coat (Lafon-Placette and Köhler, 2014). In *Arabidopsis*, the embryo undergoes a series of cell division and organ differentiation, leading to its patterning establishment and maturation. The embryo reaches the pre-globular stage after eight rounds of cell divisions. Meanwhile, the endosperm conducts rapid and successive free nuclear divisions without cell wall formation and enters the syncytial endosperm phase (Sørensen et al., 2002). With the embryo development transition from the globular to heart stage, endosperm cellularization begins from the micropylar zone near the developing embryo with cell wall formation and moves forward the chalazal pole of seed. When the embryo enters torpedo phase, endosperm completes cellularization process (Berger, 2003; Li et al., 2013). When embryo develops from heart to torpedo stage, chloroplast biogenesis initiates and embryo turns green (Mansfield and Briarty, 1991). Endosperm cellularization process is an important milestone in endosperm development and embryogenesis, and it proceeds gradually with embryo differentiation (Wang et al., 2022).

The regulatory mechanisms of embryo differentiation and endosperm cellularization are sophisticated and involve in multiplex pathways of gene regulation. The *apetala2* (*ap2*) mutant endosperm undergoes an early expanded growth period to lead to delayed endosperm cellularization and overgrowth of the endosperm central vacuole (Ohto et al., 2009). MINI3 (WRKY10), a WRKY-type transcription factor, is specifically expressed in endosperm and embryo, and its mutant has a precocious cellularization of endosperm and produces small seeds (Luo et al., 2005). AGAMOUS-LIKE proteins (AGLs) are a large family of MADS-box proteins that form heterodimers or homodimers among the MADS-box family. In the three mutants of *agl61*, *agl62*, and *agl80*, embryogenesis and endosperm cellularization are damaged to varying degrees (Portereiko et al., 2006; Kang et al., 2008; Steffen et al., 2008). In other transcription factors, the genes of the *ARF* and *bHLH* families are expressed in seeds, and their deletions also affect development processes of embryo and endosperm, leading to seed abortion (Sun et al., 2010).

Plant hormones, important signaling molecules, play vital roles during the entire plant lifecycle. Auxin is critical in embryo and endosperm development, and its biosynthesis is activated when sperm-egg fuse to form zygotes (Figueiredo et al., 2015). However, the change in auxin homeostasis affects its signal transmission in the seed coat and embryo, eventually leading to seed abortion. Excessive accumulation of auxin delays endosperm cellularization (Batista et al., 2019). In seed formation, auxin and cytokinin are related to cell division and differentiation, as well as organ growth and development (Bishopp et al., 2011; Vanstraelen and Benkova, 2012). The sizes of embryos and seeds are increased owing to disordered cytokinin homeostasis in *ahk2 ahk3 ahk4* mutants (Riefler et al., 2006). Brassinosteroids are essential steroid hormones that play critical roles in plant reproductive growth and seed development. In the rice BR-deficient mutant *brd2* and *Arabidopsis* BR deficient mutants *dwf5* and *det2*, there are smaller seeds than WT (Hong et al., 2005; Jiang and Lin, 2013).

In addition to hormones, endosperm and embryo development are also related to cell wall properties. Methyl-esterified pectin is demethylesterified by pectin methylesterases (PMEs), which change the rigidity and stiffening of the cell wall in accordance with the pattern of dimethyl esterification (Levesque-Tremblay et al., 2015).

In 5-6 DAP *Arabidopsis brd-1* mutant, the fluorescence signal of low-esterified homogalacturonic acids (HGs) in endosperm cell wall are faint and heterogeneous, but the signal of high-esterified HGs significantly increases (Cruz-Valderrama et al., 2018). In rice, there are seven predicted invertase/PME inhibitor genes, which are exclusively expressed during seed development (French et al., 2014). The overexpression of the pectin methylesterase gene *PME5 in Arabidopsis* reduces the pectin methylesterification of flower primordia leading to the formation of ectopic primordia, whereas the overexpression of *pectin methylesterase inhibitor 3* (*PMEI3*) significantly increases the pectin methylesterification, which inhibits organ formation (Peaucelle et al., 2011). In the *Arabidopsis zou-4* mutant, the seeds retain persistent endosperm growth to lead to abnormal increases in seed size, and the embryo sheath that is located near the surface of embryo at the late heart stage cannot be detected

with a JIM12 antibody (Fourquin et al., 2016; Moussu et al., 2017).

N-terminal protein α-acetylation (NTA) is a highly prevalent protein modification that affects multiple cellular functions, such as cell multiplication, regulation of stress stimulation, and immune response (Linster et al., 2015; Xu et al., 2015; Giglione and Meinnel, 2021). Plastids are one of the places where co- and post-translational N-terminal acetylation occurs, and approximately 20-30% of all plastid proteins are affected by NTA (Bischof et al., 2011; Bienvenut et al., 2012). The process of N-terminal acetylation is that N-terminal acetyltransferases (NATs) transfer acetyl groups of acetyl-CoA to α amino acids of nascent polypeptides with 25-50 amino acids (Gautschi et al., 2003; Varland et al., 2015). N-terminal acetyltransferase A (NatA), one of the Nats (NatA to NatF), consists of the catalytic subunit NAA10 and the auxiliary subunit NAA15, and is the predominant acetyltransferase complex in yeast and humans (Arnesen et al., 2009). The auxiliary subunit NAA15 anchored to the ribosome participates in acetyltransferase activity, and/or interacts with nascent polypeptides (Ree et al., 2018).

In *Arabidopsis*, a point mutation of *NAA15* leads to immune receptor suppression of NPR1 protein instability and decreases plant immunity (Xu et al., 2015), and T-DNA insertion mutants of *NAA10* and *NAA15* in *Arabidopsis* display embryo abortion and abnormal endosperm (Feng et al., 2016; Chen et al., 2018a). In humans, truncating variants of NAA15 lead to varying degrees of growth defects, including intellectual disability and autism spectrum disorder (Cheng et al., 2018). The artificial microRNAs against *NAA10* and *NAA15* significantly downregulate their transcripts and show drought stress tolerance by activation of the abscisic acid response in *Arabidopsis* (Linster et al., 2015). However, the transcriptional regulation of *NatA* during embryo differentiation and endosperm cellularization in *Arabidopsis* is unknown to date. Here, *Arabidopsis* seeds from WT and *naa15* at three developmental stages were used for transcriptomic analyses and gene expression assays. Meanwhile, the cell wall features were analyzed by detecting the degrees of pectin methyl esterification and cellulose. In our transcriptome profile, a list of genes associated with plant hormones were also identified and analyzed. These results contributed to understanding the transcriptional regulatory information and network relationship during embryo differentiation and endosperm cellularization.

Materials and methods

Plant material and growth conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) was used as wild-type plant (WT), and the T-DNA insertion mutant *naa15-1*/+ (CS836292) was used as described previously (Chen et al., 2018a). All plants were cultivated at 22 ± 2°C with 16-h light/8-h dark in a growth chamber of Wuhan University. The genotype assay of *naa15*/+ was shown in Supplementary Figure 1 by using PCR amplification, and the gene specific primers were listed in Supplementary Table 1.

Observation of cleared seeds and cellularization endosperm

Seeds were quickly and carefully dissected from siliques and then cleared in Hoyer's solution (chloral hydrate: glycerol: water, 8:1:2) 12 hours at 4°C (Chen et al., 2018a). The transparent seeds were placed onto slides and observed under a Leica SP8 microscope (Leica, Germany) equipped with differential interference contrast (DIC) optics. To observe endosperm cellularization, the seeds were fixed in 4% glutaraldehyde in PBS (pH 7.0) under vacuum conditions for 1 h and then transferred into freshly 4% glutaraldehyde in PBS (pH 7.0) and incubated overnight at 4°C. The samples were progressively dehydrated according to a series of alcohol gradients, 15%, 30%, 50%, 70%, 90%, 100%, and then rehydrated using alcohol gradients of 90%, 70%, 50%, 30%, 15%, each gradient for 20 min. The samples were cleared with Hoyer's solution overnight at 4°C, and observed under a Leica TCS SP8 confocal microscope. The excitation and emission wavelengths were 488 and 505–535 nm, respectively (Li et al., 2017).

RNA isolation and Illumina sequencing

For RNA-seq material, WT seeds at 3, 4, and 5 days after pollination (DAP) (3W, 4W, and 5W) and the homozygous white aborted *naa15* seeds at 5, 6, and 7 DAP (5M, 6M, and 7M) were separated under dissecting microscope, respectively. Transcriptome sequencing was performed by Majorbio Biotech Co., Ltd. (Shanghai, China). For RNA-sequencing (RNA-seq), the total RNA of samples was extracted using an OminiPlant RNA Kit (CWBIO, Beijing, China) and digested with DNaseI. RNA quality was analyzed by 1.2% agarose gel electrophoresis. The mRNA was purified using poly-T-oligo-attached magnetic beads. The purified mRNA was broken into 300 bp (± 20 bp) paired-end reads. Raw data were obtained and analyzed on an Illumina HiSeq 2500 platform. More than 6.73 Gb of high-quality reads were generated from each cDNA library.

Gene sequence assembly and functional annotations

To obtain high-quality clean reads, we removed some reads with over 10% uncertain bases (N) and the low sequencing quality reads of the raw reads were also cleared. Transcriptome

assembly was completed using Cufflinks 2.2.1 (http://cole-trapnell-lab.github.io/cufflinks/) and StringTie software (http://ccb.jhu.edu/software/stringtie/). Principal component analysis (PCA) of RNA-seq samples was implemented to detect the expression variances using the plotPCA function in RSEM software (http://deweylab.github.io/RSEM/). Gene functions were annotated using the NR, Swiss-Prot, Pfam, EggNOG, GO, and KEGG databases.

Identification of differentially expressed genes

The sequencing reads were aligned to the Arabidopsis TAIR 10.0 reference genome (http://plants.ensembl.org/Arabidopsisthaliana/Info/Index) using HISAT2 (http://ccb.jhu.edu/software/hisat2/index.shtml). The expression values of each transcript were calculated with the Fragments Per Kilobases per Million reads (FPKM) method. Differentially expressed genes (DEGs) were identified by comparing the gene expression levels between different groups using the DESeq R package. A $\log_2$ (Fold Change) ≥ 1 or ≤ -1 and P-value < 0.05 were set as the thresholds to discriminate DEGs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed to investigate the putative functions of DEGs. The top 30 terms of GO enrichment analysis of enriched DEGs were counted with the enriched P-value < 0.05. Venn diagrams were constructed using Venny 2.1 (http://bioinfogp.cnb.csic.es/tools/venny/) to obtain the intersection targets. The hierarchical clustering of heatmap analyses was generated using the online tool (http://www.heatmapper.ca/expression/) (Babicki et al., 2016). Pairwise distances were measured using Pearson correlation, and average linkage was set as the clustering method in heatmap analyses (Karunakaran et al., 2021; Lemke et al., 2021).

Validation of DEGs by qPCR

Total RNA was extracted using an OminiPlant RNA Kit (CWBIO, Beijing, China), and cDNA was synthesized using an EasyScript® One-Step gDNA Removal and cDNA Synthesis SuperMix Kit (TransGen, Beijing, China). Quantitative real-time PCR (qPCR) was conducted using Transtart Top Green qPCR SuperMix (TransGen Biotech, Beijing, China) and automatically analyzed by the Bio-Rad CFX Manager 3.1 system (Bio-Rad, USA). *AtGAPDH* was used as an internal reference to normalize the gene expression levels. The specific gene primers were designed by Primer Premier 5.0 and were listed in Supplementary Table 1. For qPCR analysis, the sample was independently repeated in triplicate, and the relative expression values were calculated by the delta-delta Ct (ΔΔCT) method as described previously (Li et al., 2017).

Semithin sections, histological and immunohistochemical analyses

For semithin sections, seeds of WT and the white aborted seeds of *naa15*/+ were fixed in 4% (v/v) paraformaldehyde buffered with 100 mM phosphate buffered saline (PBS, pH 7.4), placed under vacuum for 30 min, and then fixed overnight at 4°C. After fixation, the samples were washed four times with PBS. The samples were gradually dehydrated in a series of alcohol gradients (Zou et al., 2021). After dehydration, the samples were embedded in LR-White resin and polymerized at 50°C for two days. The samples were cut into sections with 1 μm thickness under an ultramicrotome (EM UC7, Leica) and then placed on slides.

To observe cell wall features of seeds, the samples were incubated in a mixture of calcofluor white (Sigma–Aldrich, St. Louis, USA) and 10% KOH (1:1, v/v) for 5 min at room temperature, and their fluorescent signals were observed under a Leica SP8 microscope. For immunofluorescent analysis of seeds, the samples were washed with distilled water and treated with the blocking solution (5% BSA buffered with PBS, pH 7.4) for 3 h. The primary rat monoclonal antibodies John Innes Monoclonal 5 (JIM5, antibody of low methyl-esterified HGs) and JIM7 (antibody of high methyl-esterified HGs) (diluted 1:10 in blocking solution) were added to the semithin sections at 4°C for 12 h. After incubation, the sections were washed with PBS solution three times to remove the primary rat antibody solution (Ma et al., 2019). The fluorochrome secondary antibody anti-mouse IgG DyLight 488 (Abbkine, Beijing, China) was added, and the sections were covered for 2 h at room temperature. Finally, the samples were observed under a Leica SP8 microscope with the same microscope magnification and fluorescence intensity. Immunofluorescence assays were conducted under the same conditions. The fluorescence signals were quantitated under a Leica LAS AF Lite v2.6.3. The fluorescence intensity of the cell walls was measured at five to ten points in each seed, and the final value was an average represented by a black dot.

Statistical analysis

All statistical tests and graphic plotting were performed using GraphPad Prism Software version 8.0 (GraphPad, San Diego, USA). Values were expressed as the mean ± standard deviation (SD). The experiments were conducted with three independent biological replicates. Statistical significance was established according to Student's *t*-test analysis, and the significance levels were established at $^{*}P<0.05$, $^{**}P<0.01$, and $^{***}P < 0.001$.

Results

Phenotypic characteristics of embryo differentiation and endosperm cellularization in WT and *naa15*

The embryo and endosperm are the most important components of seeds. In our early research, we obtained an *Arabidopsis naa15* mutant that possesses the deficient embryo and endosperm (Chen et al., 2018a). To investigate the characteristics of endosperm cellularization and embryo differentiation in *Arabidopsis* seeds, we observed the development processes of seeds at different stages. In 3W, the embryo reached the globular stage, and the endosperm presented dispersed free nuclei (Figures 1A, D). Subsequently, the embryos developed to the heart and torpedo stages at 4 DAP and 5 DAP, respectively (Figures 1B, C). In 4W, the cell wall of free endosperm nuclei formed in the middle and micropylar end of seeds, indicating that the endosperm was undergoing a process of cellularization (Figure 1E). In 5W, the cell walls of the free endosperm nuclei around the chalazal end were formed, which implied that the endosperm basically completed the process of cellularization (Figure 1F). We observed the embryo and endosperm phenotypic features of the *naa15* mutant using spontaneous fluorescence detection and semithin sections. In 5-7M, embryos stagnated at abnormal globular stages (Figures 1G–I), indicating that embryo differentiation was severely disrupted. In 5M, the endosperm was still in the syncytial stage, showing that its cellularization failed to be normally initiated. In 6 and 7 M, the partial free endosperm nuclei formed cell walls, but seeds produced cavities due to abnormal cellularization. The results indicated that the mutation of *NAA15* led to incomplete endosperm cellularization (Figures 1H, I). Therefore, we called this abnormal development occurring in the embryo and endosperm of the *naa15* seeds as "Globular embryo abnormal development" and "Endosperm delayed development". Based on the phenotypic characteristics of endosperm and embryo development described above, the 5M had an abnormal globular embryo and syncytial endosperm (Figure 1G), which was similar to the 3W. The 6M formed the partial endosperm cell wall and was similar to 4W (Figure 1H). The 5W completed cellularization, but the *naa15* mutant still had a large cavity and incomplete cellularization. Therefore, the *naa15* mutant was a well-experimental material for investigating the transcriptional regulation and gene expression in embryo differentiation and endosperm cellularization of *Arabidopsis*.

Global analysis of the RNA-seq data

The seeds of 5 DAP WT turned green from translucent, while the *naa15* seeds at 5DAP were still translucent and white in *naa15*/+ siliques (Supplementary Figure 1A). To investigate the transcriptional levels during endosperm cellularization and embryo differentiation, we collected normal seeds of 3-5 DAP WT (3W, 4W, and 5W, respectively) and white aborted homozygous seeds of 5-7 DAP *naa15* (5M, 6M, and 7M, respectively) as the tested materials and conducted RNA sequencing. The results showed that a total of 137.14 Gb of clean reads was obtained, and the mapping rate to the *Arabidopsis* genome was over 97%. The statistics of the clean reads in each sample were listed in Supplementary Table 2. The FPKM values of genes were shown in Supplementary Table 3 and the gene expression levels in each library were calculated and represented as the $\log_{10}$ (FPKM) method (Supplementary Figure 2A). The principal components of the PCA score plot (PC1 and PC2) were calculated to be 43.76% and 16.22%, respectively (Supplementary Figures 2B, C). The six sample groups were clearly separated, and three biological replicates of each group were clustered together (Supplementary Figure 2C).

We performed comparisons between 4W vs 3W (heart stage vs globular stage) and 5W vs 4W (torpedo stage vs heart stage) to uncover the changes in the transcriptome profiles during the early stages of embryo and endosperm development. A total of 2040 DEGs (1271 up- and 769 down-regulated) in 4W vs 3W and 3630 DEGs (1922 up- and 1708 down-regulated) in 5W vs 4W groups were obtained, respectively (Figures 2B, C). Venn diagrams showed that 1102 DEGs were shared between the 4W vs 3W and the 5W vs 4W groups (Figure 2A), and these DEGs were distributed in four clusters using the K-means approach, indicating that they had similar expression patterns in 3-5 DAP WT (Figures 2D-G and Supplementary Tables 4-7). Many transcription factors, such as growth-regulating proteins, MYBs, and bZIPs were in cluster I (560 DEGs), showing that their expression levels were distinctly upregulated with seed development (Figure 2D and Supplementary Table 8). In contrast, the transcript levels of auxin response factors (*ARF12*, *ARF13*, *ARF14*, *ARF15*, *ARF20*, *ARF21*, and *ARF22*) and MADS-box transcription factors (*AGL36*, *AGL48*, *AGL58*, *AGL62*, *AGL90*, and *AGL 96*) were significantly downregulated from 3W to 5W, and were all clustered in cluster II (423 DEGs) (Figure 2E and Supplementary Table 9). The numbers of DEGs in clusters III and IV were 83 and 36, respectively. The genes involved in heat shock proteins and chaperones were enriched in cluster III, while many pectin acetylesterase, methyltransferase, and protein kinase enzymes were concentrated in cluster IV (Supplementary Tables 6, 7). Therefore, we proposed that the differential transcriptional regulation of genes may contribute to early seed development.

Functional enrichment analysis of DEGs based on KEGG and GO databases

To reveal the regulatory network of seed morphogenesis, DEGs were classified and annotated using GO and KEGG databases. The up-regulated DEGs of KEGG enrichment contained indole alkaloid

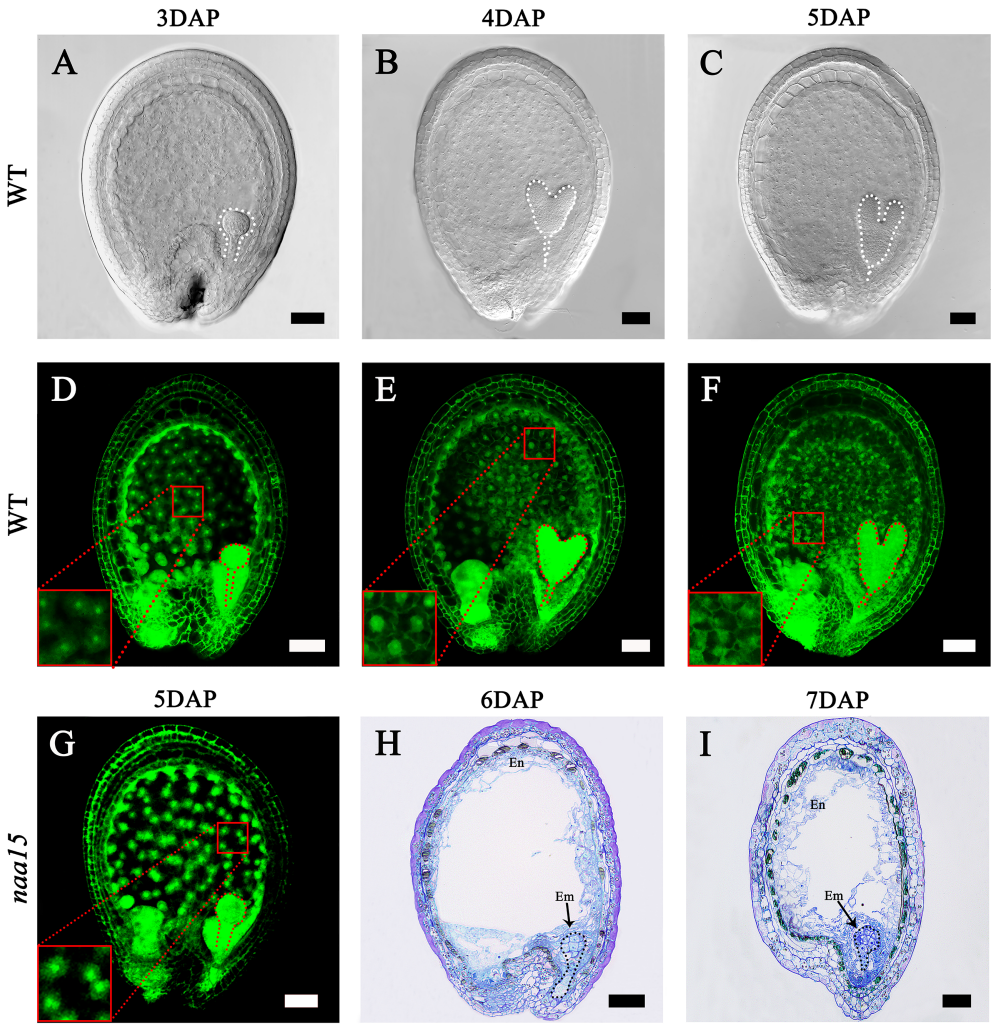


FIGURE 1
Embryo differentiation and endosperm cellularization in WT and *naa15* of *Arabidopsis*. **(A–F)** Seeds in WT; **(G-I)** Seeds in *naa15*. **(A)** A seed with a globular embryo at 3 DAP. **(B)** The seed with a heart embryo at 4 DAP. **(C)** The seed with a torpedo embryo at 5 DAP. **(D)** Free nucleus dispersed in syncytium endosperm at 3 DAP. **(E)** Free endosperm nucleus began to form cell wall at 4 DAP. **(F)** Endosperm cellularization at 5 DAP basically completed. **(G)** The seed showed an abnormal globular embryo with free endosperm nucleus without cell walls at 5 DAP. **(H, I)** Semi-thin sections showed abnormal globular embryo and partially endosperm cells at 6 DAP and 7 DAP, respectively. Embryos were marked with discontinuous lines of different colors. The small red boxes were magnified and shown on the lower left of the corresponding images. Em, embryo; En, endosperm. Scale bars = 50 μm.

biosynthesis, plant hormone signal transduction and phenylpropanoid biosynthesis in the 4W vs 3W and 5W vs 4W groups (Figures 3A, B and Supplementary Table 10). A lot of terms related-cell division, the RNA metabolism, the transcription regulator activity, cell cycle terms, and cytoskeleton-related terms, were significantly enriched in up-regulated DEGs of 4W vs 3W (Supplementary Figure 3A), indicating that 3 to 4 DAP seeds might undergo vigorous cell divisions. The photosynthesis, photosynthesis - antenna proteins, starch and sucrose metabolism, and oxidative phosphorylation pathways were significantly enriched in 5W vs 4W (Figure 3B). Within the biological process category of GO enrichment analyses, many up-regulated DEGs were also significantly enriched in photosynthesis terms (GO:0019684, GO:0009768, and GO:0015979) (Supplementary Figure 3B), indicating that they might participate in plastid transformation to the chloroplast process during early seed development and involve in the embryonic transfer from the heterotrophic stage to the autotrophic stage.

In addition, the terms related to cell wall (cell wall organization or biogenesis, plant-type cell wall organization or biogenesis, and cell wall) were also significantly enriched in cellular component using GO enrichment analysis of 4W vs 3W

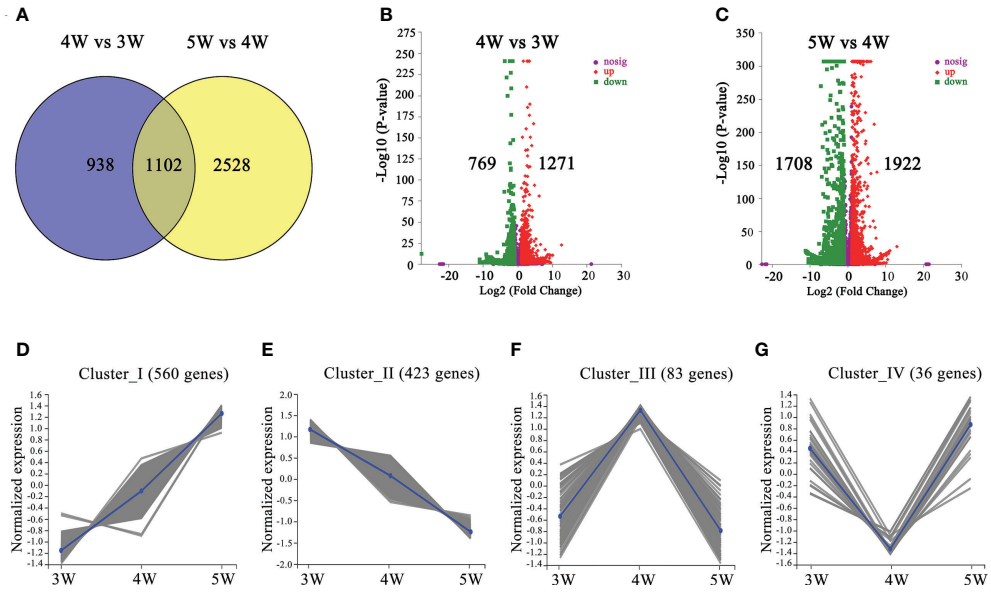


FIGURE 2
The number and functional annotation of differentially expressed genes (DEGs) during embryo differentiation and endosperm cellularization in WT of *Arabidopsis*. **(A)** The number of DEGs from venn diagram analysis in 4W vs 3W and 5W vs 4W groups, respectively. **(B, C)** Volcano maps of DEGs in 4W vs 3W and 5W vs 4W groups, respectively. Log_2 (Fold Change)≥ 1 or ≤ −1, and P-value < 0.05 were used as a standard for screening DEGs. The "-Log_{10} (P-value)" indicated the statistical significance level, and the larger value showed that the divergence of gene expression was more significant. Significantly up- or down-regulated genes were represented by using red dots or green dots, respectively. **(D–G)** Clustered gene expression profiles of 1102 DEGs from 4W vs 3W and 5W vs 4W using the statistical method of K-means.

and 5W vs 4W groups (Supplementary Figures 3B,C), which mainly included the genes related to cell wall modification (pectinesterase and pectin methylesterase inhibitor), cellulose synthase, and expansins (Supplementary Table 11). The hormone signaling pathways were significantly enriched in up- and down-regulated DEG of GO enriched terms and KEGG enriched pathways during early seed development at 3-5 DAP (Figures 3A–D). These DEGs mainly participated in the auxin, cytokinin, and abscisic acid pathways, indicating that there was wide crosstalk among hormones. The mRNA levels of auxin transport genes (*PIN6* and *LAX2*), auxin response factors (*ARF4* and *ARF11*), and small auxin upregulated RNA 78 (*SAUR78*) were markedly upregulated from 3 to 5 DAP. Particularly, cytokinin synthesis genes *LONELY GUY* 4 and 5 (*LOG4* and *LOG5*) were mainly upregulated in 4W vs 3W (Supplementary Table 11), speculating that they might be involved in the formation of embryo organ primordia and cellularization of free endosperm nuclei.

Candidate regulation pathways during endosperm cellularization and embryo differentiation of *naa15*

We performed the transcriptome comparative analysis in both WT and *naa15*, and identified a total of 9633 DEGs in 5M vs 3W, 6M vs 4W and 7M vs 5W. Among these DEGs, the 2195 DEGs were shared in three comparison groups (Figure 4A). The number of DEGs was 5988 DEGs in 5M vs 3W (containing 2743 up- and 3245 down-regulated), 5595 DEGs in 6M vs 4W (containing 2530 up- and 3065 down-regulated), and 5598 DEGs in 7M vs 5W (containing 3063 up- and 2535 down-regulated), respectively (Figures 4B–D). KEGG enriched analysis showed that the up-regulated DEGs of WT and *naa15* were significantly enriched in phenylpropanoid biosynthesis, protein processing in endoplasmic reticulum, starch and sucrose metabolism, and amino sugar and nucleotide sugar metabolism (Figures 4E–G and Supplementary Tables 12A–C). The GO enrichment analysis displayed that the up-regulated DEGs of WT and *naa15* were evidently aggregated in the terms of cell cycle process, cell division, cytoskeleton part, and related-microtubule terms (including microtubule binding, microtubule-based process, and microtubule-based movement) (Supplementary Figure 6).

The down-regulated DEGs of KEGG enrichment analyses in WT and *naa15* were mainly enriched in plant hormone signal transduction photosynthesis, phenylpropanoid biosynthesis, and starch and sucrose metabolism pathways (Figures 4H–J and Supplementary Tables 12D–F). Totally, we identified 137 DEGs involved in auxin, cytokinin, and ethylene signaling pathways in three comparison groups between WT and *naa15* (Supplementary Table 13). As shown in Figures 5A–C, the

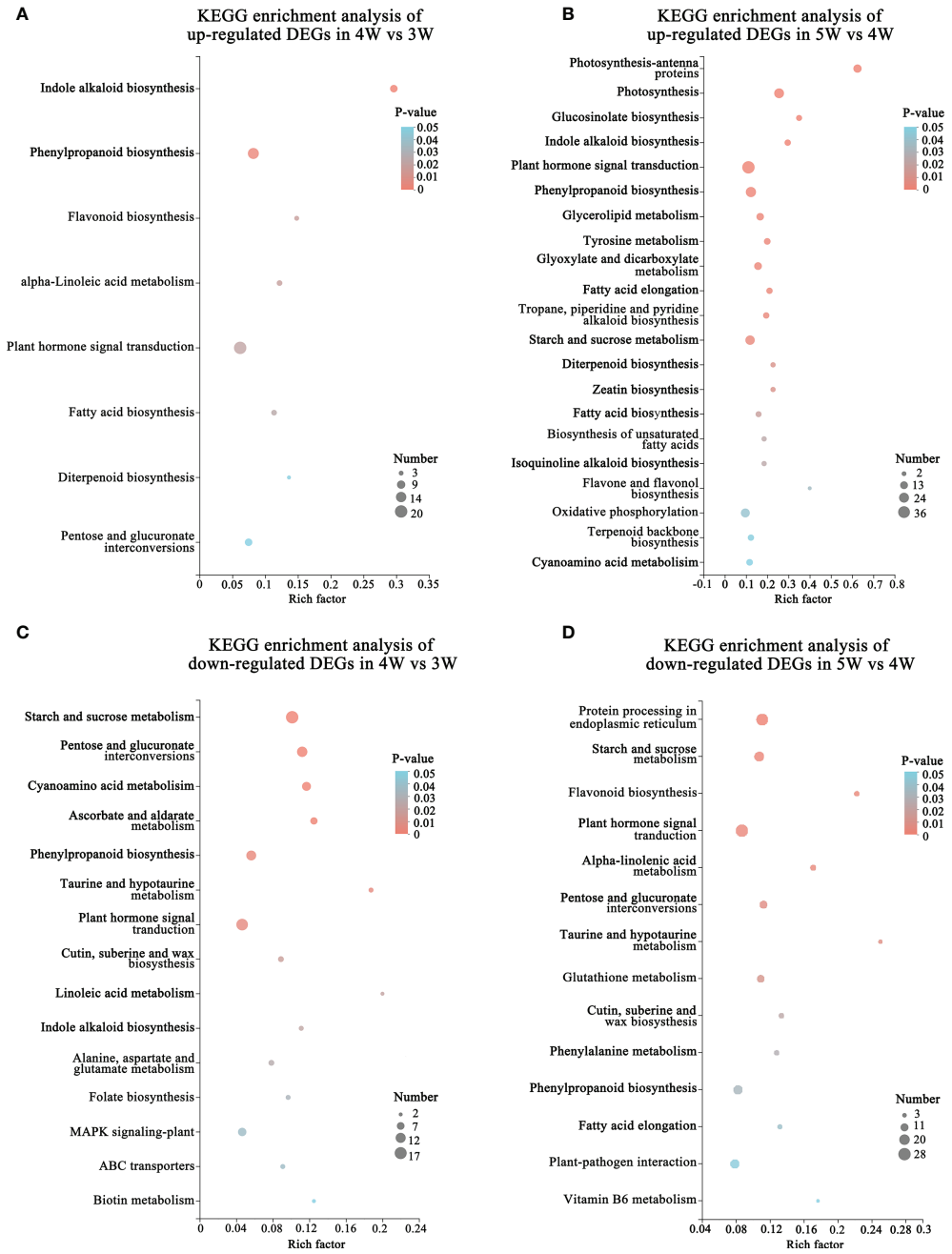


FIGURE 3
Functional annotation and classification of DEGs during embryo differentiation and endosperm cellularization in WT of *Arabidopsis*. **(A, B)** KEGG enrichment analysis of up-regulated DEGs in 4W vs 3W and 5W vs 4W groups, respectively. The "Rich factor" represented the enrichment degree of DEGs in each pathway. The number of the enriched DEGs was represented by the size of each circle. The larger the number, the bigger the circle. **(C, D)** Diagrams of KEGG enrichment analysis of down-regulated DEGs in 4W vs 3W and 5W vs 4W, respectively.

genes annotated to auxin-responsive proteins (IAA), PIN, and auxin response factors (ARF) were identified, and most of them were downregulated in *naa15*. qPCR analyses showed that the expression of auxin transporter *PIN1* was significantly increased in 3W-5W. In contrast, the *PIN1* transcript level was markedly decreased in *naa15* (Figure 5J). The mRNA levels of *ARF5* showed a downregulated pattern in WT and were decreased in 5-7M (Figures 5C, K).

In addition, the heatmap showed that the genes related to the cytokinin pathway, cytokinin receptors (*ARABIDOPSIS*

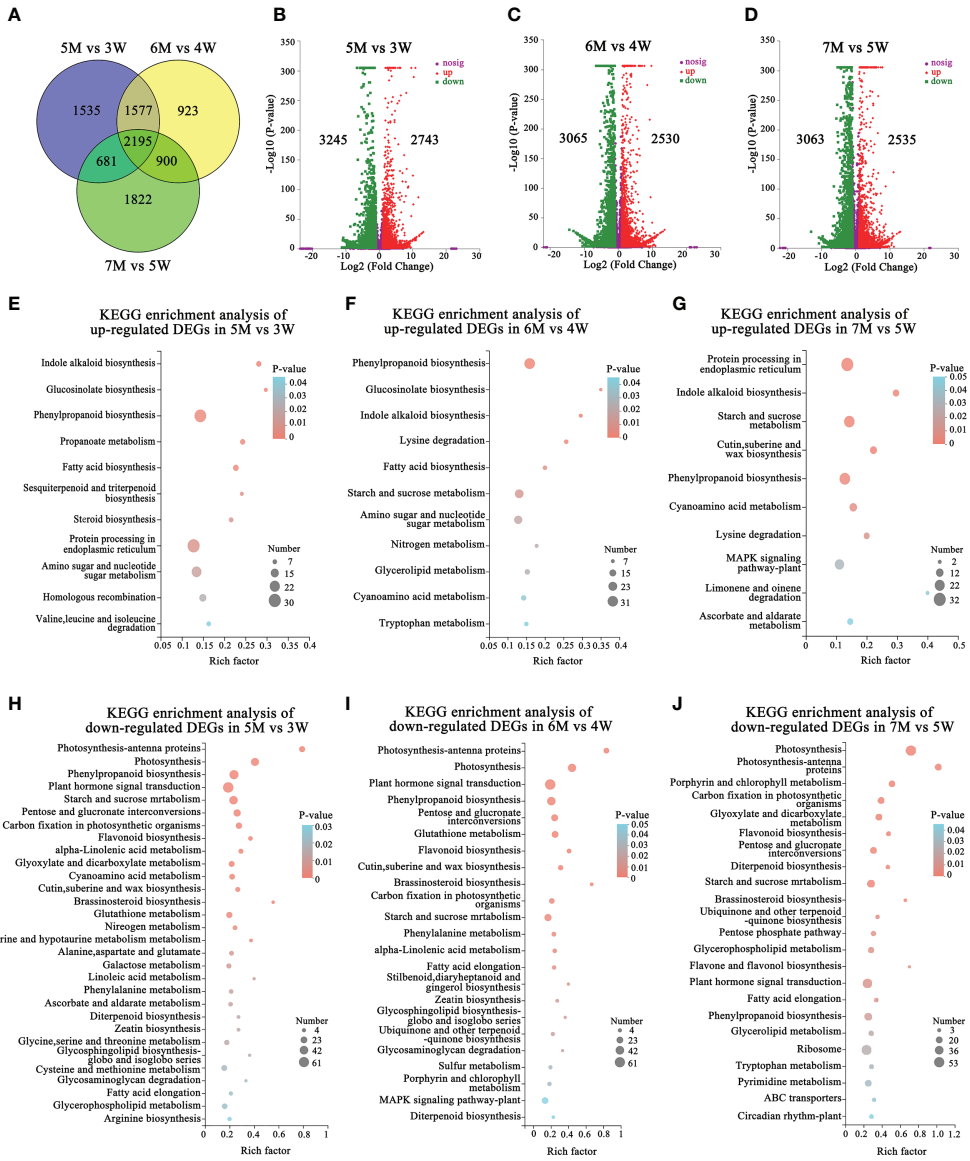


FIGURE 4
The number and KEGG enrichment pathways of DEGs between *Arabidopsis* WT and *naa15* in 5M vs 3W, 6M vs 4W, and 7M vs 5W, respectively. **(A)** Venn diagram of DEGs in three comparison groups between WT and *naa15*. **(B–D)** The volcano maps of DEGs in three comparison groups between WT and *naa15*. **(E–G)** The KEGG enrichment pathways of up-regulated DEGs in 5M vs 3W, 6M vs 4W, and 7M vs 5W, respectively. **(H–J)** The KEGG enrichment pathways of down-regulated DEGs in 5M vs 3W, 6M vs 4W, and 7M vs 5W, respectively.

HISTIDINE KINASE, *AHK3*), cytokinin dehydrogenases (*CYTOKININ OXIDASE/DEHYDROGENASE*, *CKX1*, *CKX2*, *CKX5*, and *CKX6*), and cytokinin-activating enzymes (*LONELY GUY*, *LOG4*, *LOG5*, and *LOG7*) were also downregulated in *naa15* (Figure 5D). The cytokinin receptor *AHK3* in 3-5 W was gradually downregulated, while its transcript was distinctly decreased in 5-7 M (Figure 5L). *CKX2*, encoding a limiting enzyme in cytokinin metabolism, is activated by WRKY10 to promote endosperm growth (Li et al., 2013). In our results, the expression of *CKX2* was upregulated at 3W and 4W, but down-regulated at 5W; however, it remained at low levels in 5-7M (Figure 5M). These results implied that the gene expression of *NAA15* was involved in regulating the hormone signaling pathways during embryo and endosperm development.

Transcription factors (TFs), including MYB, bHLH, WRKY, ERF, and GATA, were identified in GO and KEGG enriched analyses in comparison between WT and *naa15*. In 5M vs 3W group, there were 89 transcription factors (TFs) (53 down- and 36 up-regulated DEGs), and the most downregulated TFs was

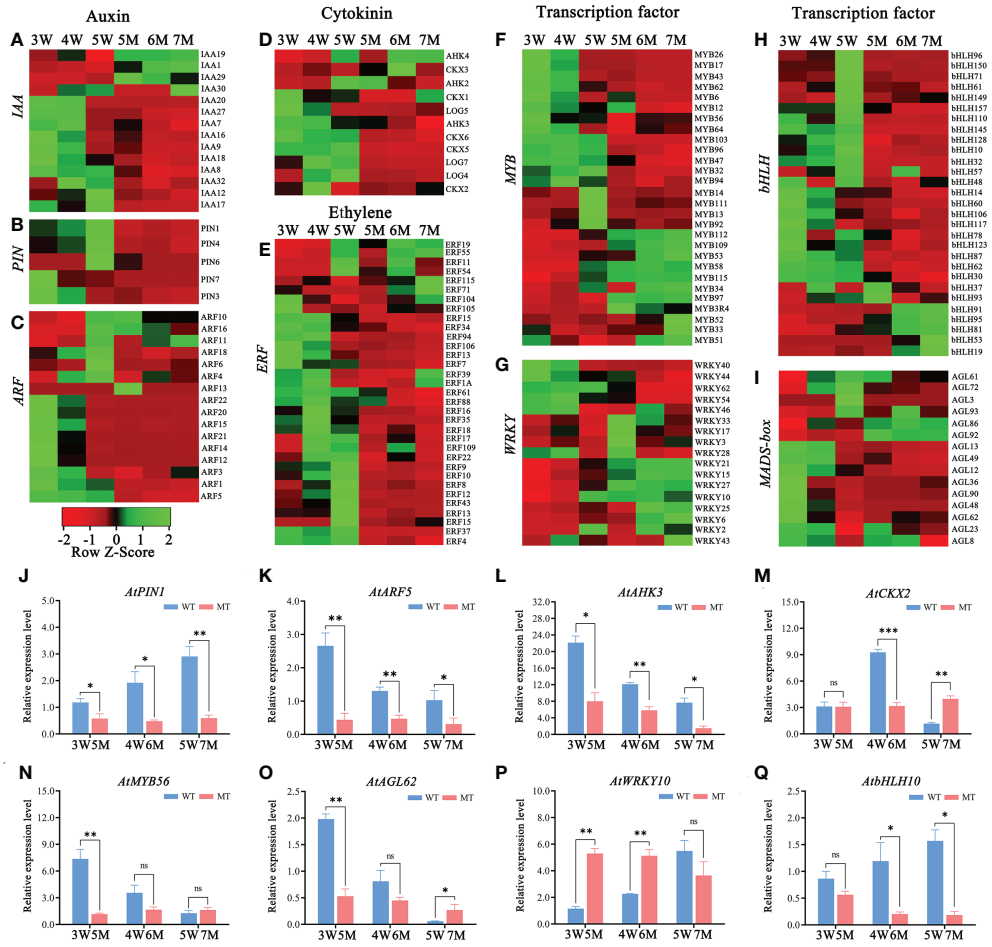


FIGURE 5
Heatmaps and qPCR analyses of DEGs related to hormones and transcription factors in WT and *naa15* seeds of *Arabidopsis*. **(A–D)** The relative expression patterns of DEGs involved in auxin and cytokinin signal pathways. **(E–I)** The relative expression patterns of DEGs involved in transcription factors including ERF, MYB, WRKY, bHLH, and MADS-box. FPKM values were standardized with Z-Score for the heatmap analysis. Red represented down-regulation of gene expression, and green represented up-regulation of gene expression. **(J–Q)** The qPCR analyses of DEGs related to hormones and transcription factors. Asterisks denoted significant differences basing on Student's *t*-test (*P<0.05, **P<0.01, and ***P < 0.001).

the bHLH family (18 DEGs), followed by ERF family (16 DEGs). In 7M vs 5W group, the downregulated TFs were mainly in the MYB, ERF, and bHLH families (Supplementary Table 14). The heatmaps showed that ERF (23 DEGs), MYB (16 DEGs) and bHLH (35 DEGs) were significantly down-regulated in 5-7M (Figures 5E, F, H). The gene expressions of *WRKY* (17 DEGs) and *AGL* (15 DEGs) were also significantly changed in 5-7M (Figures 5G, I). qPCR analyses showed that the expressions of *AtMYB56* and *AtAGL62* gradually decreased in 3-5W, and their expressions distinctly deceased in 5M compared to WT (Figures 5N, O). *AtbHLH10* and *AtWRKY10* were markedly increased in 3-5W (Figures 5P, Q). Nevertheless, *AtWRKY10* in 5 and 6M had higher expression than in 3W and 4W (Figure 5P). Hence, we hypothesized that *NAA15* impacted embryo differentiation and endosperm cellularization by directly or indirectly mediating the transcription levels of these genes.

Sucrose controls a wide range of developmental and metabolic processes including cell division, cotyledon development, far-red light signaling, and tuber development (Yoon et al., 2021). In our research, the starch and sucrose metabolism pathway was enriched in the KEGG enriched analysis of 3W-5W (Figures 3B–D). And the expressions of two sucrose synthetases (*suc1* and *suc2*) and eight sucrose transporter genes (*sweet3*, *sweet4*, *sweet5*, *sweet7*, *sweet9*, *sus1*, *sus6*, and *sus4*) were significantly decreased in *naa15* (Supplementary Figure 4A). Sucrose stimulates the expression of transcription factors such as *WRKY* and *MYB*, which function upstream of sucrose-responsive genes. However, the molecular mechanism of sucrose as a signal regulating embryo and

endosperm development remains unknown. Our transcriptome data provided some related to sucrose genes that might play vital functions during early seed development.

The photosynthesis pathway was also significantly enriched in the KEGG and GO enriched analyses up-regulated DEGs in 5W vs 4W group (Figure 3B and Supplementary Figure 3B). The GO enrichment analysis of down-regulated DEGs in *naa15* was mainly enriched in photosystem and the relevant chloroplast genes (Supplementary Figure 5). The gene expressions from six *photosystem I-antenna* (*LHCI*) genes and 12 *chlorophyll a/b binding protein complex II* were decreased in *naa15* (Supplementary Figure 4B). Interestingly, Lhca5, Lhcb1.1, PORB, and PORC have been considered potential substrates of *Arabidopsis* N-terminal acetyltransferase A (NatA), which consists of NAA10 and NAA15 (Linster et al., 2015). Hence, we speculated that plant hormones, photosynthesis, and carbohydrate pathways were critical during early seed development. However, the molecular regulatory function of *NAA15* in these pathways was unknown and requires further investigation.

In addition, the abnormal globular embryo in *naa15* implied that embryo differentiation was disrupted. Thus, we inspected the expression levels of the genes reported to be associated with embryo differentiation in 3 DAP WT and 5-7 DAP *naa15*. WOX family had an essential role in body axis of embryo. Among these DEGs from our transcriptome profiles, *WOX2*, *WOX5*, and *WOX8* in *naa15* were significantly downregulated (Supplementary Tables 2, 4), and the qPCR results also showed that the expressions of *AtWOX2* and *AtWOX5* in 5-7M were significantly decreased compared with that in 3W (Figures 6A, B), implying that the downregulated expression of WOX genes in *naa15* might lead to the formation of abnormal globular embryos. *CUP-SHAPED COTYLEDON* (*CUC*) and HD-ZIP TFs are involved in bilateral symmetry and SAM establishment. The qPCR results showed that *AtCUC2* and *ATHB8* were significantly decreased in *naa15* (Figures 6C, D). *DORNRÖSCHEN* (*DRN*) is an AP2/ERF-type TF that interacts with HD-ZIP III TFs. The expression of *AtDRN* also decreased significantly in *naa15* (Figure 6E). *PLETHORA1* (*PLT1*) and *SHORT-ROOT* (*SHR*) are required for hypophysis divided asymmetrically and quiescent center (QC) formation in late globular embryos, respectively (Aida et al., 2004).

In the results, *PLT1* and *SHR* showed high expressions at 3 DAP in WT, but down-regulated in 5-7M (Figures 6F, G), suggesting that it may be one cause of the abnormal globular embryo formation of *naa15*. Besides, the expressions of *APETALA2* (*AP2*) and *EOD1*, which influence the sizes of seeds, also significantly decreased in *naa15* (Figures 6H, I). *YABBY2* (*YAB2*), *FILAMENTOUS FLOWER* (*FIL*), and *REVOLUTA* (*REV*) played essential roles in the establishment of abaxial patterning. In our study, the expressions of *YAB2* and *FIL* were significantly downregulated, while the *REV* was evidently upregulated (Figures 6J-L). In addition, *DOF6*, *ATHB2*, *LEC1*, and *AGO10* were observably up-regulated in *naa15* (Figures 6M-P), suggesting that the abnormal expression may affect the embryo differentiation of *naa15*. Therefore, we postulated that the gene expression changes related to embryo differentiation may be one of the causes of abnormal embryonic differentiation of *naa15*, and the detailed relationship between these DEGs and *NAA15* needs to be further investigated.

Transcriptional regulation of cell wall formation and modification during early embryo and endosperm

According to GO and KEGG enrichment analyses from two comparison groups in 3-5W, many DEGs were enriched in cell wall-related pathways (e.g., cell wall organization or biogenesis and plant-typed cell wall) (Figure 3 and Supplementary Figure 3). However, these DEGs were abnormal expressions in *naa15*. Heatmap analysis showed that 13 genes encoding expansin (EXPA), which functions as the modified proteins of cell wall, were downregulated in *naa15* (Figure 7A). The mRNA level of *AtEXPA15* was increased in 3W-5W, whereas it was significantly down-regulated in 5M-7M (Figure 7G). qPCR analyses showed that *AtEXPB4* was gradually downregulated in 3W-5W, and was also significantly downregulated in 5M and 6M compared to 3W and 4W (Figure 7H), hinting that the decrease in *AtEXPA15* and *AtEXPB6* mRNA levels in *naa15* might be one reason for abnormal endosperm cellularization. In addition, we also found that cell wall modification genes, including 15 xyloglucan endotransglycosylase/hydrolase (XTH) and 27 pectin esterase/pectin methylesterase (PME) genes, were downregulated in *naa15* (Figures 7B, C). qPCR analyses demonstrated that the expression of *PME12* and *PME19* in 5M were downregulated with 7.9-fold and 9.0-fold compared to 3W, respectively (Figures 7K, L). The expressions of *XTH6* and *XTH14* in 5M and 6M were downregulated compared to 3W and 4W (Figures 7I, J). The genes encoding proteins of cell wall, such as arabinogalactan proteins (AGPs, *AGP23/7/6/11/12/40/41/26/22/17/9*) and fasciclin-like arabinogalactan proteins (FLAs, *FLA10/4/8/12/13/2/1*), were also downregulated in *naa15* (Figures 7D, E). qPCR analyses confirmed that *AGP9*, *AGP17*, and *FLA10* were significantly downregulated in 5M and 6M compared to 3W and 4W (Figures 7M–O), which was consistent with the RNA sequencing data (Figures 7D, E). However, the expressions of several cell wall related-genes CESA5, CESA8, CSLD5, CSLB3, CSLD6, and FLA10 were up-regulated in naa15 (Figures 7F, P–R), which might play important role in abnormal cell walls of the endosperm.

The features of cell walls during early embryo differentiation and endosperm cellularization

Many DEGs from our transcriptome analysis were involved in cell wall formation and modification; therefore, we analyzed

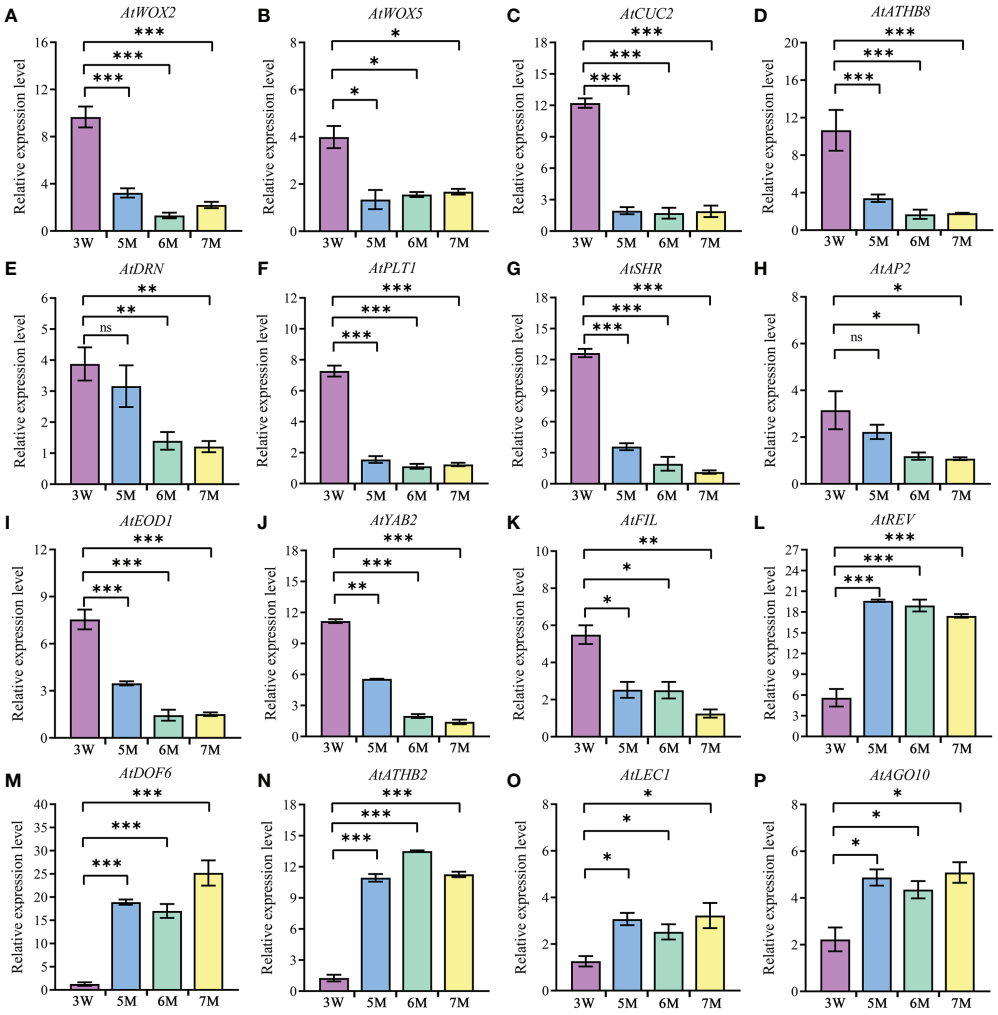


FIGURE 6
The expression of genes related to embryo differentiation in WT and *naa15*. **(A–P)** The qPCR analyses of DEGs in 3W, 5M, 6M, and 7M. 3W, 3 DAP WT; 5M, 5 DAP *naa15*; 6M, 6 DAP *naa15*; 7M, 7 DAP *naa15*. The "ns" represented no significant differences. Asterisks denoted significant differences basing on Student's *t*-test $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

the features of cell wall in the embryo and endosperm of WT and *naa15*. Histological staining and immunofluorescence assays were performed in the seeds of 4W, 5W, 6M, and 7M using calcofluor white (CW) staining and pectin methyl-esterification monoclonal antibodies of (JIM5 and JIM7). The CW fluorescence gradually increased in cell walls of 4W and 5W embryos, while the signals in 7M embryonic cell walls significantly decreased compared to 5W (Figures 8A–E). The cell wall of *naa15* endosperm at 6M and 7M exhibited decreased fluorescence intensity compared to 5W (Figures 8A–D, F), inferring that the decreased cellulose in *naa15* might affect the formation of endosperm cell walls.

Using the JIM5 antibody to detect the low methyl-esterified pectin, the fluorescence signals of embryo cells decreased gradually from 4W to 5W and from 6M to 7M (Figures 8G–K). In cell walls of endosperm at 4W and 5W, JIM5 signals decreased following endosperm cellularization, which might contribute to the rapid development of subsequent embryos (Figures 8G, H, L). JIM5 fluorescence in the cell walls of *naa15* endosperm at 6M significantly decreased compared to 4W and 5W (Figures 8G–J, L), while the signal in 7M was similar to 5W. The results indicated the delayed and impaired endosperm growth and development in *naa15*

Using the JIM7 antibody to detect high methyl-esterified pectin, the results showed that there was no apparent change in the fluorescence signals in 6M and 7M compared to embryos at 4W and 5W (Figures 8M–Q). The fluorescence signals in the cell walls of endosperm at 4W and 5W were obviously increased, while the signal in 6M and 7M evidently decreased (Figures 8M–P, R), indicating that cell wall formation in embryo and

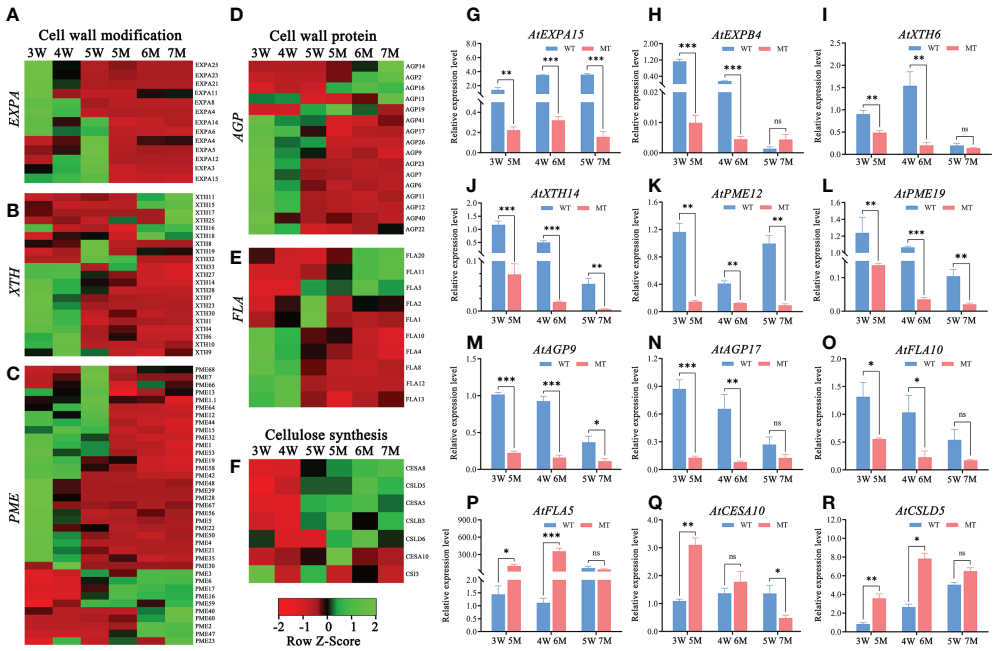


FIGURE 7
Heatmaps and qPCR analyses of DEGs related to cell wall formation between WT and *naa15* in *Arabidopsis* seeds at different development stages. **(A–C)** Heatmaps of the cell wall-modification proteins genes including the types of *EXP*, *XTH*, and *PME*/*PMEI*. **(D, E)** Heatmaps of the cell wall-protein genes including the types of *AGP* and *FLA*. **(F)** The heatmap of genes related to cellulose synthesis. **(G–R)** The qPCR analyses of DEGs related to formation of cell wall. Asterisks denoted significant differences basing on Student's t-test $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$. The "ns" represented no significant difference. AGP, arabinogalactan proteins; EXP, expansin; FLA, fasciclin-like arabinogalactan protein; PME/PMEI, pectinesterase/pectin methylesterase inhibitor; XTH, xyloglucan- endotransglycosylases/hydrolases.

endosperm of *naa15* was disrupted, resulting in the aberrant embryo differentiation and incomplete endosperm cellularization.

Discussion

It is well known that seed formation depends on well-coordinated development between endosperm and embryo. The endosperm is considered as a nutrition source to support embryo development, and its cellularization process undergoes nuclear proliferation, formation of endosperm cell wall, and cell differentiation, all of which are essential for embryo growth (Jo et al., 2019; Song et al., 2021). Nevertheless, extensive efforts are required to understand the process of embryo differentiation and endosperm cellularization, including the fundamental regulatory mechanisms of the initial endosperm cellularization process, and the relationship between embryo and endosperm. In this research, we employed RNA-sequencing analyses and observed cell histological features to identify candidate genes and expression pathways involved in embryo differentiation and endosperm cellularization, which provide new insights into seed morphogenesis.

Cell wall formation and modification during embryo differentiation and endosperm cellularization

It has been reported that the appearance of endosperm cell wall in the micropyle zone of seeds indicates that endosperm development transforms from the syncytial endosperm to cellularization stages (Otegui and Staehelin, 2000). Cell walls in plants contain various polysaccharides (cellulose, hemicellulose, and pectins), glycoproteins, and aromatic or aliphatic compounds, which form the mechanically strong and extensile network (Caffall and Mohnen, 2009). Cell walls provide mechanical strength for cells and fix the cell shape, which requires the modification and remodeling of the cell wall. The modification gene *XTH31*, coding a xyloglucan endotransglycosylase/hydrolase, is expressed in endosperm of *Arabidopsis* and can reinforce their cell walls during seed germination (Endo et al., 2012). ZHOUPI is an endosperm-specific gene in *Arabidopsis* and affects embryo and endosperm development by regulating the expression of pectin-modifying enzymes. In endosperm cell wall of *zou* mutant, (1-5)-α-L-arabinan epitopes immunized with the LM6 antibody were abundant, its endosperm cells are more robust than WT, and

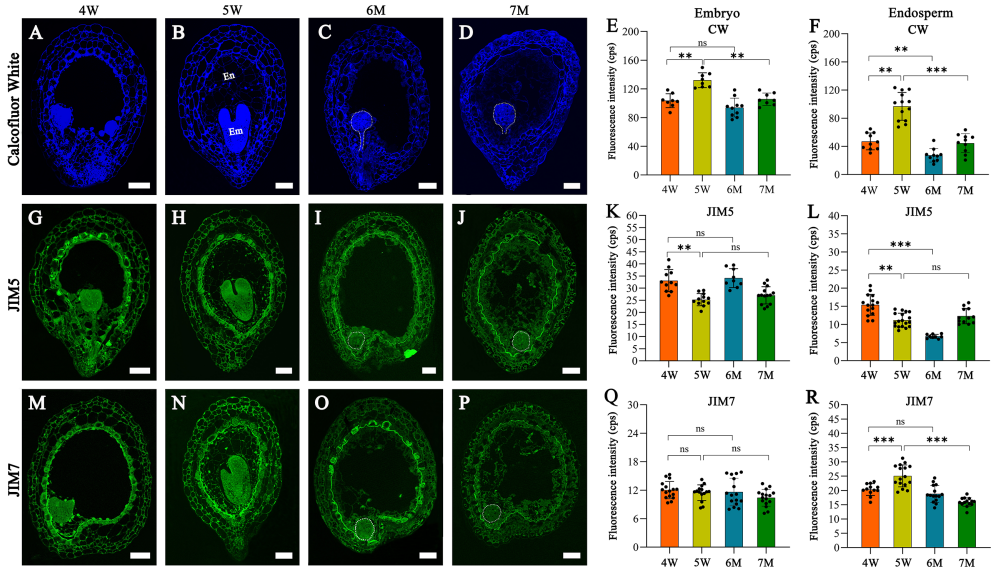


FIGURE 8
The images of histochemistry and immunofluorescence localization in WT and *naa15* seeds at different development stages. **(A–D)** Semi-thin sections of seeds were stained with calcofluor white in WT and *naa15*. **(E, F)** Fluorescence intensity with calcofluor white staining in embryo and endosperm of WT and *naa15*. **(G–J)** Semithin sections of seeds were immunolabeled using JIM5 antibody in WT and *naa15*. **(K, L)** Fluorescence intensity with JIM5 immunolabelling in embryo and endosperm of WT and *naa15*. **(M–P)** Semi-thin sections of seeds were immunolabeled using JIM7 antibody in WT and *naa15*. **(Q, R)** Fluorescence intensity with JIM7 immunolabelling of WT and *naa15* in embryo and endosperm, respectively. The p-value was calculated by using the Student *t*-test. Asterisks denoted significant differences, $^{**}P < 0.01$ and $^{***}P < 0.001$. The "ns" represented no significant difference. Each black dot represented a seed. Em, embryo; En, endosperm; cps: counts per second. Scale bars = 50 μm.

the cell wall fail to breakdown, resulting in the abnormal size of embryo (Fourquin et al., 2016). The focus of investigation of seed cell walls is mainly the roles of cell wall during seed germination; however, up to date, the transcriptional regulation of cell wall formation during endosperm cellularization is still unclear.

AGPs and FLAs are important proteoglycans that form covalent bonds with both hemicellulose and pectin in cell walls. In rice, AGP protein is widely distribution in tapetum cells, pollen mother cells, and mature pollens (Ma et al., 2019). In *FLA3* RNAi transgenic plants of *Arabidopsis*, pollen grains are shrunken and wrinkled, and cellulose abnormal distribution in pollen walls to lead to reduced male fertility (Li et al., 2010). APAP1 (ARABINOXYLAN PECTIN ARABINOGALACTAN PROTEIN1) encodes an AGP protein, and contributes to covalent linkages between cell wall proteins, pectins, and xylans. In *apap1*, the covalent linkages of pectin and xylan in cell walls of stem are less tightly (Tan et al., 2013). In our study, cell wall organization/biogenesis and cell wall terms were significantly enriched in embryo differentiation and endosperm cellularization of 3-5 DAP WT (Supplementary Figure 2 and Supplementary Table 10). These DEGs were mainly annotated to genes related to proteins of cell walls (*AGP* and *FLA*) and its modification (Supplementary Table 11). Among them, 27 DEGs annotated to *pectin esterase/pectin methylesterases* (*PME*) showed lower expression levels in *naa15* (Figure 7C). This indicated that these DEGs might be involved in the formation of the cell wall during early seed development.

Meanwhile, through analyzing the distribution of esterified pectin during embryo and endosperm development, the immunolabeled results showed that low-esterification pectin of endosperm cell wall gradually decreased by using JIM5 antibody; in contrast, by using JIM7 antibody, the high-esterification pectin was increased in 4 and 5W (Figures 8K, L, R). Thus, we proposed that decreased low-esterification pectin and increased high-esterification pectin in cell walls might be largely related to endosperm cellularization, which increased the rigidity and stiffness of cell wall. In the endosperm cell walls of *naa15*, JIM5 and JIM7 signals showed lower fluorescence intensities than in WT (Figures 8L, R), but the JIM7 signals of embryos in WT and *naa15* had no significant differences (Figure 8Q). Based on the results, we speculated that the esterification feature change of the endosperm cell wall in *naa15* might be one of the causes of the abnormal endosperm. We hypothesized that the defective embryo and endosperm of the *naa15* mutant might be due to impairing the cell characteristics, which in turn affected embryo differentiation and endosperm cellularization. It has been reported that the embryo morphology and mucilage

extrusion of seeds in *Arabidopsis hms-1* mutant were altered owing to methyl esterification degree of cell walls (Levesque-Tremblay et al., 2015). Hence, cell wall formation and modification were vital for embryo and endosperm during early seed development.

Starch and sucrose metabolism and photosynthesis-related DEGs in early embryo and endosperm development

In GO and KEGG enriched analysis of these DEGs, starch and sucrose metabolism pathways were enriched in 3-5W (Figures 3B-D). Sucrose is imported into developing embryo by sucrose transporters (suc/sut) and sucrose exporter proteins (SWEET). The *sweet11;12;15* triple mutant in *Arabidopsis* disrupts sugar transport, resulting in delayed embryo development (Chen et al., 2015). In our research, the transcriptome levels of *SWEET5*, *SWEET11*, and *SWEET15* in 3-5W were upregulated, and sucrose metabolism genes such as *AT1G19450*, *SFP1*, and *SPS3* were also upregulated. However, their roles in seed development remain unclear. The starch and sucrose pathway was also significantly enriched using KEGG analysis in *naa15* compared with WT (Figures 4E–I). The heatmap of sucrose synthesis and sucrose transporter genes showed that *SUS1*, *SUS4*, *SUC1*, *SWT3*, *SWT4*, *SWT7*, and *SWT9* were significantly down-regulated in *naa15* compared with WT. Taken together, these results implied that sucrose metabolism may provide nutrients and energy for promoting normal cell growth of early seed development. However, the role of *NAA15* in sucrose metabolism remained unclear and needed to be further investigated.

From globular-embryo stage, the organ differentiation of embryo begins to occur, and the seed is transformed from the heterotrophic stage to the autotrophic stage. In embryonic cells, proplastids are converted into chloroplasts, accompanied by the biosynthesis of chlorophyll (Tejos et al., 2010). Therefore, lots of gene expressions related to photosynthesis are regulated during embryo differentiation and endosperm cellularization. Transcriptome GO analysis shows that photosynthesis term is significantly enriched in soybean seeds at globular, heart, and cotyledon stages, implying that photosynthesis might play important role in the embryo development (Sun et al., 2020). In many homozygous mutants, the abnormal development of chloroplasts leads to albino seeds and embryonic lethality. EMB1990 and EMB2726 are plastid-localized proteins, and the chloroplasts of *EMB1990* and *EMB2726* knockout mutants are severely impaired, resulting in defective embryos (Chen et al., 2018b; Li et al., 2021). It remains poorly understood what are intriguing biological functions of embryonic chloroplast in *Arabidopsis*.

In our transcriptome data, KEGG and GO analyses showed that many DEGs of 4W vs 5W group were enriched in photosynthesis term and chloroplast components (apoplast, thylakoid, and photosynthetic membrane) (Supplementary Figures 4B–D, 6). Interestingly, six *photosystem I-antenna* (*LHCI*) genes and 12 *chlorophyll a/b binding protein complex* Ⅱ genes exhibited up-regulated expression trends at 3-5W, but the photosynthetic genes were dramatically downregulated in *naa15* (Supplementary Figure 4B), indicating that embryonic photosystem of the *naa15* mutant failed to function, which might be the reason for the white aborted seeds of *naa15*. However, further studies are needed to investigate the relationship between *NAA15* and genes related to photosystems.

Phytohormone-related DEGs during embryo differentiation and endosperm cellularization

Plant hormones are crucial signaling molecules in seed development. It has been reported that auxin is involved in the communication between the endosperm, embryo, and seed coat, and proper embryo development requires the transportation of auxin between endosperm and integument (Batista et al., 2019). In the *agl62* mutant of *Arabidopsis*, the export of auxin is impaired, causing auxin accumulation in the endosperm, and cellularization prematurely occurs (Kang et al., 2008; Figueiredo et al., 2016). In the *auxin response factor 2* (*arf2*) mutant of *Arabidopsis*, extra cell proliferation occurs in the seed integuments, leading to enlarged seed coats and enhanced seed mass (Schruff et al., 2006). Additionally, it is well documented that auxin has an immediate negative control on cytokinin levels by inhibiting its biosynthesis (Nordström et al., 2004), and the ratio of auxin and cytokinin is critical for tissue proliferation and differentiation (Day et al., 2008). The cytokinin has important roles in root vascular patterning, and the septuple mutant *log1 log2 log3 log4 log5 log7 log8* shows severe embryonic vascular tissue development and patterning defects (De Rybel et al., 2014).

In our transcriptome data, the GO and KEGG analyses showed that the signal transduction pathways related to phytohormones were enriched in 3-5W (Figure 3 and Supplementary Table 10), involving in auxin, cytokinin, and abscisic acid (Supplementary Table 11). In *naa15*, DEGs related to auxin response factor (ARF), auxin-responsive protein (IAA), and auxin efflux proteins were significantly down-regulated in *naa15* (Figures 5A-C). Hence, the auxin response and transport in *naa15* was disrupted owing to deletion of *NAA15*. In addition, the genes related to several cytokinin dehydrogenases and cytokinin-activating enzymes were also down-regulated in *naa15* (Figure 5D). It has been reported that cytokinin signaling is crucial during syncytial endosperm development (Day et al., 2008), and the homeostasis of auxin and cytokinin is essential for seed formation. Enhancing auxin or reducing cytokinin can disrupt embryo development (Zhang et al., 2017).

Based on our transcriptional data, DEG-related hormones might be useful for investigating the roles of hormone regulation in these processes. However, further investigation into the relationship between *NAA15* and phytohormones (auxin and cytokinin) were required.

Transcription factors in early embryo and endosperm development

The spatiotemporal expression of different genes is an important foundation during embryo and endosperm development and is regulated by their transcription factors (TFs). In early seed development of maize, transcriptome analysis shows that 110 TFs display high temporal specificity expression, and that *MYB131*, *MYB16*, and *BZIP109* might be critical for endosperm cellularization (Yi et al., 2019). The WOX, HD-Zip, ARF, and CUC families of the transcription factors have been reported to be crucial in the establishment of embryonic patterning (Verma et al., 2022). WOX2 and WOX8 are co-expressed in the zygote of *Arabidopsis*, and WOX2 expression in the apical region is activated by WOX8 and WOX9 to regulate proper embryo development (Breuninger et al., 2008). Besides, the *rev phb* (HD-Zip TF) double mutant displays single or radially symmetric cotyledon embryo patterning defects (Prigge et al., 2005). The WRKY transcription factor *MINISEED3* (*MINI3*) is expressed in both endosperm and globular embryos; hence, the endosperm proliferation in *mini3* mutant is retarded, leading to the reduced size of seeds (Luo et al., 2005).

In our research, the transcription regulator activity term was significantly enriched in 3-5W (Supplementary Figure 3A), but genes related to embryonic differentiation, such as *WOXs* (*WOX1*, *WOX2*, *WOX5*, *WOX8*), CUCs (*CUC2*, *CUC3*), and HD-Zip III (*ATHB8*, *REV*), exhibited downregulated expression in *naa15* (Figure 6 and Supplementary Table 15). N-terminal acetylation of cellular proteins affects protein interactions, subcellular targeting, protein folding, and protein degradation (Scott et al., 2011; Yang et al., 2013). However, up to date, no correlations between *NAA15* and these genes were reported. We proposed that *NAA15* mutation in *Arabidopsis* led to NatA inactivation, which might cause some protein misfolding and excessive accumulation, resulting in abnormal expressions of these DEGs in *naa15* seeds. Additionally, 35 *MYB*, 37 *bHLH*, 30 *WYKY*, and 39 *ERFs* from our transcriptome data showed abnormal expressions in *naa15* (Supplementary Table 14). Among these transcription factors, except for *MYB56*, *bHLH95*, *WRKY10*, and *WRKY2*, which have been reported for their roles in embryo and endosperm (Ueda et al., 2011; Zhang et al., 2013; MacGregor and Zhang., 2019), there were no reports on other transcription factors. The results implied that the transcription factors might have important functions in embryo and endosperm development and could serve as potential regulators for seed morphogenesis.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI Bioproject accession number is PRJNA856854.

Author contributions

JZ designed the program of studies and experiments, guided the entire process of the study, and revised the manuscript. CL finished the material collection, carried out most of the experiments, analyzed the data, and wrote the paper. FH performed qPCR analyses. HC performed endosperm autofluorescence assay. All authors read and approved the content of the paper. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (31870303, 32170337).

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.998664/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Seed characteristics and genotype analysis of wild type and *naa15*/+ mutants. **(A)** Seed phenotypes in siliques of the wild type (WT) and *naa15*/+. The homozygous white aborted *naa15* seeds are labeled by white arrows. Bar = 1mm. **(B)** Genotypic analysis of WT and *naa15*/+ mutants. FP, forward primer; RP, reverse primer.

SUPPLEMENTARY FIGURE 2
Global analysis of the RNA-seq profile. **(A)** Violin map of gene expression distribution. The vertical axis meant the value change of the gene expression using $\log_{10}$(FPKM). The horizontal axis represented eighteen biological samples. **(B)** Principal component analysis (PCA) of eighteen biological samples. **(C)** Pearson-correlation coefficients showed the repeatability of the samples.

SUPPLEMENTARY FIGURE 3
Gene Ontology (GO) enriched terms of DEGs in 4W vs 3W and 5W vs 4W of *Arabidopsis*. **(A,B)** GO enrichment analysis of up-regulated DEGs in 4W vs 3W and 5W vs 4W W, respectively. **(C,D)** GO enrichment analysis of down-regulated DEGs in 4W vs 3W and 5W vs 4W, respectively.

SUPPLEMENTARY FIGURE 4
Heatmaps of DEGs related to sucrose and photosynthesis in WT and *naa15* of *Arabidopsis*. **(A)** Heatmap of DEGs related to sucrose synthesis and transport in WT and *naa15*. **(B)** Heatmap of DEGs related to photosynthesis in WT and *naa15*.

SUPPLEMENTARY FIGURE 5
Gene Ontology (GO) enriched terms of down-regulated DEGs from comparisons between WT and *naa15* in 5M vs 3W, 6M vs 4W, and 7M vs 5W of *Arabidopsis*.

SUPPLEMENTARY FIGURE 6
Gene Ontology (GO) enriched terms of up-regulated DEGs from comparisons between WT and *naa15* in 5M vs 3W, 6M vs 4W, and 7M vs 5W of *Arabidopsis*.

References

Aida, M., Beis, D., Heidstra, R., Willemsen, V., Blilou, I., Galinha, C., et al. (2004). The PLETHORA genes mediate patterning of the *Arabidopsis* root stem cell niche. *Cell* 119, 109–120. doi: 10.1016/j.cell.2004.09.018

Arnesen, T., Van Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., et al. (2009). Proteomics analyses reveal the evolutionary conservation and divergence of n-terminal acetyltransferases from yeast and humans. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8157–8162. doi: 10.1073/pnas.0901931106

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A., et al. (2016). Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* 44, 147–153. doi: 10.1093/nar/gkw419

Batista, R. A., Figueiredo, D. D., Santos-González, J., and Köhler, C. (2019). Auxin regulates endosperm cellularization in *Arabidopsis*. *Genes Dev.* 33, 466–476. doi: 10.1101/gad.316554.118

Berger, F. (2003). Endosperm: the crossroad of seed development. *Curr. Opin. Plant Biol.* 6, 42–50. doi: 10.1016/S1369526602000043

Bienvenut, W. V., Sumpton, D., Martinez, A., Lilla, S., Espagne, C., Meinnel, T., et al. (2012). Comparative large scale characterization of plant versus mammal proteins reveals similar and idiosyncratic n-α-acetylation features. *Mol. Cell Proteomics* 11, M111.0151311–111.0151314. doi: 10.1074/mcp.M111.015131

Bischof, S., Baerenfaller, K., Wildhaber, T., Troesch, R., Vidi, P. A., Roschitzki, B., et al. (2011). Plastid proteome assembly without Toc159: photosynthetic protein import and accumulation of n-acetylated plastid precursor proteins. *Plant Cell* 23, 3911–3928. doi: 10.1105/tpc.111.092882

Bishopp, A., Benková, E., and Helariutta, Y. (2011). Sending mixed messages: auxin-cytokinin crosstalk in roots. *Curr. Opin. Plant Biol.* 14, 10–16. doi: 10.1016/j.pbi.2010.08.014

Breuninger, H., Rikirsch, E., Hermann, M., Ueda, M., and Laux, T. (2008). Differential expression of WOX genes mediates apical-basal axis formation in the *Arabidopsis* embryo. *Dev. Cell* 14, 867–876. doi: 10.1016/j.devcel.2008.03.008

Caffall, K. H., and Mohnen, D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr. Res.* 344, 1879–1900. doi: 10.1016/j.carres.2009.05.021

Cheng, H., Dharmadhikari, A. V., Varland, S., Ma, N., Domingo, D., Kleyner, R., et al. (2018). Truncating variants in NAA15 are associated with variable levels of intellectual disability, autism spectrum disorder, and congenital anomalies. *Am. J. Hum. Genet.* 102, 985–994. doi: 10.1016/j.ajhg.2018.03.004

Chen, H., Li, S., Li, L., Hu, H., and Zhao, J. (2018b). *Arabidopsis EMB1990* encoding a plastid-targeted YlmG protein is required for chloroplast biogenesis and embryo development. *Front. Plant Sci.* 9, 181. doi: 10.3389/fpls.2018.00181

Chen, H., Li, S., Li, L., Wu, W., Ke, X., Zou, W., et al. (2018a). Nα-acetyltransferases 10 and 15 are required for the correct initiation of endosperm cellularization in *Arabidopsis*. *Plant Cell Physiol.* 59, 2113–2128. doi: 10.1093/pcp/pcy135

Chen, L. Q., Lin, I. W., Qu, X. Q., Sosso, D., McFarlane, H. E., Londoño, A., et al. (2015). A cascade of sequentially expressed sucrose transporters in the seed coat and endosperm provides nutrition for the *Arabidopsis* embryo. *Plant Cell* 27, 607–619. doi: 10.1105/tpc.114.134585

Cruz-Valderrama, J. E., Jiménez-Durán, K., Zúñiga-Sánchez, E., Salazar-Iribe, A., Márquez-Guzmán, J., and Gamboa-deBuen, A. (2018). Degree of pectin methyl esterification in endosperm cell walls is involved in embryo bending in *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 495, 639–645. doi: 10.1016/j.bbrc.2017.11.077

Day, R. C., Herridge, R. P., Ambrose, B. A., and Macknight, R. C. (2008). Transcriptome analysis of proliferating *Arabidopsis* endosperm reveals biological implications for the control of syncytial division, cytokinin signaling, and gene expression regulation. *Plant Physiol.* 148, 1964–1984. doi: 10.1104/pp.108.128108

De Rybel, B., Adibi, M., Breda, A. S., Wendrich, J. R., Smit, M. E., Novak, O., et al. (2014). Integration of growth and patterning during vascular tissue formation in *Arabidopsis*. *Science* 345, 1255215. doi: 10.1126/science.1255215

Dresselhaus, T., Sprunck, S., and Wessel, G. M. (2016). Fertilization mechanisms in flowering plants. *Curr. Biol.* 26, 125–139. doi: 10.1016/j.cub.2015.12.032

Endo, A., Tatematsu, K., Hanada, K., Duermeyer, L., Okamoto, M., Yonekura-Sakakibara, K., et al. (2012). Tissue-specific transcriptome analysis reveals cell wall metabolism, flavonol biosynthesis and defense responses are activated in the endosperm of germinating *Arabidopsis thaliana* seeds. *Plant Cell Physiol.* 53, 16–27. doi: 10.1093/pcp/pcr171

Feng, J., Li, R., Yu, J., Ma, S., Wu, C., Li, Y., et al. (2016). Protein n-terminal acetylation is required for embryogenesis in *Arabidopsis*. *J. Exp. Bot.* 67, 4779–4789. doi: 10.1093/jxb/erw257

Figueiredo, D. D., Batista, R. A., Roszak, P. J., Hennig, L., and Köhler, C. (2016). Auxin production in the endosperm drives seed coat development in *Arabidopsis*. *eLife* 5, e20542. doi: 10.7554/eLife.20542.028

Figueiredo, D. D., Batista, R. A., Roszak, P. J., and Köhler, C. (2015). Auxin production couples endosperm development to fertilization. *Nat. Plants* 1, 15184–15190. doi: 10.1038/nplants.2015.184

Fourquin, C., Beauzamy, L., Chamot, S., Creff, A., Goodrich, J., Boudaoud, A., et al. (2016). Mechanical stress mediated by both endosperm softening and embryo growth underlies endosperm elimination in *Arabidopsis* seeds. *Development* 143, 3300–3305. doi: 10.1242/dev.137224

French, S., Abu-Zaitoon, Y., Uddin, M., Bennett, K., and Nonhebel, H. (2014). Auxin and cell wall invertase related signaling during rice grain development. *Plants* 3, 95–112. doi: 10.3390/plants3010095

Gautschi, M., Just, S., Mun, A., Ross, S., Rücknagel, P., Dubaquié, Y., et al. (2003). The yeast n(alpha)-acetyltransferase NatA is quantitatively anchored to the ribosome and interacts with nascent polypeptides. *Mol. Cell Biol.* 23, 7403–7414. doi: 10.1128/MCB.23.20.7403-7414.2003

Giglione, C., and Meinnel, T. (2021). Evolution-driven versatility of n terminal acetylation in photoautotrophs. *Trends Plant Sci.* 26, 375–391. doi: 10.1016/j.tplants.2020.11.012

Hong, Z., Ueguchi-Tanaka, M., Fujioka, S., Takatsuto, S., Yoshida, S., Hasegawa, Y., et al. (2005). The rice brassinosteroid-deficient *dwarf2* mutant, defective in the rice homolog of *Arabidopsis* DIMINUTO/DWARF1, is rescued by the endogenously accumulated alternative bioactive brassinosteroid, dolichosterone. *Plant Cell* 17, 2243–2254. doi: 10.1105/tpc.105.030973

Jiang, W. B., and Lin, W. H. (2013). Brassinosteroid functions in *Arabidopsis* seed development. *Plant Signal. Behav.* 8, e259281–e259285. doi: 10.4161/psb.25928

Jo, L., Pelletier, J. M., and Harada, J. J. (2019). Central role of the LEAFY COTYLEDON1 transcription factor in seed development. *J. Integr. Plant Biol.* 61, 564–580. doi: 10.1111/jipb.12806

Kang, I. H., Steffen, J. G., Portereiko, M. F., Lloyd, A., and Drews, G. N. (2008). The AGL62 MADS domain protein regulates cellularization during endosperm development in *Arabidopsis*. *Plant Cell* 20, 635–647. doi: 10.1105/tpc.107.055137

Karunakaran, K. B., Amemori, S., Balakrishnan, N., Ganapathiraju, M. K., and Amemori, K. I. (2021). Generalized and social anxiety disorder interactomes show distinctive overlaps with striosome and matrix interactomes. *Sci. Rep.* 11, 18392–18417. doi: 10.1038/s41598-021-97418-w

Lafon-Placette, C., and Köhler, C. (2014). Embryo and endosperm, partners in seed development. *Curr. Opin. Plant Biol.* 17, 64–69. doi: 10.1016/j.pbi.2013.11.008

Lemke, M. D., Fisher, K. E., Kozlowska, M. A., Tano, D. W., and Woodson, J. D. (2021). The core autophagy machinery is not required for chloroplast singlet oxygen-mediated cell death in the *Arabidopsis thaliana* plastid ferrochelatase two mutants. *BMC Plant Biol.* 21, 342–362. doi: 10.1186/s12870-021-03119-x

Levesque-Tremblay, G., Müller, K., Mansfield, S. D., and Haughn, G. W. (2015). HIGHLY METHYL ESTERIFIED SEEDS is a pectin methyl esterase involved in embryo development. *Plant Physiol.* 167, 725–737. doi: 10.1104/pp.114.255604

Li, J., Nie, X., Tan, J. L., and Berger, F. (2013). Integration of epigenetic and genetic controls of seed size by cytokinin in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A* 110, 15479–15484. doi: 10.1073/pnas.1305175110

Linster, E., Stephan, I., Bienvenut, W. V., Maple-Grødem, J., Myklebust, L. M., Huber, M., et al. (2015). Downregulation of n-terminal acetylation triggers ABA-mediated drought responses in *Arabidopsis*. *Nat. Commun.* 6, 7640–7653. doi: 10.1038/ncomms8640

Li, C., Shang, J. X., Qiu, C., Zhang, B., Wang, J., Wang, S., et al. (2021). Plastid-localized *EMB2726* is involved in chloroplast biogenesis and early embryo development in *Arabidopsis*. *Front. Plant Sci.* 12, 675–838. doi: 10.3389/fpls.2021.675838

Li, J., Yu, M., Geng, L. L., and Zhao, J. (2010). The fasciclin-like arabinogalactan protein gene, FLA3, is involved in microspore development of *Arabidopsis*. *Plant J.* 64, 482–497. doi: 10.1111/j.1365-313X.2010.04344.x

Li, G., Zou, W., Jian, L., Qian, J., Deng, Y., and Zhao, J. (2017). Non-SMC elements 1 and 3 are required for early embryo and seedling development in *Arabidopsis thaliana*. *J. Exp. Bot.* 68, 1039–1054. doi: 10.1093/jxb/erx016

Luo, M., Dennis, E. S., Berger, F., Peacock, W. J., and Chaudhury, A. (2005). MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A* 102, 17531–17536. doi: 10.1073/pnas.0508418102

MacGregor, D. R., and Zhang, N. (2019). ICE1 and ZOU determine the depth of primary seed dormancy in *Arabidopsis* independently of their role in endosperm development. *Plant J.* 98, 277–290. doi: 10.1111/tpj.14211

Ma, T., Dong, F., Luan, D., Hu, H., and Zhao, J. (2019). Gene expression and localization of arabinogalactan proteins during the development of anther, ovule, and embryo in rice. *Protoplasma* 256, 909–922. doi: 10.1007/s00709-019-01349-3

Mansfield, S. G., and Briarty, L. G. (1991). Early embryogenesis in *Arabidopsis thaliana*. II. the developing embryo. *Can. J. Bot.* 69, 461–476. doi: 10.1139/b91-063

Moussu, S., Doll, N. M., Chamot, S., Brocard, L., Creff, A., Fourquin, C., et al. (2017). ZHOUPI and KERBEROS mediate embryo/endosperm separation by promoting the formation of an extracuticular sheath at the embryo surface. *Plant Cell* 29, 1642–1656. doi: 10.1105/tpc.17.00016

Nordström, A., Tarkowski, P., Tarkowska, D., Norbaek, R., Astot, C., Dolezal, K., et al. (2004). Auxin regulation of cytokinin biosynthesis in *Arabidopsis thaliana*: a factor of potential importance for auxin-cytokinin-regulated development. *Proc. Natl. Acad. Sci. U.S.A* 101, 8039–8044. doi: 10.1073/pnas.0402504101

Ohto, M. A., Floyd, S. K., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (2009). Effects of APETALA2 on embryo, endosperm, and seed coat development determine seed size in *Arabidopsis*. *Sex Plant Reprod.* 22, 277–289. doi: 10.1007/s00497-009-0116-1

Otegui, M., and Staehelin, L. A. (2000). Syncytial-type cell plates: a novel kind of cell plate involved in endosperm cellularization of *Arabidopsis*. *Plant Cell* 12, 933–947. doi: 10.1105/tpc.12.6.933

Peaucelle, A., Braybrook, S. A., Le Guillou, L., Bron, E., Kuhlemeier, C., and Höfte, H. (2011). Pectin-induced changes in cell wall mechanics underlie organ initiation in *Arabidopsis*. *Curr. Biol.* 21, 1720–1726. doi: 10.1016/j.cub.2011.08.057

Portereiko, M. F., Lloyd, A., Steffen, J. G., Punwani, J. A., Otsuga, D., and Drews, G. N. (2006). AGL80 is required for central cell and endosperm development in *Arabidopsis*. *Plant Cell* 18, 1862–1872. doi: 10.1105/tpc.106.040824

Prigge, M. J., Otsuga, D., Alonso, J. M., Ecker, J. R., Drews, G. N., and Clark, S. E. (2005). Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development. *Plant Cell* 17, 61–76. doi: 10.1105/tpc.104.026161

Ree, R., Varland, S., and Arnesen, T. (2018). Spotlight on protein n-terminal acetylation. *Exp. Mol. Med.* 50, 1–13. doi: 10.1038/s12276-018-0116-z

Riefler, M., Novak, O., Strnad, M., and Schmülling, T. (2006). *Arabidopsis* cytokinin receptor mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *Plant Cell* 18, 40–54. doi: 10.1105/tpc.105.037796

Sørensen, M. B., Mayer, U., Lukowitz, W., Robert, H., Chambrier, P., Jürgens, G., et al. (2002). Cellularisation in the endosperm of *Arabidopsis thaliana* is coupled to mitosis and shares multiple components with cytokinesis. *Development* 129, 5567–5576. doi: 10.1242/dev.00152

Schruff, M. C., Spielman, M., Tiwari, S., Adams, S., Fenby, N., and Scott, R. J. (2006). The AUXIN RESPONSE FACTOR 2 gene of *Arabidopsis* links auxin signalling, cell division, and the size of seeds and other organs. *Development* 133, 251–261. doi: 10.1242/dev.02194

Scott, D. C., Monda, J. K., Bennett, E. J., Harper, J. W., and Schulman, B. A. (2011). N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. *Science* 334, 674–678. doi: 10.1126/science.1209307

Song, J. P., Xie, X., Cui, Y., and Zou, J. T. (2021). Endosperm-embryo communications: recent advances and perspectives. *Plants* 10, 2511–2520. doi: 10.3390/plants10112511

Steffen, J. G., Kang, I. H., Portereiko, M. F., Lloyd, A., and Drews, G. N. (2008). AGL61 interacts with AGL80 and is required for central cell development in *Arabidopsis*. *Plant Physiol.* 148, 259–268. doi: 10.1104/pp.108.119404

Sun, X., Shantharaj, D., Kang, X., and Ni, M. (2010). Transcriptional and hormonal signaling control of *Arabidopsis* seed development. *Curr. Opin. Plant Biol.* 13, 611–620. doi: 10.1016/j.pbi.2010.08.009

Sun, S., Yi, C., Ma, J., Wang, S., Peirats-Llobet, M., Lewsey, M. G., et al. (2020). Analysis of spatio-temporal transcriptome profiles of soybean (*Glycine max*) tissues during early seed development. *Int. J. Mol. Sci.* 21, 7603–7624. doi: 10.3390/ijms21207603

Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., et al. (2013). An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* 25, 270–287. doi: 10.1105/tpc.112.107334

Tejos, R. I., Mercado, A. V., and Meisel, L. A. (2010). Analysis of chlorophyll fluorescence reveals stage specific patterns of chloroplast-containing cells during *Arabidopsis* embryogenesis. *Biol. Res.* 43, 99–111. doi: 10.4067/S0716-97602010000100012

Ueda, M., Zhang, Z., and Laux., T. (2011). Transcriptional activation of *Arabidopsis* axis patterning genes WOX8/9 links zygote polarity to embryo development. *Dev. Cell.* 20, 264–270. doi: 10.1016/j.devcel.2011.01.009

Vanstraelen, M., and Benkova, E. (2012). Hormonal interactions in the regulation of plant development. *Annu. Rev. Cell Dev. Biol.* 28, 463–487. doi: 10.1146/annurev-cellbio-101011-155741

Varland, S., Osberg, C., and Arnesen, T. (2015). N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. *Proteomics* 15, 2385–2401. doi: 10.1002/pmic.201400619

Verma, S., Attuluri, V., and Robert, H. S. (2022). Transcriptional control of *Arabidopsis* seed development. *Planta* 255, 90–99. doi: 10.1007/s00425-022-03870-x

Wang, W., Xiong, H., Sun, K., Zhang, B., and Sun, M. X. (2022). New insights into cell-cell communications during seed development in flowering plants. *J. Integr. Plant Biol.* 64, 215–229. doi: 10.1111/jipb.13170

Xu, F., Huang, Y., Li, L., Gannon, P., Linster, E., Huber, M., et al. (2015). Two n-terminal acetyltransferases antagonistically regulate the stability of a nod-like receptor in *Arabidopsis*. *Plant Cell* 27, 1547–1562. doi: 10.1105/tpc.15.00173

Yang, D., Fang, Q., Wang, M., Ren, R., Wang, H., He, M., et al. (2013). Na-Acetylated Sir3 stabilizes the conformation of a nucleosome-binding loop in the BAH domain. *Nat. Struct. Mol. Biol.* 20, 1116–1118. doi: 10.1038/nsmb.2637

Yi, F., Gu, W., Chen, J., Song, N., Gao, X., Zhang, X., et al. (2019). High temporal-resolution transcriptome landscape of early maize seed development. *Plant Cell* 31, 974–992. doi: 10.1105/tpc.18.00961

Yoon, J., Cho, L. H., Tun, W., Jeon, J. S., and An, G. (2021). Sucrose signaling in higher plants. *Plant Sci.* 302, 110703–110712. doi: 10.1016/j.plantsci.2020.110703

Zhang, Y., Liang, W., Shi, J., Xu, J., and Zhang, D. (2013). *MYB56* encoding a R2R3 MYB transcription factor regulates seed size in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* 55, 1166–1178. doi: 10.1111/jipb.12094

Zhang, Z., Tucker, E., Hermann, M., and Laux, T. (2017). A molecular framework for the embryonic initiation of shoot meristem stem cells. *Dev. Cell* 40, 264–277. doi: 10.1016/j.devcel.2017.01.002

Zou, W. X., Li, G., Jian, L., Qian, J., Liu, Y., and Zhao, J. (2021). *Arabidopsis* SMC6A and SMC6B have redundant function in seed and gametophyte development. *J. Exp. Bot.* 72, 4871–4887. doi: 10.1093/jxb/erab181

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 25 October 2022
DOI 10.3389/fpls.2022.1020506

OPEN ACCESS

EDITED BY
Igor Cesarino,
University of São Paulo, Brazil

REVIEWED BY
Vincenzo Lionetti,
Sapienza University of Rome, Italy
Yoselin Benitez-Alfonso,
University of Leeds, United Kingdom

*CORRESPONDENCE
E. M. Bayer
emmanuelle.bayer@u-bordeaux.fr
A. Voxeur
aline.voxeur@inrae.fr

†These authors share last authorship

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

RECEIVED 16 August 2022
ACCEPTED 04 October 2022
PUBLISHED 25 October 2022

CITATION
Paterlini A, Sechet J, Immel F, Grison MS, Pilard S, Pelloux J, Mouille G, Bayer EM and Voxeur A (2022) Enzymatic fingerprinting reveals specific xyloglucan and pectin signatures in the cell wall purified with primary plasmodesmata. *Front. Plant Sci.* 13:1020506. doi: 10.3389/fpls.2022.1020506

Enzymatic fingerprinting reveals specific xyloglucan and pectin signatures in the cell wall purified with primary plasmodesmata

A. Paterlini [1], J. Sechet [2], F. Immel [1], M. S. Grison [1], S. Pilard [3], J. Pelloux [4], G. Mouille [2], E. M. Bayer [1*†] and A. Voxeur [2*†]

[1]Laboratoire de Biogenèse Membranaire, Unité mixte de recherche (UMR5200), Université Bordeaux, Centre national de la recherche scientifique (CNRS), Villenave d'Ornon, France, [2]Institut Jean-Pierre Bourgin (IJPB), Université Paris-Saclay, Institut National de Recherche pour l'Agriculture, l'alimentation et l'Environnement (INRAE), AgroParisTech, Versailles, France, [3]Plateforme Analytique, Université de Picardie, Amiens, France, [4]UMRT (Unité Mixte de Recherche Transfrontaliére) INRAE (Institut National de recherche pour l'Agriculture, l'alimentation et l'Environnement) 1158 BioEcoAgro – BIOPI Biologie des Plantes et Innovation, Université de Picardie, Amiens, France

Plasmodesmata (PD) pores connect neighbouring plant cells and enable direct transport across the cell wall. Understanding the molecular composition of these structures is essential to address their formation and later dynamic regulation. Here we provide a biochemical characterisation of the cell wall co-purified with primary PD of *Arabidopsis thaliana* cell cultures. To achieve this result we combined subcellular fractionation, polysaccharide analyses and enzymatic fingerprinting approaches. Relative to the rest of the cell wall, specific patterns were observed in the PD fraction. Most xyloglucans, although possibly not abundant as a group, were fucosylated. Homogalacturonans displayed short methylated stretches while rhamnogalacturonan I species were remarkably abundant. Full rhamnogalacturonan II forms, highly methyl-acetylated, were also present. We additionally showed that these domains, compared to the broad wall, are less affected by wall modifying activities during a time interval of days. Overall, the protocol and the data presented here open new opportunities for the study of wall polysaccharides associated with PD.

KEYWORDS

plasmodesmata, cell wall, *Arabidopsis thaliana*, enzymatic fingerprinting, xyloglucans, homogalacturonans, rhamnogalacturonan I, rhamnogalacturonan II

Introduction

Cellular domains with specific arrays of resident macromolecules are well documented in many organisms (Scorrano et al., 2019). Plasma membrane domains of various sizes have received particular attention and have been shown to carry a plethora of functions (Gronnier et al., 2018). The domain concept can also extend to the surrounding cell wall, in biological systems where this structure is present.

Plasmodesmata (PD), membrane-lined pores across the wall of neighbouring plant cells, offer an example of this (Li et al., 2020). These conduits enable cell-cell transport and communication essential for plant growth and development (Zambryski and Crawford, 2000). PD can be regarded as membrane nanodomains in reason of their specific lipid composition (Grison et al., 2015; Liu et al., 2020) and protein population (Bayer et al., 2006; Fernandez-Calvino et al., 2011; Brault et al., 2019). In addition, their proximal wall also carries unique polysaccharide signatures (Northcote et al., 1989; Turner et al., 1994; Casero and Knox, 1995; Roy et al., 1997; Orfila and Knox, 2000; Faulkner et al., 2008).

Relevantly, the tight developmental and environmental control over the conductive status of PD is largely mediated by polysaccharides present in the local wall (Knox and Benitez-Alfonso, 2014). Callose, a polysaccharide solely composed of linear β-(1→3)-linked glucose residues, is enriched around PD (Northcote et al., 1989; Turner et al., 1994) and its amount negatively correlates with PD permeability (Amsbury et al., 2018). Callose synthase and glucanase proteins localise to PD and dynamically control the levels of this polysaccharide (Levy et al., 2007; Vatén et al., 2011). The composition of the PD wall microdomain is therefore both context dependent and differentially regulated compared to the bulk wall (via local modifying proteins). Callose might also impart mechanical properties to the PD wall domain through interactions with other wall components (Abou-Saleh et al., 2018), highlighting an additional level of regulation.

While callose has a relatively unique standing in terms of classification (Scheller and Ulvskov, 2010), cell wall polysaccharides have been traditionally divided into three groups: cellulose, hemicelluloses and pectins (Lampugnani et al., 2018). Cellulose consists of linear β-(1→4)-linked glucan chains and provides much of the mechanical strength and rigidity of plant walls (Kerstens et al., 2001). A depletion of cellulose seems to characterise PD positions in tobacco (Faulkner et al., 2008). Speculatively, a pliable wall might better accommodate dynamic changes in PD aperture or enable the modification of PD morphology itself. However, significant amounts of cellulose might actually delimit PD in onion peels (Liesche et al., 2013). Pectins, composed of homogalacturonans (HGs) and rhamnogalacturonans (RG) - I and II, and hemicelluloses such as xyloglucans, xylans, mannans and glucomannans are more heterogeneous and can present complex branched structures. As a general simplification, pectins contain large amounts of galacturonic acid residues (GalA) while hemicelluloses have β-(1→4)-linked backbones of glucose, mannose or xylose (Voragen et al., 2009; Scheller and Ulvskov, 2010). Pectin polysaccharides also display specific patterns at PD. Enrichments in (1 → 5)-α-arabinan-containing RG-I species (Orfila and Knox, 2000; Faulkner et al., 2008) and depletions of (1 → 4)-β-galactan ones (Roy et al., 1997; Orfila and Knox, 2000) have been established in tobacco, tomato and apple species. Low-esterified HGs are also abundant at PD in tomatoes and apples (Casero and Knox, 1995; Roy et al., 1997; Orfila and Knox, 2000). Pectin methylesterase (PME) proteins, which would influence this class of polysaccharides, have been localised at PD in flax and tobacco (Morvan et al., 1998; Chen et al., 2000). Their activity has been associated with the systemic spread of viruses (Dorokhov et al., 1999; Chen and Citovsky, 2003; Lionetti et al., 2014a). Overall, these examples highlight how PD wall composition might be both structurally and functionally important. The data also strongly supports the idea of PD as a spatially defined wall subdomain.

However, we still have an incomplete picture of the polysaccharide species present at PD. Most of the mentioned PD wall microenvironment data has been obtained from antibody and stain approaches. While these provide extremely valuable *in-situ* information, they are target restricted and dependent on the accessibility of the epitopes in the intact walls. We therefore considered alternative approaches to obtain broader (and more high-throughput) information at PD.

PD purification from *Arabidopsis thaliana* cell cultures has been successfully employed to produce lists of proteins closely associated with these structures (Bayer et al., 2006; Fernandez-Calvino et al., 2011; Brault et al., 2019) and to describe their lipid environment (Grison et al., 2015; Liu et al., 2020). The cells in the culture display clear PD morphologies (Bayer et al., 2004; Nicolas et al., 2017) that are likely "primary" in nature. Primary is a term used to indicate PD that have formed during cell division (Ehlers and Kollmann, 2001). The PD purification protocol involves isolation of a wall pellet (containing PD) and a subsequent mild enzymatic digestion with cell wall degrading enzymes, releasing PD-derived structures (Faulkner and Bayer, 2017). The final "PD fraction" is highly enriched in PD membrane compartments. However, the wall in the immediate proximity of PD is also likely to be co-purified. A number of extracellular cell wall modifying proteins could be indeed identified in proteomic analyses based on this approach (Knox and Benitez-Alfonso, 2014).

We reasoned that rather than treating the polysaccharides as undesired compounds co-purified with the membranous PD, the composition of the wall around PD could be studied exploiting these fractions. Xyloglucans and pectins were of particular interest as they have been speculated to be of relevance for PD function (Burch-Smith et al., 2011; Li et al., 2020; Kirk et al., 2022). In line with previous proteomic and lipidomic studies, we

employed liquid cultured cells of *A. thaliana* Landberg erecta ecotype. Two ages were considered, 4 and 7 days, as changes in PD morphology had been observed between those stages and remodelling of the PD-wall environment was suggested to contribute to the process (Nicolas et al., 2017).

Combining subcellular fractionation with biochemical analyses and enzymatic fingerprinting approaches, we highlight a number of specific polysaccharide signatures in the wall fraction purified with PD. Speculations on some functional roles for the same are also provided.

Results

A biochemical strategy to isolate and characterise wall components in cell culture fractions

The biochemical rationale for PD purification has never been explicitly discussed (Levy et al., 2007; Fernandez-Calvino et al., 2011) as it was the result of experimental trials validated by PD protein enrichments. We address the purification logic here, as it carries relevance for the specific results of this paper (Figure 1A).

The protocol relies on the solubilisation of a cell wall domain, that of PD, *via* a mild treatment with a commercial cellulase mixture (Cellulase "Onozuka" R-10 from *Trichoderma viride*). The mixture contains hemicellulase and pectinase activities in addition to the titular cellulase one (Beldman et al., 1985). The adjective mild is used in comparison to protoplasting approaches, which aim to fully digest the cell wall and usually employ additional enzymes (Pasternak et al., 2021). A shorter exposure and a lower concentration of enzyme (per amount of substrate) are employed during PD purification. These differences relate to the aim of solubilising a wall domain, rather than specific polysaccharides. The latter endeavour would require chelators and other specific substances, which are not employed. The cell wall digestion is also intentionally partial as a longer digestion could solubilise most of the cell wall.

The assumption would be that the "Onozuka" R-10 mixture digests cellulose, releasing PD structures. Secondary enzymatic activities in the mixture (in addition to cellulase) have received less attention by the PD community. However, by treating pure commercial xyloglucans and pectins with the "Onozuka" R10 mixture, a number of polysaccharide fragments and by-products became visible upon High Pressure Size Exclusion chromatography (HPSEC) combined with high resolution mass spectrometry (HRMS) (Figures 1B, C; Supplementary Figures S1A, B). Our analysis shows that the mixture can actually easily degrade low methyl-esterified pectins in di-, tri- and tetrameres of GalA. PD are embedded in low-esterified HGs wall domains (Roy et al., 1997; Orfila and Knox, 2000) so, in our view, processing of these highly digestible components is more likely to explain the release of PD structures than cellulose digestion (Figure 1A).

Once the attachment sites are partially digested, PD (membranes and associated wall components) can be released from the rest of the wall. These small suspended domains containing PD are then separated from the bulk wall by a low speed centrifugation (Figure 1A). To then isolate PD from soluble digestion products (oligosaccharides released by the "Onozuka" R-10 enzymatic mixture), an ultracentrifugation step is performed. PD are expected to be the biggest product because of the presence of callose - which would not be digested by this enzymatic mixture - and of retained membranous components. These structures can therefore be easily separated from other degradation products (Figure 1A)

The standard PD purification strategy described in Faulkner and Bayer (2017) and applied in a number of papers in the field (Fernandez-Calvino et al., 2011; Brault et al., 2019) was re-deployed in this study just changing the fractions being collected. Cells were passed through a N_2 disruptor device and the wall pellet of the lysed cells was collected. It was defined as the non-digested wall (NDW) fraction. The pellet (bulk wall) after mild "Onozuka" R-10 cellulase enzymatic digestion was defined as the digested wall fraction (DIW) and the pellet obtained from the supernatant after ultracentrifugation as the PD wall (PDW) fraction (Figure 1A).

Transmission electron microscopy (TEM) observations provided early indications that cell wall polysaccharides are indeed retained during PD purification (Fernandez-Calvino et al., 2011). Microfibril structures are visible among fixed PD derived vesicles labelled with antibodies against Plasmodesmata Located Protein 1 (PDLP1) (Thomas et al., 2008) and colloidal gold particles (Figures 2A, B). A range of polysaccharides, among which callose and cellulose (Cifuentes et al., 2010; Ding et al., 2014) has this specific appearance. The presence of callose in PDW was directly confirmed by immunolabeling with antibodies against Plasmodesmata callose binding proteins (PDCB1) (Simpson et al., 2009) and callose itself, combined with colloidal gold (Figure 2C). While the extracellular domains of PDLPs' display structural homology to fungal lectins, capable of binding a range of carbohydrates (Vaattovaara et al., 2019), the extracellular domains of PDCBs specifically recognise callose (Barral et al., 2005; Simpson et al., 2009). A reduction in the number of microfibril structures can be observed comparing PDW to DIW fractions (Supplementary Figure S2). Extensive anamorphous-looking material is also visible (Figures 2B, C) and might correspond to non-crystalline hemicelluloses and/or pectins, presenting less defined appearances under TEM (Chen et al., 2016; Wathoni et al., 2019).

The NDW, DIW and PDW fractions were subjected to the analyses depicted in the lower part of Figure 3. We performed a monosaccharide analysis (Fang et al., 2016) to obtain broad information on the cell wall sugars (Figure 3). Specific structural features were then defined using enzymatic fingerprinting

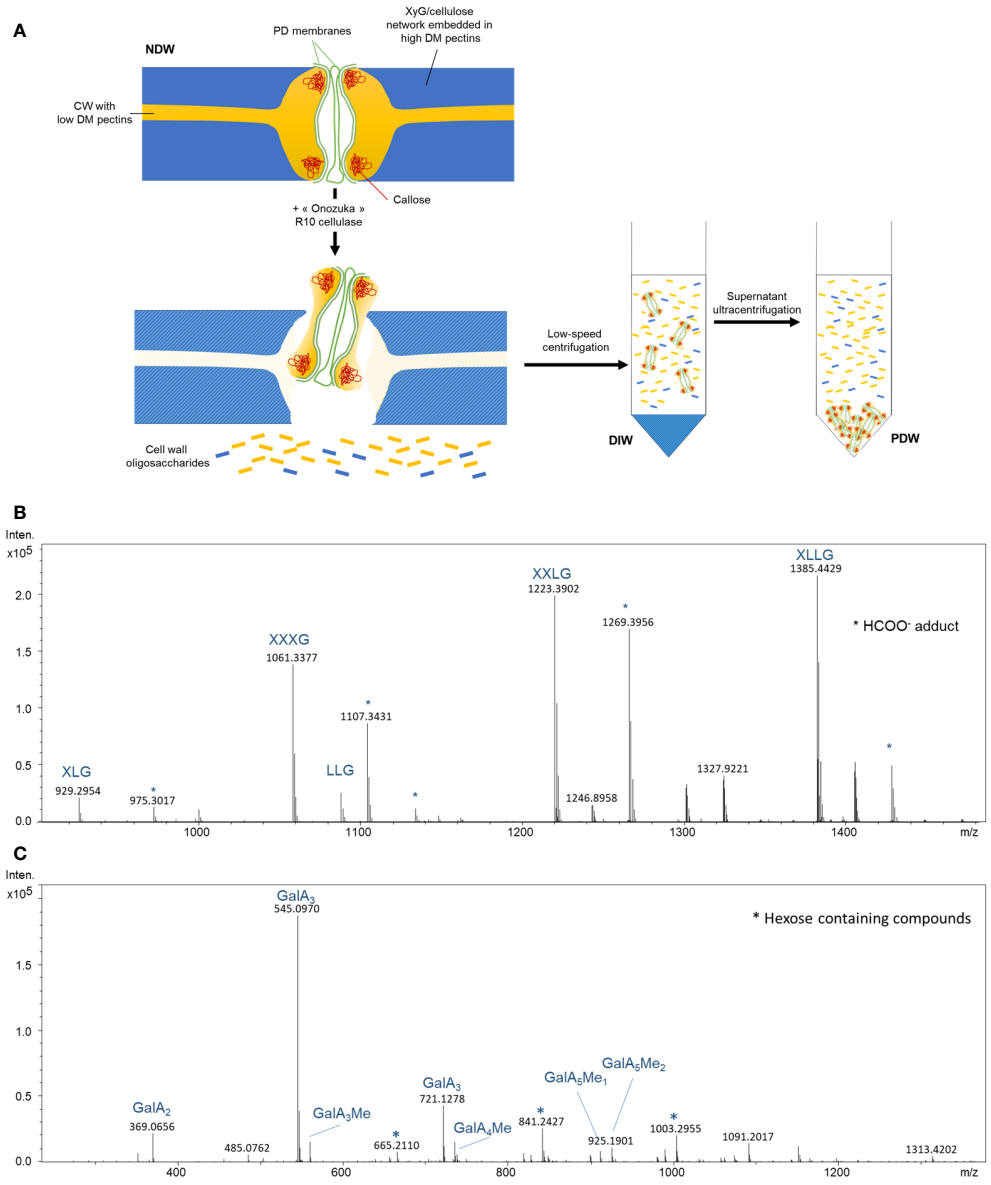


FIGURE 1
Biochemical steps involved in the release and isolation of PD from the cell wall. **(A)** Suggested model for the activity of "Onozuka" R 10 cellulase treatment on wall polysaccharides nearby PD. Subsequent key steps involved in PD purification are also depicted. DM, degree of methylesterification; XyG, Xyloglucan; NDW, non-digested wall fraction; DIW, digested wall fraction; PDW, PD wall fraction. Exposure of commercial tamarind xyloglucan **(B)** and citrus pectins **(C)** to "Onozuka" R 10 cellulase results in the production of polysaccharide fragments. MS spectrum obtained from treated samples. Intens: signal intensity; m/z: mass to charge ratio. Xyloglucan oligosaccharides follow the nomenclature described in Fry et al. (1993): G stands for glucose; X for xylose-glucose; L for galactose-xylose-glucose. OGs are named GalAxMeyAcz. Numbers indicate the degree of polymerization and the number of methyl ester groups. GalA, galacturonic acid; Me methyl-ester group. * Adduct.

approaches. By digesting cell walls with enzymes possessing well-defined cleavage specificities, a unique fingerprint of oligosaccharides can be produced. The subsequent detailed analysis of these products by HPSEC-HRMS (Voxeur et al., 2019) gives structural information on the original polysaccharide (Limberg et al., 2000). The validity of the approach has been previously confirmed using immunolabeling approaches (Jobert et al., 2022 - bioRxiv; Cavalier et al., 2008)

An endo-cellulase treatment, which cuts the xyloglucan backbone after non-substituted glucose residues (Figure 3; Lerouxel et al., 2002), was employed to describe xyloglucan structures. The xyloglucan oligosaccharides produced are

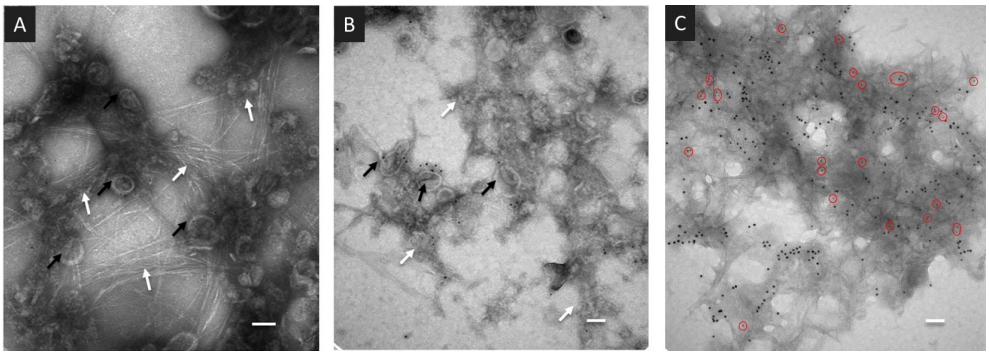


FIGURE 2
Wall-derived material is present in PDW fractions purified from *A.thaliana* cell cultures. **(A, B)** TEM images of the PDW fractions containing PD-derived vesicles (black arrows) labelled with antibodies against the PDLP1 protein and colloidal gold particles. Fibrillar **(A)** and anamorphous **(B)** structures are visible (white arrows). **(C)** TEM images of the PDW fraction labelled with antibodies against the PDCB1 protein (15nm - larger - colloidal gold particles) and callose (5 nm - smaller - gold particles highlighted by red circles). Panels **(A, B)** have scale bars of 50 nm; panel **(C)** has a scale bar of 200 nm.

characterised by their monosaccharide sequence, their branching and their acetylation status (Fry et al., 1993).

Digestion with endo-polygalacturonase (cutting between two non-methyl-esterified galacturonic acids of HGs) (Figure 3; Pedrolli et al., 2009) combined with oligosaccharide analysis revealed the structure of oligogalacturonan (OG) stretches derived from the digestible HG backbone. OGs are linear and are characterised by their degree of polymerisation

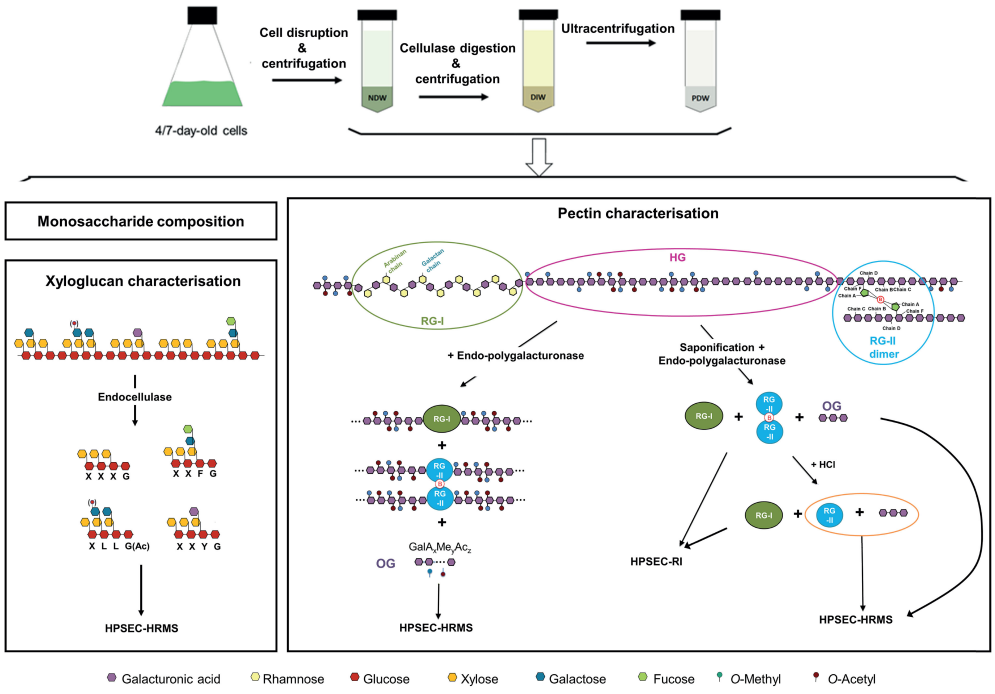


FIGURE 3
Purification steps required to obtain cell wall fractions and biochemical analyses to identify the polysaccharides present in the same. NDW, Non digested cell wall; DIW, Digested cell wall; PDW, plasmodesmata cell wall; RG-I, Rhamnogalacturonan I; RG-II, Rhamnogalacturonan II; HG, Homogalacturonans; HPSEC, High-pressure size exclusion chromatography; HRMS, High resolution mass spectrometry; RI, refractive index; OG, Oligogalacturonides; Ac, Acetyl-ester group; Me, Methyl-ester group; OGs are named GalAxMeyAcz. Numbers indicate the degree of polymerisation (DP) and the number of methyl- and acetyl-ester groups, respectively.

(DP) and their acetylation (Ac) and methylation status (Me). They are named GalAxMeyAcz where the numbers indicate the degree of polymerization and the number of methyl- and acetyl-ester groups, respectively (Voragen et al., 2009). A similar analysis, was performed after saponification which largely strips pectins of their methyl- and acetyl-ester chemical modifications (Séveno et al., 2009) and creates a fully digestible HG backbone. Digestion with endo-polygalacturonase of these saponified pectins results in the release of OGs, RG-I and RG-II dimers, which can be next separated by size-exclusion chromatography and detected by refractive index (RI) (Figure 3; Ishii et al., 1999; O'Neill et al., 2001). Since RG-II dimers, unlike monomers, cannot be detected by mass spectrometry, we performed a hydrochloric acid (HCl) treatment, which releases RG-II monomers from boron cross-linking (O'Neill et al., 1996). RG-II monomer and OGs were then analysed by HPSEC-HRMS (Figure 3).

"Onozuka" R10 cellulase treatment solubilizes fucosylated and galacturonic acid-containing xyloglucans, HG, RG-I and RG-II from the wall

As discussed, treatment with a commercial cellulase mixture is necessary to release PD from the rest of the wall (Figures 1, 3; Faulkner and Bayer, 2017). Polysaccharides lost from the NDW fraction will therefore include those present in the proximity of PD (PDW fraction) plus those directly processed by the enzymes. A detailed comparison between NDW and DIW fractions was performed to better define which polysaccharides are lost upon "Onozuka" R10 cellulase treatment.

In the monosaccharide analysis we focused on fucose, galactose, xylose, mannose, rhamnose, arabinose, glucose and galacturonic acid residues, which would be derived from hemicelluloses and pectin components (Voragen et al., 2009; Scheller and Ulvskov, 2010). No robust changes in the levels of any of these sugars were observed upon "Onozuka" R10 cellulase digestion, in both fractions derived from 4- or 7- day-old cell cultures (Supplementary Figure S3). This suggests that cellulose, pectins and hemicelluloses are roughly equally digested upon the PD purification process.

Upon xyloglucan fingerprinting, we observed a significant loss of fucosylated and galacturonic acid-containing xyloglucans (such as XXFG and XXYG) (Figure 4) in DIW fractions. This loss, together with cellulose degradation, causes increases in the abundance of other xyloglucan species (e.g. the galactose containing XLLG and XXLG forms) (Figure 4), as individual peak areas are scaled by the alcohol insoluble residue (AIR) weight.

HG fingerprinting also revealed that methyl-esterified pectins were released upon "Onozuka" R10 cellulase digestion. The NDW and DIW enzymatic fingerprinting highlighted solubilisation of HG containing monomeric galacturonic acid, GalA2 and methyl-esterified stretches (DPx>4Mey and DPx>4MeyAcz) (Figure 5A) upon PD extraction. A full profile of the OGs detected in NDW and DIW is provided in Supplementary Figure S4. As HG methyl-esterification prevents endo-polygalacturonase digestion (Séveno et al., 2009), we performed saponification on the fractions prior to digestion in order to assess the total amount of pectins released upon PD extraction. An increase of GalA3 was observed in fractions from both 4- and 7-day-old cultures. This suggests that GalA3 is likely derived from non and/or lowly methyl-esterified wall regions, obtained by saponification. The amount of this OG is indeed negligible before treatment (Figures 5A, B). GalA/GalA2, conversely, are more likely to correspond to short de-methyl-esterified stretches present among highly methyl-esterified ones prior to saponification (Figure 5A). This analysis confirmed that methyl-esterified pectins were released upon the PD purification process.

Saponification (when combined with endo-polygalacturonase digestion) additionally allows pure RG-II and RG-I isolation (Sun et al., 2019). RG-II is a complex polysaccharide with multiple different side branches. Six conserved types (Supplementary Figure S5) have been described to date (O'Neill et al., 2004; Ndeh et al., 2017).

In the HPSEC-HRMS analysis of saponified and PG-digested wall fractions we observed a main three-charged ion at monoisotopic m/z 1518.7 (Figure 6A - NDW as an example) of which the retention time matched those obtained for RG-II monomers (Supplementary Figure S5). Since RG-II is a high molecular weight molecule, the most abundant species in the pattern was actually the di-isotopic one, containing two ^{13}C atoms (m/z 1519.4; Figure 6A). Specifically, the ion corresponds to a molecular formula of $C_{166}H_{246}O_{145}$, fitting a dehydrated RG-II backbone containing 8 GalA substituted with the highly conserved side chains A, C and D, a B side chain hexasaccharide and two distinct α-L-arabinofuranose substituents (E and F; Supplementary Figure S5). We attributed other ions to mono or di-methyl-etherified GalA (m/z 1523.3 and 1528.0) and mono-acetylated side chain A (m/z 1532.7) (Figure 6A). A complete RG-II form is therefore present in the walls of *A. thaliana* cell cultures. After HCl treatment, which releases the RG-II monomers from boron cross-linking (O'Neill et al., 1996; Supplementary Figure S6), the main ions were observed as chloride adducts and more acetylated RG-IIs were detected in the fractions (Figure 6B - using NDW as an example). RG-II dimers have been reported to be more acetylated than monomers (O'Neill et al., 2020) (Figure 6B).

Most of these RG-II pectins remained in the DIW fraction after "Onozuka" R10 cellulase digestion (Figure 6C). Only a specific population of methyl-acetylated forms was solubilized (Figure 6C). As individual peak areas are scaled by the AIR weight, this results in the increase of RG-II content relative to other RG-II species in DIW (Figure 6C).

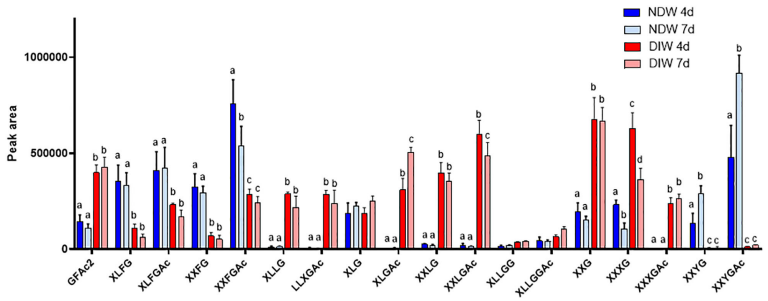


FIGURE 4
Enzymatic fingerprinting of xyloglucans in NDW and DIW fractions. Histograms display the peak area of each xyloglucan oligosaccharide released upon endo-cellulase digestion of NDW and DIW fractions obtained from 4- and 7-day-old *A.thaliana* cell cultures. Data are the mean of three biological replicates, error bars are the standard deviation. Lowercase letters indicate significant differences in the amounts of specific xyloglucans between fractions (two-way ANOVA, $p < 0.05$). Oligosaccharide nomenclature follows that described in Fry et al. (1993): G stands for glucose; X for xylose-glucose; L for galactose-xylose-glucose; F for fucose-galactose-xylose-glucose; Y for galacturonic acid-xylose-glucose; Ac: acetyl ester group.

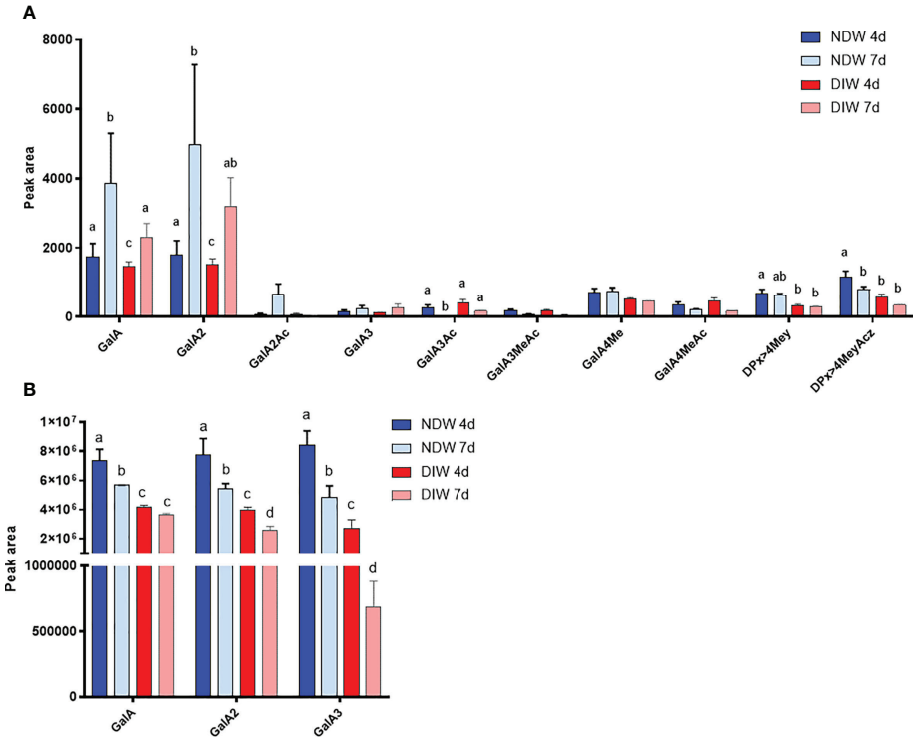


FIGURE 5
Enzymatic fingerprinting of HG in NDW and DIW fractions. Histograms display the peak area of each OGs analysed by HPSEC-HRMS and released upon endo-polygalacturonase digestion of non-saponified **(A)** and saponified **(B)** NDW and DIW fractions obtained from 4- and 7-day-old *A thaliana* cell cultures. Data are the mean of three biological replicates, error bars are the standard deviation. Lowercase letters indicate significant differences in the amounts of specific OGs between fractions (two-way ANOVA, $p < 0.05$). OGs are named GalAxMeyAcz. Numbers indicate the degree of polymerization and the number of methyl/acetyl ester groups. GalA, galacturonic acid; Me: methyl-ester group; Ac, acetyl-ester group; DP, degree of polymerisation. OGs larger than GalA4 are grouped together. z and y represent potential higher levels of acetylation/methylation in those longer OGs.

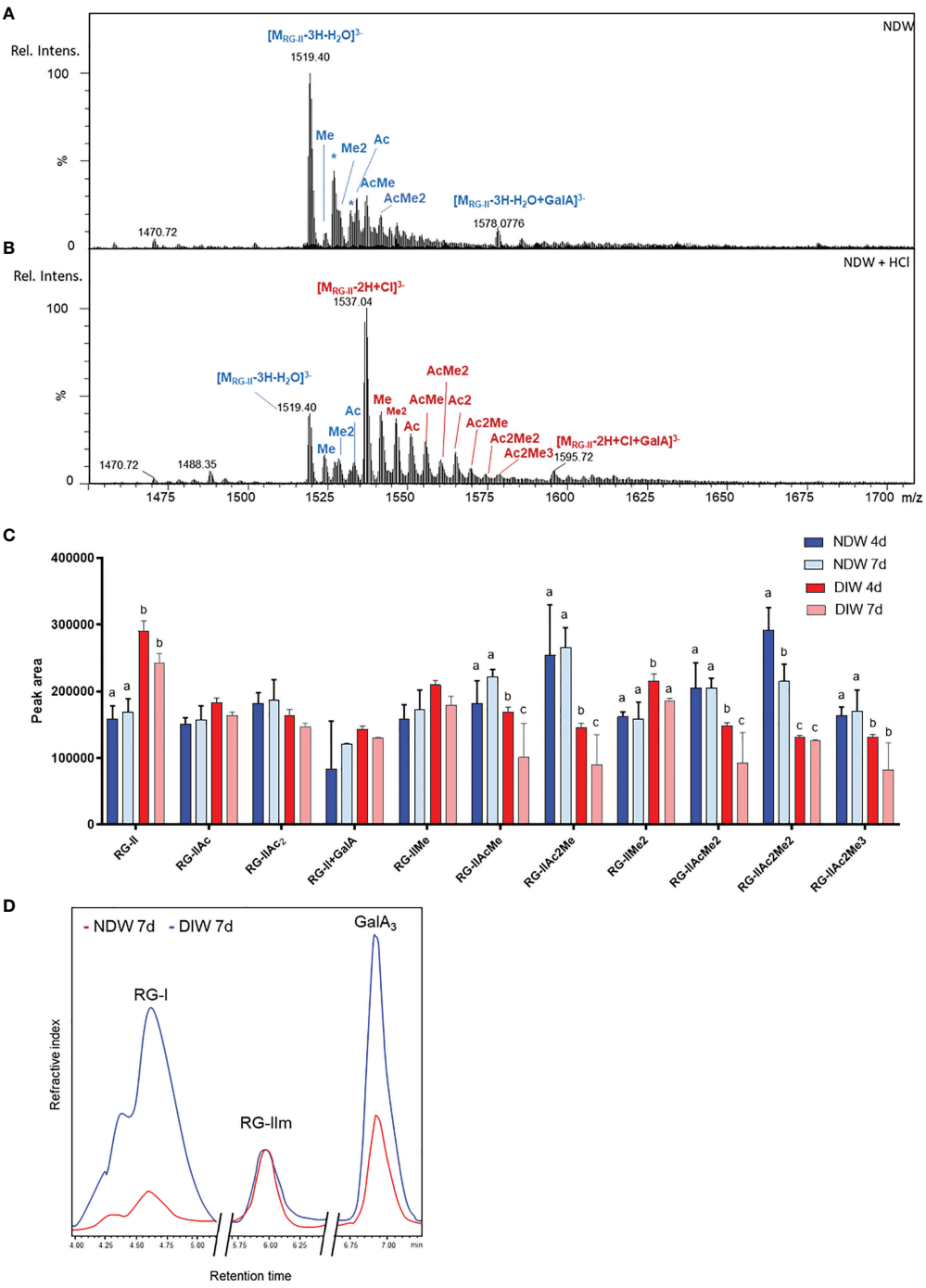


FIGURE 6
Rhamnogalacturonan analysis of NDW and DIW fractions. Mass spectrum of complete RG-II forms in NDW fractions derived from 7-day-old *A.thaliana* cell cultures before **(A)** and after **(B)** HCl treatment. Chloride adducts are represented in red and dehydrated forms in blue. Asterisks correspond to Na adducts. RG-II: rhamnogalacturonan-II; GalA; galacturonic acid; Rel. Intens.: relative signal intensity; m/z: mass-to-charge ratio; Me and Ac denote methylation and acetylation in RG-II species. **(C)** Histograms displaying the peak area of each RG-II form released upon endo-polygalacturonase digestion of saponified and HCl-treated NDW and DIW fractions obtained from 4- and 7-day-old *A thaliana* cell cultures. Data are the mean of three biological replicates, error bars are the standard deviation. Lowercase letters indicate significant differences in the amounts of specific RG-II species between fractions (two-way ANOVA, $p < 0.05$). **(D)** HPSEC-RI quantification of RG-I, RG-II monomers and GalA3 in HCl-treated NDW and DIW fractions derived from 7-day-old *A.thaliana* cell cultures. The x axis is broken.

HPSEC-RI on wall material derived from 7-day-old cell cultures also confirmed that the amount of RG-II is largely unaltered between NDW and DIW fractions (Figure 6D). The loss of low methyl-esterified pectins (GalA3) observed in Figure 5B was similarly supported. HPSEC-RI additionally displayed a significant loss of RG-I-containing pectins upon "Onozuka" R10 cellulase digestion (Figure 6D).

Overall, this cellulase digestion treatment seems to result in loss of fucosylated and galacturonic acid-containing xyloglucans; methyl-esterified HG; RG-I; and methyl-acetylated RG-II forms from the wall of *A. thaliana* cells.

The PDW fraction presents fucosylated xyloglucans, low methyl-esterified HGs, RG-I with limited branching and acetylated RG-II

To determine which of the solubilised polysaccharides belonged to the proximal PD wall and which were the result of direct enzymatic activities, we analysed PDW fractions. PDW pellets are obtained from purification of polysaccharides released upon "Onozuka" R10 cellulase treatment (Figures 1, 3; Faulkner and Bayer, 2017). Comparing these fractions with NDW ones (representative of the general wall of plant cell cultures) enables us to determine the presence of unique polysaccharide signatures at PD.

As the amount of PD material isolated is very low, PDW samples cannot be weighed and the results of enzymatic analyses have to be represented as relative data. Clear polysaccharide signals were nonetheless obtained from the monosaccharide analysis, indicating efficient retention of wall material in PDW fractions (Figure 7). This observation matches the preliminary TEM images (Figures 2A–C). A significant exception in detection was xylose, which we did not identify in any PDW sample (Figure 7). A low response factor or limited amounts of xyloses and xyloglucans, sources of this monosaccharide (Scheller and Ulvskov, 2010) might be the cause.

Clear differences between NDW and PDW fractions were also observed. Galacturonic acid, a sugar highly present in pectins (Figure 3; Voragen et al., 2009), was relatively more abundant in PDW (Figure 7). Galactan and arabinan monosaccharides, derived from side chains of RG-I (Figure 3; Zablackis et al., 1995) and/or arabinogalactan proteins (Ellis et al., 2010), instead displayed the opposite pattern (Figure 7). This might indicate low-substituted RG-I in PDW fractions.

Very few types of xyloglucans - mostly acetylated/non-acetylated-fucose-containing forms (XXFG and XLFG) - were detected in the PDW fraction (Figure 8). A loss of these species had indeed been observed from NDW fractions upon "Onozuka" R10 cellulase digestion (Figure 4). Galacturonic acid containing species, whose abundance was similarly reduced in DIW, were conversely not present in PDW. Loss of the latter might be attributable to interactions with cellulose (Park and Cosgrove, 2015) or by direct galacturonidase activity in the enzymatic mixture. Unlike NDW, galactosylated (L) but non-fucosylated (F) species were absent in PDW (Figure 8).

Upon HG enzymatic fingerprinting, PDW fractions yielded significantly higher amounts of GalA3 and lower amounts of GalA compared to NDW. Surprisingly, no GalA2 was produced upon poly-galacturonase treatment (Figure 9A). Based on the association of GalA3 with non and/or lowly methyl-esterified wall regions and GalA/GalA2 with methyl-esterified ones, the HG in the PD environment might have low levels of substitution. Higher levels of galacturonic acid in the monosaccharide analysis (Figure 7) also suggest more pectins. Nonetheless, low methyl-esterified regions, of which the degradation leads to the release of GalA4Me, were clearly present in PDW fractions (Figure 9A). A full profile of OGs detected in PDW is provided in Supplementary Figure S7. As fewer OGs extend beyond GalA9Me4 in PDW compared to NDW, the methyl-esterified HG stretches at PD might be shorter.

Upon saponification, PDW fractions continued to release more GalA3 than NDW ones (Figure 9B) confirming that PD are enriched in HGs compared to NDW. Since GalA was detected in treated PDW fractions, the continued absence of GalA2 must be due to specific structural motifs such as the presence of side chains and/or of PG hydrolysable HG stretches amongst RG-I backbone rather than HG methylesterication. Large losses of methyl-esterified HG (GalA- and GalA2-associated) had been observed upon "Onozuka" R10 cellulase treatment in DIW fraction (Figure 5B). The PDW polysaccharide content does not seem to fully mirror this composition (Figures 9A, B).

Due to the small amount of material in PDW fractions, it was not always possible to detect RG-II in those samples. A qualitative, rather than quantitative assessment, is therefore provided here. Upon HCl treatment, HPSEC-HRMS detected highly methyl-acetylated RG-II monomers, RG-IIMe_2Ac_2 being the main compound (Figure 10A). The degree of substitution was higher than that observed in NDW (Figure 6A). Loss of methyl-acetylated RG-II species had been observed upon "Onozuka" R10 cellulase digestion (Figure 6B), matching this PDW population.

The HPSEC-RI further supported the presence of RG-II monomers in PDW (Figure 10B). RG-II monomers and GalA3 were respectively slightly less and more abundant in PDW, compared to NDW (Figure 10B). The increased GalA3 abundance matches the previous OGs results (Figures 9A, B). HPSEC-RI also highlighted a high abundance of RG-I in PDW (Figure 10B), matching the loss observed in DIW (Figure 6D). A population of RG-I, corresponding to the first of the two peaks in the chromatogram, seems particularly prominent (Figure 10A).

Overall, the PDW fraction displayed a unique pattern of fucosylated xyloglucans; poorly methyl-esterified HGs with specific structures; potentially low-substituted RG-I species; and acetylated RG-II forms.

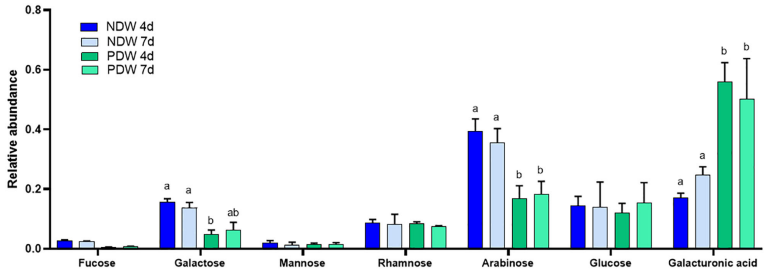


FIGURE 7
Monosaccharide analysis of NDW and PDW fractions. Histograms display the relative abundance of monosaccharides released upon AIR hydrolysis with 2M trifluoroacetic acid of NDW and PDW fractions obtained from 4- and 7-day-old *A.thaliana* cell cultures. Data are the mean of three biological replicates, error bars are the standard deviation. Lowercase letters indicate significant differences in the amounts of specific monosaccharides between fractions (two-way ANOVA, $p < 0.05$).

PDW fractions displays limited age-related polysaccharide remodelling

As plant tissues mature, their cell walls (including those proximal to PD) can display morphological changes (Nicolas et al., 2017 as an example). To determine potential shifts in wall composition (involving xyloglucans or HGs), we compared fractions derived from 4- and 7-day-old cultures.

Monosaccharide abundances in PDW and NDW fractions did not seem to display temporal patterns. (Supplementary Figure S3, Figure 7), suggesting that the main classes of polysaccharide are stable in relative amounts. Limited remodelling of xyloglucans (mostly involving galactosylation - L species) was observed in PDW fractions between the two time points assessed (Figure 8). A more pronounced change in galacturonydation (Y species) was conversely detected in NDW xyloglucans (Figure 4).

Pectin-derived OGs were the polysaccharides displaying the most significant changes between 4- and 7-day-old cultures: digestible pectic patterns that lead to GalA and GalA2 production significantly increased within the interval (Figure 5A; Supplementary Figure S4). Concomitantly, a reduction in digestible pectic patterns leading to larger methyl-acetylated OGs (DPx>4MeyAcz) also occurred (Figure 5A; Supplementary Figure S4). Enzymes mediating *in muro* de-methyl-acetylation likely underpin these processes. Saponification, conversely, made apparent that the total HG content (from which GalA3 is produced after saponification) decreased with time (Figure 5B).

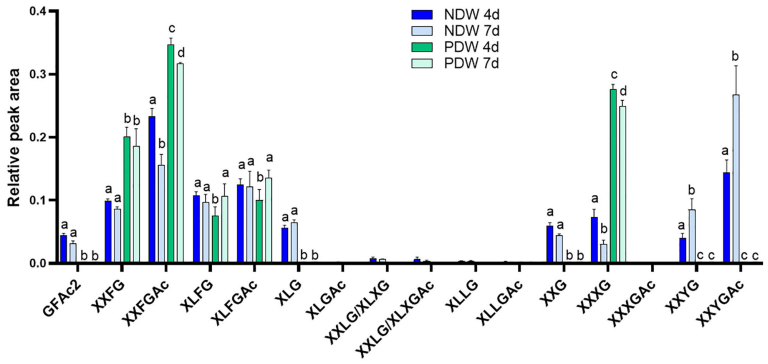


FIGURE 8
Enzymatic fingerprinting of xyloglucans in NDW and PDW fractions. Histograms display the relative peak area of each xyloglucan oligosaccharides released upon endo-cellulase digestion of NDW and PDW fractions obtained from 4- and 7-day-old *A. thaliana* cell cultures. Data are the mean of three biological replicates, error bars are the standard deviation. Lowercase letters indicate significant differences in the amounts of specific xyloglucans between fractions (two-way ANOVA, $p < 0.05$). Oligosaccharide nomenclature follows that described in Fry et al. (1993): G stands for glucose; X for xylose-glucose; L for galactose-xylose-glucose; F for fucose-galactose-xylose-glucose; Y for galacturonic acid-xylose-glucose; Ac: acetyl ester group.

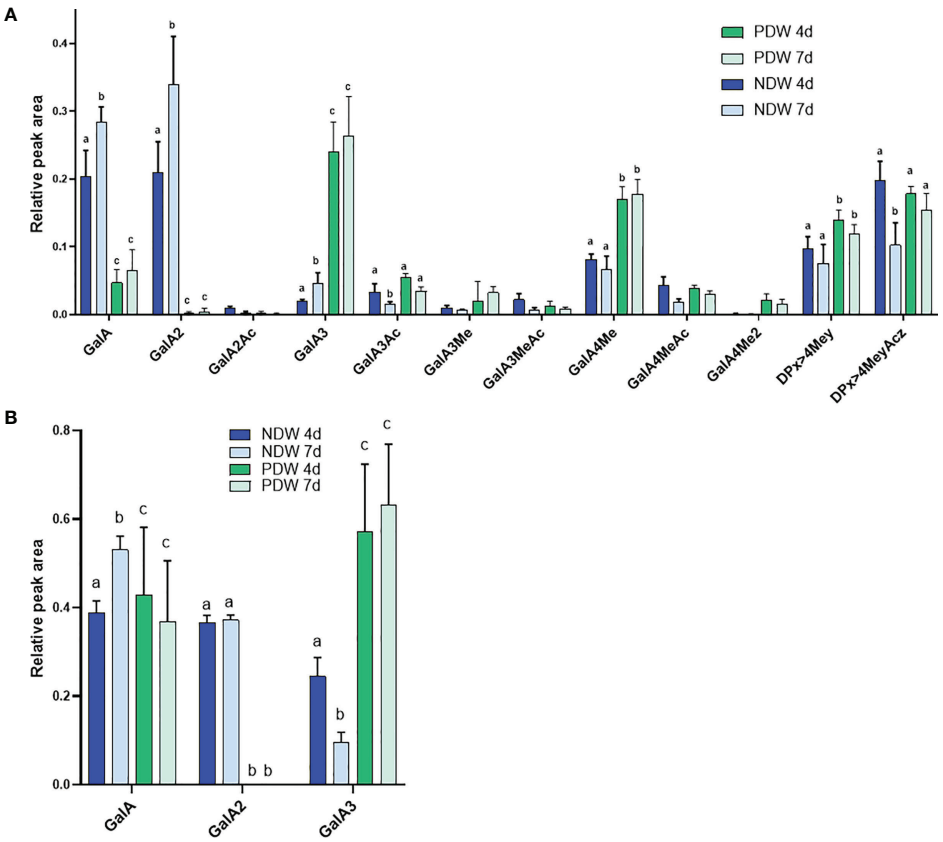


FIGURE 9
Enzymatic fingerprinting of HG in NDW and PDW fractions. Histograms display the peak area of each OGs analysed by HPSEC-HRMS and released upon endo-polygalacturonase digestion of non-saponified **(A)** and saponified **(B)** NDW and DIW fractions obtained from 4- and 7-day-old *A thaliana* cell cultures. Data are the mean of three biological replicates, error bars are the standard deviation. Lowercase letters indicate significant differences in the amounts of specific OGs between fractions (two-way ANOVA, $p < 0.05$). OGs are named GalAxMeyAcz. Numbers indicate the degree of polymerization and the number of methyl/acetyl ester groups. GalA, galacturonic acid; Me, methyl-ester group; Ac, acetyl-ester group; DP, Degree of polymerisation. OGs larger than GalA4 are grouped together. z and y represent potential higher levels of acetylation/methylation in those longer OGs.

The PDW pectins from 4-day-old cultures were already highly de-methyl-esterified to start with. HG fingerprinting shows stable contents of the various OG classes between 4 and 7 days both before (Figure 9A) and after (Figure 9B) saponification. The only exception is GalA3Ac production which decreased in the time interval (Supplementary Figure S7).

Overall, compared to the broad cell wall, the PDW fraction seems to display more limited remodelling within the time window assessed.

Discussion

In the last two decades, PD purification approaches have significantly increased our knowledge of lipids (Grison et al., 2015; Liu et al., 2020) and proteins (Fernandez-Calvino et al., 2011; Brault et al., 2019) closely associated with these structures. While most of the data has been obtained in cell culture systems, it effectively translates to intact plants (Brault et al., 2019). The third molecular component of PD, that of wall polysaccharides, had not yet benefited from the purification approach.

Here we provide a proof-of-concept attempt to address this aspect. We show that wall polysaccharides are co-purified in PD fractions (Figures 2, 7–10) and that they can be effectively analysed with a combination of biochemical strategies (Figure 3). The polysaccharides in the PDW fractions most likely derive from the cell wall in the proximity of PD. Some of the observed signatures have indeed been independently reported by *in-situ* studies in the literature. However, we acknowledge the difficulty of inferring spatial proximity and *in-muro* composition from digestion products alone. Similar limitations might also apply to the absolute purity of the fractions: some PD (and associated polysaccharides) might not be released by the "Onozuka" R10 cellulase mixture and remain

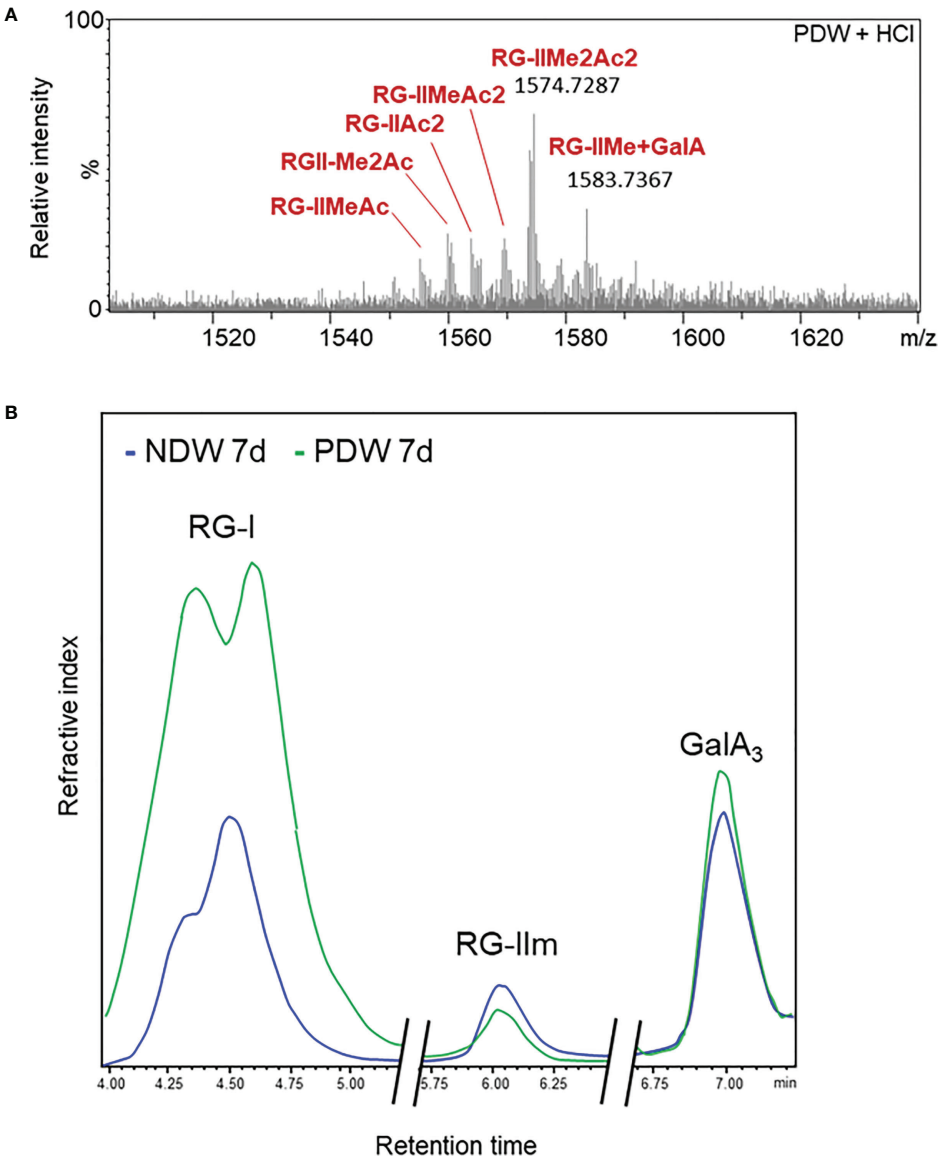


FIGURE 10
Rhamnogalacturonan characterisation of PDW. **(A)** Mass spectrum profile of complete monomerised RG-II forms in PDW fractions derived from 7-day-old *A.thaliana* cell cultures. Me and Ac denote methylation and acetylation in RG-II species. GalA; galacturonic acid. m/z: mass-to-charge ratio. **(B)** HPSEC-RI quantification of RG-I, RG-II monomers (RGIIm) and GalA3 in NDW and PDW fractions derived from 7-day-old *A.thaliana* cell cultures. The x axis is broken.

in DIW fractions. That said, their specific signatures would be highly diluted.

In this manuscript we focused in particular on xyloglucans and pectin components, which together account for around 60% of the plant cell wall (Zablackis et al., 1995). These components are also suitable for enzymatic fingerprinting, an approach with increased descriptive power and quantitative precision (Jermendi et al., 2022). These aspects make fingerprinting an ideal approach to study wall microenvironments. Polysaccharide microdomains - with the exception of xylem pits (Wang et al., 2022) - remain poorly defined to date.

In PDW fractions the failed detection of xylose might indicate a lower abundance of xyloglucans and xylans, sources of this monosaccharide (Figure 7). Immunolabeling studies in different plant species similarly reported that PD positions appeared to have low amounts of xyloglucans (Vaughn et al., 1996; Sutherland et al., 1999). Here we provide details on the composition of the xyloglucan species that are present, which

happen to be highly fucosylated in the PDW fraction (Figure 8). Xyloglucans participate in interactions with cellulose (Park and Cosgrove, 2015) and fucosylated species seem to play a specific role in this interaction (Levy et al., 1991). The loss of galacturonic acid containing xyloglucans upon "Onozuka" R10 cellulase treatment, is also compatible with this notion (Figure 4, NDW versus DIW). Interestingly, the PDW fraction, despite experiencing the same cellulase treatment, retained fucosylated xyloglucan profiles. These forms might remain unaffected (or protected) at PD because they do not interact with cellulose (which is depleted around PD - Faulkner et al., 2008) or because they engage with other local wall polysaccharides. This is a relevant area of future work that could be addressed in a manner similar to Abou-Saleh et al., 2018. Differential abundance of fucosylated species might be ultimately biologically relevant. Accumulation of xylans/non-fucosylated xyloglucans at PD seemed to occur in compatible plant-pathogen interactions (Otulak-Kozieł et al., 2018). Lastly, a more diverse range of xyloglucan patterns was observed in the cell culture fractions compared to the cell wall of specialised tissues in plants. For example, in our NDW fractions we identified acidic forms containing galacturonic acid, xylose, and glucose (YXXG/YXXGAc), previously only reported in root hair cell walls (Peña et al., 2012) (Figure 4).

Enrichments in low-esterified HGs at PD have been described in multiple species using antibodies (Roy et al., 1997; Orfila and Knox, 2000; Faulkner et al., 2008; Giannoutsou et al., 2013). The high levels of GalA3 we quantified in PDW fractions (Figure 9A; Supplementary Figure S7) are in good agreement with this. Based on the comprehensive list of OGs released upon polygalacturonase action, we also show that methylation is present in short stretches (Supplementary Figure S7). This feature could be important to provide wall flexibility in accommodating PD aperture changes. Altering the level of pectin methylation at PD - *via* PME or PME inhibitor proteins - indeed influences plant-virus interactions (Lionetti et al., 2014b). In addition, highly de-methyl esterified HG could be indicative of a high pH in the PD wall environment. PMEs are indeed processive under those conditions (Hocq et al., 2021 - bioRxiv). A member of another group of pH dependent cell-wall modifying proteins, that of expansins, has also been localised to PD (Park et al., 2017).

The persisting lack of GalA2 release in the saponified PDW fractions (Figure 5B) could be linked to the existence of short HG stretches inside an RG-I backbone (Rha-GalA-GalA-GalA-GalA-Rha) (Yapo et al., 2007). The absence of such a domain in PDW would avoid the polygalacturonase to hydrolyze RG-I in smaller subdomains. Consistent with this assumption, we observed two populations of RG-I and PDW fractions are relatively enriched in the larger one (Figure 10B). The lower levels of arabinose and galactose monosaccharides in PDW, compared to the other fractions (Figure 7), do not mirror the reported abundance of arabinan-containing pectins and depletion of galactan-containing ones (Roy et al., 1997; Orfila and Knox, 2000; Faulkner et al., 2008). Gal/Ara ratios are actually roughly equal across all fractions (Figure 7; Supplementary Figure S3). Differences between tissues, species and the non-pectin-specific nature of the monosaccharide analysis could all be plausible explanations for this. Interestingly, RG-I has been shown to be an important determinant of cell-cell adhesion in wood cell walls (Yang et al., 2020). As PD might be viewed as mechanical weakness points in the otherwise solid cell wall, it is tantalising to speculate that the RG-I there might also carry similar functional roles.

The discrepancy between NDW-DIW and PDW composition in terms of OGs (Figures 5, 9) and galacturonic-acid-containing xyloglucans (e.g. YXXG) (Figures 4, 8) indicates that a fraction of HG and xyloglucans might have been directly digested by enzymatic activities in the mixture. Loss of HG could support our initial speculation that these polysaccharides act as the attachment sites that retain PD in the wall (Figure 1). Overall, in assessing the results presented in this paper, it is important to keep in mind that lack of detection of specific polysaccharides in DIW or PDW fractions does not imply absolute absence *in muro*. In the future, PD might be released from the wall with the use of more specific enzymes matching the local polysaccharide composition rather than with the "Onozuka" R10 cellulase treatment. We hope that the data in this manuscript will help the field move in such a direction.

The highly methyl-acetylated RG-II species detected at PD could suggest a non-random distribution of this polysaccharide in the plant cell wall (Figure 10A). This would be a novel concept. The lower extraction efficiency could also indicate reduced RG-II presence (in proportion to the small amount of PD material) or reflect tighter binding of these pectins in those fractions, causing inefficient extraction. RG-II have indeed been shown to directly bind GIPCs lipids (Voxeur and Fry, 2014), postulating interactions between the wall and the outer leaflet of the plasma membrane. Relevantly, those same lipid components are known to be enriched at PD (Grison et al., 2015; Liu et al., 2020). RG-II could act as structural elements that anchor PM-Wall and maintain the stability of PD in face of pressures favouring a collapse of PD membranes.

Our data additionally suggest that pectin and xyloglucan signatures linked to PD are present early on and experience limited changes, as the cell wall (and its PD) remodels (Nicolas et al., 2017) (Figures 8, 9; Supplementary Figure S7). This is in contrast with the rest of the wall, which is more dynamic in methylation and acetylation modifications (Figure 5; Supplementary Figure S4). A unique PD wall environment might emerge already during phragmoplast and primary plasmodesmata formation, when numerous vesicles containing wall material are being delivered to the forming wall (Seguí-Simarro et al., 2004). Targeting of content-specific vesicles to determined positions (or exclusion from the same) would be

required. An alternative explanation for the polysaccharide patterns would be that PD resident proteins alter the wall composition *in muro*, shortly after the establishment of primary PD (Hepler, 1982). This has been described in the context of callose (Levy et al., 2007; Vatén et al., 2011). Several pectin modifying proteins have also been detected in the PD proteome (Knox and Benitez-Alfonso, 2014).

Overall, we hope that the PD biology community will test and perfect this strategy to characterise PD wall polysaccharides. Applying it to a range of developmental, environmental and genetic situations will bring further robustness to association claims and might eventually unravel relevant aspects of cell-cell transport regulation and PD structural integrity.

Methods

PDW purification

PDW fractions were obtained from *A.thaliana* suspension cultured cells of the ecotype *Landsberg erecta*. Four- and seven-day-old cells (3 flasks of 200 mL each) were spun at 800 g for 5 min and resuspended with 25 ml of cold wall-preparation buffer (100 mM Tris-HCl, pH 8.0, 100 mM KCl, 10% v/v glycerol, 10 mM ethylenediaminetetraacetic acid, 0.45 M mannitol, and a complete protease inhibitor cocktail). The cells were passed four times through a N_2 disruption bomb: the preparation was equilibrated for 5 min under N_2 at 120 bars and then passed slowly through the release valve. After centrifugation at 400 g at 4°C for 5 min the pelleted walls were ground to a fine powder in liquid N_2, then washed with cold wall-washing buffer (10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10% v/v glycerol, 10 mM ethylenediaminetetraacetic acid) by sequential centrifugations at 400 g for 5 min. The wall fraction obtained was named NDW.

These purified cell walls were then digested with 0.7% w/v cellulase "Onozuka" R10 (Karlan) in digestion buffer (10 mM MES, pH 5.5, and 4.4% mannitol) containing 1 mM phenylmethylsulfonyl fluoride and complete protease inhibitor cocktail for 1h30 h at 37°C with 200 rpm shaking. After centrifugation (5850 g) for 5 min at 4°C, the supernatant and pellet fractions were collected separately. The supernatant, which contain PD-structures, was process through ultracentrifugation at 110,000 g for 40 min at 4°C. The resulting pellet, called PDW, was washed in an excess volume of cold Tris-buffered saline (TBS; 20 mM Tris-HCl, 0.14 M NaCl, and 2.5 mM KCl, pH 7.4), spinned down again at at 110,000 g for 40 min at 4°C. Finally, the pellet was resuspended in cold 1X TBS containing protease inhibitors. Approximately 600 mL of cultured cells was used to obtain 200 µg protein-equivalent of PDW. Protein amount was determined with a bicinchoninic acid protein assay using BSA as standard.

TEM microscopy

For immunogold labelling of the PDW fraction, 5µl of purified membranes, at a protein concentration of 0.1 mg/mL, were pipetted onto plastic- and carbon-coated grids. Excess liquid was removed, and the grids were incubated with 10 mL of PBS blocking buffer (0.15 M NaCl, 7.5 mM Na_2HPO_4, and 0.25 mM NaH_2PO_4 containing 5% bovine serum albumin, 5% normal goat serum, and 0.1% cold water fish skin) for 1 h and then washed with 1X PBS containing 0.05% Tween 20, before being incubated for 90 min with the primary antibodies diluted in PBS 0.1% acetylated BSA. Dilutions of 1:100, 1:30, and 1:30 were used for PDCB1 (Simpson et al., 2009), PDLP1 (Thomas et al., 2008), and callose (Biosupplies monoclonal antibody) antisera. Controls were performed with pre-immune and second antibodies only. After six washes (5 min each with PBS 0.05% Tween 20) antibody binding was detected using 5 nm and 10 nm gold-conjugated goat anti-rabbit and anti-mouse antibodies diluted at 1:30. After 1 h of incubation at room temperature, the grids were washed six times with PBS 0.05% Tween 20 and three times with 0.2 µm filtered water and then negatively stained with 2% (w/v) uranyl acetate. PDW and DIW observations were carried out on a FEI TECNAI Spirit 120 kV electron microscope.

Cell wall monosaccharide composition

NDW, DIW and PDW fractions were submerged in 96% (v/v) ethanol and boiled at 70°C for 10 min. The pellets were collected by centrifugation (13000 g for 10 min) and dried in a speed vacuum concentrator at 30°C overnight. The monosaccharide compositions of the non-cellulosic fractions were determined by hydrolysis of 100 µg AIR with 2 M trifluoroacetic acid for 1 h at 120°C. After cooling and centrifugation, the supernatant was dried under a vacuum, resuspended in 200 µL of water and analysed by high-performance anion-exchange chromatography/pulsed amperometric detection on a Dionex ICS5000 instrument (ThermoFisher Scientific) as described in Fang et al., 2016.

Enzymatic fingerprinting of pectins and xyloglucans

The PDW, DIW and NDW fractions were dried in a speed vacuum concentrator at 30°C overnight. Samples were digested with 1 U/mg DW sample of *Aspergillus aculeatus* endo-polygalacturonase M2 or endo-cellulase (Megazyme, Bray, Ireland) (Lerouxel et al., 2002) in 50 mM ammonium acetate buffer (pH 5) at 37°C for 18 h. Samples were then centrifuged at 13000 rpm for 10 min and 100 µL of the supernatants were transferred into vials. For MS analysis, 10 µl of each fraction were injected in the machine.

Oligosaccharide characterization and quantification by LC/HRMS analysis

The oligosaccharides released from digestion were separated according to Voxeur et al., 2019. Chromatographic separation was performed on an ACQUITY UPLC Protein BEH SEC Column (125A, 1.7 µm, 4.6 mm X 300 mm, Waters Corporation, Milford, MA, USA) coupled with guard Column BEH SEC Column (125A, 1.7 µm, 4.6 mm X 30 mm). Elution was performed in 50 mM ammonium formate, 0.1% formic acid at a flow rate of 0.4 mL/min and with a column oven temperature of 40°C. The injection volume was set to 10 µl. The quantitative evaluation of xyloglucan fragments was made using an HPLC system (UltiMate 3000 RS HPLC system, Thermo Scientific, Waltham, MA, USA). The system was coupled to an Impact II Ultra-High Resolution Qq-Time-Of-Flight (UHR-QqTOF) spectrometer (Bruker Daltonics, Bremen, Germany) equipped with an electrospray ionisation (ESI) source in negative mode with the end plate offset set voltage to 500 V, capillary voltage to 4000 V, nebulizer to 40 psi, dry gas to 8 l/min and dry temperature of 180°C. The Compass 1.8 software (Bruker Daltonics) was used to acquire the data.

The Mzmine 2.53 software was used to analyse data according to Pluskal et al., 2010. To perform integration, the filter noise level for mass detection was set to 500. The ADAP Chromatogram Builder (Myers et al., 2017) was used with the following parameters: range of 6 - 10 min; min group size of scan 10; group intensity threshold of 1000; min highest peak of 500; and m/z tolerance of 0.01 or 10 ppm. The chromatogram peaks were deconvoluted using aBaseline cut-off of 300. Chromatograms were manually de-isotoped and peaks aligned (m/z tolerance of 0.01 or 5 ppm; retention time tolerance of 0.1).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

EMB and AV designed the experiments. AP wrote the manuscript with input from all other authors. AP, FI, MSG extracted wall and PD fractions from cell cultures. EMB performed TEM microscopy. JS and GM performed monosaccharide analysis. SP performed mass spectrometry analysis. AV performed all other analyses. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No 772103-BRIDGING) to E.M.B; the EMBO Young Investigator Program to E.M.B; the AAP INNOVATION et ENTREPRENARIAT- PREMATURATION 2018 to A.V (Grant agreement No CDE-2018-002330-IRE 2018-0024, OGome). This work also benefited from IJPB's Plant Observatory technological platforms, which are supported by Saclay Plant Sciences-SPS (ANR-17-EUR-0007).

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1020506/full#supplementary-material

References

Abou-Saleh, R. H., Hernandez-Gomez, M. C., Amsbury, S., Paniagua, C., Bourdon, M., Miyashima, S., et al. (2018). Interactions between callose and cellulose revealed through the analysis of biopolymer mixtures. *Nat. Commun.* 9, 1–13. doi: 10.1038/s41467-018-06820-y

Amsbury, S., Kirk, P., and Benitez-Alfonso, Y. (2018). Emerging models on the regulation of intercellular transport by plasmodesmata-associated callose. *J. Exp. Bot.* 69, 105–115. doi: 10.1093/jxb/erx337

Barral, P., Suárez, C., Batanero, E., Alfonso, C., Alché, J. D. D., Rodríguez-García, M. I., et al. (2005). An olive pollen protein with allergenic activity, ole e 10, defines a novel family of carbohydrate-binding modules and is potentially implicated in pollen germination. *Biochem. J.* 390, 77–84. doi: 10.1042/BJ20050456

Bayer, E. M., Bottrill, A. R., Walshaw, J., Vigouroux, M., Naldrett, M. J., Thomas, C. L., et al. (2006). Arabidopsis cell wall proteome defined using multidimensional protein identification technology. *Proteomics* 6, 301–311. doi: 10.1002/pmic.200500046

Bayer, E., Thomas, C. L., and Maule, A. J. (2004). Plasmodesmata in *Arabidopsis thaliana* suspension cells. *Protoplasma* 223, 93–102. doi: 10.1007/s00709-004-0044-8

Beldman, G., Searle-Van Leeuwen, M. F., Rombouts, F. M., and Voragen, F. G. (1985). The cellulase of trichoderma viride: Purification, characterization and comparison of all detectable endoglucanases, exoglucanases and β-glucosidases. *Eur. J. Biochem.* 146, 301–308. doi: 10.1111/j.1432-1033.1985.tb08653.x

Brault, M. L., Petit, J. D., Immel, F., Nicolas, W. J., Glavier, M., Brocard, L., et al. (2019). Multiple C2 domains and transmembrane region proteins (MCTP s) tether membranes at plasmodesmata. *EMBO Rep.* 20, e47182. doi: 10.15252/embr.201847182

Burch-Smith, T. M., Brunkard, J. O., Choi, Y. G., and Zambryski, P. C. (2011). Organelle–nucleus cross-talk regulates plant intercellular communication *via* plasmodesmata. *Proc. Natl. Acad. Sci.* 108, E1451–E1460. doi: 10.1073/pnas.1117226108

Casero, P. J., and Knox, J. P. (1995). The monoclonal antibody JIM5 indicates patterns of pectin deposition in relation to pit fields at the plasma-membrane-face of tomato pericarp cell walls. *Protoplasma* 188, 133–137. doi: 10.1007/BF01276804

Cavalier, D., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., et al. (2008). Disrupting two arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell* 20, 1519–1537. doi: 10.1105/tpc.108.059873

Chen, M. H., and Citovsky, V. (2003). Systemic movement of a tobamovirus requires host cell pectin methylesterase. *Plant J.* 35, 386–392. doi: 10.1046/j.1365-313x.2003.01818.x

Chen, M. H., Sheng, J., Hind, G., Handa, A. K., and Citovsky, V. (2000). Interaction between the tobacco mosaic virus movement protein and host cell pectin methylesterases is required for viral cell-to-cell movement. *EMBO J.* 19, 913–920. doi: 10.1093/emboj/19.5.913

Chen, J., Wang, L., Su, X., Wang, K., Wu, X., Chen, L., et al. (2016). Structure, morphology, thermostability and irradiation-mediated degradation fractions of hemicellulose treated with γ-irradiation. *Waste Biomass valorization* 7, 1415–1425. doi: 10.1007/s12649-016-9489-1

Cifuentes, C., Bulone, V., and Emons, A. M. C. (2010). Biosynthesis of callose and cellulose by detergent extracts of tobacco cell membranes and quantification of the polymers synthesized *in vitro*. *J. Integr. Plant Biol.* 52, 221–233. doi: 10.1111/j.1744-7909.2010.00919.x

Ding, S. Y., Zhao, S., and Zeng, Y. (2014). Size, shape, and arrangement of native cellulose fibrils in maize cell walls. *Cellulose* 21 (2), 863–871. doi: 10.1007/s10570-013-0147-5

Dorokhov, Y. L., Mäkinen, K., Frolova, O. Y., Merits, A., Saarinen, J., Kalkkinen, N., et al. (1999). A novel function for a ubiquitous plant enzyme pectin methylesterase: the host-cell receptor for the tobacco mosaic virus movement protein. *FEBS Lett.* 461, 223–228. doi: 10.1016/s0014-5793(99)01447-7

Ehlers, K., and Kollmann, R. (2001). Primary and secondary plasmodesmata: structure, origin, and functioning. *Protoplasma* 216, 1–30. doi: 10.1007/BF02680127

Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010). Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol.* 153, 403–419. doi: 10.1104/pp.110.156000

Fang, L., Ishikawa, T., Rennie, E. A., Murawska, G. M., Lao, J., Yan, J., et al. (2016). Loss of inositol phosphorylceramide sphingolipid mannosylation induces plant immune responses and reduces cellulose content in arabidopsis. *Plant Cell* 28, 2991–3004. doi: 10.1105/tpc.16.00186

Faulkner, C., Akman, O. E., Bell, K., Jeffree, C., and Oparka, K. (2008). Peeking into pit fields: a multiple twinning model of secondary plasmodesmata formation in tobacco. *Plant Cell* 20, 1504–1518. doi: 10.1105/tpc.107.056903

Faulkner, C., and Bayer, E. M. (2017). "Isolation of plasmodesmata," in *Isolation of plant organelles and structures* (New York, NY: Humana Press), 187–198.

Fernandez-Calvino, L., Faulkner, C., Walshaw, J., Saalbach, G., Bayer, E., Benitez-Alfonso, Y., et al. (2011). Arabidopsis plasmodesmal proteome. *PloS One* 6, e18880. doi: 10.1371/journal.pone.0018880

Fry, S. C., York, W. S., Albersheim, P., Darvill, A., Hayashi, T., Joseleau, J. P., et al. (1993). An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiol Plantarum* 89, 1–3. doi: 10.1111/j.1399-3054.1993.tb01778.x

Giannoutsou, E., Sotiriou, P., Apostolakos, P., and Galatis, B. (2013). Early local differentiation of the cell wall matrix defines the contact sites in lobed mesophyll cells of zea mays. *Ann. Bot.* 112, 1067–1081. doi: 10.1093/aob/mct175

Grison, M. S., Brocard, L., Fouillen, L., Nicolas, W., Wewer, V., Dörmann, P., et al. (2015). Specific membrane lipid composition is important for plasmodesmata function in arabidopsis. *Plant Cell* 27, 1228–1250. doi: 10.1105/tpc.114.135731

Gronnier, J., Gerbeau-Pissot, P., Germain, V., Mongrand, S., and Simon-Plas, F. (2018). Divide and rule: plant plasma membrane organization. *Trends Plant Sci.* 2, 899–917. doi: 10.1016/j.tplants.2018.07.007

Hepler, P. K. (1982). Endoplasmic reticulum in the formation of the cell plate and plasmodesmata. *Protoplasma* 111 (2), 121–133. doi: 10.1007/BF01282070

Hocq, L., Habrylo, O., Voxeur, A., Pau-Roblot, C., Safran, J., Sénéchal, F., et al. (2021). Arabidopsis AtPME2 has a pH-dependent processivity and control cell wall mechanical properties bioRxiv. doi: 10.1101/2021.03.03.433777

Ishii, T., Matsunaga, T., Pellerin, P., O'Neill, M. A., Darvill, A., and Albersheim, P. (1999). The plant cell wall polysaccharide rhamnogalacturonan II self-assembles into a covalently cross-linked dimer. *J. Biol. Chem.* 274, 13098–13104. doi: 10.1074/jbc.274.19.13098

Jermendi, É., Beukema, M., van den Berg, M. A., de Vos, P., and Schols, H. A. (2022). Revealing methyl-esterification patterns of pectins by enzymatic fingerprinting: Beyond the degree of blockiness. *Carbohydr. Polymers* 277, 118813. doi: 10.1016/j.carbpol.2021.118813

Jobert, F., Guénin, S., Voxeur, A., Lee, K. J., Bouton, S., Sénéchal, F., et al. (2022). Pectin remodeling belongs to a homeostatic system and triggers transcriptomic and hormonal modulations. *bioRxiv*. doi: 10.1101/2021.07.22.453319

Kerstens, S., Decraemer, W. F., and Verbelen, J. P. (2001). Cell walls at the plant surface behave mechanically like fiber-reinforced composite materials. *Plant Physiol.* 127, 381–385. doi: 10.1104/pp.010423

Kirk, P., Amsbury, S., German, L., Gaudioso-Pedraza, R., and Benitez-Alfonso, Y. (2022). A comparative meta-proteomic pipeline for the identification of plasmodesmata proteins and regulatory conditions in diverse plant species. *BMC Biol.* 20, 1–21. doi: 10.1186/s12915-022-01331-1

Knox, J. P., and Benitez-Alfonso, Y. (2014). Roles and regulation of plant cell walls surrounding plasmodesmata. *Curr. Opin. Plant Biol.* 22, 93–100. doi: 10.1016/j.pbi.2014.09.009

Lampugnani, E. R., Khan, G. A., Somssich, M., and Persson, S. (2018). Building a plant cell wall at a glance. *J. Cell Sci.* 131, jcs207373. doi: 10.1242/jcs.207373

Lerouxel, O., Choo, T. S., Séveno, M., Usadel, B., Faye, L., Lerouge, P., et al. (2002). Rapid structural phenotyping of plant cell wall mutants by enzymatic oligosaccharide fingerprinting. *Plant Physiol.* 130, 1754–1763. doi: 10.1104/pp.011965

Levy, A., Erlanger, M., Rosenthal, M., and Epel, B. L. (2007). A plasmodesmata-associated β-1, 3-glucanase in arabidopsis. *Plant J.* 49, 669–682. doi: 10.1111/j.1365-313X.2006.02986.x

Levy, S., York, W. S., Stuike-Prill, R., Meyer, B., and Staehelin, L. A. (1991). Simulations of the static and dynamic molecular conformations of xyloglucan. the role of the fucosylated sidechain in surface-specific sidechain folding. *Plant J.* 1, 195–215. doi: 10.1111/j.1365-313X.1991.00195.x

Liesche, J., Ziomkiewicz, I., and Schulz, A. (2013). Super-resolution imaging with pontamine fast scarlet 4BS enables direct visualization of cellulose orientation and cell connection architecture in onion epidermis cells. *BMC Plant Biol.* 13, 226. doi: 10.1186/1471-2229-13-226

Limberg, G., Korner, R., Buchholt, H. C., Christensen, T. M., Roepstorff, P., and Mikkelsen, J. D. (2000). Analysis of different de-esterification mechanisms for pectin by enzymatic fingerprinting using endopectin lyase and endopolygalacturonase II from aspergillus niger. *Carbohydr. Res.* 327, 293–307. doi: 10.1016/s0008-6215(00)00067-7

Lionetti, V., Raiola, A., Cervone, F. , and Bellincampi, D. (2014a). Transgenic expression of pectin methylesterase inhibitors limits tobamovirus spread in tobacco and arabidopsis. *Mol. Plant Pathol.* 15, 265–274. doi: 10.1111/mpp.12090

Lionetti, V., Raiola, A., Cervone, F. , and Bellincampi, D. (2014b). How do pectin methylesterases and their inhibitors affect the spreading of tobamovirus? *Plant Signaling Behav.* 9, e972863. doi: 10.4161/15592316.2014.972863

Li, Z. P., Paterlini, A., Glavier, M., and Bayer, E. M. (2020). Intercellular trafficking *via* plasmodesmata: molecular layers of complexity. *Cell. Mol. Life Sci.* 78, 799–816. doi: 10.1007/s00018-020-03622-8

Liu, N. J., Zhang, T., Liu, Z. H., Chen, X., Guo, H. S., Ju, B. H., et al. (2020). Phytosphinganine affects plasmodesmata permeability *via* facilitating PDLP5-stimulated callose accumulation in arabidopsis. *Mol. Plant* 13, 128–143. doi: 10.1016/j.molp.2019.10.013

Morvan, O., Quentin, M., Jauneau, A., Mareck, A., and Morvan, C. (1998). Immunogold localization of pectin methylesterases in the cortical tissues of flax hypocotyl. *Protoplasma* 202, 175–184. doi: 10.1007/BF01282545

Myers, O. D., Sumner, S. J., Li, S., Barnes, S., and Du, X. (2017). One step forward for reducing false positive and false negative compound identifications from mass spectrometry metabolomics data: new algorithms for constructing extracted ion chromatograms and detecting chromatographic peaks. *Anal Chem.* 89 (17), 8696–8703. doi: 10.1021/acs.analchem.7b00947

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A. S., Basle, A., Gray, J., et al. (2017). Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* 544, 65–70. doi: 10.1038/nature21725

Nicolas, W. J., Grison, M. S., Trépout, S., Gaston, A., Fouché, M., Cordelières, F. P., et al. (2017). Architecture and permeability of post-cytokinesis plasmodesmata lacking cytoplasmic sleeves. *Nat. Plants* 3, 17082. doi: 10.1038/nplants.2017.82

Northcote, D. H., Davey, R., and Lay, J. (1989). Use of antisera to localize callose, xylan and arabinogalactan in the cell-plate, primary and secondary walls of plant cells. *Planta* 178, 353–366. doi: 10.1007/BF00391863

O'Neill, M. A., Eberhard, S., Albersheim, P., and Darvill, A. G. (2001). Requirement of borate cross-linking of cell wall rhamnogalacturonan II for arabidopsis growth. *Science* 294, 846–849. doi: 10.1126/science.1062

O'Neill, M. A., Ishii, T., Albersheim, P., and Darvill, A. G. (2004). Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annu. Rev. Plant Biol.* 55, 109. doi: 10.1146/annurev.arplant.55.031903.141750

O'Neill, M. A., Warrenfeltz, D., Kates, K., Pellerin, P., Doco, T., Darvill, A. G., et al. (1996). Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester. *In vitro* conditions for the formation and hydrolysis of the dimer. *J. Biol. Chem.* 271, 22923–22930. doi: 10.1074/jbc.271.37.22923

O'Neill, M. A., Black, I., Urbanowicz, B., Bharadwaj, V., Crowley, M., Koj, S., et al. (2020). Locating methyl-etherified and methyl-esterified uronic acids in the plant cell wall pectic polysaccharide rhamnogalacturonan II. *SLAS TECHNOL: Translating Life Sci. Innovation* 25, 329–344. doi: 10.1177/2472630320923321

Orfila, C., and Knox, J. P. (2000). Spatial regulation of pectic polysaccharides in relation to pit fields in cell walls of tomato fruit pericarp. *Plant Physiol.* 122, 775–782. doi: 10.1104/pp.122.3.775

Otulak-Kozieł, K., Kozieł, E., and Bujarski, J. J. (2018). Spatiotemporal changes in xylan-1/xyloglucan and xyloglucan xyloglucosyl transferase (Xth-Xet5) as a step-in of ultrastructural cell wall remodelling in potato–potato virus y (PVYntn) hypersensitive and susceptible reaction. *Int. J. Mol. Sci.* 19, 2287. doi: 10.3390/ijms19082287

Park, Y. B., and Cosgrove, D. J. (2015). Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol.* 56, 180–194. doi: 10.1093/pcp/pcu204

Park, S. H., Li, F., Renaud, J., Shen, W., Li, Y., Guo, L., et al. (2017). NbEXPA1, an α-expansin, is plasmodesmata-specific and a novel host factor for potyviral infection. *Plant J.* 92, 846–861. doi: 10.1111/tpj.13723

Pasternak, T., Paponov, I. A., and Kondratenko, S. (2021). Optimizing protocols for arabidopsis shoot and root protoplast cultivation. *Plants* 10, 375. doi: 10.3390/plants10020375

Pedrolli, D. B., Monteiro, A. C., Gomes, E., and Carmona, E. C. (2009). Pectin and pectinases: Production, characterization and industrial application of microbial pectinolytic enzymes. *Open Biotechnol. J.* 3, 9–18. doi: 10.2174/1874070700903010009

Peña, M. J., Kong, Y., York, W. S., and O'Neill, M. A. (2012). A galacturonic acid–containing xyloglucan is involved in arabidopsis root hair tip growth. *Plant Cell* 24, 4511–4524. doi: 10.1105/tpc.112.103390

Pluskal, T., Castillo, S., Villar-Briones, A., and Orešič, M. (2010). MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinf.* 11, 1–11. doi: 10.1186/1471-2105-11-395

Roy, S., Watada, A. E., and Wergin, W. P. (1997). Characterization of the cell wall microdomain surrounding plasmodesmata in apple fruit. *Plant Physiol.* 114, 539–547. doi: 10.1104/pp.114.2.539

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

Scorrano, L., De Matteis, M. A., Emr, S., Giordano, F., Hajnóczky, G., Kornmann, B., et al. (2019). Coming together to define membrane contact sites. *Nat. Commun.* 10, 1–11. doi: 10.1038/s41467-019-09253-3

Seguí-Simarro, J. M., Austin, J. R., White, E. A., and Staehelin, L. A. (2004). Electron tomographic analysis of somatic cell plate formation in meristematic cells of arabidopsis preserved by high-pressure freezing. *Plant Cell* 16, 836–856. doi: 10.1105/tpc.017749

Séveno, M., Voxeur, A., Rihouey, C., Wu, A. M., Ishii, T., Chevalier, C., et al. (2009). Structural characterisation of the pectic polysaccharide rhamnogalacturonan II using an acidic fingerprinting methodology. *Planta* 230, 947–957. doi: 10.1007/s00425-009-0996-1

Simpson, C., Thomas, C., Findlay, K., Bayer, E., and Maule, A. J. (2009). An arabidopsis GPI-anchor plasmodesmal neck protein with callose binding activity and potential to regulate cell-to-cell trafficking. *Plant Cell* 21, 581–594. doi: 10.1105/tpc.108.060145

Sun, L., Ropartz, D., Cui, L., Shi, H., Ralet, M. C., and Zhou, Y. (2019). Structural characterization of rhamnogalacturonan domains from panax ginseng CA Meyer. *Carbohydr. polymers* 203, 119–127. doi: 10.1016/j.carbpol.2018.09.045

Sutherland, P., Hallett, I., Redgwell, R., Benhamou, N., and MacRae, E. (1999). Localization of cell wall polysaccharides during kiwifruit (*Actinidia deliciosa*) ripening. *Int. J. Plant Sci.* 160, 1099–1109. doi: 10.1086/314196

Thomas, C. L., Bayer, E. M., Ritzenthaler, C., Fernandez-Calvino, L., and Maule, A. J. (2008). Specific targeting of a plasmodesmal protein affecting cell-to-cell communication. *PloS Biol.* 6, e7. doi: 10.1371/journal.pbio.0060007

Turner, A., Wells, B., and Roberts, K. (1994). Plasmodesmata of maize root tips: structure and composition. *J. Cell Sci.* 107, 3351–3361. doi: 10.1242/jcs.107.12.3351

Vaattovaara, A., Brandt, B., Rajaraman, S., Safronov, O., Veidenberg, A., Luklová, M., et al. (2019). Mechanistic insights into the evolution of DUF26-containing proteins in land plants. *Commun. Biol.* 2, 1–18. doi: 10.1038/s42003-019-0306-9

Vatén, A., Dettmer, J., Wu, S., Stierhof, Y.-D., Miyashima, S., Yadav, S. R., et al. (2011). Callose biosynthesis regulates symplastic trafficking during root development. *Dev. Cell* 21, 1144–1155. doi: 10.1016/j.devcel.2011.10.006

Vaughn, K. C., Hoffman, J. C., Hahn, M. G., and Staehelin, L. A. (1996). The herbicide dichlobenil disrupts cell plate formation: immunogold characterization. *Protoplasma* 194, 117–132. doi: 10.1007/BF01882020

Voragen, A. G., Coenen, G. J., Verhoef, R. P., and Schols, H. A. (2009). Pectin, a versatile polysaccharide present in plant cell walls. *Struct. Chem.* 20, 263–275. doi: 10.1007/s11224-009-9442-z

Voxeur, A., and Fry, S. C. (2014). Glycosylinositol phosphorylceramides from Rosa cell cultures are boron-bridged in the plasma membrane and form complexes with rhamnogalacturonan II. *Plant J.* 79, 139–149. doi: 10.1111/tpj.12547

Voxeur, A., Habrylo, O., Guénin, S., Miart, F., Soulié, M. C., Rihouey, C., et al. (2019). Oligogalacturonide production upon *Arabidopsis thaliana–botrytis cinerea* interaction. *Proc. Natl. Acad. Sci.* 116 (39), 19743–19752. doi: 10.1073/pnas.1900317116

Wang, H., Yang, H., Wen, Z., Gao, C., Gao, Y., Tian, Y., et al. (2022). Xylan-based nanocompartments orchestrate plant vessel wall patterning. *Nat. Plants* 8, 295–306. doi: 10.1038/s41477-022-01113-1

Wathoni, N., Shan, C. Y., Shan, W. Y., Rostinawati, T., Indradi, R. B., Pratiwi, R., et al. (2019). Characterization and antioxidant activity of pectin from Indonesian mangosteen (*Garcinia mangostana l.*) rind. *Heliyon* 5, e02299. doi: 10.1016/j.heliyon.2019.e02299

Yang, H., Benatti, M. R., Karve, R. A., Fox, A., Meilan, R., Carpita, N. C., et al. (2020). Rhamnogalacturonan-I is a determinant of cell–cell adhesion in poplar wood. *Plant Biotechnol. J.* 18, 1027–1040. doi: 10.1111/pbi.13271

Yapo, B. M., Lerouge, P., Thibault, J. F., and Ralet, M. C. (2007). Pectins from citrus peel cell walls contain homogalacturonans homogenous with respect to molar mass, rhamnogalacturonan I and rhamnogalacturonan II. *Carbohydr. Polymers* 69, 426–435. doi: 10.1016/j.carbpol.2006.12.024

Zablackis, E., Huang, J., Muller, B., Darvill, A. G., and Albersheim, P. (1995). Characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves. *Plant Physiol.* 107, 1129–1138. doi: 10.1104/pp.107.4.1129

Zambryski, P., and Crawford, K. (2000). Plasmodesmata: gatekeepers for cell-to-cell transport of developmental signals in plants. *Annu. Rev. Cell Dev. Biol.* 16, 393–421. doi: 10.1146/annurev.cellbio.16.1.393

frontiers | Frontiers in Plant Science

TYPE Perspective
PUBLISHED 25 November 2022
DOI 10.3389/fpls.2022.1043429

OPEN ACCESS

EDITED BY
Howard Scott Neufeld,
Appalachian State University,
United States

REVIEWED BY
Kim Johnson,
La Trobe University, Australia
Heather Joesting,
Georgia Southern University,
United States

*CORRESPONDENCE
Marc Fradera-Soler
mfs@plen.ku.dk
Olwen M. Grace
o.grace@kew.org

†These authors have contributed equally to this work and share first authorship

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

RECEIVED 13 September 2022
ACCEPTED 31 October 2022
PUBLISHED 25 November 2022

CITATION
Fradera-Soler M, Leverett A, Mravec J, Jørgensen B, Borland AM and Grace OM (2022) Are cell wall traits a component of the succulent syndrome? *Front. Plant Sci.* 13:1043429. doi: 10.3389/fpls.2022.1043429

Are cell wall traits a component of the succulent syndrome?

Marc Fradera-Soler[1,2*†], Alistair Leverett[3,4†], Jozef Mravec[1,5], Bodil Jørgensen[1], Anne M. Borland[3] and Olwen M. Grace[2*]

[1]Department of Plant and Environmental Sciences, University of Copenhagen, Frederiksberg, Denmark, [2]Department of Accelerated Taxonomy, Royal Botanic Gardens, Kew, Richmond, Surrey, United Kingdom, [3]School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, United Kingdom, [4]School of Life Sciences, University of Essex, Colchester, United Kingdom, [5]Institute of Plant Genetics and Biotechnology, Slovak Academy of Sciences, Plant Science and Biodiversity Center, Nitra, Slovakia

Succulence is an adaptation to low water availability characterised by the presence of water-storage tissues that alleviate water stress under low water availability. The succulent syndrome has evolved convergently in over 80 plant families and is associated with anatomical, physiological and biochemical traits. Despite the alleged importance of cell wall traits in drought responses, their significance in the succulent syndrome has long been overlooked. Here, by analyzing published pressure–volume curves, we show that elastic adjustment, whereby plants change cell wall elasticity, is uniquely beneficial to succulents for avoiding turgor loss. In addition, we used comprehensive microarray polymer profiling (CoMPP) to assess the biochemical composition of cell walls in leaves. Across phylogenetically diverse species, we uncover several differences in cell wall biochemistry between succulent and non-succulent leaves, pointing to the existence of a 'succulent glycome'. We also highlight the glycomic diversity among succulent plants, with some glycomic features being restricted to certain succulent lineages. In conclusion, we suggest that cell wall biomechanics and biochemistry should be considered among the characteristic traits that make up the succulent syndrome.

KEYWORDS

succulence, plant diversity, cell walls, cell wall elasticity, CoMPP, glycomics, turgor

Abbreviations: Physiological parameters: ε, Bulk modulus of cell wall elasticity; *P*, Turgor pressure; RWC, Relative water content; TLP_{Ψ}, Turgor loss point, i.e. water potential at which turgor is lost; Ψ, Water potential; π_O, Osmotic potential of tissue at full hydration. Cell wall polymers: AGP, Arabinogalactan protein; HG, Homogalacturonan; RG-I, Rhamnogalacturonan I; DM, Degree of methyl-esterification; DP, Degree of polymerization.

Introduction

Climate change-induced aridity is expected to increase across much of the globe in the future (Sheffield and Wood, 2008; Jiao et al., 2021). Consequently, it has become imperative that we understand the ways in which plants cope with drought (Choat et al., 2018; Trueba et al., 2019). Recently, plant scientists have begun to pay renewed attention to the drought adaptations found in succulent plants (Heyduk et al., 2016; Males, 2017; Fradera-Soler et al., 2021; Leverett et al., 2021). Succulence is defined by the presence of water stores, in the leaf, stem and/or roots, which can be mobilized when a plant is dehydrated (Ogburn and Edwards, 2010). Typically, succulent tissues (i.e. the tissues responsible for water storage) arise due to the development of enlarged cells, either in the photosynthetic tissue (chlorenchyma), in a specialized achlorophyllous water-storage tissue (hydrenchyma), or a combination of the two (Eggli and Nyffeler, 2009; Borland et al., 2018; Heyduk, 2021; Leverett et al., 2022). If water stored in large cells can be mobilized during drought, succulent plants can dehydrate whilst maintaining water potentials (Ψ) at safe, stable levels. By buffering plant Ψ, succulence prevents a number of detrimental processes from occurring, such as the closing of stomata, the buckling of cells and the formation of emboli in the xylem (Brodribb et al., 2016; Vollenweider et al., 2016; Zhang et al., 2016; Henry et al., 2019). The benefits conferred by succulence have resulted in the succulent syndrome being found in plants across the globe, following adaptive radiations into the world's arid and semi-arid ecosystems (Arakaki et al., 2011).

The adaptive benefits of succulence have recently drawn the attention of synthetic biologists, who have begun to recognize the potential this adaptation could have for food security and bioenergy in a drying world under climate change scenarios (Borland et al., 2009; Grace, 2019). Both modelling and field trials have assessed the value of growing succulent *Agave* and *Opuntia* in dry marginal and underused lands (Owen and Griffiths, 2014; Davis et al., 2017; Hartzell et al., 2021; Neupane et al., 2021). Furthermore, progress has been made to synthetically produce succulence in non-succulent species. The introduction of an exogenous transcription factor gene into *Arabidopsis thaliana* led to increased tissue succulence and higher water-use efficiency (Lim et al., 2018; Lim et al., 2020). These findings strongly suggest that bioengineering succulence has the potential to enhance drought resistance in crops. Whilst some work has been done to understand the genetic programs controlling the development of succulence (Heyduk, 2021), a great deal more research is needed if we are to fully utilize this adaptation in agricultural settings. In addition, we must appreciate every important trait that makes up the succulent syndrome. Beyond the genetic control of cell size, succulent species often exhibit a number of other co-adaptive traits, such as 3D vascular patterning, crassulacean acid metabolism (CAM) and waxy cuticles (Griffiths and Males, 2017). Cell walls have recently been postulated as an often-overlooked key component of the succulent syndrome (Ahl et al., 2019; Fradera-Soler et al., 2022), yet the precise mechanistic relevance of cell walls in succulent tissues remains largely speculative. In the present study, we analyse cell wall biomechanics and biochemistry in diverse succulent species and propose that these traits should be considered among the characteristic components of the succulent syndrome.

Cell wall biomechanics in succulents

All plant cells are encased in a lattice-like structure, the cell wall (Popper et al., 2011). Primary, extensible cell walls are complex and dynamic systems composed largely of polysaccharides, polyphenols and certain types of glycoproteins (Carpita et al., 2015). When plant cells are hydrated, an osmotic gradient exists across the plasma membrane which results in water moving into the protoplasm (Beadle et al., 1993). This intake of water causes the plasma membrane to push against the cell wall, generating a positive pressure called turgor (*P*). The bulk modulus of cell wall elasticity (ε) relates to *P* according to the equation:

$$\varepsilon = \frac{\delta P}{\delta RWC} \quad (1)$$

where relative water content (RWC) is the percentage of total water present in a tissue. Higher values of ε indicate greater cell wall rigidity and thus more resistance for the plasma membrane to push against, with changes in RWC resulting in large changes in *P*. Conversely, when ε is low and cell walls are highly elastic, changes to RWC have a lower impact on *P*, because cell walls can stretch and provide less resistance.

For succulent plants, ε has the potential to affect the point at which turgor is lost. As plant tissues dehydrate, Ψ falls, which results in a linear drop in *P* (Beadle et al., 1993). Eventually, Ψ falls to a point where $P = 0$, meaning there has been a total loss of turgor. When this turgor loss point (TLP_Ψ) has been reached, leaves will typically wilt and cells will begin to experience damage (Trueba et al., 2019). Consequently, it is beneficial for plants to avoid reaching their TLP_Ψ (Kunert et al., 2021). Bartlett et al. (2012) found that the TLP_Ψ can be estimated by:

$$TLP_\Psi = \frac{\pi_O \times \varepsilon}{\pi_O + \varepsilon} \quad (2)$$

where π_O is the osmotic potential of fully hydrated tissues (a more negative π_O corresponds to a higher concentration of osmotically active solutes). Modifying ε or π_O are named elastic and osmotic adjustment, respectively, and can be used to alter the TLP_Ψ in order to allow cells to maintain turgor at more negative water potentials. Lower ε could result in cell walls capable of changing shape and folding as the protoplasm within

shrinks (Ahl et al., 2019; Fradera-Soler et al., 2022). This would prevent the catastrophic disruption of the membrane-wall continuum and other forms of irreversible damage due to mechanical stress which occur when the TLP_Ψ is reached. However, studies of non-succulent species have found that ε is generally so high that changes to this trait are inconsequential for the TLP_Ψ (Bartlett et al., 2012). Put differently, in non-succulent species, cell walls are quite rigid, which means that even substantial changes to their elastic properties will not affect their TLP_Ψ. This can be visualized by considering Figure 1A. If π_O is held constant and ε is allowed to vary, the TLP_Ψ can be simulated using Equation 2. This simulation forms a curve, and in non-succulent tissues the true value of ε intersects at the flat portion of the curve. Consequently, the phenotypic space inhabited by non-succulent species is one where changes to ε have no effect on the TLP_Ψ.

The primary cell walls in succulent tissues are generally very thin and elastic (Goldstein et al., 1991; Ogburn and Edwards, 2010). Thus, the true value of ε for succulent species more often falls on the curved portion of the line (Figure 1B). This means that for many succulent tissues, changes to cell wall biomechanics through elastic adjustment would have a much more substantial effect on the TLP_Ψ than in non-succulent plants. We sought to quantify this effect of ε on TLP_Ψ by repeating the simulation in Figures 1A, B for several species. Ogburn and Edwards (2012) studied the relationship between parameters derived from pressure–volume curves and measures of succulence in 25 species in the Caryophyllales, an angiosperm order comprising many succulent-rich groups with a broad range of tissue succulence. Using their published data, π_O was held constant for each of the 25 species and ε was allowed to vary in order to simulate the TLP_Ψ according to Equation 2. Then, for each species, we found the derivative of the curve, at the true value of ε (i.e. where the dashed line intersects the curve). This derivative, f′(ε), is a quantitative estimate of the extent to which changing ε affects the TLP_Ψ. As ε values become very low in highly succulent species, f′(ε) becomes exponentially higher (Figure 1C). Finally, we explored the relationship between f′(ε) and saturated water content (SWC), as the latter has been shown to be a powerful metric to quantify succulence in the Caryophyllales (Ogburn and Edwards, 2012). Log-transformed estimates of f′(ε) correlated significantly with SWC, using a linear regression model (Figure 1D).

Together, our data show that unlike non-succulent species, succulent plants occupy a phenotypic space in which increases in cell wall elasticity during drought (i.e. elastic adjustment) can

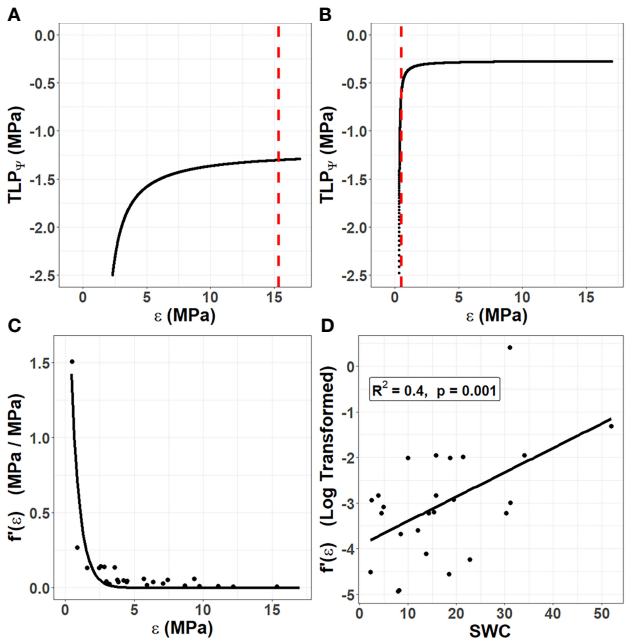


FIGURE 1
The unique role of cell wall biomechanics in succulent species. Using data from 25 species in the Caryophyllales published by Ogburn and Edwards (2012), the turgor loss point (TLP_Ψ) was simulated according to Equation 2 by holding the osmotic potential at full hydration (π_O) constant for each species and varying the bulk modulus of elasticity (ε). **(A)** In non-succulent species, such as *Calandrinia colchaguensis*, the true value of ε (dashed line) intersects at the flat portion of the curve. Hence, changes to ε have little to no effect on the TLP_Ψ. **(B)** In some succulent species, such as *Grahamia bracteata*, the true value of ε intersects at the curved potion of the line, meaning changes to ε affect the TLP_Ψ. A quantitative estimate of the extent to which changing ε affects the TLP_Ψ was generated by finding the derivative of the curve at the point where the dashed line intersects [f′(ε)]. **(C)** Among the 25 species, lower values of ε result in exponentially higher values of f′(ε). An exponential curve still fit these data well when the species with the highest f′(ε) value was removed (data not shown). **(D)** Among the 25 species, saturated water content (SWC) correlates with the f′(ε), after this value has been log transformed.

result in substantial decreases in TLP_{Ψ}. Furthermore, once a succulent species moves into this phenotypic space, decreasing ε has an exponential effect on f'(ε), so that alterations to cell wall biomechanics become an increasingly efficient means of controlling the TLP_{Ψ}. This agrees with the recently observed drought-induced modifications of pectic polysaccharides in hydrenchyma cell walls of *Aloe* (Ahl et al., 2019), which are believed to be a form of elastic adjustment that allows them to fold as cells shrink during dehydration.

Cell wall biochemistry in succulents

One way to assess the biochemical composition of cell walls is to investigate the extracellular glycome, which encompasses the entirety of extracellular carbohydrates in a tissue, organ or plant, and the majority of which corresponds to the cell wall. Characterizing glycomic profiles across different plant species can indicate which cell wall components have been favored under different environmental conditions. Whilst the glycomes of some economically important succulent taxa, such as *Agave*, *Aloe* and *Opuntia*, have recently been analyzed (Ginestra et al., 2009; Li et al., 2014; Ahl et al., 2018; Jones et al., 2020), little has been done to compare the cell wall composition of other distantly related succulent species. Hence, we sought to test the hypothesis that the extracellular glycome of phylogenetically diverse succulent species will exhibit some differences from those of non-succulents, so that a common 'succulent glycome' emerges. To this end, we sampled leaf material from 10 species with succulent leaves and 10 with non-succulent leaves, representing diverse lineages within the angiosperms (Figure 2A and Table S1). Using the succulence index (SI) from Ogburn and Edwards (2010) as a proxy for the degree of succulence (see Suppl. Methods), these two groups differed significantly ($p < 0.01$) (Figure 2B and Table S2). We used comprehensive microarray polymer profiling (CoMPP) to estimate and compare the relative polysaccharide contents of leaves from these species (see Suppl. Methods) (Moller et al., 2007; Ahl et al., 2018). We used whole leaves for comparability across species, assuming that mesophyll tissues would dominate the results. In the current study we used three extraction steps: water (targeting soluble unbound or loosely bound polysaccharides), CDTA (targeting primarily pectins) and NaOH (targeting primarily hemicelluloses). CoMPP relies on antibody-based molecular probes, so we used 49 monoclonal

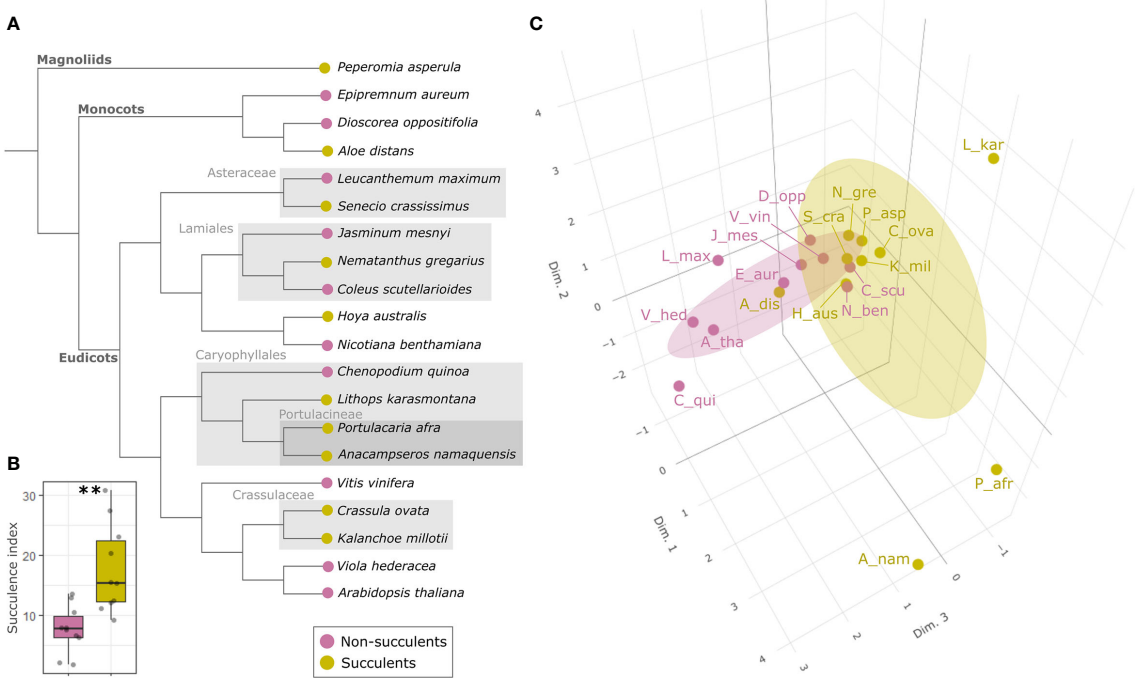


FIGURE 2
The succulent glycome: succulents and non-succulents occupy different phenotypic spaces in terms of glycomic profiles. **(A)** Cladogram representing the phylogenetic relationship (according to Baker et al., 2022) among the 20 species included in this study; some major clades are indicated in grey for clarity. **(B)** Boxplot of succulence index (SI) values for all the species; the two groups differ significantly (**$p < 0.01$) according to Welch's *t*-test. **(C)** 3D score plot of the first three MFA dimensions (17%, 12.1% and 10.8% of total variance respectively) of glycomic data from 10 leaf succulents and 10 non-succulents (see Table S1 for abbreviations), with concentration ellipsoids for each group. Succulents and non-succulents occupy distinct phenotypic spaces, particularly along dimension 3.

antibodies (mAbs) to target the majority of known cell wall polymer motifs (Moller et al., 2007; Rydahl et al., 2018) (Tables S3–S5). No representatives of commelinid monocots were included, given that their type-II cell wall biochemistry is particularly distinct from that of the rest of angiosperms (Carpita et al., 2015). CoMPP results in heatmap format can be found in Suppl. Data.

CoMPP results were analyzed using multiple factor analysis (MFA) (see Suppl. Methods), which indicated that succulent species occupy a distinct phenotypic space different from non-succulent species (Figure 2C). Of particular note is MFA dimension 3, along which succulents and non-succulents differed significantly ($p < 0.01$) and was driven mostly by glycoprotein- and pectin-targeting mAbs (Figure S1). Three succulent species were "pulling" along dimension 1 and fell far from the main cluster, but even when omitting these three outliers from the MFA, the results still showed a significant difference between succulents and non-succulents (Figure S2). We observed a higher signal for rhamnogalacturonan I (RG-I) backbones in succulents compared to non-succulents (Figure 3). RG-I and its side chains (i.e. arabinans, galactans and/or arabinogalactans) have been linked to increased cell wall elasticity (Harholt et al., 2010; Carroll et al., 2022) and have been postulated as cell wall plasticizers, which is a crucial feature for cells undergoing structural wall changes during dehydration and rehydration (Moore et al., 2013). Furthermore, we observed a higher signal for homogalacturonans (HGs) with a high degree of methyl-esterification (DM) in succulents (Figure 3), which may indicate highly elastic cell walls. In contrast, non-succulents had a higher signal for low-DM HGs (Figure 3), which may indicate stiffer cell walls. According to the textbook model, low-DM HGs can cross-link in the presence of Ca^{2+} and stiffen the cell wall, whereas a high DM prevents the formation of these cross-links and renders the wall more elastic (Willats et al., 2001; Cosgrove, 2005). However, several factors (e.g. pH, Ca^{2+} availability, different enzymatic activities) can influence the outcome, so caution should be taken when using DM as a proxy for cell wall mechanics (Palin and Geitmann, 2012; Bidhendi and Geitmann, 2016; Chebli and Geitmann, 2017; Hocq et al., 2017). Together, MFA of the CoMPP results suggests that fundamental differences exist between the cell wall composition of diverse succulent and non-succulent species. The three outlying succulent species (*Anacampseros namaquensis*, *Lithops karasmontana* and *Portulacaria afra*) belong to the core Caryophyllales, and two of them (*A. namaquensis* and *P. afra*) to suborder Portulacineae. These species showed remarkably high signal for RG-I and its side chains and for glucuronoxylans (Figure 3), which most likely reflects the presence of highly hydrophilic apoplastic mucilage in succulents in the Caryophyllales, particularly those in the Portulacineae (Figure 4)

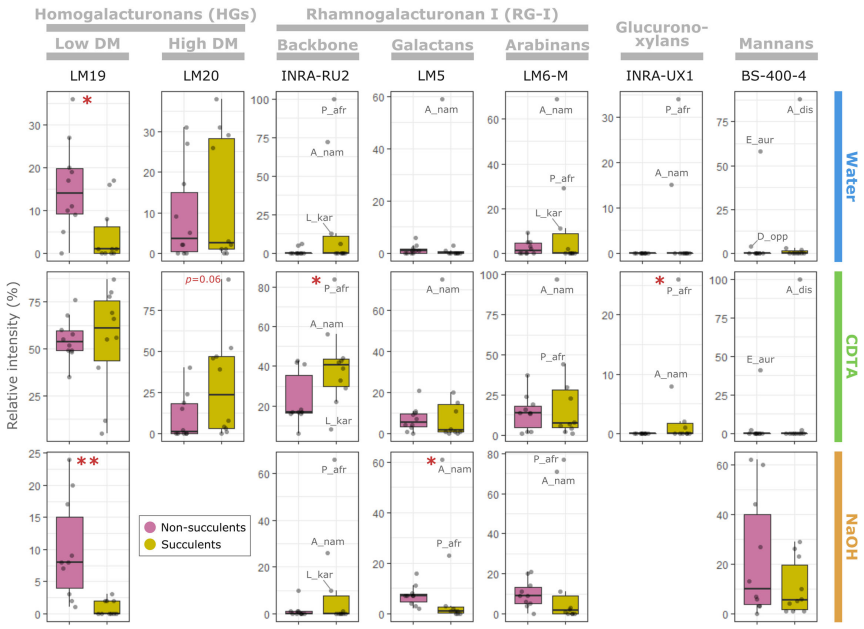


FIGURE 3
The succulent glycome: main glycomic differences between succulents and non-succulents. Selection of antibodies depicting the main glycomic differences between succulents and non-succulents for pectins (homogalacturonans and rhamnogalacturonan I), glucuronoxylans and mannans (see Suppl. Data for raw data). Each column corresponds to a specific antibody and rows represent the three fractions (i.e. CoMPP extraction steps); the y-axes represent relative intensity of signal within a specific fraction. Some outlying species have been labelled (see Table S1 for abbreviations). Significant differences between the two groups, assessed using either Welch's *t*-test (if both are normally distributed) or Wilcoxon test, are indicated by asterisks (*$p \leq 0.05$; **$p \leq 0.01$).

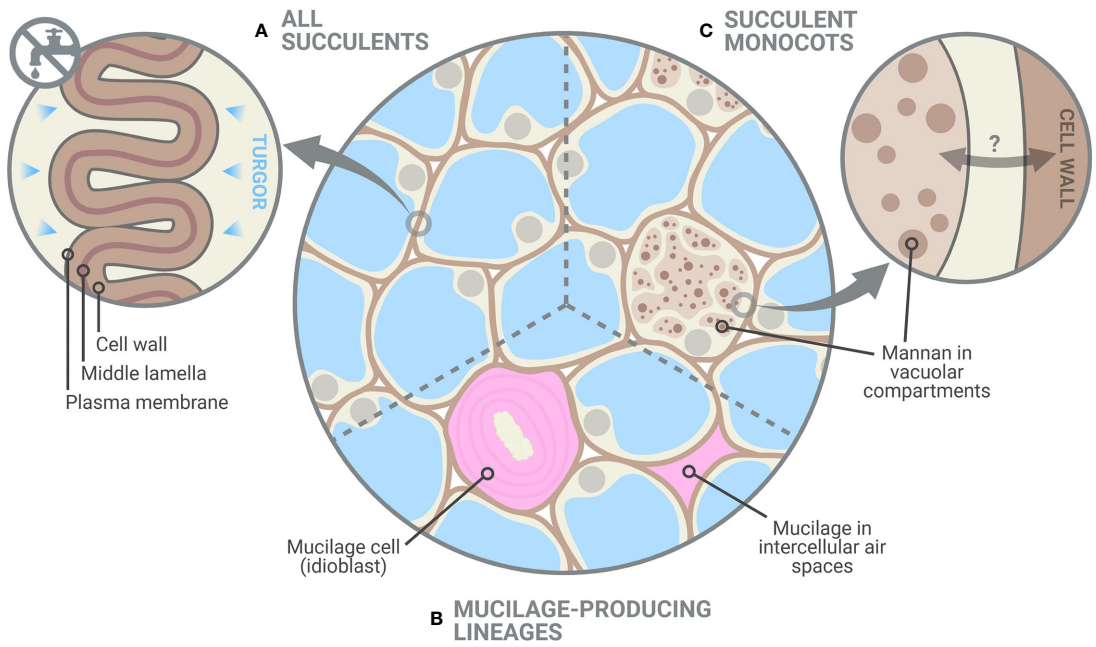


FIGURE 4
Glycomic diversity among succulent plants. **(A)** Succulent tissues have thin and highly elastic cell walls and, as shown in this study, elastic adjustment through cell wall remodelling likely plays a crucial role in preventing turgor loss during dehydration. Despite the clear differences between succulents and non-succulents, we also noted considerable glycomic diversity among succulents. **(B)** Mucilage-producing succulent lineages, mostly those in the Caryophyllales and particularly in the Portulacineae, accumulate pectin-rich mucilage in the periplasmic space of mucilage cells and/or in intercellular spaces, which boosts their water-storage capacity (Mauseth, 2005; Ogburn and Edwards, 2009). **(C)** Storage mannans can be found in vegetative tissues of many monocot lineages, often stored within vacuolar compartments. In succulent monocots, mobilization of these mannans may be part of the drought response, as seen in *Aloe* (Ahl et al., 2019), in which cell wall-associated mannan is remobilized into the protoplasm. However, the dynamics between cell-wall associated and vacuolar mannans in monocots remain largely unexplored. Created with BioRender.com.

(Cárdenas et al., 1997; Vignon and Gey, 1998; Hernandes-Lopes et al., 2016; Cole, 2022).

In addition to MFA, we used a random forest (RF) algorithm to determine whether glycomic profiles can be used to predict if a species is succulent or non-succulent (see Suppl. Methods). Based on CoMPP data alone, the RF algorithm was able to classify species in their respective categories with 90% accuracy (Suppl. Data). The variable importance plot from the RF algorithm identified several cell wall components driving this classification (Figure S3), namely arabinogalactan proteins (AGPs), xylans, low-DM HGs and RG-I (incl. arabinan and galactan side chains). Regarding HGs and RG-I, these results agree with the differences between succulents and non-succulents mentioned above. We also observed drastically lower levels of xylans in the succulents studied, compared to the non-succulent species (Suppl. Data). However, xylans are often found in lignified support tissues (Zhong et al., 2013), and small-stature succulent species such as the ones we studied generally lack these tissues, relying primarily on turgor for support (Niklas, 1992; Gibson, 1996; Bobich and North, 2009). Thus, such differences may not hold for larger succulents. An interesting observation concerns AGPs, a notoriously complex group of cell wall glycoproteins with many suggested functions (Seifert and Roberts, 2007; Silva et al., 2020). LM14 and MAC207, two mAbs that recognize the same or structurally related AGP epitopes (Marzec et al., 2015; Yan et al., 2015), did not yield any signal among succulents despite being present in most non-succulent species tested (Suppl. Data). In contrast, other AGP-targeting mAbs (e.g. JIM13) showed comparable levels between the two groups, likely reflecting the diversity of AGPs and their numerous alleged functions. For instance, periplasmic AGPs have been postulated as stabilizers of the membrane-cell wall continuum and may also act as cell wall plasticizers when they are released from their membrane anchors (Gens et al., 2000; Knox, 2006; Lamport et al., 2006; Liu et al., 2015). The striking differences in signal intensity of the AGP-targeting mAbs we used warrant further exploration into the specific epitopes that they recognize and their functions.

The mobilization of soluble mannans has been suggested as a general drought response among succulents, based on studies of succulent leaves of *Aloe* and succulent-like storage organs of orchids and monocot geophytes (Ranwala and Miller, 2008; Wang et al., 2008; Chua et al., 2013; Ahl et al., 2019). However, our CoMPP data showed no clear difference between the

mannans of succulents and non-succulents (Figure 3). Instead, two species exhibited remarkably high signal for loosely bound soluble mannans, *Aloe distans* (leaf succulent) and *Epipremnum aureum* (non-succulent), with *Dioscorea oppositifolia* (non-succulent) also showing above-average levels. These three species are the only non-commelinid monocots included in this study. Among angiosperms, the presence of storage mannans in vegetative tissues is believed to be restricted to monocots, with mannans being stored in granular or highly hydrated mucilaginous form within vacuolar cell compartments (Meier and Reid, 1982; He et al., 2017). Soluble mannans may therefore be uniquely important to monocots, being repurposed for drought response in succulent monocots (e.g. Ahl et al., 2019; Figure 4), and not a component of a more general succulent glycome.

Conclusions and future directions

Cell wall biomechanics and biochemistry of succulent leaves exhibit distinct differences from non-succulent species. In non-succulent species, highly rigid cell walls prevent elastic adjustment from having a physiologically meaningful impact on the TLP_{Ψ} (Bartlett et al., 2012). However, many succulent species have highly elastic cell walls, and our modelling indicates that even slight increases in cell wall elasticity (i.e. decreases in ε) in these species can have a large exponential effect on the TLP_{Ψ}. Therefore, succulent plants use elastic adjustment advantageously during dehydration to acclimate to declining Ψ. In addition to biomechanical differences, our glycomic data show several similarities across phylogenetically diverse succulent species, namely a higher degree of HG methyl-esterification and a greater abundance of RG-I. These biochemical differences likely contribute to the high elasticity in the cell walls of succulent organs, which in turn facilitates the folding process during dehydration (Fradera-Soler et al., 2022). Interestingly, some glycomic features seem to be restricted to certain succulent lineages, pointing to some glycomic diversity among succulent plants: succulent monocots may have co-opted soluble mannans for drought response, whereas succulents in the Caryophyllales contain pectin-rich apoplastic mucilage which boosts their water-storage capacity. Together, our data demonstrate that succulent plants occupy a unique phenotypic space regarding both cell wall biomechanics and biochemistry. We suggest that cell wall traits should be regarded as one of the core components of the adaptations that make up the succulent syndrome.

Looking forward, it will be valuable to explore cell wall biology among closely related succulent taxa and considering cell wall trait heterogeneity within succulent organs. Cell wall thickness and elasticity are known to differ between hydrenchyma and chlorenchyma in some succulent organs (Goldstein et al., 1991; Nobel, 2006; Leverett et al., 2022), but further examination of cell wall biomechanics and biochemistry is needed to fully understand how these traits aid in whole-plant survival during drought. Ultimately, further research is needed into the dynamic nature of cell walls in succulent plants and to determine whether cell wall traits are indeed regulated during drought. Besides high-throughput methods based on immune-profiling such as CoMPP, our understanding of cell wall composition, structure and assembly in succulents can also be advanced using visualization with fluorescent probes (Rydahl et al., 2018; Bidhendi et al., 2020), high-resolution microscopy techniques (Zhao et al., 2019; DeVree et al., 2021), and nuclear magnetic resonance (NMR) (Zhao et al., 2020).

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

MF-S and AL conceived the study and wrote the manuscript. AL designed and conducted modelling and data analysis regarding cell wall biomechanics. MF-S designed and conducted experiments and data analysis regarding cell wall biochemistry. JM, BJ, AB and OG contributed with supervision, review and editing of the manuscript. All authors read and approved the final manuscript.

Funding

This research was partially funded by Newcastle University's R. B. Cook Scholarship. JM is supported by a grant from the Slovak Academy of Sciences (project number IM-2021-23). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 801199.

Acknowledgments

The authors would like to thank Matthew Ogburn and Erika Edwards for the data that was used in the biomechanical modelling (see Ogburn and Edwards, 2012). We would also like to thank Theodor E. Bolsterli, Morten L. Stephensen, Ouda Khammy, Davide Visintainer and Luu Trinh (PLEN, University of Copenhagen) and the Kakteen-Haage nursery (Erfurt, Germany) for providing plant material. A special thanks to Sylwia Głazowska and Jeanett Hansen (PLEN, University of Copenhagen) for support during laboratory work, and to Karen S. Nissen for helpful discussion regarding statistical analyses.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1043429/full#supplementary-material

References

Ahl, L. I., Grace, O. M., Pedersen, H. L., Willats, W. G. T., Jørgensen, B., and Rønsted, N. (2018). Analyses of *Aloe* polysaccharides using carbohydrate microarray profiling. *J. AOAC Int.* 101, 1720–1728. doi: 10.5740/jaoacint.18-0120

Ahl, L. I., Mravec, J., Jørgensen, B., Rudall, P. J., Rønsted, N., and Grace, O. M. (2019). Dynamics of intracellular mannan and cell wall folding in the drought responses of succulent *Aloe* species. *Plant Cell Environ.* 42, 2458–2471. doi: 10.1111/pce.13560

Arakaki, M., Christin, P.-A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R. M., et al. (2011). Contemporaneous and recent radiations of the world's major succulent plant lineages. *PNAS* 108, 8379–8384. doi: 10.1073/pnas.1100628108

Baker, W. J., Bailey, P., Barber, V., Barker, A., Bellot, S., Bishop, D., et al. (2022). A comprehensive phylogenomic platform for exploring the angiosperm tree of life. *Syst. Biol.* 71, 301–319. doi: 10.1093/sysbio/syab035

Bartlett, M. K., Scoffoni, C., and Sack, L. (2012). The determinants of leaf turgor loss point and prediction of drought tolerance of species and biomes: A global meta-analysis. *Ecol. Lett.* 15, 393–405. doi: 10.1111/j.1461-0248.2012.01751.x

Beadle, C. L., Ludlow, M. M., and Honeysett, J. L. (1993). "Water relations," in *Photosynthesis and production in a changing environment: A field and laboratory manual*. Eds. D. O. Hall, J. M. O. Scurlock, H. R. Bolhàr-Nordenkampf, R. C. Leegood and S. P. Long (Dordrecht: Springer Netherlands), 113–128. doi: 10.1007/978-94-011-1566-7_8

Bidhendi, A. J., Chebli, Y., and Geitmann, A. (2020). Fluorescence visualization of cellulose and pectin in the primary plant cell wall. *J. Microsc.* 278, 164–181. doi: 10.1111/jmi.12895

Bidhendi, A. J., and Geitmann, A. (2016). Relating the mechanics of the primary plant cell wall to morphogenesis. *J. Exp. Bot.* 67, 449–461. doi: 10.1093/jxb/erv535

Bobich, E. G., and North, G. B. (2009). "Structural implications of succulence: architecture, anatomy, and mechanics of photosynthetic stem succulents, pachycauls, and leaf succulents," in *Perspectives in biophysical plant ecophysiology: A tribute to park s. Nobel*. Eds. E. de la Barrera and W. K. Smith (México: UNAM), 3–37.

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

Borland, A. M., Leverett, A., Hurtado-Castano, N., Hu, R., and Yang, X. (2018). "Functional anatomical traits of the photosynthetic organs of plants with crassulacean acid metabolism," in *The leaf: A platform for performing photosynthesis*. Eds. W. W. Adams III and I. Terashima (Cham: Springer International Publishing), 281–305. doi: 10.1007/978-3-319-93594-2_10

Brodribb, T. J., Skelton, R. P., McAdam, S. A. M., Bienaimé, D., Lucani, C. J., and Marmottant, P. (2016). Visual quantification of embolism reveals leaf vulnerability to hydraulic failure. *New Phytol.* 209, 1403–1409. doi: 10.1111/nph.13846

Cárdenas, A., Higuera-Ciapara, I., and Goycoolea, F. M. (1997). Rheology and aggregation of cactus (*Opuntia ficus-indica*) mucilage in solution. *J. Prof. Assoc. Cactus Dev.* 2, 152–159. doi: 10.56890/jpacd.v2i.181

Carpita, N. C., Ralph, J., and McCann, M. C. (2015). "The cell wall," in *Biochemistry and molecular biology of plants*. Eds. B. B. Buchanan, W. Gruissem and R. L. Jones (Chichester: John Wiley & Sons), 45–110.

Carroll, S., Amsbury, S., Durney, C. H., Smith, R. S., Morris, R. J., Gray, J. E., et al. (2022). Altering arabinans increases *Arabidopsis* guard cell flexibility and stomatal opening. *Curr. Biol.* 32, 3170–3179. doi: 10.1016/j.cub.2022.05.042

Chebli, Y., and Geitmann, A. (2017). Cellular growth in plants requires regulation of cell wall biochemistry. *Curr. Opin. Cell Biol.* 44, 28–35. doi: 10.1016/j.ceb.2017.01.002

Choat, B., Brodribb, T. J., Brodersen, C. R., Duursma, R. A., López, R., and Medlyn, B. E. (2018). Triggers of tree mortality under drought. *Nature* 558, 531–539. doi: 10.1038/s41586-018-0240-x

Chua, M., Hocking, T. J., Chan, K., and Baldwin, T. C. (2013). Temporal and spatial regulation of glucomannan deposition and mobilization in corms of *Amorphophallus konjac* (Araceae). *Am. J. Bot.* 100, 337–345. doi: 10.3732/ajb.1200547

Cole, T. C. H. (2022). *Caryophyllales Phylogeny Poster (CPP), 2022/V3.*

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

Davis, S. C., Kuzmick, E. R., Niechayev, N., and Hunsaker, D. J. (2017). Productivity and water use efficiency of *Agave americana* in the first field trial as bioenergy feedstock on arid lands. *GCB Bioenergy* 9, 314–325. doi: 10.1111/gcbb.12324

DeVree, B. T., Steiner, L. M., Głazowska, S., Ruhnow, F., Herburger, K., Persson, S., et al. (2021). Current and future advances in fluorescence-based visualization of plant cell wall components and cell wall biosynthetic machineries. *Biotechnol. Biofuels* 14, 1–26. doi: 10.1186/s13068-021-01922-0

Eggli, U., and Nyffeler, R. (2009). Living under temporarily arid conditions – succulence as an adaptive strategy. *Bradleya* 2009, 13–36. doi: 10.25223/brad.n27.2009.a10

Fradera-Soler, M., Grace, O. M., Jørgensen, B., and Mravec, J. (2022). Elastic and collapsible: Current understanding of cell walls in succulent plants. *J. Exp. Bot.* 73, 2290–2307. doi: 10.1093/jxb/erac054

Fradera-Soler, M., Rudall, P. J., Prychid, C. J., and Grace, O. M. (2021). Evolutionary success in arid habitats: Morpho-anatomy of succulent leaves of *Crassula* species from southern Africa. *J. Arid. Environ.* 185, 104319. doi: 10.1016/j.jaridenv.2020.104319

Gens, J. S., Fujiki, M., and Pickard, B. G. (2000). Arabinogalactan protein and wall-associated kinase in a plasmalemmal reticulum with specialized vertices. *Protoplasma* 212, 115–134. doi: 10.1007/BF01279353

Gibson, A. C. (1996). *Structure-function relations of warm desert plants* (Berlin, Heidelberg: Springer-Verlag).

Ginestra, G., Parker, M. L., Bennett, R. N., Robertson, J., Mandalari, G., Narbad, A., et al. (2009). Anatomical, chemical, and biochemical characterization of cladodes from prickly pear [*Opuntia ficus-indica* (L.) mill.]. *J. Agric. Food Chem.* 57, 10323–10330. doi: 10.1021/jf9022096

Goldstein, G., Andrade, J. L., and Nobel, P. S. (1991). Differences in water relations parameters for the chlorenchyma and the parenchyma of *Opuntia ficus-indica* under wet versus dry conditions. *Funct. Plant Biol.* 18, 95–107. doi: 10.1071/pp9910095

Grace, O. M. (2019). Succulent plant diversity as natural capital. *Plants People Planet.* 1, 336–345. doi: 10.1002/ppp3.25

Griffiths, H., and Males, J. (2017). Succulent plants. *Curr. Biol.* 27, R890–R896. doi: 10.1016/j.cub.2017.03.021

Harholt, J., Suttangkakul, A., and Scheller, H. V. (2010). Biosynthesis of pectin. *Plant Physiol.* 153, 384–395. doi: 10.1104/pp.110.156588

Hartzell, S., Bartlett, M. S., Inglese, P., Consoli, S., Yin, J., and Porporato, A. (2021). Modelling nonlinear dynamics of crassulacean acid metabolism productivity and water use for global predictions. *Plant Cell Environ.* 44, 34–48. doi: 10.1111/pce.13918

Henry, C., John, G. P., Pan, R., Bartlett, M. K., Fletcher, L. R., Scoffoni, C., et al. (2019). A stomatal safety-efficiency trade-off constrains responses to leaf dehydration. *Nat. Commun.* 10, 3398. doi: 10.1038/s41467-019-11006-1

Hernandes-Lopes, J., Oliveira-Neto, M. A., and Melo-de-Pinna, G. F. A. (2016). Different ways to build succulent leaves in portulacineae (Caryophyllales). *Int. J. Plant Sci.* 177, 198–208. doi: 10.1086/684178

He, C., Wu, K., Zhang, J., Liu, X., Zeng, S., Yu, Z., et al. (2017). Cytochemical localization of polysaccharides in *Dendrobium officinale* and the involvement of *DoCSLA6* in the synthesis of mannan polysaccharides. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00173

Heyduk, K. (2021). The genetic control of succulent leaf development. *Curr. Opin. Plant Biol.* 59, 101978. doi: 10.1016/j.pbi.2020.11.003

Heyduk, K., McKain, M. R., Lalani, F., and Leebens-Mack, J. (2016). Evolution of a CAM anatomy predates the origins of crassulacean acid metabolism in the agavoideae (Asparagaceae). *Mol. Phylogenet. Evol.* 105, 102–113. doi: 10.1016/j.ympev.2016.08.018

Hocq, L., Pelloux, J., and Lefebvre, V. (2017). Connecting homogalacturonan-type pectin remodeling to acid growth. *Trends Plant Sci.* 22, 20–29. doi: 10.1016/j.tplants.2016.10.009

Jiao, W., Wang, L., Smith, W. K., Chang, Q., Wang, H., and D'Odorico, P. (2021). Observed increasing water constraint on vegetation growth over the last three decades. *Nat. Commun.* 12, 3777. doi: 10.1038/s41467-021-24016-9

Jones, A. M., Zhou, Y., Held, M. A., and Davis, S. C. (2020). Tissue composition of *Agave americana* l. yields greater carbohydrates from enzymatic hydrolysis than advanced bioenergy crops. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00654

Knox, J. P. (2006). Up against the wall: arabinogalactan-protein dynamics at cell surfaces. *New Phytol.* 169, 443–445. doi: 10.1111/j.1469-8137.2006.01640.x

Kunert, N., Zailaa, J., Herrmann, V., Muller-Landau, H. C., Wright, S. J., Pérez, R., et al. (2021). Leaf turgor loss point shapes local and regional distributions of evergreen but not deciduous tropical trees. *New Phytol.* 230, 485–496. doi: 10.1111/nph.17187

Lamport, D. T. A., Kieliszewski, M. J., and Showalter, A. M. (2006). Salt stress upregulates periplasmic arabinogalactan proteins: using salt stress to analyse AGP function. *New Phytol.* 169, 479–492. doi: 10.1111/j.1469-8137.2005.01591.x

Leverett, A., Castaño, N. H., Ferguson, K., Winter, K., and Borland, A. M. (2021). Crassulacean acid metabolism (CAM) supersedes the turgor loss point (TLP) as an important adaptation across a precipitation gradient, in the genus *Clusia*. *Funct. Plant Biol.* 48, 703–716. doi: 10.1071/FP20268

Leverett, A., Hartzell, S., Winter, K., Garcia, M., Aranda, J., Virgo, A., et al. (2022). Dissecting succulence: Crassulacean acid metabolism and hydraulic capacitance are independent adaptations in *clusia* leaves. *bioRxiv*. doi: 10.1101/2022.03.30.486278

Lim, S. D., Mayer, J. A., Yim, W. C., and Cushman, J. C. (2020). Plant tissue succulence engineering improves water-use efficiency, water-deficit stress attenuation and salinity tolerance in *Arabidopsis*. *Plant J.* 103, 1049–1072. doi: 10.1111/tpj.14783

Lim, S. D., Yim, W. C., Liu, D., Hu, R., Yang, X., and Cushman, J. C. (2018). A *Vitis vinifera* basic helix–loop–helix transcription factor enhances plant cell size, vegetative biomass and reproductive yield. *Plant Biotechnol. J.* 16, 1595–1615. doi: 10.1111/pbi.12898

Li, H., Pattathil, S., Foston, M. B., Ding, S.-Y., Kumar, R., Gao, X., et al. (2014). *Agave* proves to be a low recalcitrant lignocellulosic feedstock for biofuels production on semi-arid lands. *Biotechnol. Biofuels* 7, 50. doi: 10.1186/1754-6834-7-50

Liu, Z., Persson, S., and Sánchez-Rodríguez, C. (2015). At The border: the plasma membrane–cell wall continuum. *J. Exp. Bot.* 66, 1553–1563. doi: 10.1093/jxb/erv019

Males, J. (2017). Secrets of succulence. *J. Exp. Bot.* 68, 2121–2134. doi: 10.1093/jxb/erx096

Marzec, M., Szarejko, I., and Melzer, M. (2015). Arabinogalactan proteins are involved in root hair development in barley. *J. Exp. Bot.* 66, 1245–1257. doi: 10.1093/jxb/eru475

Mauseth, J. D. (2005). Anatomical features, other than wood, in subfamily opuntioideae (Cactaceae). *Haseltonia* 2005 (11), 113–125. doi: 10.2985/1070-0048(2005)11[113:AFOTWI]2.0.CO;2

Meier, H., and Reid, J. S. G. (1982). "Reserve polysaccharides other than starch in higher plants," in *Plant carbohydrates I*. Eds. F. A. Loewus and W. Tanner (Berlin, Heidelberg: Springer), 418–471. doi: 10.1007/978-3-642-68275-9_11

Moller, I., Sørensen, I., Bernal, A. J., Blaukopf, C., Lee, K., Øbro, J., et al. (2007). High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *Plant J.* 50, 1118–1128. doi: 10.1111/j.1365-313X.2007.03114.x

Moore, J. P., Nguema-Ona, E. E., Vicré-Gibouin, M., Sørensen, I., Willats, W. G. T., Driouich, A., et al. (2013). Arabinose-rich polymers as an evolutionary strategy to plasticize resurrection plant cell walls against desiccation. *Planta* 237, 739–754. doi: 10.1007/s00425-012-1785-9

Neupane, D., Mayer, J. A., Niechayev, N. A., Bishop, C. D., and Cushman, J. C. (2021). Five-year field trial of the biomass productivity and water input response of cactus pear (*Opuntia* spp.) as a bioenergy feedstock for arid lands. *GCB Bioenergy* 13, 719–741. doi: 10.1111/gcbb.12805

Niklas, K. J. (1992). *Plant biomechanics: An engineering approach to plant form and function* (Chicago: University of Chicago Press).

Nobel, P. S. (2006). Parenchyma–chlorenchyma water movement during drought for the hemiepiphytic cactus *Hylocereus undatus*. *Ann. Bot.* 97, 469–474. doi: 10.1093/aob/mcj054

Ogburn, R. M., and Edwards, E. J. (2009). Anatomical variation in cactaceae and relatives: trait lability and evolutionary innovation. *Am. J. Bot.* 96, 391–408. doi: 10.3732/ajb.0800142

Ogburn, R. M., and Edwards, E. J. (2010). "The ecological water-use strategies of succulent plants," in *Advances in botanical research*. Eds. J.-C. Kader and M. Delseny (Burlington, MA: Academic Press), 179–225. doi: 10.1016/B978-0-12-380868-4.00004-1

Ogburn, R. M., and Edwards, E. J. (2012). Quantifying succulence: a rapid, physiologically meaningful metric of plant water storage. *Plant Cell Environ.* 35, 1533–1542. doi: 10.1111/j.1365-3040.2012.02503.x

Owen, N. A., and Griffiths, H. (2014). Marginal land bioethanol yield potential of four crassulacean acid metabolism candidates (*Agave fourcroydes*, *Agave salmiana*, *Agave tequilana* and *Opuntia ficus-indica*) in Australia. *GCB Bioenergy* 6, 687–703. doi: 10.1111/gcbb.12094

Palin, R., and Geitmann, A. (2012). The role of pectin in plant morphogenesis. *Biosystems* 109, 397–402. doi: 10.1016/j.biosystems.2012.04.006

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

Ranwala, A. P., and Miller, W. B. (2008). Analysis of nonstructural carbohydrates in storage organs of 30 ornamental geophytes by high-performance anion-exchange chromatography with pulsed amperometric detection. *New Phytol.* 180, 421–433. doi: 10.1111/j.1469-8137.2008.02585.x

Rydahl, M. G., Hansen, A. R., Kračun, S. K., and Mravec, J. (2018). Report on the current inventory of the toolbox for plant cell wall analysis: proteinaceous and small molecular probes. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00581

Seifert, G. J., and Roberts, K. (2007). The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* 58, 137–161. doi: 10.1146/annurev.arplant.58.032806.103801

Sheffield, J., and Wood, E. F. (2008). Global trends and variability in soil moisture and drought characteristics 1950–2000, from observation-driven simulations of the terrestrial hydrologic cycle. *J. Clim.* 21, 432–458. doi: 10.1175/2007JCLI1822.1

Silva, J., Ferraz, R., Dupree, P., Showalter, A. M., and Coimbra, S. (2020). Three decades of advances in arabinogalactan-protein biosynthesis. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.610377

Trueba, S., Pan, R., Scoffoni, C., John, G. P., Davis, S. D., and Sack, L. (2019). Thresholds for leaf damage due to dehydration: declines of hydraulic function, stomatal conductance and cellular integrity precede those for photochemistry. *New Phytol.* 223, 134–149. doi: 10.1111/nph.15779

Vignon, M. R., and Gey, C. (1998). Isolation, ^{1}H and ^{13}C NMR studies of (4-*O*-methyl-D-glucurono)-D-xylans from luffa fruit fibres, jute bast fibres and mucilage of quince tree seeds. *Carbohydr. Res.* 307, 107–111. doi: 10.1016/S0008-6215(98)00002-0

Vollenweider, P., Menard, T., Arend, M., Kuster, T. M., and Günthardt-Goerg, M. S. (2016). Structural changes associated with drought stress symptoms in foliage of central European oaks. *Trees* 30, 883–900. doi: 10.1007/s00468-015-1329-6

Wang, C.-Y., Chiou, C.-Y., Wang, H.-L., Krishnamurthy, R., Venkatagiri, S., Tan, J., et al. (2008). Carbohydrate mobilization and gene regulatory profile in the pseudobulb of *Oncidium* orchid during the flowering process. *Planta* 227, 1063–1077. doi: 10.1007/s00425-007-0681-1

Willats, W. G. T., McCartney, L., Mackie, W., and Knox, J. P. (2001). Pectin: cell biology and prospects for functional analysis. *Plant Mol. Biol.* 47, 9–27. doi: 10.1023/A:1010662911148

Yan, Y., Takáč, T., Li, X., Chen, H., Wang, Y., Xu, E., et al. (2015). Variable content and distribution of arabinogalactan proteins in banana (*Musa* spp.) under low temperature stress. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00353

Zhang, Y.-J., Rockwell, F. E., Graham, A. C., Alexander, T., and Holbrook, N. M. (2016). Reversible leaf xylem collapse: A potential "circuit breaker" against cavitation. *Plant Physiol.* 172, 2261–2274. doi: 10.1104/pp.16.01191

Zhao, W., Fernando, L. D., Kirui, A., Deligey, F., and Wang, T. (2020). Solid-state NMR of plant and fungal cell walls: a critical review. *Solid State Nucl. Magn. Reson.* 107, 101660. doi: 10.1016/j.ssnmr.2020.101660

Zhao, Y., Man, Y., Wen, J., Guo, Y., and Lin, J. (2019). Advances in imaging plant cell walls. *Trends Plant Sci.* 24, 867–878. doi: 10.1016/j.tplants.2019.05.009

Zhong, R., McCarthy, R. L., Haghighat, M., and Ye, Z.-H. (2013). The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PloS One* 8, e69219. doi: 10.1371/journal.pone.0069219

frontiers | Frontiers in Plant Science

TYPE Perspective
PUBLISHED 11 January 2023
DOI 10.3389/fpls.2022.1076298

OPEN ACCESS

EDITED BY
Rowan Andrew Craig Mitchell,
Rothamsted Research,
United Kingdom

REVIEWED BY
Wei Zeng,
Zhejiang Agriculture and Forestry University, China

*CORRESPONDENCE
Paul Dupree
pd101@cam.ac.uk

†PRESENT ADDRESS
Louis Frederick Lundy Wilson,
Howard Hughes Medical Institute, University of Virginia School of Medicine, Charlottesville, VA, United States
Mathias Sorieul,
Scion, Rotorua, New Zealand

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

RECEIVED 21 October 2022
ACCEPTED 22 December 2022
PUBLISHED 11 January 2023

CITATION
Anders N, Wilson LFL, Sorieul M, Nikolovski N and Dupree P (2023) β-1,4-Xylan backbone synthesis in higher plants: How complex can it be? *Front. Plant Sci.* 13:1076298. doi: 10.3389/fpls.2022.1076298

β-1,4-Xylan backbone synthesis in higher plants: How complex can it be?

Nadine Anders, Louis Frederick Lundy Wilson†, Mathias Sorieul†, Nino Nikolovski and Paul Dupree*

Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom

Xylan is a hemicellulose present in the cell walls of all land plants. Glycosyltransferases of the GT43 (IRX9/IRX9L and IRX14/IRX14L) and GT47 (IRX10/IRX10L) families are involved in the biosynthesis of its β-1,4-linked xylose backbone, which can be further modified by acetylation and sugar side chains. However, it remains unclear how the different enzymes work together to synthesize the xylan backbone. A xylan synthesis complex (XSC) has been described in the monocots wheat and asparagus, and co-expression of asparagus *Ao*IRX9, *Ao*IRX10 and *Ao*IRX14A is required to form a catalytically active complex for secondary cell wall xylan biosynthesis. Here, we argue that an equivalent XSC exists for the synthesis of the primary cell wall of the eudicot *Arabidopsis thaliana*, consisting of IRX9L, IRX10L and IRX14. This would suggest the existence of distinct XSCs for primary and secondary cell wall xylan synthesis, reminiscent of the distinct cellulose synthesis complexes (CSCs) of the primary and secondary cell wall. In contrast to the CSC, in which each CESA protein has catalytic activity, the XSC seems to contain proteins with non-catalytic function with each component bearing potentially unique but crucial roles. Moreover, the core XSC formed by a combination of IRX9/IRX9L, IRX10/IRX10L and IRX14/IRX14L might not be stable in its composition during transit from the endoplasmic reticulum to the Golgi apparatus. Instead, potential dynamic changes of the XSC might be a means of regulating xylan biosynthesis to facilitate coordinated deposition of tailored polysaccharides in the plant cell wall.

KEYWORDS

Xylan biosynthesis, enzyme complex, primary cell wall, IRX9/IRX9L, IRX10/IRX10L, IRX14/IRX14L

Introduction

The plant cell wall is a complex matrix of polysaccharides, proteins and other polymers. The primary cell wall formed around the growing cell mostly consists of cellulose, hemicellulose and pectin, and is generally thin and flexible. By contrast, the secondary cell wall is produced in fully grown specialized cells and forms a thick and more rigid structure with differing polysaccharide composition. Incorporation of lignin rigidifies and reduces the water permeability of the secondary cell wall. Xylan is present in both primary and secondary cell walls (Scheller and Ulvskov, 2010) and is crucial to cell wall strength and biomass recalcitrance. It consists of a linear backbone of β-1,4-linked D-xylosyl residues. The backbone can be substituted with decorations such as acetyl groups, α-D-glucuronic acid, 4-*O*-methyl-α-D-glucuronic acid, α-L-arabinofuranose or β-D-xylose, and the sugar side chains themselves can be further modified. The type, pattern and degree of substitution vary considerably depending on species and tissue, which has been recently reviewed in Smith et al., 2017 and Ye and Zhong, 2022. In addition, differences in primary and secondary xylan structure have been described. In Arabidopsis primary cell wall, glucuronic acid is further substituted with a pentosyl modification, suggested to be α-1,2-L-Ara*p* (Chong et al., 2015; Mortimer et al., 2015). An α-1,2-L-Ara*p* modification on glucuronic acid of xylan was also identified in non-commelinid monocots, though, while present in growing Asparagus tips, it was not detectable in stem xylan (Pena et al., 2016), suggesting that α-1,2-L-Ara*p* modifications on glucuronic acid could be primary cell wall-specific. In contrast to the primary cell wall, the majority of Arabidopsis secondary cell wall xylan carries acetate or glucuronic acid decorations on even-numbered, alternating xylose residues, while other xylan regions have clustered glucuronic acid substitutions (Bromley et al., 2013; Busse-Wicher et al., 2014; Grantham et al., 2017).

Over the years, a plethora of proteins have been implicated in the biosynthesis of xylan. Many are known to modify the xylan backbone or are proposed to produce the essential reducing end oligosaccharide found in eudicots; others are of yet unknown function. An overview of xylan biosynthesis can be found in Smith et al., 2017 and Ye and Zhong, 2022. Here, we would like to focus on three functionally non-redundant groups of proteins that are essential for the β-1,4-xylan backbone synthesis in a wide range of plants (Zhong et al., 2019), comprising two groups assigned to CAZy glycosyltransferase family GT43 (IRX9/IRX9L and IRX14/IRX14L) and one to family GT47 (IRX10/IRX10L) (Drula et al., 2022). Early on, the identification of *irx9* and *irx14* as xylan-deficient Arabidopsis mutants with reduced xylan synthesis activity led to the idea that both enzymes might act together in a protein complex (Brown et al., 2007; Lee et al., 2007; Lee et al., 2012a). More recent genetic evidence supporting the existence of a XSC describes the dominant negative effect of IRX10 point mutants (Brandon et al., 2020). Biochemical analysis of wheat and asparagus proteins further supports the existence of a XSC in monocots (Zeng et al., 2010; Jiang et al., 2016; Zeng et al., 2016). In addition, simultaneous expression of the three asparagus proteins *Ao*IRX9, *Ao*IRX10 and *Ao*IRX14A in tobacco is required to obtain a catalytically active complex in the Golgi.

How three functionally non-redundant putative GTs work together to facilitate the β-1,4-linkage of the xylan backbone, however, remained unanswered. *In vitro* xylosyltransferase or xylan synthase activity has only been successfully demonstrated for isolated GT47 proteins from Arabidopsis, rice, Plantago and Physcomitrium (formerly Physcomitrella), as well as for an IRX10 ortholog from the streptophyte alga *Klebsormidium nitens* (formerly *Klebsormidium flaccidum*) (Jensen et al., 2014; Urbanowicz et al., 2014; Jensen et al., 2018; Wang et al., 2022). The most recent experimental data suggest two *in vitro* activities of rice IRX10: one for elongating the xylan chain and the other for initiating xylan synthesis (Wang et al., 2022), although only elongating activity was detected in Arabidopsis IRX10L (Urbanowicz et al., 2014). Lee et al. showed β-1,4-xylosyltransferase activity of IRX9 and IRX14 when co-expressed in tobacco (Lee et al., 2012a; Lee et al., 2012b)), however, a positive effect of the co-expressed proteins on endogenous IRX10/IRX10L activity cannot be excluded, especially as the activity of IRX9 and IRX14 alone could not be shown (Lee et al., 2012a; Lee et al., 2012b; Urbanowicz et al., 2014).

GT43s adopt a GT-A structural fold; enzymes of this class typically possess an essential catalytic DxD motif, which is critical in coordinating divalent metal ions required for nucleotide-binding (Lairson et al., 2008). Indeed, mutations in the DxD (or DD) motif of *Ao*IRX14A, *Ao*IRX14B and *At*IRX14 result in a loss of function, suggesting that IRX14 could be catalytically active or require UDP-Xyl binding for function (Ren et al., 2014; Zeng et al., 2016). Mutations in the DxD motif of *Ao*IRX9L and *At*IRX9L, however, have no effect on the function of the protein and the motif is not conserved in IRX9. These findings, taken together with the prediction that some IRX10s and IRX10Ls lack a transmembrane domain (Jensen et al., 2014), led to the hypothesis that, rather than a catalytic role, IRX9 and IRX9L might serve a structural role in xylan synthesis (Ren et al., 2014; Zeng et al., 2016). Yet, the functional role of the non-catalytic GTs in the XSC may be more significant in that they could control active complex formation and consequently regulate xylan biosynthesis.

Evolution of paralogs for primary and secondary cell wall xylan synthesis

Streptophyte algae exhibit both β-1,4-xylan and GT43 xylan synthesis genes. Furthermore, at least for *Klebsormidium nitens*, IRX10 xylan synthesis function has been demonstrated (Taujale and Yin, 2015; Jensen et al., 2018; Hsieh and Harris, 2019).

However, in this species, only one ortholog of each group of IRX9/IRX9L, IRX10/IRX10L and IRX14/IRX14L is present. This is also the case in the early vascular model plant *Selaginella moellendorffii*. Gene duplication has taken place in the moss *Physcomitrium patens*, which has only one catalytically active IRX10/IRX10L (Jensen et al., 2014), but two IRX9/IRX9L and three IRX14 paralogs (Hornblad et al., 2013; Haghighat et al., 2016); however, these duplicated genes might not encode functional orthologs of the two paralogous gene groups seen in higher plants. The paralogous genes seen in higher plants emerged at different stages during evolution: the divergence of IRX9 from IRX9L most likely occurred earlier than that of IRX10/IRX10L, or that of IRX14/IRX14L, with separate homologs for IRX9 and IRX9L in monocots and in contrast low sequence diversion of IRX14 and IRX14L in eudicots (Wu et al., 2010; Ren et al., 2014; Wang et al., 2022).

An early step forward in understanding the duplication of xylan biosynthesis enzymes in higher plants came with the realization that each protein is partially redundant with respect to its closest homolog (with the IRX-L mutants displaying less severe phenotypes); the Arabidopsis double mutants of each paralogous pair (*irx9 irx9l, irx10 irx10l* and *irx14 irx14l*) manifest significantly more severe phenotypes than either of the single mutants, but they can functionally replace each other in overexpression and promoter swap experiments (Brown et al., 2009; Wu et al., 2009; Keppler and Showalter, 2010; Lee et al., 2010; Wu et al., 2010; Mortimer et al., 2015). This suggests that differences predominantly arise due to differential expression rather than functional divergence. Supporting this, secondary cell wall xylan in Arabidopsis is mostly dependent on IRX9, IRX10 and IRX14, while IRX9L, IRX10L, and IRX14 are essential for primary cell wall xylan synthesis (Mortimer et al., 2015). Gene expression analyses in asparagus, poplar and rice further support the idea of different sets of IRX homologs being required for xylan biosynthesis in the primary versus the secondary cell wall (Chiniquy et al., 2013; Ratke et al., 2015; Song et al., 2015). The separation into distinct enzyme groups in higher plants is reminiscent of the primary and secondary wall-specific CSCs (Kumar and Turner, 2015).

The primary cell wall XSC in eudicots

Although research in Arabidopsis has greatly advanced our understanding of xylan synthesis, with genetic data hinting at an interaction of XSC enzymes, protein interaction has not been shown in eudicots. To investigate the existence and composition of a XSC in Arabidopsis, we generated *irx14 irx14l* double-mutant plants expressing IRX14-GFP under its endogenous promoter, creating functional tagged IRX14 (Supplementary Figure S1). To investigate XSC formation in primary cell wall-rich tissue, we performed anti-GFP immunoprecipitation in Golgi-enriched microsomal fractions of root callus culture, which we generated from the homozygous transgenic plants. Callus from plants expressing STL1-GFP, a Golgi protein partially co-localizing with IRX9L (Zhang et al., 2016), was used as a negative control. Liquid Chromatography with tandem Mass Spectrometry (LC-MS/MS) analysis was used to detect interacting proteins, showing that IRX14 interacts with IRX9L and IRX10L, but not with STL1 (Figure 1A; Supplementary Table S1). The immunoprecipitation using STL1-GFP control showed no interaction with IRX9L, IRX10L or IRX14.

These data suggest that IRX9L, IRX10L and IRX14 form a XSC in eudicot primary cell wall-rich tissues, comparable to the XSC synthesizing the secondary cell wall of monocots. Unique peptides were only found for IRX9L and IRX10L, but not for their homologs IRX9 or IRX10 (Supplementary Table S2), most likely due to a lack of expression of IRX9 and IRX10 in callus tissue (Mortimer et al., 2015). The detection of IRX9L and IRX10L in the callus XSC is consistent with the genetic evidence that these proteins function together in xylan synthesis of primary cell walls (Mortimer et al., 2015). To support the structural feasibility of a XSC in Arabidopsis, we also note that AlphaFold-Multimer predicts a heterotrimer of the IRX9L, IRX10L and IRX14 globular domains with high intermolecular confidence scores when provided with one copy of each protein (Figure 1B; Supplementary Figure S2A). The model predicts a pseudo-symmetric interaction of IRX9L and IRX14 and a potential stabilizing role of a conserved proline-rich region/putative α-helix that sits N-terminally of the GT43 domain of IRX14 (Figure 1B; Supplementary Figures S2A, F). An essentially identical model can be generated from the equivalent secondary cell wall components (Supplementary Figure S2B).

The core of the complex

Genetic and protein interaction studies suggest that xylan is synthesized by a protein complex, however, the protein composition of this complex is not uniformly described. Surprisingly, an IRX9 homolog was not detected in the wheat XSC (Zeng et al., 2010; Jiang et al., 2016), raising the question as to whether all three proteins are required in the active complex.

In *Plantago ovata*, at least four different *IRX10* genes are highly expressed in the seed mucilaginous layer, while IRX9 and IRX14 orthologs are only expressed at very low levels (Jensen et al., 2013; Jensen et al., 2014). Therefore, Jensen et al. (2014) suggested that IRX10 might be solely responsible for mucilage xylan synthesis, whereas in stem tissues of *Plantago*, IRX9 and IRX14 are relevant.

In Arabidopsis, the only known player to synthesize the backbone of seed mucilage xylan is IRX14, whereas, surprisingly, IRX10 appears not to be important (Voiniciuc et al., 2015; Hu

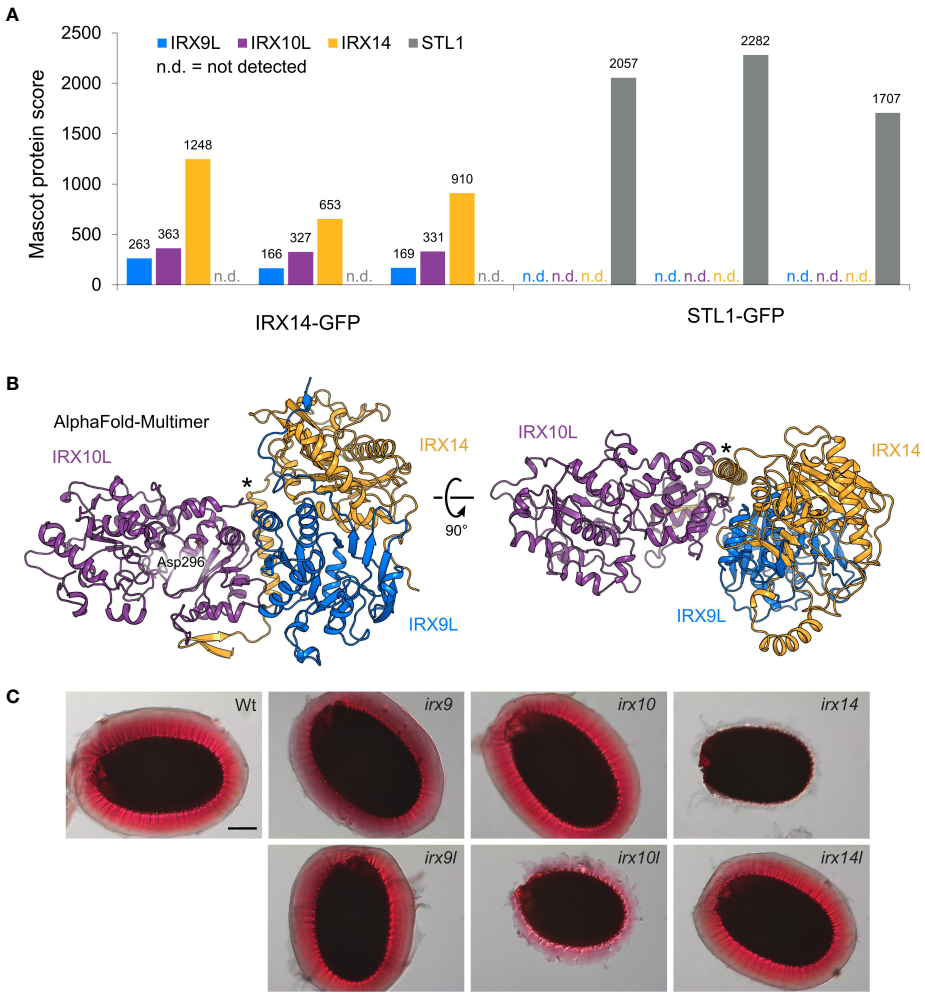


FIGURE 1
The XSC in primary cell wall xylan synthesis in Arabidopsis. **(A)** Anti-GFP immunoprecipitation of the XSC in Arabidopsis root callus. Mascot protein scores for IRX9L, IRX10L, IRX14 and STL1 (control) are shown for three biological replicates using IRX14-GFP (*left*) or STL1-GFP (control, *right*) as bait, respectively. **(B)** AlphaFold-Multimer model of a heterotrimer of the IRX9L (blue), IRX10L (violet), and IRX14 (yellow) globular domain. A putative nucleotide-binding residue, Asp296 (Wilson et al., 2022), is shown in IRX10L. Note the predicted α-helix of IRX14 at the interaction surfaces of both IRX9L and IRX10L (asterisk). **(C)** Phenotype of the adherent mucilage of XSC mutants. Mucilage is stained with ruthenium red. Wt: Columbia Col-0 wildtype control. Scale bar = 100 µm.

et al., 2016). To analyze whether IRX14 is indeed the only XSC protein with a function in xylan biosynthesis in Arabidopsis mucilage, we stained the adherent mucilage of wildtype, *irx9*, *irx9l*, *irx10*, *irx10l*, *irx14* and *irx14l* mutant seeds with ruthenium red. In the absence of xylan, the mucilage does not adhere to the seed (Ralet et al., 2016). The mucilage released from the *irx9*, *irx9l*, *irx10* and *irx14l* mutants had a similar appearance to the mucilage of wildtype seeds, whereas *irx10l* seed mucilage was clearly defective, although the phenotype was less severe than in *irx14* (Figure 1C). These data show that IRX10L is also required for xylan biosynthesis in Arabidopsis mucilage. The finding is consistent with the Arabidopsis gene-chip data available on SeedGeneNetwork (Belmonte et al., 2013), showing that IRX10L and IRX14 are highly expressed in the general seed coat at the linear cotyledon stage (the stage of mucilage biosynthesis gene expression; Supplementary Figure S3A) (Francoz et al., 2015). An involvement of IRX9/IRX9L in mucilage xylan biosynthesis is unclear due to a lack of an obvious mucilage phenotype in the ruthenium red staining. However, the impact of gene redundancy of IRX9/IRX9L could not easily be assessed, as the double mutants do not produce seeds. Strong expression of IRX9, although possibly delayed to the maturation green stage, might suggest a role of IRX9, whereas IRX9L expression remains low throughout these stages of seed development (Supplementary Figure S3).

Taken together, there could be a certain degree of variability even in the three core components of the XSC, depending on species and tissue. Nevertheless, a critical component of xylan

synthase activity seems to be IRX10/IRX10L, which is in line with their biochemical activity.

Homo- and hetero-oligomerization of the XSC components

The apparent size of the asparagus XSCs in native gels is slightly smaller than the marker size of 242 kDa (Zeng et al., 2016). Therefore, the complex could consist of one protein of IRX9/IRX9L, IRX10/IRX10L and IRX14/IRX14L forming a heterotrimeric complex, as shown for the primary and secondary cell wall XSCs in our AlphaFold models (Figure 1B; Supplementary Figures S2A, B). In tobacco, *Ao*IRX9 and *Ao*IRX14A interact with each other in bimolecular fluorescence complementation experiments (Zeng et al., 2016).

Interestingly, however, the GT43 proteins *Ta*GT43-4, *Ao*IRX9 and *Ao*IRX14A were each shown to interact with themselves (Jiang et al., 2016; Zeng et al., 2016), which is in line with the finding that mammalian GT43 β-1,3-glucuronyltransferases form homodimers (Terayama et al., 1998; Ouzzine et al., 2000). Structural analysis revealed that the three human isoforms homodimerize *via* conserved interaction surfaces (Pedersen et al., 2000; Kakuda et al., 2004; Shiba et al., 2006). As Arabidopsis GT43s are predicted to adopt a similar secondary structure to the human enzymes (Taujale and Yin, 2015), it is possible that homodimerization occurs through similar surfaces. This is supported by AlphaFold modelling, (Supplementary Figures S2C, D). This means homodimeric interactions would presumably compete with pseudo-symmetric heterodimerization of IRX9/IRX9L with IRX14/IRX14L by the same interface. However, a trimeric complex of a GT43 homodimer with IRX10L appears less likely, based on the modelling scores. Our AlphaFold modelling of the globular domains alone of two copies of each IRX9L, IRX10L and IRX14 is predicted to form two separate heterotrimers (Supplementary Figure S2E). Also, this modelling does not suggest IRX10 dimerization, which is consistent with the results for wheat *Ta*GT47-13 by bimolecular fluorescence complementation (Jiang et al., 2016). On the other hand, bimolecular fluorescence complementation showed self-interaction of asparagus *Ao*IRX10 (Zeng et al., 2016).

In plants, alternative homo- or hetero-oligomerization of GTs has been reported for the *N*-glycan processing Arabidopsis α-mannosidase I and *Nicotiana tabacum* β-1,2-*N*-acetyglucosaminyltransferase I (Schoberer et al., 2013). Another example is the xylosyltransferase XXT2, which can homodimerize, but also interact with XXT5 or XXT1 (Chou et al., 2012). Interestingly, while this homodimerization involves disulfide bridges, heterodimerization seems not to. It is unclear whether plant GT43s only interact *via* the conserved interaction surface of the GT domain, meaning homo- and hetero-oligomerization would be mutually exclusive or whether an additional interaction surface has evolved to allow both simultaneously. Expression in *Pichia pastoris* led to a complex that appeared to contain two *Ta*GT43-4 proteins and one *Ta*GT47-13 (Jiang et al., 2016). This suggests that at least IRX14 can interact with other components of the XSC while forming a homo-oligomer. This was also suggested in the model, proposing a XSC consisting of IRX9 and IRX14 interacting homodimers, which indirectly interact with an IRX10 homodimer (Zeng et al., 2016). In summary, the stoichiometry of the XSC is yet to be understood.

Golgi GTs are typically type II membrane proteins. Hence, in addition to the luminal GT domain, most exhibit a short cytosolic tail, a single transmembrane domain as well as a stem region (CTS). This CTS region can also mediate protein–protein interactions, including *via* disulfide bridges in the transmembrane helix (TMH) and stem (Tu and Banfield, 2010; Kellokumpu et al., 2016). Interestingly, we found that IRX9/IRX9L orthologs harbor a highly conserved CFxxGxxxG motif in their predicted TMH (Figure 2A), resembling a canonical GxxxG motif, which could act as a GAS_{right} helix oligomerization motif like that in glycophorin A (Mueller et al., 2014). The previously identified WxxxHxxCCxxSxxLGxRFS motif of IRX14/IRX14L orthologs (Jiang et al., 2016) could be considered a GAS_{right} variant since it contains a SxxxGxxxS motif (Figure 2A). Indeed, AlphaFold modelling supports the notion that the IRX9 and IRX14 TMHs could form a disulfide-linked homo- or heterodimer with the GAS_{right} motifs situated at the helix interface (Figure 2B), providing a potential mechanism of interaction. This motif might provide an additional dimerization surface, allowing larger asparagus XSC complex formation such as that additionally detected in the native gel analysis running just below the 480kDa marker (Zeng et al., 2016).

Subcellular localization of the XSC and dynamic changes of its composition

As well as mediating protein interactions, the CTS region of Golgi GTs can determine correct localization to distinct Golgi cisternae (Schoberer and Strasser, 2011; Schoberer et al., 2013). The cytoplasmic tail of IRX14 and IRX14L also contains a conserved cytosolic di-arginine motif (Jiang et al., 2016), a motif that has been described in governing ER retrieval or Golgi localization (Welch and Munro, 2019). Proteins in assembled complexes with an ER-retaining motif can pass to the Golgi, possibly through masking of the motif (Michelsen et al., 2005; Banfield, 2011). Ren et al., 2014 noted that the milder phenotype of *irx9-2* (versus *irx9-1*) might be due to the presence of the truncated protein, assigning a potential function to the N-terminus of IRX9.

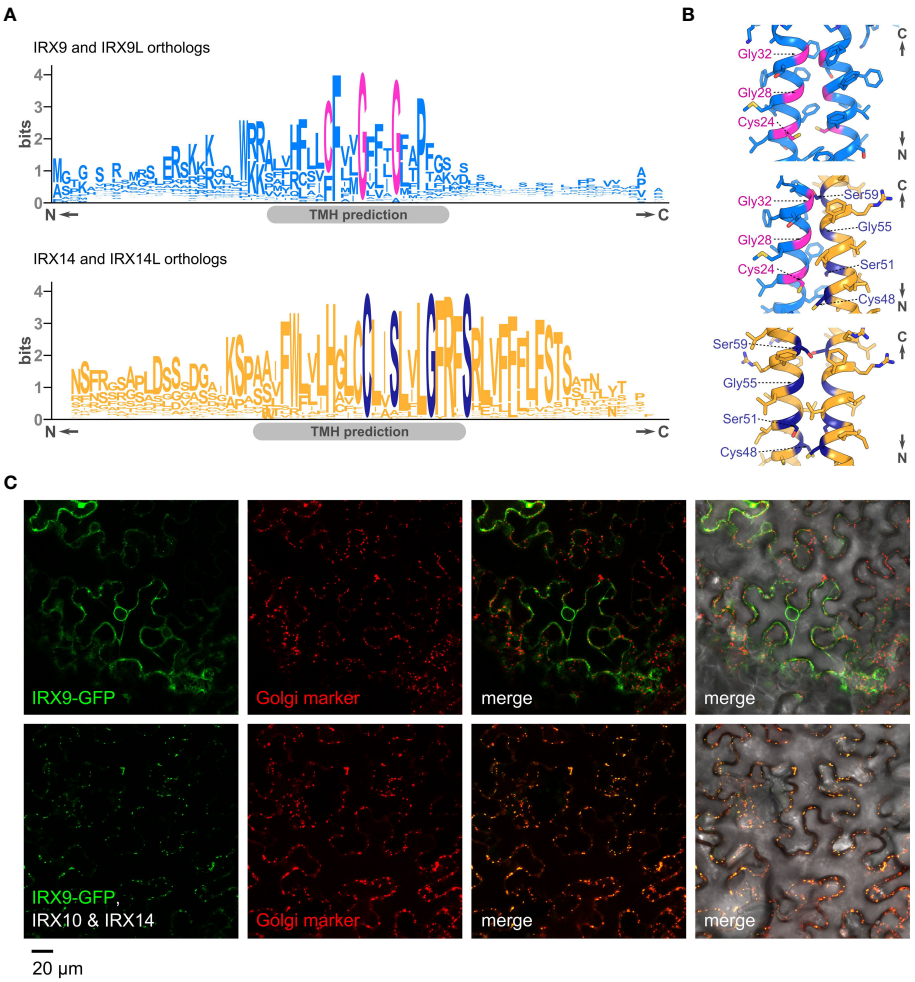


FIGURE 2
The CTS of GT43s, their potential interactions and impact of interactions on Golgi localization. **(A)**. Sequence logos showing conserved motifs in the predicted transmembrane helices (TMH) of orthologs of IRX9/IRX9L (*top*) and IRX14/IRX14L (*bottom*). **(B)** Close-ups of AlphaFold-Multimer models of the transmembrane dimers for an IRX9 homodimer (*top*), an IRX9–IRX14 heterodimer (*middle*) and an IRX14 homodimer (*bottom*). Residues of the GAS_{right} motif and conserved cysteines are highlighted in magenta or dark blue, respectively. **(C)** Subcellular localization of Arabidopsis IRX9-GFP, transiently expressed in tobacco leaves. Top panel (*left to right*), IRX9-GFP (*green*), ManI-mCherry from soy (Golgi marker, *red*), merge of the two former, and merge including differential interference contrast (DIC) image. Bottom panel shows the same localization analysis in presence of co-expressed, untagged IRX10 and IRX14. Note the change of localization of IRX9-GFP in presence of IRX10 and IRX14.

Notably, the interaction of *Ta*GT47-13 and *Ta*GT43-4 already occurs in the ER and their co-expression is required for ER to Golgi translocation of the complex in tobacco (Jiang et al., 2016). Using a similar approach, co-expression of *Ao*IRX9, *Ao*IRX10 and *Ao*IRX14A was shown to be required for catalytic activity and XSC transition from the ER to the Golgi (Zeng et al., 2016). This mechanism seems not unique to monocots; we have also observed that co-expression of IRX10 and IRX14 is required for the Golgi-localization of IRX9-GFP in tobacco (Figure 2C). Co-expression with IRX10 or IRX14 alone were not tested.

Sub-Golgi localization of GTs in Arabidopsis using free-flow electrophoresis followed by mass-spectrometry analysis revealed that IRX14 and the glucuronyltransferase GUX3 mostly locate to the *trans*-Golgi, while IRX10L was predominantly found in *medial*-Golgi (Parsons et al., 2019). Quantitative immuno-transmission electron microscopy also shows that IRX9 localizes predominantly to the *medial*-Golgi (Meents et al., 2019). Hence, the three core components of the XSC seem not consistently co-localized, suggesting they do not always exist in the same complex. Similar to the sequential distribution of GTs involved in *N*-glycosylation of proteins, recent models suggest that GTs involved in polysaccharide biosynthesis of the plant cell wall are distributed in different Golgi-cisternae *cis, medial* to *trans*, depending on their functional activity (Parsons et al., 2019; Hoffmann et al., 2021). Supporting this idea is the reported absence of IRX9 and presence of two putative UDP-sugar

mutases, *Ta*GT75-3 and *Ta*GT75-4, and two non-GTs, *Ta*GLP and *Ta*VER2, in the wheat XSC (Zeng et al., 2010; Jiang et al., 2016), although these interactions have not been described in other systems. The XSC isolated from Golgi-enriched membranes of wheat also exhibited arabinosyl- and glucuronyltransferase activity, indicating that the XSC might harbor xylan modifying enzymes, although the respective enzymes were not detected (Zeng et al., 2010; Jiang et al., 2016). Similarly, in our immunoprecipitation study we did not detect co-precipitation of the glucuronyltransferase GUX3. Nevertheless, it is tempting to think that a specialized primary cell wall XSC might differ in its interactions compared with the secondary cell wall XSC, leading to the differences reported in GlcA modification patterns of xylan (Mortimer et al., 2010; Bromley et al., 2013). The interaction of the core XSC proteins with xylan backbone-modifying enzymes would be expected to occur in the *trans*-Golgi. Other potential interactors (for example the UDP-Xyl transporter UXT1, which is suggested to channel the UDP-Xyl substrate for xylan biosynthesis (Ebert et al., 2015)) might interact with the XSC early in the Golgi. Hence, the XSC might not have a uniform composition, but rather might change while transiting from the ER to the Golgi and through the Golgi cisternae, reflecting its functional activity.

Discussion

All three components of the XSC, IRX9/9L, IRX14/IRX14L and IRX10/IRX10L, are necessary for xylan synthesis in most systems. Despite this, the emerging picture is that IRX10/IRX10L is the catalytically active enzyme essential for the biochemical function of the XSC. In contrast, the role of IRX9/9L and IRX14/IRX14L in the XSC remains poorly understood and future research will have to establish why IRX9/IRX9L and IRX14/IRX14L are essential for xylan synthesis, albeit not directly involved in the catalytic reaction. Diverse non-catalytic functions of the GT-like proteins have been suggested over the years. These range from membrane anchoring of or UDP-Xyl channeling to IRX10/IRX10L, to serving as a scaffold for XSC assembly or assembly-dependent trafficking (Ren et al., 2014; Jiang et al., 2016; Zeng et al., 2016).

In mammalian Golgi *N*-glycosyltransferases, homomers are disassembled and heteromers formed depending on pH-changes during transition through the Golgi (Hassinen et al., 2010; Hassinen and Kellokumpu, 2014). It is unknown whether such a mechanism plays a role in plants; however, it is important to keep in mind that interactions between proteins might not be static. Thus, one step to understanding the role of the individual XSC components might be through analysis of their sub-Golgi localization, complex formation and interaction with xylan-modifying enzymes. Finally, changes in the protein composition of the XSC through homo- and heterodimerization of its components could provide a mechanistic tool to regulate the localization of the XSC and its activity in response to environmental cues, allowing dynamic adjustment of the plant cell wall.

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 author.

Author contributions

MS created transgenic lines, NN and MS performed immunoprecipitation analysis; NN, NA, and MS performed the mucilage staining. LW conducted transient expression in tobacco, sequence analysis, and AlphaFold modelling. NA conceptualized the manuscript, collated the data and wrote the manuscript with support of LW and PD. All authors contributed to the article and approved the submitted version.

Funding

This project was supported through funding by the BBSRC (BB/K005537/1) (NA); by the BBSRC OpenPlant (BB/L014130/1) (LW); by the BBSRC Sustainable Bioenergy Centre cell wall sugars programme (BB/G016240/1) and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 707911 (NN); by the European Community's Seventh Framework Programme SUNLIBB (FP7/2007–2013) under grant agreement #251132 (MS).

Acknowledgments

The authors thank Alberto Echevarría-Poza for critical reading of the manuscript, Henry Temple for supporting the confocal microscopy and Joel Wurman-Rodrich for supporting the immunoprecipitation analysis. Rita Delgado Marques and Xiaolan Yu for technical assistance and George Lomonossoff (John Innes Centre, UK), who developed the pEAQ-HyperTrans expression system used in this study. Plant Bioscience Limited supplied the pEAQ-HT vector.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1076298/full#supplementary-material

References

Banfield, D. K. (2011). Mechanisms of protein retention in the golgi. *Cold Spring Harbor Perspect. Biol.* 3. doi: 10.1101/cshperspect.a005264

Belmonte, M. F., Kirkbride, R. C., Stone, S. L., Pelletier, J. M., Bui, A. Q., Yeung, E. C., et al. (2013). Comprehensive developmental profiles of gene activity in regions and subregions of the arabidopsis seed. *Proc. Natl. Acad. Sci. U.S.A.* 110, E435–E444. doi: 10.1073/pnas.1222061110

Brandon, A. G., Birdseye, D. S., and Scheller, H. V. (2020). A dominant negative approach to reduce xylan in plants. *Plant Biotechnol. J.* 18, 5–7. doi: 10.1111/pbi.13198

Bromley, J. R., Busse-Wicher, M., Tryfona, T., Mortimer, J. C., Zhang, Z., Brown, D. M., et al. (2013). GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *Plant J.* 74, 423–434. doi: 10.1111/tpj.12135

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., 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

Busse-Wicher, M., Gomes, T. C., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N. J., et al. (2014). The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of arabidopsis thaliana. *Plant J.* 79, 492–506. doi: 10.1111/tpj.12575

Chiniquy, D., Varanasi, P., Oh, T., Harholt, J., Katnelson, J., Singh, S., et al. (2013). Three novel rice genes closely related to the arabidopsis IRX9, IRX9L, and IRX14 genes and their roles in xylan biosynthesis. *Front. Plant Sci.* 4, 83. doi: 10.3389/fpls.2013.00083

Chong, S. L., Koutaniemi, S., Juvonen, M., Derba-Maceluch, M., Mellerowicz, E. J., and Tenkanen, M. (2015). Glucuronic acid in arabidopsis thaliana xylans carries a novel pentose substituent. *Int. J. Biol. Macromol* 79, 807–812. doi: 10.1016/j.ijbiomac.2015.05.056

Chou, Y. H., Pogorelko, G., and Zabotina, O. A. (2012). Xyloglucan xylosyltransferases XXT1, XXT2, and XXT5 and the glucan synthase CSLC4 form golgi-localized multiprotein complexes. *Plant Physiol.* 159, 1355–1366. doi: 10.1104/pp.112.199356

Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 50, D571–D577. doi: 10.1093/nar/gkab1045

Ebert, B., Rautengarten, C., Guo, X., Xiong, G., Stonebloom, S., Smith-Moritz, A. M., et al. (2015). Identification and characterization of a golgi-localized UDP-xylose transporter family from arabidopsis. *Plant Cell* 27, 1218–1227. doi: 10.1105/tpc.114.133827

Francoz, E., Ranocha, P., Burlat, V., and Dunand, C. (2015). Arabidopsis seed mucilage secretory cells: regulation and dynamics. *Trends Plant Sci.* 20, 515–524. doi: 10.1016/j.tplants.2015.04.008

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

Haghighat, M., Teng, Q., Zhong, R., and Ye, Z. H. (2016). Evolutionary conservation of xylan biosynthetic genes in selaginella moellendorffii and physcomitrella patens. *Plant Cell Physiol.* 57, 1707–1719. doi: 10.1093/pcp/pcw096

Hassinen, A., and Kellokumpu, S. (2014). Organizational interplay of golgi n-glycosyltransferases involves organelle microenvironment-dependent transitions between enzyme homo- and heteromers. *J. Biol. Chem.* 289, 26937–26948. doi: 10.1074/jbc.M114.595058

Hassinen, A., Rivinoja, A., Kauppila, A., and Kellokumpu, S. (2010). Golgi n-glycosyltransferases form both homo- and heterodimeric enzyme complexes in live cells. *J. Biol. Chem.* 285, 17771–17777. doi: 10.1074/jbc.M110.103184

Hoffmann, N., King, S., Samuels, A. L., and Mcfarlane, H. E. (2021). Subcellular coordination of plant cell wall synthesis. *Dev. Cell* 56, 933–948. doi: 10.1016/j.devcel.2021.03.004

Hornblad, E., Ulfstedt, M., Ronne, H., and Marchant, A. (2013). Partial functional conservation of IRX10 homologs in physcomitrella patens and arabidopsis thaliana indicates an evolutionary step contributing to vascular formation in land plants. *BMC Plant Biol.* 13, 3. doi: 10.1186/1471-2229-13-3

Hsieh, Y. S. Y., and Harris, P. J. (2019). Xylans of red and green algae: What is known about their structures and how they are synthesised? *Polymers (Basel)* 11(2), 354. doi: 10.3390/polym11020354

Hu, R. B., Li, J. L., Wang, X. Y., Zhao, X., Yang, X. W., Tang, Q., et al. (2016). Xylan synthesized by irregular xylem 14 (IRX14) maintains the structure of seed coat mucilage in arabidopsis. *J. Exp. Bot.* 67, 1243–1257. doi: 10.1093/jxb/erv510

Jensen, J. K., Busse-Wicher, M., Poulsen, C. P., Fangel, J. U., Smith, P. J., Yang, J. Y., et al. (2018). Identification of an algal xylan synthase indicates that there is functional orthology between algal and plant cell wall biosynthesis. *New Phytol.* 218, 1049–1060. doi: 10.1111/nph.15050

Jensen, J. K., Johnson, N., and Wilkerson, C. G. (2013). Discovery of diversity in xylan biosynthetic genes by transcriptional profiling of a heteroxylan containing mucilaginous tissue. *Front. Plant Sci.* 4, 183. doi: 10.3389/fpls.2013.00183

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

Jiang, N., Wiemels, R. E., Soya, A., Whitley, R., Held, M., and Faik, A. (2016). Composition, assembly, and trafficking of a wheat xylan synthase complex. *Plant Physiol.* 170, 1999–2023. doi: 10.1104/pp.15.01777

Kakuda, S., Shiba, T., Ishiguro, M., Tagawa, H., Oka, S., Kajihara, Y., et al. (2004). Structural basis for acceptor substrate recognition of a human glucuronyltransferase, GlcAT-p, an enzyme critical in the biosynthesis of the carbohydrate epitope HNK-1. *J. Biol. Chem.* 279, 22693–22703. doi: 10.1074/jbc.M400622200

Kellokumpu, S., Hassinen, A., and Glumoff, T. (2016). Glycosyltransferase complexes in eukaryotes: long-known, prevalent but still unrecognized. *Cell Mol. Life Sci.* 73, 305–325. doi: 10.1007/s00018-015-2066-0

Keppler, B. D., and Showalter, A. M. (2010). IRX14 and IRX14-LIKE, two glycosyl transferases involved in glucuronoxylan biosynthesis and drought tolerance in arabidopsis. *Mol. Plant* 3, 834–841. doi: 10.1093/mp/ssq028

Kumar, M., and Turner, S. (2015). Plant cellulose synthesis: CESA proteins crossing kingdoms. *Phytochemistry* 112, 91–99. doi: 10.1016/j.phytochem.2014.07.009

Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* 77, 521–555. doi: 10.1146/annurev.biochem.76.061005.092322

Lee, C., O'neill, M. A., Tsumuraya, Y., Darvill, A. G., and Ye, Z. H. (2007). The irregular xylem9 mutant is deficient in xylan xylosyltransferase activity. *Plant Cell Physiol.* 48, 1624–1634. doi: 10.1093/pcp/pcm135

Lee, C., Teng, Q., Huang, W., Zhong, R., and Ye, Z. H. (2010). The arabidopsis family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. *Plant Physiol.* 153, 526–541. doi: 10.1104/pp.110.155309

Lee, C., Zhong, R., and Ye, Z. H. (2012a). 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

Lee, C., Zhong, R., and Ye, Z. H. (2012b). Biochemical characterization of xylan xylosyltransferases involved in wood formation in poplar. *Plant Signal Behav.* 7, 332–337. doi: 10.4161/psb.19269

Meents, M. J., Motani, S., Mansfield, S. D., and Samuels, A. L. (2019). Organization of xylan production in the golgi during secondary cell wall biosynthesis. *Plant Physiol.* 181, 527–546. doi: 10.1104/pp.19.00715

Michelsen, K., Yuan, H. B., and Schwappach, B. (2005). Hide and run - arginine-based endoplasmic-reticulum-sorting motifs in the assembly of heteromultimeric membrane proteins. *EMBO Rep.* 6, 717–722. doi: 10.1038/sj.embor.7400480

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

Mortimer, J. C., Miles, G. P., Brown, D. M., Zhang, Z., Segura, M. P., Weimar, T., et al. (2010). Absence of branches from xylan in arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17409–17414. doi: 10.1073/pnas.1005456107

Mueller, B. K., Subramaniam, S., and Senes, A. (2014). A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical c alpha-h hydrogen bonds. *Proc. Natl. Acad. Sci. United States America* 111, E888–E895. doi: 10.1073/pnas.1319944111

Ouzzine, M., Gulberti, S., Netter, P., Magdalou, J., and Fournel-Gigleux, S. (2000). Structure/function of the human Ga1beta1,3-glucuronosyltransferase. dimerization and functional activity are mediated by two crucial cysteine residues. *J. Biol. Chem.* 275, 28254–28260. doi: 10.1074/jbc.M002182200

Parsons, H. T., Stevens, T. J., Mcfarlane, H. E., Vidal-Melgosa, S., Griss, J., Lawrence, N., et al. (2019). Separating golgi proteins from cis to trans reveals underlying properties of cisternal localization. *Plant Cell* 31, 2010–2034. doi: 10.1105/tpc.19.00081

Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., and Negishi, M. (2000). Heparan/chondroitin sulfate biosynthesis. structure and mechanism of human glucuronyltransferase I. *J. Biol. Chem.* 275, 34580–34585. doi: 10.1074/jbc.M007399200

Pena, M. J., Kulkarni, A. R., Backe, J., Boyd, M., O'neill, M. A., and York, W. S. (2016). Structural diversity of xylans in the cell walls of monocots. *Planta* 244, 589–606. doi: 10.1007/s00425-016-2527-1

Ralet, M. C., Crepeau, M. J., Vigouroux, J., Tran, J., Berger, A., Salle, C., et al. (2016). Xylans provide the structural driving force for mucilage adhesion to the arabidopsis seed coat. *Plant Physiol.* 171, 165–178. doi: 10.1104/pp.16.00211

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

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

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

Schoberer, J., Liebminger, E., Botchway, S. W., Strasser, R., and Hawes, C. (2013). Time-resolved fluorescence imaging reveals differential interactions of n-glycan processing enzymes across the golgi stack in planta. *Plant Physiol.* 161, 1737–1754. doi: 10.1104/pp.112.210757

Schoberer, J., and Strasser, R. (2011). Sub-Compartmental organization of golgi-resident n-glycan processing enzymes in plants. *Mol. Plant* 4, 220–228. doi: 10.1093/mp/ssq082

Shiba, T., Kakuda, S., Ishiguro, M., Morita, I., Oka, S., Kawasaki, T., et al. (2006). Crystal structure of GlcAT-s, a human glucuronyltransferase, involved in the biosynthesis of the HNK-1 carbohydrate epitope. *Proteins* 65, 499–508. doi: 10.1002/prot.21118

Smith, P. J., Wang, H. T., York, W. S., Pena, M. J., and Urbanowicz, B. R. (2017). Designer biomass for next-generation biorefineries: leveraging recent insights into xylan structure and biosynthesis. *Biotechnol. Biofuels* 10, 286. doi: 10.1186/s13068-017-0973-z

Song, L., Zeng, W., Wu, A., Picard, K., Lampugnani, E. R., Cheetamun, R., et al. (2015). Asparagus spears as a model to study heteroxylan biosynthesis during secondary wall development. *PloS One* 10, e0123878. doi: 10.1371/journal.pone.0123878

Taujale, R., and Yin, Y. (2015). Glycosyltransferase family 43 is also found in early eukaryotes and has three subfamilies in charophycean green algae. *PloS One* 10, e0128409. doi: 10.1371/journal.pone.0128409

Terayama, K., Seiki, T., Nakamura, A., Matsumori, K., Ohta, S., Oka, S., et al. (1998). Purification and characterization of a glucuronyltransferase involved in the biosynthesis of the HNK-1 epitope on glycoproteins from rat brain. *J. Biol. Chem.* 273, 30295–30300. doi: 10.1074/jbc.273.46.30295

Tu, L., and Banfield, D. K. (2010). Localization of golgi-resident glycosyltransferases. *Cell Mol. Life Sci.* 67, 29–41. doi: 10.1007/s00018-009-0126-z

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

Voiniciuc, C., Gunl, M., Schmidt, M. H., and Usadel, B. (2015). Highly branched xylan made by IRREGULAR XYLEM14 and MUCILAGE-RELATED21 links mucilage to arabidopsis seeds. *Plant Physiol.* 169, 2481–2495. doi: 10.1104/pp.15.01441

Wang, H., Yang, H., Wen, Z., Gao, C., Gao, Y., Tian, Y., et al. (2022). Xylan-based nanocompartments orchestrate plant vessel wall patterning. *Nat. Plants* 8, 295–306. doi: 10.1038/s41477-022-01113-1

Welch, L. G., and Munro, S. (2019). A tale of short tails, through thick and thin: investigating the sorting mechanisms of golgi enzymes. *FEBS Lett.* 593, 2452–2465. doi: 10.1002/1873-3468.13553

Wilson, L. F. L., Dendooven, T., Hardwick, S. W., Echevarria-Poza, A., Tryfona, T., Krogh, K., et al. (2022). The structure of EXTL3 helps to explain the different roles of bi-domain exostosins in heparan sulfate synthesis. *Nat. Commun.* 13, 3314. doi: 10.1038/s41467-022-31048-2

Wu, A. M., Hornblad, 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

Wu, A. M., Rihouey, C., Seveno, M., Hornblad, E., Singh, S. K., Matsunaga, T., et al. (2009). The arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J.* 57, 718–731. doi: 10.1111/j.1365-313X.2008.03724.x

Ye, Z. H., and Zhong, R. (2022). Outstanding questions on xylan biosynthesis. *Plant Sci.* 325, 111476. doi: 10.1016/j.plantsci.2022.111476

Zeng, W., Jiang, N., Nadella, R., Killen, T. L., Nadella, V., and Faik, A. (2010). A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. *Plant Physiol.* 154, 78–97. doi: 10.1104/pp.110.159749

Zeng, W., Lampugnani, E. R., Picard, K. L., Song, L., Wu, A. M., Farion, I. M., et al. (2016). Asparagus IRX9, IRX10, and IRX14A are components of an active xylan backbone synthase complex that forms in the golgi apparatus. *Plant Physiol.* 171, 93–109. doi: 10.1104/pp.15.01919

Zhang, Y., Nikolovski, N., Sorieul, M., Vellosillo, T., Mcfarlane, H. E., Dupree, R., et al. (2016). Golgi-localized STELLO proteins regulate the assembly and trafficking of cellulose synthase complexes in arabidopsis. *Nat. Commun.* 7, 11656. doi: 10.1038/ncomms11656

Zhong, R., Cui, D., and Ye, Z. H. (2019). Secondary cell wall biosynthesis. *New Phytol.* 221, 1703–1723. doi: 10.1111/nph.15537

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 16 January 2023
DOI 10.3389/fpls.2022.1125003

OPEN ACCESS

EDITED BY
Wagner Rodrigo De Souza,
Federal University of ABC, Brazil

REVIEWED BY
Yunjun Zhao,
Center for Excellence in Molecular Plant Sciences (CAS), China
Leonard Blaschek,
University of Copenhagen, Denmark

*CORRESPONDENCE
Claire Halpin
c.halpin@dundee.ac.uk

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

RECEIVED 15 December 2022
ACCEPTED 22 December 2022
PUBLISHED 16 January 2023

CITATION
Shafiei R, Hooper M, McClellan C, Oakey H, Stephens J, Lapierre C, Tsuji Y, Goeminne G, Vanholme R, Boerjan W, Ralph J and Halpin C (2023) Downregulation of barley ferulate 5-hydroxylase dramatically alters straw lignin structure without impact on mechanical properties. *Front. Plant Sci.* 13:1125003. doi: 10.3389/fpls.2022.1125003

Downregulation of barley ferulate 5-hydroxylase dramatically alters straw lignin structure without impact on mechanical properties

Reza Shafiei[1], Matthew Hooper[1], Christopher McClellan[1], Helena Oakey[1,2], Jennifer Stephens[3], Catherine Lapierre[4], Yukiko Tsuji[5,6], Geert Goeminne[7], Ruben Vanholme[8,9], Wout Boerjan[8,9], John Ralph[6] and Claire Halpin[1]*

[1]Division of Plant Sciences, School of Life Sciences, University of Dundee at the James Hutton Institute, Dundee, United Kingdom, [2]Faculty of Sciences, School of Agriculture, Food and Wine, University of Adelaide, Adelaide, SA, Australia, [3]Cell And Molecular Sciences, James Hutton Institute, Dundee, United Kingdom, [4]INRAE AgroParisTech, Université Paris Saclay, IJPB, Versailles, France, [5]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, United States, [6]Department of Energy's Great Lakes Bioenergy Research Center, The Wisconsin Energy Institute, University of Wisconsin-Madison, Madison, WI, United States, [7]VIB-UGent, Metabolomics Core, Ghent, Belgium, [8]Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, [9]VIB-UGent, Center for Plant Systems Biology, Ghent, Belgium

Barley is a major cereal crop for temperate climates, and a diploid genetic model for polyploid wheat. Cereal straw biomass is an attractive source of feedstock for green technologies but lignin, a key determinant of feedstock recalcitrance, complicates bio-conversion processes. However, manipulating lignin content to improve the conversion process could negatively affect agronomic traits. An alternative approach is to manipulate lignin composition which influences the physical and chemical properties of straw. This study validates the function of a barley ferulate 5-hydroxylase gene and demonstrates that its downregulation using the RNA-interference approach substantially impacts lignin composition. We identified five barley genes having putative ferulate 5-hydroxylase activity. Downregulation of *HvF5H1* substantially reduced the lignin syringyl/guaiacyl (S/G) ratio in straw while the lignin content, straw mechanical properties, plant growth habit, and grain characteristics all remained unaffected. Metabolic profiling revealed significant changes in the abundance of 173 features in the *HvF5H1*-RNAi lines. The drastic changes in the lignin polymer of transgenic lines highlight the plasticity of barley lignification processes and the associated potential for

manipulating and improving lignocellulosic biomass as a feedstock for green technologies. On the other hand, our results highlight some differences between the lignin biosynthetic pathway in barley, a temperate climate grass, and the warm climate grass, rice, and underscore potential diversity in the lignin biosynthetic pathways in grasses.

KEYWORDS

ferulate 5-hydroxylase, lignin, barley, RNAi, S/G, straw, biosynthetic pathway, bioeconomy

1 Introduction

Lignocellulosic feedstocks are attractive commodities for producing a variety of bio-products due to their renewable nature and smaller environmental footprint compared to petrochemicals. However, conversion of lignocellulosic biomass into high value bio-products and second-generation biofuels is challenging, partly due to the presence, heterogeneity, and structural complexity of lignin. This has led to a focus on researching the biosynthesis, properties, and manipulation of lignin to increase both fundamental understanding, and to explore opportunities for biomass improvement.

Lignin is produced *via* the phenylpropanoid pathway which starts with the enzyme phenylalanine ammonia-lyase (PAL) and ultimately synthesizes the H (*p*-hydroxyphenyl), G (guaiacyl), and S (syringyl) monolignol units. Lignification begins when these monolignols are oxidised by cell-wall laccases and/or peroxidases resulting in radicals that are combinatorially coupled through ether or carbon-carbon bonds leading to assembly of lignin polymers. The outcome of the polymerization is heavily influenced by the type and amount of phenolic monomers that are translocated into the cell wall and that are competent for free-radical polymerization (for review, see Tobimatsu et al., 2013). Grass lignin is mainly made up of S and G units, with H units constituting a small but significant percentage. Lignin composition influences the physical and chemical properties of biomass feedstocks and has been one of the key targets of cell wall engineering studies (Halpin, 2019).

Recent research has highlighted some of the unique features of the lignin pathway in grasses compared to dicots. In dicots, lignin biosynthesis starts with the deamination of phenylalanine (Phe) into cinnamic acid and subsequent conversion to *p*-coumaric acid (*p*CA) *via* cinnamate 4-hydroxylase (C4H) activity. Grasses are able to synthesize *p*CA from Phe or tyrosine (Tyr) using bifunctional phenylalanine/tyrosine ammonia-lyases (PTAL), with the TAL activity contributing to nearly half of the cell wall total lignin in Brachypodium (Barros et al., 2016). The metabolites derived from Tyr seem to be preferentially incorporated into S lignin units and into cell-wall-bound *p*CA, another characteristic feature of grass cell walls (Withers et al., 2012; Barros et al., 2016). Some of that wall-bound *p*CA is attached to lignin through the involvement of *p*-coumaroyl-CoA:monolignol acyltransferase (PMT) in the acylation of a proportion of monolignols to produce *p*-coumaroylated S monolignol conjugates (Marita et al., 2014; Petrik et al., 2014). The conventional lignin biosynthetic pathway enzyme, ferulate 5-hydroxylase (F5H, also known as coniferaldehyde 5-hydroxylase or Cald5H), is essential to S lignin biosynthesis in dicots, but relatively few studies have examined its role in grasses. Recent work in rice has suggested that F5H may only catalyse production of the nonacylated portion of lignins in grasses (Takeda et al., 2019a). This has led to the hypothesis that a grass specific lignin pathway may exist for the production of *p*-coumaroylated S lignin which is independent of the co-existing conventional lignin biosynthesis pathway (Takeda et al., 2019a; Barros and Dixon, 2020). However, the genes and enzymes of the hypothetical grass specific lignin pathway remain to be identified and it is still unclear whether this pathway, if it exists, is a feature of lignin biosynthesis unique to rice or common to other grasses. Consequently, the role of F5H in dictating the overall S lignin content of grasses is still an open question.

The proportions of G and S units in lignin critically influence the structure of the polymer and the ease of enzymatic biomass processing to release sugars (saccharification) for downstream uses. In the conventional S lignin biosynthetic pathway, the C5 position on the G-lignin precursor units (coniferaldehyde or coniferyl alcohol) is hydroxylated and methylated by F5H and caffeic acid *O*-methyltransferase (COMT) enzymes, respectively. *F5H* knockout plants of Arabidopsis lack S-lignins (Chapple et al., 1992). The availability of C5 on G-lignin precursors can lead to pretreatment-recalcitrant β–5 or 5–5' linkages within lignin and trigger inter-chain connections during lignin polymerization, such that G-rich lignins are more chemically complicated with larger molecular weights and higher melting points compared to S rich lignins (Vanholme et al., 2010; Ralph et al., 2019). Structurally, a lignin polymer that is extreme in its content of S units, methoxylated at both C3 and C5, is mainly

composed of more homogeneous linear chains, which are shorter relative to H- or G-rich lignins and have a high proportion of chemically-labile β–*O*–4 bonds (Stewart et al., 2009; Vanholme et al., 2010). Many studies in various dicot species in which F5H expression has been manipulated illustrate that reduced *F5H* expression reduces S lignin units and can lower sugar release after biomass pretreatment, whereas increased *F5H* expression increases S lignin units and may improve sugar release (e.g., Ciesielski et al., 2014; Fan et al., 2020). However, the few studies in grasses that link F5H expression and S lignin content to saccharification efficiency have yielded contradictory results. Research performed on Brachypodium yielded results consistent with those from dicots (Sibout et al., 2017), whereas work on switchgrass revealed no change to saccharification on F5H up- or down-regulation (Wu et al., 2019), and work on sugarcane surprisingly indicated that reduced F5H expression and reduced S units in lignin might increase sugar release upon saccharification after mild acid pretreatment (Bewg et al., 2016). The roles of F5H and S lignin content in influencing saccharification and biomass processing in grasses after various pretreatments therefore remains an open question and it is possible that not all grasses behave in the same way.

In order to clarify both the role of F5H in influencing the S lignin content of grass lignins, and to explore the impact of *F5H* downregulation on saccharification efficiency, research is needed in a wider variety of grasses. Barley (*Hordeum vulgare*) is the fourth largest cereal crop globally and a close relative of wheat (*Triticum aestivum*), the second most important global crop. Together, barley and wheat crops are the greatest source of straw biomass in temperate regions. In this study, we identified five *F5H* genes in barley and by RNAi downregulation we validated that one of these genes, *HvF5H1*, plays a dominant role in S lignin production and in determining the lignin S/G ratio in straw. Subsequently, we studied the impact of *F5H* downregulation on saccharification recalcitrance, agronomic traits, and straw mechanical properties. Our data suggest differences in overall S lignin pathways between barley and what has been reported for rice and highlights potential diversity in lignin biosynthesis in grasses with implications for straw improvement strategies.

2 Results

2.1 Barley *F5H* gene discovery

F5H is a cytochrome P450 monooxygenase enzyme and a member of the CYP84 family (Meyer et al., 1998). Putative barley F5Hs were shortlisted in two steps. First, we blast-searched the barley whole-genome Morex assembly and used multiple sequence alignment to identify proteins that had all the conserved structures and signature motifs in F5Hs as described previously (Meyer et al., 1996; Bak et al., 2011). Next, using phylogenetic analysis, candidate sequences from barley were compared with functionally characterised *F5H* genes in Arabidopsis (Meyer et al., 1998) and grasses including rice (Takeda et al., 2017), switchgrass (Wu et al., 2019), and sugarcane (Bewg et al., 2016). This led to identification of five putative *F5H* genes in barley, named *HvF5H1-5* (HORVU1Hr1G047220, HORVU3Hr1G014930, HORVU2Hr1G109440, HORVU7Hr1G084500, and HORVU2Hr1G092360, respectively). The derived amino acid sequences of these genes are shown in Figure S1.

A phylogenetic tree was constructed using the Maximum Likelihood method and a JTT matrix-based model that shows the relationships between these five putative barley F5H amino acid sequences and those from other C3 and C4 grasses (Figure 1). Spike moss (*Selaginella moellenodorffii*) smF5H was brought into this analysis as an outlier (Weng et al., 2008). The phylogenetic tree shows that the majority of putative F5H proteins in grasses group closely together, forming a large cluster that encompasses all previously functionally characterised grass F5Hs (Figure 1 red clade). This cluster includes protein sequences for barley *HvF5H1*, *HvF5H2*, and *HvF5H3* genes, hypothetically supporting a direct role for these in lignin biosynthesis. The presence of three barley F5Hs in this clade represents an expansion of the gene family in barley compared to some other grasses that have only one F5H in the clade (e.g., Brachypodium and rice). By comparison, wheat has two or three TaF5H counterparts for each barley sequence representing the three ancestral genomes that compose modern hexaploid wheat. Collectively these data suggest duplications of the *F5H* gene in barley and wheat and possibly all Triticeae after their divergence from the Brachypodieae lineage. Predictably, *Sm*F5H and *At*FAH1 proteins appeared to be less closely related to the grass F5Hs and are located in isolated distant clades. We obtained similar phylogenetic results using the UPGMA (Unweighted Pair Group Method with Arithmetic mean) and Neighbour Joining methods.

The phylogenetic tree also encompasses five other small clades of amino acid sequences that carry F5H signature motifs but have unknown functions potentially distinct from lignin biosynthesis. The *OsF5H2* (CYP84A6) gene was reported to be expressed in rice ovary 3 days after flowering (Takeda et al., 2017) and in mature leaves (Kim et al., 2006) but is not considered to be an essential gene in lignifying tissues (Takeda et al., 2017). Similarly, in Arabidopsis, duplication and neofunctionalisation of *AtFAH1* has produced a paralogous gene, *AtCYP84A4*, of which the corresponding enzyme has a widely altered catalytic activity in α-pyrone metabolite biosynthesis (Weng et al., 2012).

The *Hv*F5Hs show 60-65% amino acid sequence identity to *At*FAH1 and 65-81% identity to orthologs in the Poaceae family. Out of the five identified F5H homologs in barley, *Hv*F5H1 shared the highest amino acid sequence identity to the

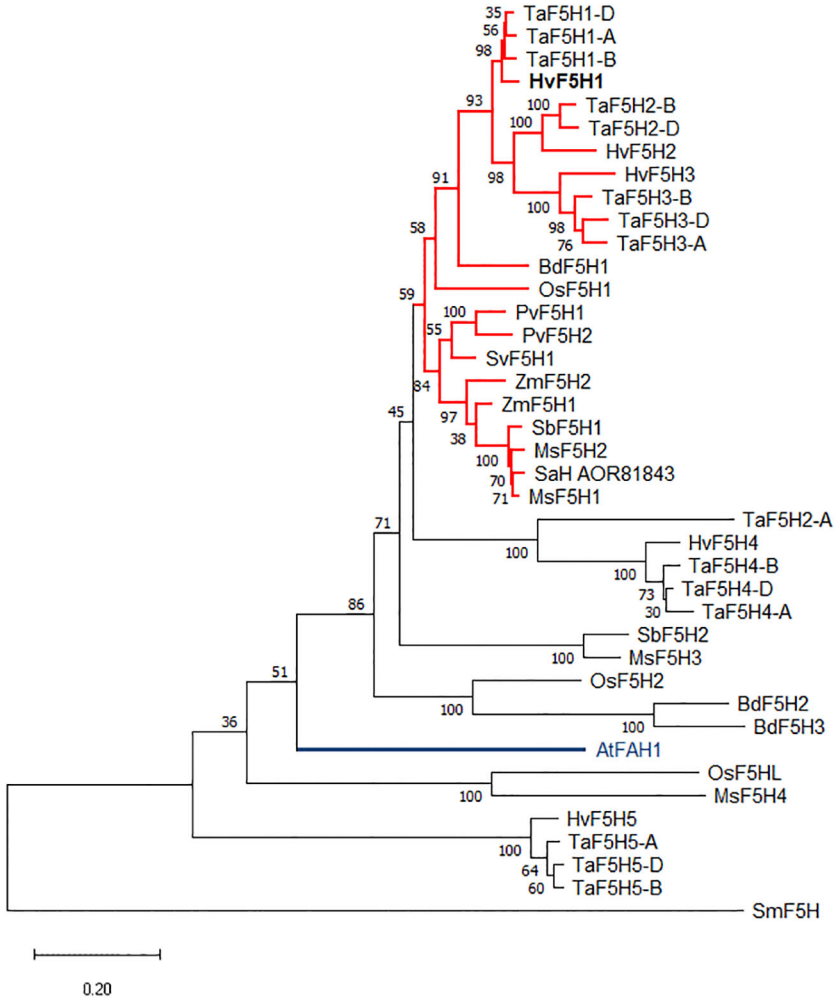


FIGURE 1
The rooted phylogenetic tree was constructed using the Maximum Likelihood method and JTT matrix-based model, with the F5H protein sequences from *Selaginella moellenodorffii* (Sm), barley (Hv), wheat (Ta), rice (Os), maize (Zm), Brachypodium (Bd), *Miscanthus* (Ms), sugarcane (SaH), sorghum (Sb), switchgrass (Pv), green foxtail (Sv), and Arabidopsis (At). The Arabidopsis *AtFAH1* is highlighted in blue. All functionally characterised grass F5H proteins with activity in lignin biosynthesis belong to the red coloured clade although not all the proteins in this clade have been characterised. Bootstrapping values from 100 trials are shown next to the branches. The tree is drawn to scale and the scale bar represents 0.2 amino acid substitutions per site. See Table S1 for the accession numbers and further information.

functionally characterised F5Hs from Arabidopsis (*At*FAH1, 65%), rice (*Os*F5H1, 80.7%), switchgrass (*Pv*F5H1, 80.8%), and sugarcane (*SaH*_AOR81843, 81%). Furthermore, barley RNAseq data (Mascher et al., 2017) revealed that the expression pattern of *HvF5H1* is characteristically different from the other *HvF5H*s (Figure S2). *HvF5H1* is the only *HvF5H* expressed to substantial levels with the highest transcript abundances found in lignifying tissues as well as leaves (Figure S2). This is particularly important as it demonstrates *HvF5H1* is highly active in lignocellulosic biomass but not in the grain or inflorescence tissues. *HvF5H3* shows some expression only in embryo and roots, whereas *HvF5H4* expression was found only in roots. Very limited expression was detected for *HvF5H2* or *HvF5H5* in any of the 16 tissues presented in the dataset. The expression data combined with phylogenetic analysis clearly suggest the *HvF5H1* gene as the strongest candidate for an involvement in lignin biosynthesis.

2.2 Downregulating *HvF5H1* in barley

To manipulate lignin composition and structure and investigate the influence of *Hv*F5H1 on S unit biosynthesis, we used the RNAi approach to produce transgenic plants with suppressed *HvF5H1* expression. A fragment of 669 bp (Figure S3) was used to produce a hairpin inverted repeat sequence in the plasmid pBract207 under the control of the constitutive maize ubiquitin promoter. Subsequently, 23 T0 independent transgenic plants were

produced following the transformation of the Golden Promise cultivar (Figure S4A). Southern blot analysis identified plants with a single T-DNA locus and qPCR was employed to assay plants for reduced *HvF5H1* expression relative to wild-type plants (Figure S4B). The qPCR was carried out on samples of second internode taken at growth stage 32 on the Zadoks scale (Zadoks et al., 1974). Of 10 T0 plants that had a single insertion site of the T-DNA locus, nine also exhibited at least 60% reduction in *HvF5H1* expression compared to the wild type and greater than 67% reduction compared to the empty vector (EV) plants; the best four lines showed approximately 95% or more reduction in *HvF5H1* expression. From the pool of nine T0 plants, we selected seven lines to develop into the T1 generation. These lines were named B, I, K, M, T, W, Y. T1 generation plants were screened to identify *HvF5H1*-RNAi homozygote and azygote plants that had lost the transgene through segregation. Consequently, we selected three independent lines B, T, and W, to analyse further in the T2 generation. For this purpose, homozygous RNAi T2 plants along with control lines and wild type were grown in the greenhouse in a Randomised Complete Block design with five replications. Levels of *HvF5H1* transcript reduction were evaluated to be 90% for line B, 88% for line W and 93% for line T (Figure S5). Using Mäule staining, the second internodes of the T2 plants were investigated for a change in S lignin relative to control plants. Whereas all control plants (wild type, empty-vector controls, and azygote plants that had lost the RNAi transgene through segregation) showed bright red staining in lignified tissues, in the RNAi lines the vascular bundles and cortical fibre tissues exhibited more of a brown colour, indicating a reduction in the proportion of S lignin (Figure S6).

2.3 Altered lignin composition in downregulated plants

To investigate the consequence of suppressed *HvF5H1* expression on lignin content and composition/structure, extract-free cell wall residue (CWR) was prepared from the main tiller of fully-senesced T2 generation RNAi homozygote plants and related control plants. We removed leaf sheaths from the tillers as their cell wall structure and lignin content differs from culms. Klason lignin determinations on tiller CWR showed no significant differences in Klason lignin content between the *HvF5H1*-RNAi plants and the controls (Figure 2A).

Analytical thioacidolysis and 2D NMR (HSQC) were employed to study lignin composition of *HvF5H1* RNAi lines in more detail. Thioacidolysis degradation releases H, G and S

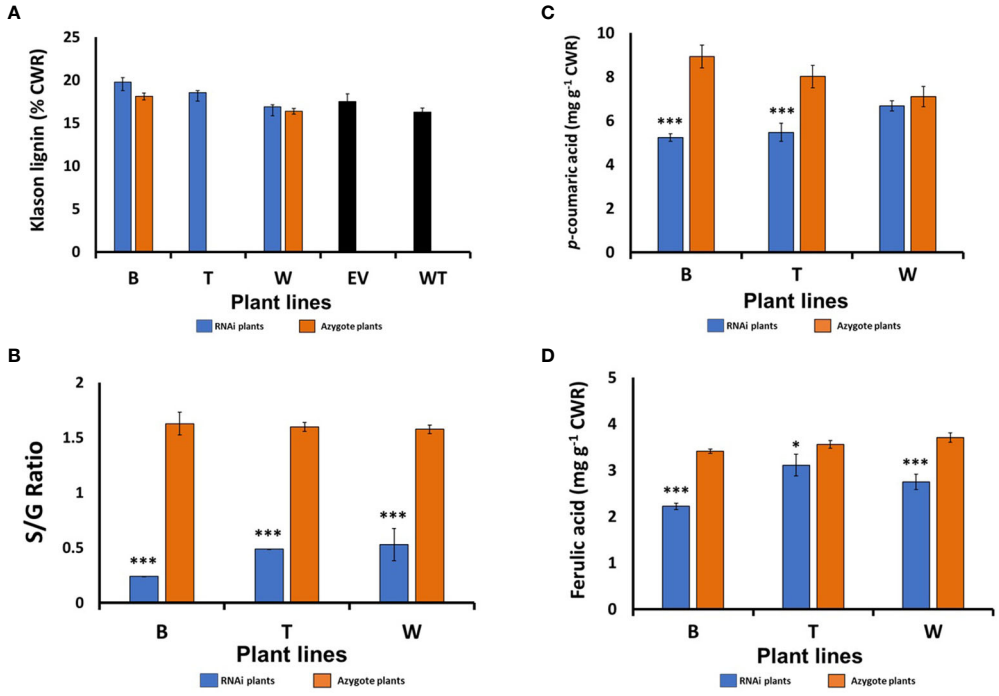


FIGURE 2
Cell wall compositional analysis of the senesced barley straw from the T2 generation *HvF5H1*-RNAi lines and controls: **(A)** Klason lignin content, **(B)** S/G ratio, **(C)** Ester-linked *p*-coumaric acid (*p*CA), **(D)** ester-linked ferulic acid (FA), as measured by mild alkaline hydrolysis. Lines marked with * & *** are significantly different from their controls which were the respective azygote plants in the case of lines B and W; for line T where no azygous controls were available, comparisons were made with empty vector (EV) plants for Klason lignin or with all azygote plants for S/G ratio and ester-linked *p*CA and FA, Student's t-test $P<0.05$ and $P<0.001$, respectively for * and ***. The error bars represent standard errors between 3-7 biological replicates (i.e. independent plants). B, T, and W; RNAi lines; EV: Empty vector; WT: wild type.

thioethylated monomers only from H, G and S lignin units linked by β-aryl ether (β–O–4) bonds to reflect relative abundance of these monomers (Lapierre et al., 1986). The thioacidolysis yield (in μmol/g Klason lignin) was reduced by up to 27% in the *HvF5H1*-RNAi lines indicating a reduction in the proportion of units linked by β-aryl ether bonds. The S/G ratio of thioacidolysis released monomers in *HvF5H1*-RNAi lines decreased by up to 85% from 1.62 in line B azygotes to 0.24 in the corresponding RNAi plants ($P<0.01$; Figure 2B) reflecting a shift from 60% S units and 37% G units in lignin in azygote plants to 19% S units and 78% G units in lignin in RNAi plants (Figure S7). Changes in RNAi lines W and T were similar but slightly less severe. These findings are in line with F5H's role in hydroxylating G-lignin precursor units (coniferaldehyde or coniferyl alcohol) to produce S lignin monomers. Our analyses also revealed a significant reduction ($P < 0.01$) of 28% and 41% in the amount of ester-linked *p*CA released by mild alkaline hydrolysis in *HvF5H1*-RNAi lines T and B respectively (Figure 2C), although the 6% reduction in line W was not significant. Ester-linked FA was also significantly reduced in all RNAi lines (Figure 2D). Reductions in ester-linked *p*CA are consistent with the reduction of S lignin given that much of the cell wall *p*CA in grasses is ester-linked to S units (Ralph et al., 1994).

A 2D-NMR (HSQC) analysis of line B cell wall material further verified the lignin compositional changes in the *HvF5H1*-RNAi plants that had been demonstrated through analytical thioacidolysis (Figures 3A, B). The greatest alterations were observed in S and G units. The G units in the RNAi line increased proportionally to 86% compared to 44% in the control. On the other hand, suppression of *HvF5H1* led to reduction of S unit content to 12% in the RNAi plant compared to 53% in the control plant. The H-unit content slightly decreased in the RNAi plant (2%) relative to control (3%). In agreement with thioacidolysis, NMR detected 12% of *p*CA in the RNAi plant versus 20% in the control and 4% tricin in the RNAi plant versus 3% in the control; such mobile end-units are, however, over-quantified by NMR (Mansfield et al., 2012).

The aliphatic subregion of the 2D-NMR spectra shows the signals that are derived from interunit linkages in the CWR lignin polymer (Figures 3C, D). In both RNAi and control plants, the dominant units are β-aryl ethers followed by substantially lower percentages of phenylcoumaran (β–5) and resinol (β–β). The proportion of β-aryl ether bonds, which are the main intermolecular linkages in native lignin, were reduced from 93% in control to 83% in the RNAi line, consistent with the thioacidolysis yield reduction which also indicated reduced β-aryl ether linkages. On the other hand, the share of

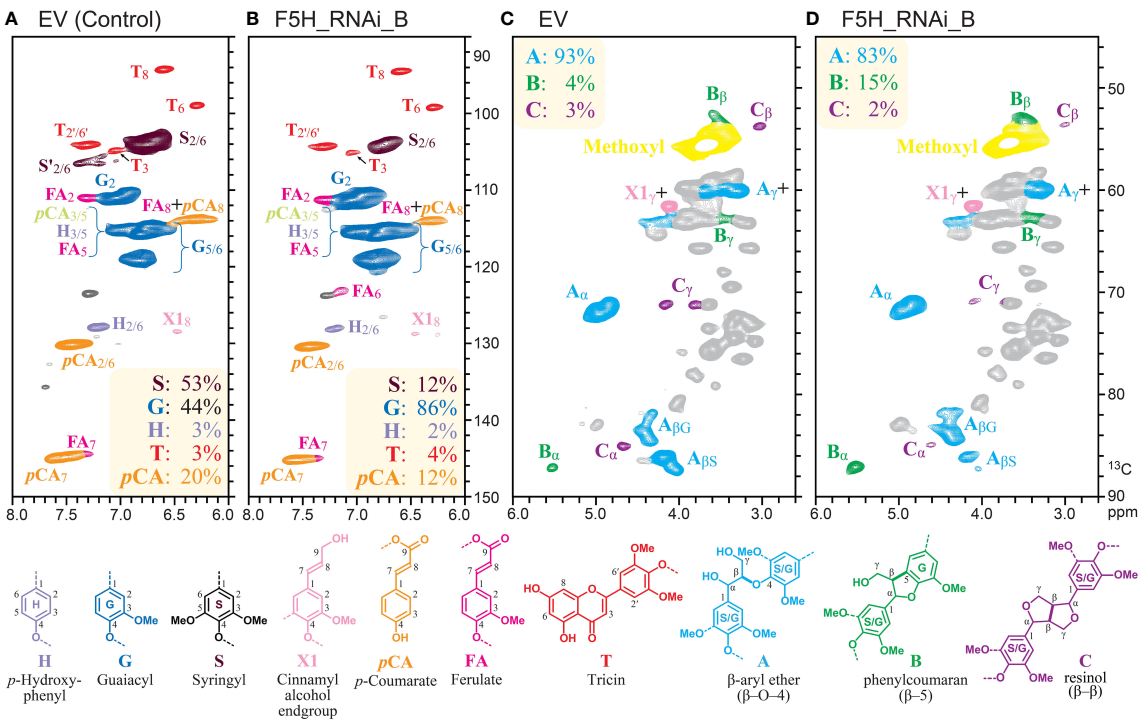


FIGURE 3
Aromatic and aliphatic subregions of 2D-NMR (HSQC) spectra of stem cell walls from *HvF5H1*-RNAi plants (T2 generation) and EV. Summary data in the boxes color-coded to match their respective assignments, and represent percentages of; **(A, B)** major lignin subunits to the total amount of H, G, and S units; and **(C, D)** characterised inter-monomeric bonds to the total number of the lignin interunit linkages.

phenylcoumaran linkages in the RNAi line increased to 15% compared to 4% in the control. There was little change in the proportion of resinol bonds. These data are all consistent with the higher proportion of G units in the *HvF5H1*-downregulated plants.

2.4 Metabolic profiling of *HvF5H1* RNAi lines

To study the impact of *HvF5H1* suppression on intracellular phenolic metabolites, we sampled the first three internodes of plants from two *HvF5H1*-RNAi lines (lines B and W) to compare with wild-type and EV control plants. UHPLC-MS was employed to analyse phenolic metabolites extracted from the internodes. Initially, our analysis detected 7486 features with an abundance of more than 100 counts in at least one sample. Features were selected for further investigation if their abundance was significantly (P<0.01) different from the controls for both *HvF5H1*-RNAi transgenic lines. This led to the identification of 173 features (i.e., *m/z* traces) with at least three-fold change in both RNAi lines compared to control values. Of these, 95 features showed higher intensity in the *Hv*F5H1 RNAi plants whereas 78 *m/z* traces exhibited lower intensity (Table S2). Taking into account the accurate *m/z*, retention time, and MS/MS fragmentation, we characterised 47 metabolites. The 21 identified metabolites with higher intensity in the RNAi lines were mainly small oligomers composed of G units only, or in combination with units derived from ferulic acid or tricin. In contrast, the metabolites with reduced abundance were sinapyl alcohol, three conjugates with syringic acid, and 22 di- and tri-lignols containing at least one S unit.

2.5 Transcriptomic profiling in *HvF5H1*-RNAi lines

We performed a transcript profiling in order to evaluate the effect of *HvF5H1* downregulation on the expression of other genes in stems. For this purpose, the first three internodes were sampled from two *HvF5H1*-RNAi lines (lines B and W) and compared with wild-type and EV control. To identify differentially regulated genes, expression data was first log-normalised and the genes that were significantly (P<0.01) different between the *HvF5H1*-RNAi lines and both controls were identified. Then we shortlisted genes that were up- or down-regulated greater than three-fold. In the process, probes that had a significant (P<0.001) expression difference between the wild-type and empty vector controls were excluded due to lack of consistency. Subsequently, fourteen genes were identified as upregulated and four genes as downregulated (Table S3) in the RNAi lines compared to controls. *HvF5H1*, the targeted gene, was downregulated 4.7-fold. Of the other *HvF5H* genes, *HvF5H2* and *HvF5H5* expression was not detected by transcript profiling and expression of both *HvF5H3* and *HvF5H4* remained unchanged. The top-downregulated gene (HORVU2Hr1G052170) belonged to the alpha/beta hydrolase superfamily and was downregulated about 7-fold.

The most differentially overexpressed genes in the RNAi lines were a DNA/RNA helicase which is involved in meiotic crossover formation (HORVU4Hr1G049830, 22-fold), and a hydroxyjasmonate sulfotransferase (HORVU2Hr1G117700, 15-fold). An NBS LRR disease resistance gene along with a MYB transcription factor were upregulated 12.4 and 6.6-fold, respectively, in the *HvF5H1*-RNAi lines.

2.6 *HvF5H1* downregulation did not affect growth habit or straw characteristics

Given the considerable modification in lignin structure of the *HvF5H1*-RNAi plants, we phenotyped them for a variety of agronomic traits to determine whether such dramatic changes to lignin might impact directly on crop productivity. Visually, the *HvF5H1*-RNAi lines did not exhibit noticeable phenotypic differences compared to the control plants (Figure 4A). Plants were phenotyped for shoot height, tiller number, straw biomass, spike number, and seed characteristics including seed length, width, and area and thousand-grain weight (TGW) (Table S4). Employing ANOVA, none of these traits in the RNAi lines was significantly different from the ones in control plants (P<0.05).

Lignin modification could impact straw strength and potentially the plant's resilience or lodging resistance, with indirect effects on productivity and yield. We therefore tested the mechanical properties of stems sampled from the RNAi lines at harvest time by quantifying the maximum stretching force straw can tolerate before it breaks, i.e. the 'Flexural Strength Stress at Yield'. For this purpose, we sampled second and fourth internodes from the main tiller and employed an Instron Universal Testing System for precision phenotyping. Predictably, in each genotype significantly (P<0.05) less force was required to break the fourth internode compared to the corresponding second internode. However, we did not find a significant difference between the RNAi lines and the controls (Figure 4C, Table S5) when either the second or fourth internodes were analysed, suggesting that the S/G ratio reduction did not impact the mechanical strength of barley straw.

2.7 Saccharification of *Hv*F5H1-RNAi lines

To investigate the consequence of lignin structural changes on straw recalcitrance, we carried out saccharification assays

10.3389/fpls.2022.1125003

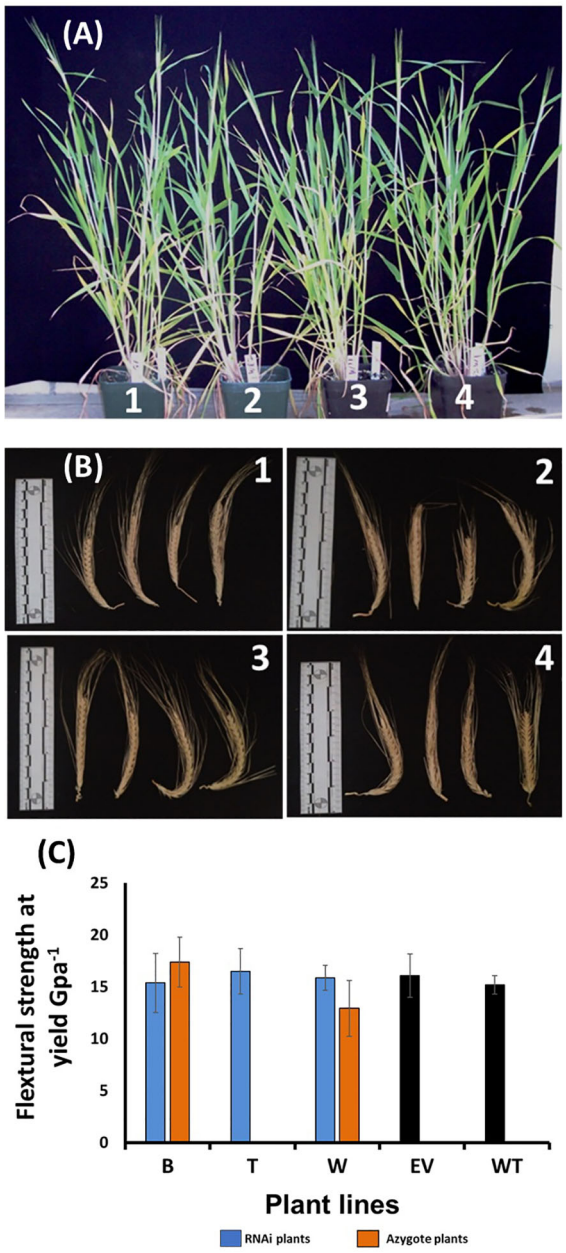


FIGURE 4
Phenotypic characterization of the *HvF5H1*-RNAi lines and controls from the T2 generation; **(A)** plant stature and **(B)** spikes at harvest time, where 1,2,3,4 are wild type, empty vector, azygote and *HvF5H1* RNAi plants respectively; **(C)** mechanical strength of the second and fourth internodes sampled from main tillers of the ripe senesced straw; No significant differences were identified ($P < 0.05$) following Student's t-test. B, T, W are RNAi lines and B and W have corresponding azygote lines. For line T where no azygotes were available, empty vector and wild type plants were used as controls. The error bars represent standard errors between biological triplicates (i.e. independent plants).

after subjecting the straws to hot water pretreatment. The enzymatic digestion phase lasted 96 hours and samples were collected at 4, 24, 48 and 96 hours. However, there was not a significant improvement or deterioration in saccharification efficiency when each RNAi line was compared with its most relevant controls following Student's t-test ($P < 0.05$) (Figure 5).

3 Discussion

This study demonstrates that expression of *HvF5H1* plays a major role in determining the proportions of S and G units in lignin in barley culm tissues. Transgenic plants with downregulated *HvF5H1* expression have substantially reduced frequency of S units in culm lignin and a correlatively high enrichment in G units compared to control plants. In the RNAi line with the largest shift in lignin composition, S units represented only 19% of the thioacidolysis released units as compared to 60% in the azygote control plants, representing a relative decrease of 68% in S units. Correspondingly, G units increased from 37% of the units released from azygote plants to 78% of the released units in RNAi plants, a relative increase of 111%. Concomitant with the %S decrease and %G increase, the thioacidolysis yield was reduced, which is consistent with the higher proclivity of G units to be involved in resistant interunit bonds such as phenylcoumaran or biphenyl linkages that thioacidolysis does not release monomers from. Our 2D NMR analysis independently confirmed the thioacidolysis lignin data revealing a 73% reduction in S units and a 95% increase in G units overall. Unlike thioacidolysis which releases monomers from lignin units only involved in β-aryl ether units (linked by their characteristic β–O–4-ether bonds), 2D-NMR on unfractionated cell walls can provide information on additional lignin linkages and showed that, although the proportion of β-ether units decreased from 93% in the wild type to 83% in the RNAi line, phenylcoumaran linkages proportionally increased significantly, from 4% in the wild type to 15% in the RNAi line. There was little change in resinol linkages in the RNAi line. These changes to lignin on *HvF5H1* suppression are very consistent with what has previously been seen in dicot species (e.g., Meyer et al., 1998; Reddy et al., 2005) and converse changes (i.e., an increase in the proportion of S units) have been seen when *F5H* is over-expressed (e.g., Meyer et al., 1998; Franke et al., 2000; Stewart et al., 2009). Our analyses also showed significant reductions in ester-linked *p*CA in two *HvF5H1* RNAi lines. Mild alkaline hydrolysis indicated a 28 and 41% reduction in the amount of ester-linked *p*CA in *HvF5H*1-RNAi lines T and B respectively. As much of the cell wall *p*CA in grasses is ester-linked to S units in the lignin, a reduction in esterified *p*CA when S units are reduced is not unexpected, although the proportion of *p*CA that is ester-linked to arabinoxylans would be retained. Again, 2D-

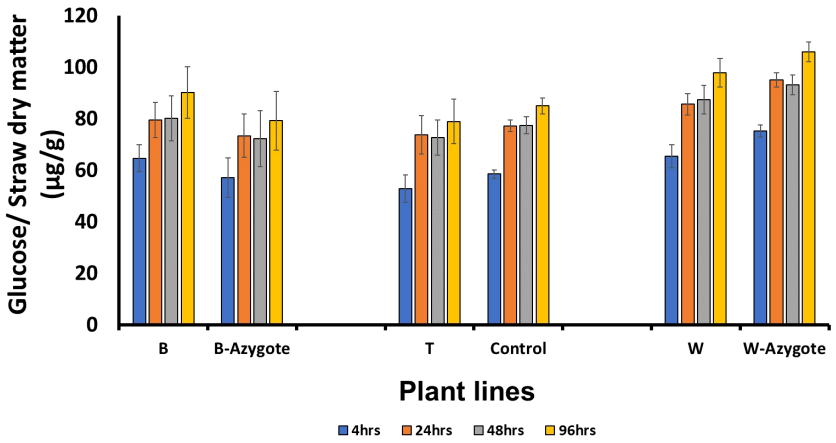


FIGURE 5
Saccharification following hot water pretreatment. No significant differences ($P < 0.05$) were identified following Student's t-test. B, T, W are RNAi lines; B-azygote, W-azygote are corresponding azygote lines; for line T where no azygotes were available, empty vector and wild type plants were used as controls. The error bars represent standard errors between biological triplicates (i.e. independent plants).

NMR confirmed a *p*CA reduction of approximately 40% suggesting that the impact of *HvF5H1* suppression is not restricted to the nonacylated portion of lignin as in rice (Takeda et al., 2019a), but indicates that the *p*-coumaroylated S lignin has been reduced.

In contrast to the many studies on F5H manipulation in dicots, only a few studies address the impact of F5H suppression in grasses. In sugarcane, *F5H* was suppressed by RNAi to varying degrees although 16-94% of control expression remained in four transgenic lines analysed (Bewg et al., 2016). Nevertheless, the best transgenic line had an altered S:G ratio of 48:52 compared to 61:39 in controls, i.e., a 21% decrease in the proportion of S units and a 33% increase in the proportion of G units, changes completely consistent with but less extreme than what we observed in barley. The NMR analysis also detected phenylcoumaran (β–5-linked) units in this transgenic plant but not in the controls (Bewg et al., 2016). Similarly, in *F5H*-RNAi suppressed switchgrass (Wu et al., 2019), plants with approximately 55% reduction in *F5H* expression, by our calculations had an altered S:G ratio of about 32:68 compared to 46:54 in controls, i.e., a 30% decrease in the proportion of S units and a 26% increase in the proportion of G units. In an initial study in rice, *F5H1* suppression by RNAi eliminated over 90% *F5H1* expression but S:G ratio determined *via* thioacidolysis was only reduced from 36:64 in control to 26:74 in the RNAi lines, i.e., a relative decrease of S units by 28% according to Takeda et al. (2017). To determine whether the high proportion of remaining S units was due to residual F5H1 activity, or to the activity of alternative unidentified hydroxylases, the same research group knocked out rice *F5H1* using CRISPR/Cas9 (Takeda et al., 2019a). Surprisingly, plants with presumed total knock-out of *F5H1* still produced considerable amounts of S units in the culm, to approximately 62–70% of the level of wild-type plants. This is very unexpected as an F5H knock-out mutant in Arabidopsis, the *fah1-2* mutant, produces no S lignin (Chapple et al., 1992). Moreover, DFRC lignin analysis of the rice *F5H1* CRISPR knock-out indicated that the changes in G and S units were restricted to the non-*p*-coumaroylated lignin units and units with *p*CA attached were almost unchanged (Takeda et al., 2019a). This discovery prompted the authors and others to propose that rice and possibly grasses in general must have two parallel pathways for making S lignin, only one of which requires F5H1 whereas monolignol units to which *p*CA is attached are produced by an undiscovered alternative pathway (Takeda et al., 2019a; Barros and Dixon, 2020).

Our results throw some of the details of this hypothesis into question at least for barley. Only *HvF5H1* was downregulated in our experiments and transcriptomic data from RNAi plants confirmed that expression of other *HvF5Hs* did not change. Nevertheless, we achieved a substantial decrease of 68-73% in S units using RNAi compared to in rice where there were only 32% and 38% S unit reductions using RNAi and CRISPR knock-outs respectively. The approximately 30% of S units remaining in our barley plants could be explained by the low level of *HvF5H1* expression that remains. The levels of esterified *p*CA in cell walls was also clearly reduced in barley by approximately 30–40% consistent with the reduction in S units and the fact that most *p*CA esterified to grass lignin is found on S units (although a significant proportion is found esterified to arabinoxylan). Consequently, although we cannot conclude that an alternative S lignin pathway does not exist in barley, our results can be explained based on the conventional lignin pathway alone. Taking into consideration the scale of intermolecular bond alterations, the high impact of *HvF5H1* RNAi downregulation

on S and G lignin, and the significant reduction of *p*CA, *HvF5H1* is a major gene for S lignin biosynthesis in barley and the conventional lignin pathway clearly dominates.

The changes in lignin in our *HvF5H1*-RNAi plants appear to have no impact on saccharification. This is consistent with most other work in the literature that either shows no change, or a reduction in saccharification, in *F5H*-suppressed plants, and an increase in saccharification when F5H is overexpressed and S lignin proportion increased. Several papers demonstrate that the predominantly G-lignin in the Arabidopsis *fah1* F5H mutant impedes saccharification compared to wild-type lignin after liquid hot water pretreatment (Li et al., 2010) or maleic acid treatment (Ciesielski et al., 2014), although weak alkaline pretreatment reveals no change to saccharification (Smith et al., 2022). On the other hand, *Atf5h1* T-DNA insertion mutants allelic to *fah1* did show an improved saccharification when no pretreatment was used (van Acker et al., 2013). In *Populus tomentosa*, downregulating *F5H* expression by manipulating miR6443, a microRNA that regulates S lignin biosynthesis, increases G units and reduces S units in lignin and lowers sugar yield after saccharification with both alkaline or acidic pretreatments (Fan et al., 2020). Conversely, increasing *F5H* expression increases S lignin and saccharification in Arabidopsis after alkaline pretreatment (Ciesielski et al., 2014; Smith et al., 2022), or maleic acid (Sakamoto et al., 2020), or liquid hot water pretreatment (Li et al., 2010) and in poplar after pretreatments with alkali (Lapierre et al., 2021) or acid/alkali (Fan et al., 2020). In the rice *F5H1* CRISPR knock-out lines, saccharification efficiency was lowered compared to wild type only when using hot water pretreatment and unchanged with dilute acidic or alkaline pretreatments (Takeda et al., 2019b), whereas in switchgrass up- or down-regulation of *F5H* did not affect saccharification (Wu et al., 2019). In Brachypodium, *F5H* overexpression increased S lignin and saccharification without pretreatment (Sibout et al., 2017). In sugarcane, one F5H-suppressed line apparently had increased sugar release after mild acidic pretreatment but others did not (Bewg et al., 2016) so the relevance of this observation which conflicts with the rest of the literature is unclear. Our data showing that barley *HvF5H1*-RNAi plants have increased resistant phenylcoumaran units and reduced esterified *p*CA (which generally occurs as free-phenolic pendent groups on lignin (Ralph et al., 1994; Lu and Ralph, 1999)) suggests a more condensed or more *p*CA-decorated lignin would not necessarily have any beneficial effect on saccharification. Incidentally, in artificial model systems where other factors could be controlled, lignin S/G composition had little effect on wall degradability (Grabber et al., 1997; Grabber et al., 2009).

Downregulation of *HvF5H1 via* RNAi led to changes in the abundance of 173 features. The majority of metabolite features have unknown identities. As detailed in Table S2, 78 compounds were less abundant in the *HvF5H1*-RNAi plants. Predictably, S lignin-related compounds such as sinapyl alcohol and syringic acid 4-*O*-hexoside were among the known compounds shown to be decreased, which is consistent with the reduction of S lignin. The α-oxidized β-ether oligomer of sinapyl alcohol, Sox(8-O-4)S, was 150-fold less abundant in the *HvF5H1*-RNAi plants, followed by several small chains of G-S lignols such as G(8-O-4)S(8-8)S. Sox(8-O-4)S was previously shown also to be reduced 10-fold in abundance in barley *HvCOMT*-RNAi plants (Daly et al., 2019). On the other hand, the abundance of 95 features increased significantly in the *HvF5H1*-RNAi plants. Although the identity of the majority of these features is unknown, we notably found that di- or tri-lignols composed of only G units were increased 10 to 73-fold in the *HvF5H1*-RNAi lines, consistent with the increased abundance of G units in lignin and reduced proportion of S units. This is very different from *HvCOMT*-RNAi lines in which the greatest increased abundance detected involved 5-hydroxyconiferyl alcohol, the product of F5H activity and substrate for subsequent COMT activity, and caffeoyl alcohol, another supposed substrate for COMT (Daly et al., 2019).

Although the G-unit-enriched lignin of *F5H*-suppressed plants may not be improved in digestibility, it may be useful for other purposes; for example the more highly-branched and less oxygenated lignin has a higher fuel-value and improved combustion properties. However, if we are to improve the lignocellulosic biomass of food crops it is crucial that changes in cell wall characteristics do not negatively affect agronomic traits. In this study, downregulation of *HvF5H1* did not lead to changes in grain characteristics or other agronomic traits and the mechanical characteristics of straw in the *HvF5H1*-RNAi lines were similar to the controls. Some previous studies have explored the relationship between lignin structure and biomechanical properties but firm conclusions have proved elusive. In poplar supressed in the lignin biosynthesis gene CAD, lignin content was associated with mechanical stiffness (Özparpucu et al., 2017), but the pronounced changes in lignin composition did not alter wood tensile properties (Özparpucu et al., 2018). However, in wheat, lignin content was not associated with straw breaking force but a lower proportion of S lignin appeared to increase it (Muhammad et al., 2020). In Arabidopsis *fah1* F5H mutants, stiffness was increased compared to wild type in apical and middle stem segments, but flexibility was unaltered between *fah1* and wild type plants (Ménard et al., 2022). Differences in stem architecture between dicots and grasses and indeed between individual species are likely to complicate any simplistic generalized interpretation of the effects of lignin characteristics on plant mechanical properties since even different morphotypes of tracheary element have different lignin chemistries and mechanical properties (Ménard et al., 2022). In our work, it is promising to see that radical changes to lignin composition and structure in HvF5H1 suppressed plants is not accompanied by any change to straw strength. Thus, F5H may represent a good target for manipulating lignin composition while maintaining crop health. Indeed, F5H suppression might improve some

agronomic properties as illustrated recently in *Brassica napus* in which four genes for F5H were knocked out by CRISPR/Cas 9 and this apparently improved resistance to *Sclerotinia* stem rot (Cao et al., 2022).

Our work highlights once again that, despite striking similarities among grasses in terms of their cell wall composition, there may be important differences between species in their lignin biosynthetic pathways. For example, we previously showed that the key lignin biosynthetic gene caffeoyl shikimate esterase (CSE), originally identified in Arabidopsis, only seems to have *bone fide* orthologues in some grasses including rice and switchgrass, but is apparently absent in barley, Brachypodium, and maize (Vanholme et al., 2013; Ha et al., 2016). Here we show that the dominant route to S lignin biosynthesis in barley proceeds though the conventional lignin pathway and F5H, and that the proposed alternative 'grass specific' pathway that apparently exists in rice, is either absent or only a minor route in barley.

4 Materials and methods

4.1 Bioinformatics and phylogenetic analysis

Putative *HvF5H* genes were initially retrieved from a public database (Mayer et al., 2012) and reconfirmed with the latest released version (Mascher et al., 2017). Barley amino acid sequences of *Hv*F5H that showed 40% or more identity to the well-studied *At*FAH1 (Meyer et al., 1996) were selected. Next, the best aligning protein sequences were blasted back into the barley database to pull out putative genes with lower identity to *At*FAH1. Subsequently, all the selected protein sequences were aligned with those of functionally characterised *F5H* genes in rice (Takeda et al., 2017), switchgrass (Wu et al., 2019), sugarcane (Bewg et al., 2016), spike moss (Weng et al., 2008), Brachypodium (Sibout et al., 2017), maize (Andersen et al., 2008) and foxtail millet (Muthamilarasan et al., 2015). In the phylogenetic analysis we also used unpublished F5H sequences that stemmed from bioinformatics screenings. These genes were identified or revisited by blasting against the latest relevant genome database and selecting those containing all five signature motifs of F5Hs as described before (Chapple, 1998).

Phylogenetic analysis and construction of the tree were carried out using MEGA X software (Kumar et al., 2018). Amino acid sequences were aligned using ClustalW and a Neighbor-Joining tree constructed tested by 100 bootstrap replications. Evolutionary distances were estimated using the Poisson model based on the number of amino acid substitutions per site.

Barley genes transcriptomic profiling data from different plant tissues were extracted from the publicly available database (Mayer et al., 2012; Milne et al., 2021).

4.2 Plant materials, growth conditions, and designation of internodes

Barley cultivar Golden Promise was used to generate transgenic plants. All plants were grown in the greenhouse with light and temperature controls. High-pressure sodium lamps provided supplementary lights when necessary. Plants were grown in seven-inch pots filled with standard compost mix (150 L Levingtons M2 compost, 75 L perlite, 500 g Osmocote Plus Mini, 125 g Celcote Certis and 75 g Horti-Wet). Whole above-ground biomass of individual plants was harvested after senescence at the Z9.2 growth stage based on the Zadoks scale (Zadoks et al., 1974). Subsequently, biomass was air-dried at 35° C overnight. Following removing the leaves, second internodes from the main tillers were used to conduct cell wall characterization assays or saccharification. Second internodes were identified as described previously (Tottman, 1987).

4.3 Generation of barley *HvF5H1* RNAi transgenics

A cDNA of *HvF5H1* (*HORVU1Hr1G047220*) was used to amplify a 669 bp fragment using primers containing Gateway AttB sites. Subsequently, the amplicon was cloned into the entry vector using BP clonase II (Invitrogen) and then the fragments were transferred from the entry clone to the RNAi vector, pBract207 (John Innes Centre) using LR Clonase (Invitrogen). The RNAi cassette was transferred into *Agrobacterium tumefaciens* AGL1 containing the helper pSOUP plasmid. Barley transgenic lines were produced by the FunGen Group at The James Hutton Institute through the transformation of embryos of the Golden Promise cultivar (Harwood et al., 2009). The T0 transgenic lines were grown to produce seeds and perform preliminary screenings. The T1 generation was screened using a hygromycin root assay to establish plants' zygosity and the southern blot assay to determine numbers of T-DNA loci. Briefly, twenty seeds were germinated on agar plates with 0.5% phytogel and 100 µg/mL of hygromycin as described by Jacobsen et al. (2006). In subsequent generations, hygromycin assays and PCR was used on germinating progeny from parental plants to identify parents homozygous for the RNAi transgene and those that had lost it through segregation (azygotes).

4.4 Expression analysis in RNAi lines

RNA extraction was performed using TRI-Reagent (Sigma-Aldrich) as described by the manufacturer. The resulting RNA was further cleaned up using the Qiagen RNeasy Mini Kit RNA Cleanup Protocol (Cat No 74536). 500 ng of RNA was used to

synthesise cDNA using qScript (Perfecta). The resulting products contain a supermix that has both OligoDT and random primers for better amplification. This was then amplified using the protocol outlined by the manufacturer. The expression analysis was conducted using the Applied Biosystems StepOne machine and related software. SnRK1 (SNF1-related kinase 1) was selected as the internal control gene and we used primers designed against the 3 different orthologs of the *HvF5H1* gene. Primers and probes were designed (allocated) through the Roche Universal Probe Library version.

4.5 Mäule staining

Freehand stem cross sections were obtained from the second internodes using a No. 11 Scalpel blade. Sections were fixed in 4% glutaraldehyde for 1 h at room temperature, rinsed with water, and immersed in 0.5% (w/v) potassium permanganate solution for 10 min at room temperature. Then they were rinsed with water several times and incubated in 10% hydrochloric acid for 5 min. Following rinsing twice with water, the sample was dipped in concentrated ammonium hydroxide under a fume hood for 2 min and mounted on the slide. A Nikon light microscope was employed to study stem sections. Images were taken using a microscope mounted Lecia DC500 camera.

4.6 Klason lignin and lignin composition

Lignin contents were quantified following modification of a previously described method (Dence, 1992). In summary, whole tillers (excluding sheaths and leaves) of three biological repeats for each line were powdered to 0.5 mm particle sizes using a Retsch ZM-200 centrifugal rotor mill and dried at 35°C overnight. Followed by an extraction process using exhaustive hot water and hot ethanol *via* Soxhlet. Samples of 300 ± 5 mg were then hydrolysed in 5 mL of 72% H_2SO_4 for 2 h. Distilled water (12.5 mL) was added and the samples were kept at room temperature for 1 h before being transferred into 250 mL Schott bottles by washing with 150 mL dH_2O. The samples were incubated at 121°C for 1 h and filtered through a crucible to collect both the solid (acid-insoluble fraction) and the liquid filtrate (acid-soluble fraction). The lignin was collected by vacuum filtration on Whatman GF/C 55 mm filter paper and subsequently washed and dried at 110°C overnight. The lignin content was calculated as a percentage of the dry weight of the starting material. Lignin values were adjusted for ash content by burning the extracted lignin at 545°C in a furnace.

Lignin structure was evaluated by the simplified thioacidolysis procedure as described previously (Lapierre et al., 2021). Ester-linked *p*-coumaric and ferulic units were measured after mild alkaline hydrolysis as described previously (Sibout et al., 2016).

4.7 Cell wall characterization by two-dimensional solution-state NMR

Cell walls were characterised without fractionation using two-dimensional (2D) solution-state NMR (Kim and Ralph, 2010; Mansfield et al., 2012). Straw (2-mm pieces) was pre-ground using a Mixer Mill MM400 (Retsch; 30/s vibrational frequency for 90–120 s). Samples were extracted three times with water, three times with 80% ethanol and once with acetone, then allowed to dry. The pre-ground extracted samples were ball-milled using a Fritsch Planetary micro mill Pulverisette 7 vibrating at 600 rpm with zirconium dioxide (ZrO_2) vessels containing ZrO_2 ball-bearings (10 mm × 10) with 5-min milling and 5-min cooling per milling cycle (cycle number depended on the amount of sample). The ball-milled samples were subjected to digestion (72 h × 2) to obtain 'enzyme lignin' (EL) by Cellulysin® cellulase from *Trichoderma viridae* (Calbiochem), at 35°C in acetate buffer (pH 5.0). The ELs were dissolved into DMSO-d_6/pyridine-d_5 (4:1 v/v) and subjected to NMR on a Bruker Biospin AVANCE-III 700 MHz spectrometer equipped with a 5-mm QCI $^{1}H/^{31}P/^{13}C/^{15}N$ cryoprobe with inverse geometry (proton coil closest to the sample). 2D-^{1}H–^{13}C HSQC spectra were acquired using Bruker's pulse program (hsqcetgpsip2.2). Bruker's Topspin 3.2 (Mac) software was used to process spectra. The central DMSO peak was used as internal references (δ_C: 39.51, δ_H: 2.49 ppm).

4.8 Transcript and metabolite profiling

Five plants from each line were grown for 61 days in a Randomised Block Design (RBD) in a greenhouse with supplementary lighting. The first three internodes of each plant were collected and ground in liquid nitrogen, subsequently 100 mg of the ground stem from each replication was used for transcriptome or metabolite analysis as described before (Daly et al., 2019).

4.9 Straw mechanical properties

Straw was sampled from the RNAi and control lines at harvest time, Zadoks stage 97 (Zadoks et al., 1974), and mechanical properties were analysed by quantifying the Flexural Strength Stress at Yield. Second and fourth internodes from the main tiller were fixed into in an Instron Universal

Testing System (https://www.instron.com/) and the force needed to break the straw was measured.

4.10 Saccharification analyses

A robotic platform (Labman Automation Ltd, UK) was employed for alkaline pre-treatment and saccharification as described before (Oakey et al., 2013). A modified version was used for hot water pre-treatment and subsequent saccharification. Briefly, 7.0 ± 0.5 mg extracted pulverised sample was pre-treated with hot water in the autoclave at 130°C for 30 min. The solid was washed 3× with deionised water. The samples were then incubated while shaking at 50°C for 96 h in the presence of an enzyme cocktail (4:1 ratio of Celluclast and Novozyme 188). The GOPOD assay kit (K-GLUC) (Megazyme, Ireland) was used to quantify the glucose released during saccharification.

4.11 Statistical analysis

Data were collected as described for each trait. Where required, a square root transformation of response was used to ensure model assumptions of normality and homogeneity of variance in residuals were met (Miller, 1997). T-test was employed when comparing two groups. To compare significance levels for three or more groups, ANOVA was employed. Where ANOVA showed significance at the 5% level, *post-hoc* Tukey's test was used for comparisons. The Minitab Statistical Software version 18.1 or R (R Core Team, 2021) was used for statistical analysis. Model-adjusted means were used for the metabolomics/transcriptomic experiment. Metabolites/probes for which the combined mean was at least threefold and significantly ($P<0.01$) different from combined controls in both *HvF5H1*-RNAi lines were considered to be significantly different.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material. Additionally, the reference transcript dataset (RTD) is available in the Zenodo repository https://doi.org/10.5281/zenodo.3360434.

Author contributions

RS performed phylogenetic analysis, designed genetic constructs and, along with MH, characterized the transgenic plants. JS produced the transgenic plants. CL performed thioacidolysis and mild alkaline hydrolysis. YT and JR performed 2D NMR. CC and HO designed the metabolomic and transcriptomic experiments. GG, RV and WB performed the metabolomics analysis. CH conceived and coordinated the experiments. RS and CH wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

CH and RS received funding from BBSRC grants BB/N023455/1 and BB/G016232/1. CH was a Royal Society Wolfson Research Merit Award holder during some of the work. YT and JR were funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-SC0018409).

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1125003/full#supplementary-material

References

Andersen, J. R., Zein, I., Wenzel, G., Darnhofer, B., Eder, J., Ouzunova, M., et al. (2008). Characterization of phenylpropanoid pathway genes within European maize (*Zea mays* l.) inbreds. *BMC Plant Biol.* 8, 2. doi: 10.1186/1471-2229-8-2

Bak, S., Beisson, F., Bishop, G., Hamberger, B., Hofer, R., Paquette, S., et al. (2011). Cytochromes p450. *Arabidopsis Book* 9, e0144. doi: 10.1199/tab.0144

Barros, J., and Dixon, R. A. (2020). Plant phenylalanine/tyrosine ammonia-lyases. *Trends Plant Sci.* 25, 66–79. doi: 10.1016/j.tplants.2019.09.011

Barros, J., Serrani-Yarce, J. C., Chen, F., Baxter, D., Venables, B. J., and Dixon, R. A. (2016). Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nat. Plants* 2, 16050. doi: 10.1038/nplants.2016.50

Bewg, W. P., Poovaiah, C., Lan, W., Ralph, J., and Coleman, H. D. (2016). RNAi downregulation of three key lignin genes in sugarcane improves glucose release without reduction in sugar production. *Biotechnol. Biofuels* 9, 270. doi: 10.1186/s13068-016-0683-y

Cao, Y., Yan, X., Ran, S., Ralph, J., Smith, R. A., Chen, X., et al. (2022). Knockout of the lignin pathway gene *BnF5H* decreases the S/G lignin compositional ratio and improves *Sclerotinia sclerotiorum* resistance in *Brassica napus*. *Plant Cell Environ.* 45, 248–261. doi: 10.1111/pce.14208

Chapple, C. (1998). Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49, 311–343. doi: 10.1146/annurev.arplant.49.1.311

Chapple, C. C., Vogt, T., Ellis, B. E., and Somerville, C. R. (1992). An arabidopsis mutant defective in the general phenylpropanoid pathway. *Plant Cell* 4, 1413–1424. doi: 10.1105/tpc.4.11.1413

Ciesielski, P. N., Resch, M. G., Hewetson, B., Killgore, J. P., Curtin, A., Anderson, N., et al. (2014). Engineering plant cell walls: tuning lignin monomer composition for deconstructable biofuel feedstocks or resilient biomaterials. *Green Chem.* 16, 2627–2635. doi: 10.1039/c3gc42422g

Daly, P., Mcclellan, C., Maluk, M., Oakey, H., Lapierre, C., Waugh, R., et al. (2019). RNAi-suppression of barley caffeic acid O-methyltransferase modifies lignin despite redundancy in the gene family. *Plant Biotechnol. J.* 17, 594–607. doi: 10.1111/pbi.13001

Dence, C. W. (1992). "The determination of lignin," in *Methods in lignin chemistry*. Eds. S. Y. Lin and C. W. Dence (Berlin, Heidelberg: Springer Berlin Heidelberg).

Fan, D., Li, C., Fan, C., Hu, J., Li, J., and Yao, S. (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

Franke, R., Mcmichael, C. M., Meyer, K., Shirley, A. M., Cusumano, J. C., and Chapple, C. (2000). Modified lignin in tobacco and poplar plants over-expressing the arabidopsis gene encoding ferulate 5-hydroxylase. *Plant J.* 22, 223–234. doi: 10.1046/j.1365-313x.2000.00727.x

Grabber, J. H., Mertens, D. R., Kim, H., Funk, C., Lu, F., and RALPH, J. (2009). Cell wall fermentation kinetics are impacted more by lignin content and ferulate cross linking than by lignin composition. *J. Sci. Food Agric.* 89, 122–129. doi: 10.1002/jsfa.3418

Grabber, J. H., Ralph, J., Hatfield, R. D., and QUIDEAU, S. (1997). *p*-hydroxyphenyl, guaiacyl, and syringyl lignins have similar inhibitory effects on wall degradability. *J. Agric. Food Chem.* 45, 2530–2532. doi: 10.1021/jf970029v

Ha, C. M., Escamilla-Trevino, L., Yarce, J. C., Kim, H., Ralph, J., CHEN, F., et al. (2016). An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in medicago truncatula. *Plant J.* 86, 363–375. doi: 10.1111/tpj.13177

Halpin, C. (2019). Lignin engineering to improve saccharification and digestibility in grasses. *Curr. Opin. Biotechnol.* 56, 223–229. doi: 10.1016/j.copbio.2019.02.013

Harwood, W. A., Bartlett, J. G., Alves, S. C., Perry, M., Smedley, M. A., Leyland, N., et al. (2009). Barley transformation using agrobacterium-mediated techniques. *Methods Mol. Biol.* 478, 137–147. doi: 10.1007/978-1-59745-379-0_9

Jacobsen, J., Venables, I., Wang, M.-B., Matthews, P., Ayliffe, M., and Gubler, F. (2006). "Barley (Hordeum vulgare l.)," in *Agrobacterium protocols*. Ed. K. Wang (Totowa, NJ: Humana Press).

Kim, H., and Ralph, J. (2010). Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6/pyridine-d_5. *Organic Biomolecular Chem.* 8, 576–591. doi: 10.1039/B916070A

Kim, J. H., Yang, D. H., Kim, J. S., Baek, M. H., Park, Y. M., Wi, S. G., et al. (2006). Cloning, characterization, and expression of two cDNA clones for a rice ferulate-5-hydroxylase gene, a cytochrome P450-dependent monooxygenase. *J. Plant Biol.* 49, 200–204. doi: 10.1007/BF03030533

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lapierre, C., Monties, B., and Rolando, C. (1986). Preparative thioacidolysis of spruce lignin - isolation and identification of main monomeric products. *Holzforschung* 40, 47–50. doi: 10.1515/hfsg.1986.40.1.47

Lapierre, C., Sibout, R., Laurans, F., Lesage-Descauses, M. C., Dejardin, A., and Pilate, G. (2021). *p*-coumaroylation of poplar lignins im ts lignin structure and improves wood saccharificationb Pac. *Plant Physiol.* 187, 1374–1386. doi: 10.1093/plphys/kiab359

Li, X., Ximenes, E., Kim, Y., Slininger, M., Meilan, R., Ladisch, M., et al. (2010). Lignin monomer composition affects arabidopsis cell-wall degradability after liquid hot water pretreatment. *Biotechnol. Biofuels* 3, 27. doi: 10.1186/1754-6834-3-27

Lu, F., and Ralph, J. (1999). Detection and determination of *p*-coumaroylated units in lignins. *J. Agric. Food Chem.* 47, 1988–1992. doi: 10.1021/jf981140j

Mansfield, S. D., Kim, H., Lu, F., and Ralph, J. (2012). Whole plant cell wall characterization using solution-state 2D NMR. *Nat. Protoc.* 7, 1579–1589. doi: 10.1038/nprot.2012.064

Marita, J. M., Hatfield, R. D., Rancour, D. M., and Frost, K. E. (2014). Identification and suppression of the *p*-coumaroyl CoA:hydroxycinnamyl alcohol transferase in *Zea mays* l. *Plant J.* 78, 850–864. doi: 10.1111/tpj.12510

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., et al. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature* 544, 427–433. doi: 10.1038/nature22043

Mayer, K. F. X., Waugh, R., Langridge, P., Close, T. J., Wise, R. P., Graner, A., et al. (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491, 711–716. doi: 10.1038/nature11543

Ménard, D., Blaschek, L., Kriechbaum, K., Lee, C. C., Serk, H., Zhu, C., et al. (2022). Plant biomechanics and resilience to environmental changes are controlled by specific lignin chemistries in each vascular cell type and morphotype. *Plant Cell* 34, 4877–4896. doi: 10.1093/plcell/koac284

Meyer, K., Cusumano, J. C., Somerville, C., and Chapple, C. C. S. (1996). Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. *Proc. Natl. Acad. Sci. United States America* 93, 6869–6874.

Meyer, K., Shirley, A. M., Cusumano, J. C., Bell-Lelong, D. A., and Chapple, C. (1998). Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in arabidopsis. *Proc. Natl. Acad. Sci. United States America* 95, 6619–6623.

Miller, R. G. (1997). Beyond ANOVA: Basics of applied statistics, Chapman and Hall/CRC. *Texts in Statistical Science Series*. doi: 10.1201/b15236

Milne, L., Bayer, M., Rapazote-Flores, P., Mayer, C. D., Waugh, R., and Simpson, C. G. (2021). EORNA, a barley gene and transcript abundance database. *Sci. Data* 8, 90. doi: 10.1038/s41597-021-00872-4

Muhammad, A., Hao, H., Xue, Y., Alam, A., Bai, S., Hu, W., et al. (2020). Survey of wheat straw stem characteristics for enhanced resistance to lodging. *Cellulose* 27, 2469–2484. doi: 10.1007/s10570-020-02972-7

Muthamilarasan, M., Khan, Y., Jaishankar, J., Shweta, S., Lata, C., and Prasad, M. (2015). Integrative analysis and expression profiling of secondary cell wall genes in C4 biofuel model *Setaria italica* reveals targets for lignocellulose bioengineering. *Front. Plant Sci.* 6, 965. doi: 10.3389/fpls.2015.00965

Oakey, H., Shafiei, R., Comadran, J., Uzrek, N., Cullis, B., Gomez, L. D., et al. (2013). Identification of crop cultivars with consistently high lignocellulosic sugar release requires the use of appropriate statistical design and modelling. *Biotechnol. Biofuels* 6, 185. doi: 10.1186/1754-6834-6-185

Özparpucu, M., Gierlinger, N., Burgert, I., Van Acker, R., Vanholme, R., Boerjan, W., et al. (2018). The effect of altered lignin composition on mechanical properties of CINNAMYL ALCOHOL DEHYDROGENASE (CAD) deficient poplars. *Planta* 247, 887–897. doi: 10.1007/s00425-017-2828-z

Özparpucu, M., Rüggeberg, M., Gierlinger, N., Cesarino, I., Vanholme, R., Boerjan, W., et al. (2017). Unravelling the impact of lignin on cell wall mechanics: a comprehensive study on young poplar trees downregulated for CINNAMYL ALCOHOL DEHYDROGENASE (CAD). *Plant J.* 91, 480–490. doi: 10.1111/tpj.13584

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

Ralph, J., Hatfield, R. D., Quideau, S., Helm, R. F., Grabber, J. H., and Jung, H. J. G. (1994). Pathway of *p*-coumaric acid incorporation into maize lignin as revealed by NMR. *J. Am. Chem. Soc.* 116, 9448–9456. doi: 10.1021/ja00100a006

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

R Core Team (2021). *R: A language and environment for statistical computing* (Vienna, Austria: Foundation for Statistical Computing). Available at: https://www.R-project.org/.

Reddy, M. S., Chen, F., Shadle, G., Jackson, L., Aljoe, H., and Dixon, R. A. (2005). Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* l.). *Proc. Natl. Acad. Sci. U.S.A.* 102, 16573–16578. doi: 10.1073/pnas.0505749102

Sakamoto, S., Kamimura, N., Tokue, Y., Nakata, M. T., Yamamoto, M., Hu, S., et al. (2020). Identification of enzymatic genes with the potential to reduce biomass recalcitrance through lignin manipulation in arabidopsis. *Biotechnol. Biofuels* 13, 97. doi: 10.1186/s13068-020-01736-6

Sibout, R., Le Bris, P., Legee, F., Cezard, L., and Renault H. & Lapierre, C. (2016). Structural redesigning arabidopsis lignins into alkali-soluble lignins through the

expression of *p*-Coumaroyl-CoA:Monolignol transferase PMT. *Plant Physiol.* 170, 1358–1366. doi: 10.1104/pp.15.01877

Sibout, R., Proost, S., Hansen, B. O., Vaid, N., Giorgi, F. M., Ho-Yue-Kuang, S., et al. (2017). Expression atlas and comparative coexpression network analyses reveal important genes involved in the formation of lignified cell wall in *Brachypodium distachyon*. *New Phytol.* 215, 1009–1025. doi: 10.1111/nph.14635

Smith, R. A., Lu, F., Muro-Villanueva, F., Cusumano, J. C., Chapple, C., and Ralph, J. (2022). Manipulation of lignin monomer composition combined with the introduction of monolignol conjugate biosynthesis leads to synergistic changes in lignin structure. *Plant Cell Physiol.* 63, 744–754. doi: 10.1093/pcp/pcac031

Stewart, J. J., Akiyama, T., Chapple, C., Ralph, J., and Mansfield, S. D. (2009). The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol.* 150, 621–635. doi: 10.1104/pp.109.137059

Takeda, Y., Koshiba, T., TobimatsU, Y., Suzuki, S., MurakamI, S., Yamamura, M., et al. (2017). Regulation of *CONIFERALDEHYDE 5-HYDROXYLASE* expression to modulate cell wall lignin structure in rice. *Planta* 246, 337–349. doi: 10.1007/s00425-017-2692-x

Takeda, Y., Suzuki, S., Tobimatsu, Y., Osakabe, K., Osakabe, Y., Ragamustari, S. K., et al. (2019a). Lignin characterization of rice *CONIFERALDEHYDE 5-HYDROXYLASE* losS of-function mutants generated with the CRISPR/Cas9 system. *Plant J.* 97, 543–554. doi: 10.1111/tpj.14141

Takeda, Y., Tobimatsu, Y., Yamamura, M., Takano, T., Sakamoto, M., and Umezawa, T. (2019b). Comparative evaluations of lignocellulose reactivity and usability in transgenic rice plants with altered lignin composition. *J. Wood Sci.* 65, 6. doi: 10.1186/s10086-019-1784-6

Tobimatsu, Y., Chen, F., Nakashima, J., Escamilla-Trevino, L. L., Jackson, L., Dixon, R. A., et al. (2013). Coexistence but independent biosynthesis of catechyl and guaiacyl/syringyl lignin polymers in seed coats. *Plant Cell* 25, 2587–2600. doi: 10.1105/tpc.113.113142

Tottman, D. R. (1987). The decimal code for the growth-stages of cereals, with illustrations. *Ann. Appl. Biol.* 110, 441–454. doi: 10.1111/j.1744-7348.1987.tb03275.x

van Acker, R., Vanholme, R., Storme, V., Mortimer, J. C., and Dupree P. & Boerjan, W. (2013). Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in arabidopsis thaliana. *Biotechnol. Biofuels* 6, 46. doi: 10.1186/1754-6834-6-46

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

Weng, J. K., Li, Y., Mo, H., and Chapple, C. (2012). Assembly of an evolutionarily new pathway for α-pyrone biosynthesis in *Arabidopsis*. *Science* 337, 960–964. doi: 10.1126/science.1221614

Weng, J. K., Li, X., Stout, J., and Chapple, C. (2008). Independent origins of syringyl lignin in vascular plants. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7887–7892. doi: 10.1073/pnas.0801696105

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

Wu, Z., Wang, N., Hisano, H., Cao, Y., Wu, F., Liu, W., et al. (2019). Simultaneous regulation of *F5H* in COMT-RNAi transgenic switchgrass alters effects of *COMT* suppression on syringyl lignin biosynthesis. *Plant Biotechnol. J.* 17, 836–845. doi: 10.1111/pbi.13019

Zadoks, J., Chang, T., and Konzak, C. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 17 January 2023
DOI 10.3389/fpls.2022.1088879

OPEN ACCESS

EDITED BY
Lei Wang,
Institute of Botany (CAS), China

REVIEWED BY
Masahiro Sakamoto,
Kyoto University, Japan
Qiao Zhao,
Shenzhen Institutes of Advanced Technology (CAS), China

*CORRESPONDENCE
Rowan A.C. Mitchell
rowan.mitchell@rothamsted.ac.uk

†These authors have contributed equally to this work and share senior authorship

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

RECEIVED 03 November 2022
ACCEPTED 28 December 2022
PUBLISHED 17 January 2023

CITATION
Chandrakanth NN, Zhang C, Freeman J, de Souza WR, Bartley LE and Mitchell RAC (2023) Modification of plant cell walls with hydroxycinnamic acids by BAHD acyltransferases.
Front. Plant Sci. 13:1088879.
doi: 10.3389/fpls.2022.1088879

Modification of plant cell walls with hydroxycinnamic acids by BAHD acyltransferases

Niharika Nonavinakere Chandrakanth [1], Chengcheng Zhang [2], Jackie Freeman [3], Wagner Rodrigo de Souza [4†], Laura E. Bartley [1†] and Rowan A.C. Mitchell [3*†]

[1]Institute of Biological Chemistry, Washington State University, Pullman, WA, United States, [2]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, United States, [3]Plant Sciences, Rothamsted Research, West Common, Harpenden, Hertfordshire, United Kingdom, [4]Center for Natural and Human Sciences, Federal University of ABC, Santo André, Brazil

In the last decade it has become clear that enzymes in the "BAHD" family of acyl-CoA transferases play important roles in the addition of phenolic acids to form ester-linked moieties on cell wall polymers. We focus here on the addition of two such phenolics—the hydroxycinnamates, ferulate and *p*-coumarate—to two cell wall polymers, glucuronoarabinoxylan and to lignin. The resulting ester-linked feruloyl and p-coumaroyl moities are key features of the cell walls of grasses and other commelinid monocots. The capacity of ferulate to participate in radical oxidative coupling means that its addition to glucuronoarabinoxylan or to lignin has profound implications for the properties of the cell wall – allowing respectively oxidative crosslinking to glucuronoarabinoxylan chains or introducing ester bonds into lignin polymers. A subclade of ~10 BAHD genes in grasses is now known to (1) contain genes strongly implicated in addition of *p*-coumarate or ferulate to glucuronoarabinoxylan (2) encode enzymes that add *p*-coumarate or ferulate to lignin precursors. Here, we review the evidence for functions of these genes and the biotechnological applications of manipulating them, discuss our understanding of mechanisms involved, and highlight outstanding questions for future research.

KEYWORDS

ferulic acid, para-coumaric acid, grasses, cell wall, xylan, lignin, plant biotechnology, bioenergy

Background - importance of cell wall hydroxycinnamates

Cell walls are integral to plant growth and development, encapsulating most cells, dictating their shape and comprising most plant biomass. Cell wall polymer composition and modifications vary across cell types and developmental stages defining the properties of the wall. Primary cell walls are deposited at the cell plate and around expanding cells, and during development must

allow for breaking of bonds within or between polymers as part of remodeling. In contrast, secondary cell walls are typically deposited only around fully expanded cells, adding strength, hydrophobicity, and a thick barrier for defense. Primary cell wall polymers during expansion are all polysaccharides which allow for different modes of remodeling; whereas, secondary cell wall polymers often include lignin where cross-links are considered irreversible. This review focuses on a particular subset of cell wall polymer modifications that occur on both polysaccharides and lignin, the abundant acylation with hydroxycinnamates that are a key feature of both primary and secondary cell walls of grasses and other commelinid monocots. Hydroxycinnamates are simple phenylpropanoid molecules, closely related to canonical lignin monomers, that share their ability to oxidatively couple and thereby cross-link polymers (Ralph et al., 1992; Ralph et al., 1995). The two most abundant cell wall phenolic esters in grasses, those derived from ferulic and *p*-coumaric acids (Figure 1A), differ greatly in this property. Feruloyl modifications (FA) have a much greater propensity than *p*-coumaroyl modifications (*p*CA) to undergo oxidative coupling (Figure 1B, D). This key difference has profound implications for the effects of these modifications on cell wall and biomass properties.

Hydroxycinnamate modification of xylan

In grass primary and secondary cell walls, hydroxycinnamate modifications of polysaccharides occur as acylation of the 5-carbon of arabinofuranosyl (Ara*f*) decoration of the xylan backbone in glucuronoarabinoxylan (GAX; Figure 2) (Ishii, 1997; Ralph et al., 1998; Ralph et al., 2004; Buanafina, 2009; Bartley et al., 2013). Xylan is the most abundant polysaccharide other than cellulose in both primary and secondary cell walls of grasses, accounting for ~50% of grass hemicellulose (Scheller and Ulvskov, 2010), which represents, for example, 20-25% of dry switchgrass biomass (David and Ragauskas, 2010). Xylan is much more abundant in primary cell walls (PCWs) of grasses than in dicots (~30% compared to 5% of cell wall, respectively), displacing pectins as the most abundant non-cellulose polysaccharide. The 3-linked Ara*f* decoration of xylan is rare or absent in dicots and the FA and *p*CA acylation of this Ara*f* in GAX are believed to be completely specific to grass and other recently evolved monocots, known as commelinids (Harris and Trethewey, 2010). Grass GAX also possesses other substitutions on Ara*f*, such as β-(1->2)Xyl-(1->2)Gal (Saulnier et al., 1995), β-(1->2)-Gal and β-(1->2)-Xyl (Wende and Fry, 1997; Chiniquy et al., 2012), and substitutions shared with dicot xylan, such as acetylation and (4-O-methyl-) glucuronosyl at the O2- position (Scheller and Ulvskov, 2010). Other HCAs also occur at lower abundance ester-linked to GAX in grasses. Recent mass spectrometry analysis of products of mild acidolysis of rice cell walls has detected caffeic acid on Ara*f* of GAX (Feijao et al., 2022), and sinapate also occurs ester-linked to arabinoxylan in cereal grain (Bunzel et al., 2003). The presence of FA on GAX in particular confers a mode of cross-linking to grass primary cell walls absent in those of dicots since FA can undergo radical oxygen-mediated coupling to form ether bonds or C-C bonds, making diferulates and triferulates that result in xylan-xylan cross-linking (Takahama and Oniki, 1994; Bunzel et al., 2008) (Figure 1C). Across studies in various species and organs (Hatfield et al., 1999;

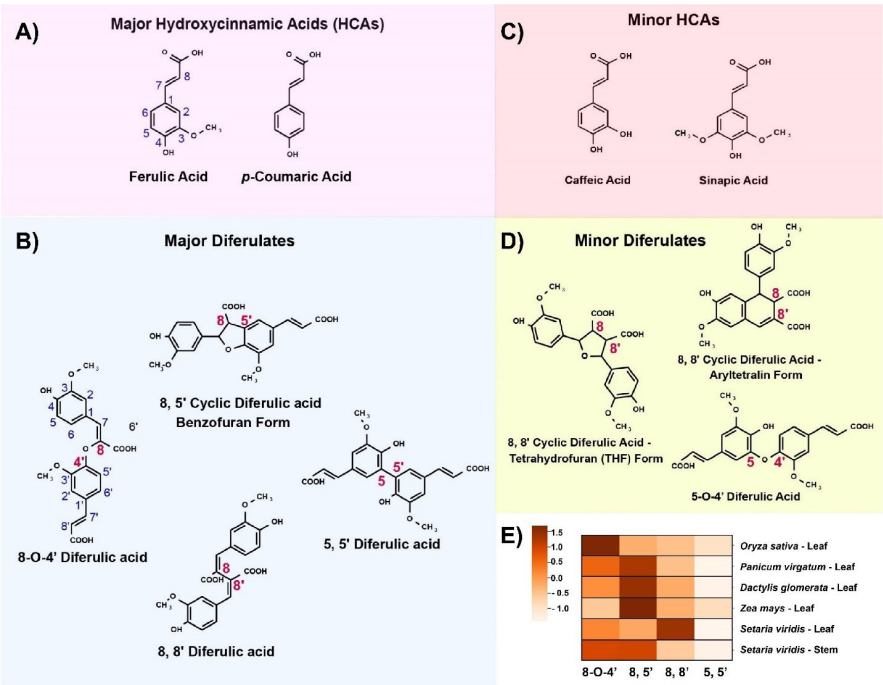


FIGURE 1
Hydroxycinnamic acid and diferulic acid structures and relative abundances. **(A–D)** Chemical structures of major and minor hydroxycinnamates and diferulates observed from grass cell walls. **(E)** Heatmap represents the relative abundances of major diferulates in *Oryza sativa* leaf tissue (Bartley et al., 2013), *Panicum virgatum* leaf, *Dactylis glomerata* leaf, *Zea mays* leaf (Hatfield et al., 1999), and *Setaria viridis* leaf and stem (de Souza et al., 2018). Relative abundances are Z-scores [(observed value−mean for a given species) /std deviation for that species)].

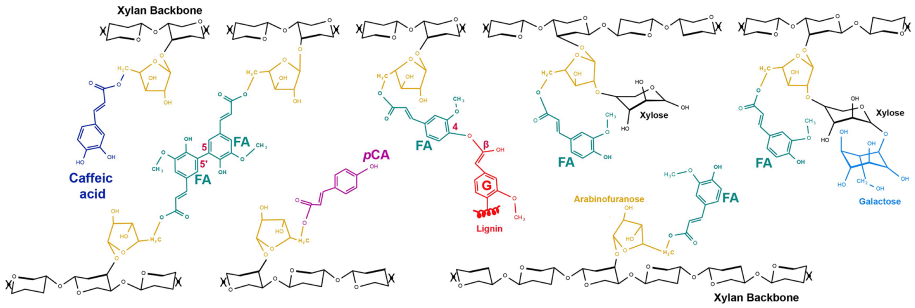


FIGURE 2
Hydroxycinnamoyl (HCA) decorations of grass glucuronoarabinoxylan (GAX). All HCA decorations occur on 5-O of Araf which is α-(1,3)-linked to xylan backbone. FA decorations of GAX (turquoise) are abundant and include forms where Araf is additionally substituted with β(1,2)-linked Xyl and this may itself be further substituted by β(1,4)-linked galactose. A FA 5,5′ dimer crosslinking GAX chains is shown as one example of a dimer that can crosslink GAX chains (see Figure 1). FA crosslinked to S lignin monomer is shown as one example of FA crosslinking GAX to lignin; FA can also link to G lignin monomers and to tricin (structures in Figure 3). *p*CA decorations (pink) of GAX can also be abundant in some tissues and smaller amounts of caffeoyl-decorations (dark blue) have recently been detected (Feijao et al., 2022). Glucuronic acid, O-methyl-glucuronic acid, acetyl substitutions, non-acylated Araf that are commonly present on GAX are not shown here.

Bartley et al., 2013; de Souza et al., 2018), the 8-5 and 8-O-4 dimers are often the most abundant diferulates (Figure 1E).

In PCWs, the functions of hydroxycinnamate modifications on GAX remain to be fully elucidated. The greater abundance of GAX and lower abundance of pectin and xyloglucan in grass PCWs compared to dicot PCW suggests grass GAX may have taken over some of the roles performed by pectin and xyloglucan, which is supported by solid-state NMR analysis of PCWs (Wang et al., 2014b). Potentially, the FA dimer and trimer cross-links on grass GAX partially substitute for the roles played by ionic cross-linking of pectin and for the oxidative cross-linking of extensin proteins in dicot PCWs. Consistent with this, the simplest form of extensins, those lacking a signaling domain, are not found in grasses; (Johnson et al., 2017). Also, FA is especially abundant per mass cell walls in very young tissue (Obel et al., 2002; Lin et al., 2016). Furthermore, abundance of cell wall FA and FA dimers was found to be negatively correlated with cell wall extensibility in wheat coleoptiles, suggesting an important role in control of PCW expansion (Wakabayashi et al., 1997). Additionally, hydroxycinnamates have antimicrobial properties (Akin, 2008) so their presence in grass primary cell walls also may serve to inhibit microbial attack and FA dimers (Figure 1B, D) may inhibit digestion (Grabber et al., 1998). Thus, the hydroxycinnamates on GAX in grass PCWs may confer evolutionary advantages by making young grass tissue with many expanding cells less readily digestible.

In lignified secondary cell walls (SCW) of both grasses and dicots, solid-state NMR suggests xylan in a twofold screw conformation (Xn^{2f}) binds to cellulose microfibrils (Simmons et al., 2016); whereas, distorted twofold or threefold screw xylan (Xn^{3f}) interacts closely with lignin (Kang et al., 2019; Duan et al., 2021). Thus, xylan bridges the two main components of SCW, although a study on sorghum SCW suggested Xn^{2f} was much less prevalent there (Gao et al., 2020). A clear difference in grass compared to dicot SCW is that the FA on GAX covalently bonds lignin *via* oxidative coupling, although the extent of this is difficult to determine (Ralph, 2010). GAX-FA is abundant on both Xn^{2f} and Xn^{3f} conformations in Brachypodium stems, and the authors proposed a model of grass SCW where FA on Xn^{2f} bound to cellulose crosslinks with other xylan FAs, and FA on Xn^{3f} covalently links to lignin (Duan et al., 2021). This model fits with several lines of evidence that show the abundance of GAX-FA and linkage of FA to lignin are correlated with recalcitrance to digestion of grass biomass (reviewed in (Buanafina, 2009; de Oliveira et al., 2015; Terrett and Dupree, 2019)). Thus, the FA-mediated linking of GAX to lignin inhibits access of hydrolytic enzymes to the cellulose to release glucose (the normal measure of digestibility).

In addition, a key role of FA in initial deposition of lignin is suggested by abundant Araf-FA coupled to coniferyl alcohol, the G-lignin monomer released from grass SCW by mild acidolysis (Lapierre et al., 2019; Feijao et al., 2022). This supports a model that GAX-FA act as the nucleation sites from which the lignin polymers grow that was developed from biomimetic studies of lignification of maize suspension culture cell walls (Grabber et al., 2002). Interestingly, this mode of nucleation with many separate sites (i.e. abundant GAX-FA) may explain the lower molecular weight of grass lignin polymers compared with those of other plants, which could allow some flexibility in developing tissue (Hatfield et al., 2017).

The role of *p*CA on GAX is less apparent than that of FA because *p*CA oxidatively couples much less readily than FA and whereas FA-GAX is found in every tissue in grasses, *p*CA-GAX has low abundance in stems (Fanelli et al., 2021; Möller et al., 2022). One possibility is that, analogous to the putative role of *p*CA on lignin discussed below, *p*CA-GAX may participate in radical transfer, thus catalyzing the oxidative coupling of neighboring FA on GAX. This is compatible with the observation that *p*CA on GAX rapidly increases in response to jasmonic acid application to Brachypodium callus (Hyde et al., 2018) which could be part of a priming of defense, allowing rapid cross-linking to occur in response to additional signals.

Hydroxycinnamate modification of lignin

Lignin biosynthesis occurs by generation of three main monolignols (*p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol) in cytosol and subsequent radical coupling of these in the apoplast (Boerjan et al.,

2003). Monolignols acylated by phenolic acids (especially *p*CA; FA; and *p*-hydroxybenzoate, a simple phenolic with two fewer carbons than hydroxycinnamates) and acetate, are now established as additional monomers of lignification in various species (Figure 3). Attachment of *p*CA to lignin has been found in a diversity of grass species (Soreng et al., 2015) including maize, bromegrass, bamboo, sugarcane, elephant grass, rice (Withers et al., 2012; Karlen et al., 2016; Takeda et al., 2017), switchgrass (Shen et al., 2009), and Brachypodium (Petrik et al., 2014). Recently, *p*-coumaryl lignin was also found in other commelinid monocots (*Zingiberales, Commelinales, and Arecales*) (Karlen et al., 2018) and in the dicot mulberry (Moracacea) (Hellinger et al., 2022). Another phenolic acid, *p*-hydoxybenzoate, also occurs ester-linked to lignin in the poplar, willows, and oil palms (de Vries et al., 2021; Zhao et al., 2021). Like *p*CA, *p*-hydoxybenzoate does not readily oxidatively couple, so terminates lignin chains (Figure 3). The acylation of monolignols by the other major hydroxycinnamate, FA, is a topic of great biotechnological interest because the FA becomes incorporated into lignin polymer *via* its propensity to oxidatively couple, thereby introducing alkaline-labile ester bonds (Figure 3) making the lignin much easier to break apart (Wilkerson et al., 2014). We discuss this further in biotechnological applications below. It is now clear that FA-lignin occurs at low abundance naturally in all commelinids examined as well as sporadically within eudicots (Karlen et al., 2016).

The role of lignin acylation by hydroxycinnamates is uncertain. One possibility is that *p*CA (and hydroxybenzoate) moieties on lignin act as "radical catalysts." Model studies of *p*CA show that it is readily oxidized. However, the fact that it has not been observed to oxidatively couple *in muro* has led to a model that oxidized *p*-coumaryl esters rapidly pass radicals to sinapyl alcohols, thereby facilitating lignin polymerization (Takahama and Oniki, 1994; Ralph,

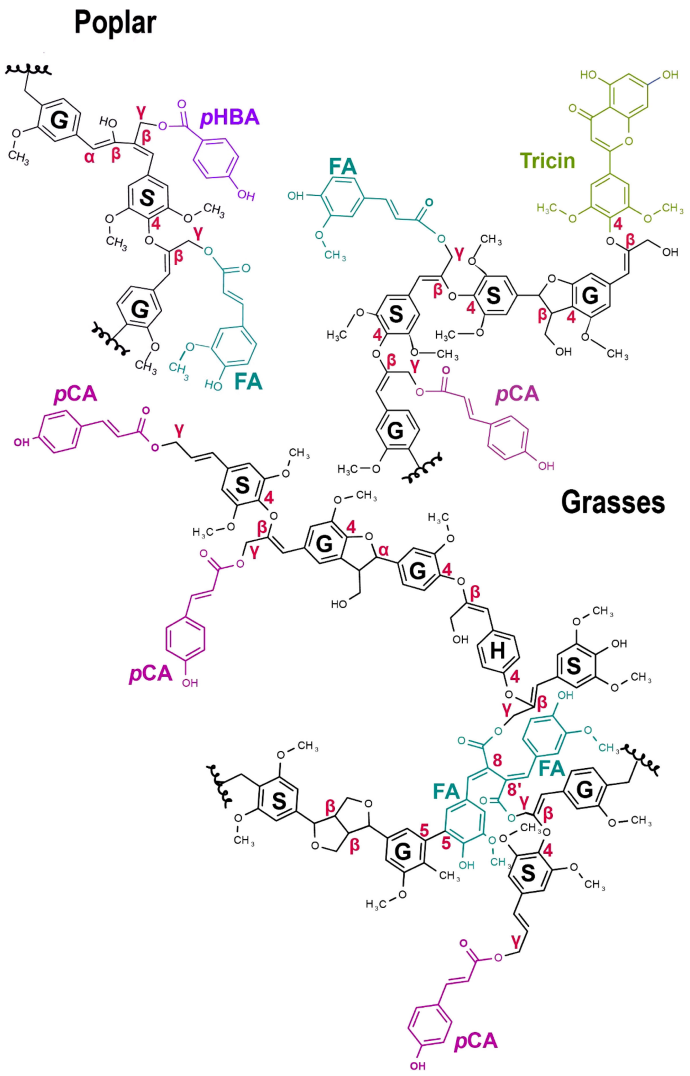


FIGURE 3
Hydroxycinnamoylated grass and poplar lignin polymer models. *p*-Coumaric acid (*p*CA - pink) and Ferulic acid (FA – turquoise) occur on lignin in commelinid grasses. Tricin (green) is a grass-specific flavonoid biosynthetic product that occurs as pendant groups on lignin. *p*-hydroxybenzoate (*p*HBA - purple) and FA (at low levels) occur in poplar and other dicots. γ-Feruloylated lignin naturally occurs in some non-commelinid grasses and dicots, generally at lower levels (Karlen et al., 2016). Here, we have only shown a FA dimer. Canonical lignin monomers include S, Syringyl lignin; G, Guaiacyl lignin; and H – *p*-hydroxyphenyl lignin. SGH monomers in the figure do not represent their actual ratios. The different C-C, ether (β-O-4), and γ-ester bonds occurring in the lignin polymer are highlighted in red.

2010). For the lower abundance acylation by FA, the biological functions are an open question.

BAHD acyl CoA transferases

We have gradually gained knowledge of enzymes responsible for the incorporation of *p*CA and FA into grass cell walls on both GAX polysaccharide and lignin. These proteins are all "BAHD" acyl-CoA acyltransferases, a large enzyme family in plants that acylate metabolites with CoA thioester donors named for the first four activities described for this family (BEAT, AHCT, HCBT, and DAT) (D'Auria, 2006). The BAHD family is divided into five clades; Clade V includes quinate hydroxycinnamoyl transferase (HCT) an enzyme in phenylpropanoid pathway for monolignol synthesis. BAHD enzymes are known for their versatility (i.e., low specificity) and often show activity with multiple acyl-CoA donors and acceptors such that their activity *in vivo* might be dictated by relative availability of substrates (D'Auria, 2006). They are also known for examples of convergent evolution as BAHD enzymes from different Clades can have the same activity (Luo et al., 2007).

Candidate BAHD enzymes for feruloylation and p-coumarylation of GAX

Looking for candidate genes for addition of FA to GAX, Mitchell et al. (2007) searched for genes that are highly expressed in grasses while the most similar genes in dicots are much less expressed and differ substantially in protein sequence, since feruloylation is abundant in every grass tissue and absent in dicots. They found a small subclade of BAHD genes that met these criteria in Clade V [Clade Va of Tuominen et al. (2011)] and as acyl transferases these were postulated as involved in feruloylation. Furthermore, some of these BAHD grass genes are co-expressed with other genes responsible for GAX synthesis (Mitchell et al., 2007; Molinari et al., 2013). To facilitate communication about these grass BAHD acyltransferases, Bartley et al. (2013) called the group of 20 rice genes the "Mitchell Clade" and identified subclade i and subclade ii containing, *Oryza sativa* (Os) acyltransferases (AT), OsAT1-OsAT10, and OsAT11-OsAT20, respectively. Figure 4 shows subclade i for selected model and economically relevant grass species. Alternative names were proposed of the form BAHD01-BAHD20 (Molinari et al., 2013) and are used in some publications; here we show the equivalent names in Figure 4 but will use the AT nomenclature in the text. As discussed in greater detail below, grasses generally possess 8-10 subclade i ATs per haploid genome (Figures 4, 5). The function of Mitchell subclade ii genes, which have undergone grass species-/tribe-specific expansions/deletions (Karlen et al., 2016) remains unknown. Due to the absence of studies about them and their generally low expression (Bartley et al., 2013) we have excluded the subclade ii genes from this review and use "Mitchell subclade" to refer exclusively to subclade i.

There is now good evidence that several Mitchell subclade ATs are indeed involved in acylation of GAX, as predicted; while others acylate monolignols. Genetic manipulation of the Mitchell subclade was first achieved by Piston et al. (2010) by simultaneous downregulation of OsAT7, OsAT8, OsAT9, and OsAT10 in rice which resulted in decreased amounts of ester-linked FA in a cell-wall enriched fraction from leaves. The authors observed 2- to 3-fold reductions in gene expression of these ATs and an average of 20% reduction in cell wall FA content in the leaves, but due to use of constructs targeting multiple ATs could not determine which of the silenced genes were responsible and did not demonstrate that the FA was attached to GAX. Bartley et al. (2013) provided the first genetic evidence on single ATs involved in addition of hydroxycinnamates to GAX. Using mild acidolysis to break glycosidic bonds, they were able to show effects on a five-carbon sugar-esterified hydroxycinnamates, likely Ara-*p*CA and Ara-FA, released from rice cell walls (rather than saponification to release ester-linked FA and *p*CA from cell wall polymers, in general). They designated OsAT10 as a putative *p*-coumaroyl CoA arabinofuranose transferase (PAT) since overexpression of *OsAT10, via* an activation tagged line in rice (*OsAT10-D1*), induced a 5-fold increase in *p*CA levels in young green tissues. The observed increased saccharification yields, in the transgenic line were possibly due to concomitant 50% decrease in FA linked to GAX (Bartley et al., 2013). Since then, other studies also achieved several-fold increases in *p*CA-GAX by heterologous expression of *OsAT10* in switchgrass (Li et al., 2018) and sorghum (Tian et al., 2021) and of sugarcane *AT10* (*ScAT10*) in maize (Fanelli et al., 2021). As in Bartley et al. (2013) this was sometimes (Li et al., 2018; Fanelli et al., 2021), but not universally, accompanied by a decrease in FA-AX. Assumed to be an indirect effect, the mechanism of the alternating abundance of *p*CA-AX and FA remains an open question. Since then, a complete knock out mutants of *OsAT10* in rice have been generated using CRISPR/Cas9 rice plants, leading to an almost complete lack of *p*CA-GAX, which was found to be most abundant in rice husks, compared to mature leaf and stems, of wild-type plants (Möller et al., 2022). The gene edited AT10 lines also exhibited an increase in FA, but no differences in cell wall composition or digestibility. Giving further support for the AT10 PAT function across species, a natural allele of the *AT10* ortholog in barley (*HORVU7Hr1G085100*; Figure 4), predicted to encode a defective enzyme, has less ester-linked *p*CA and more ester-linked FA in grain cell walls (Houston et al., 2020).

Genetic analysis has provided various strengths of support for several ATs acting as putative feruloyl arabinofuranose transferases (FATs). RNAi silencing of Brachypodium *BdAT1* showed an approximately 25% reduction in FA amounts and BdAT1 overexpression resulted in an approximately 15% increase in FA in leaves and stems (Buanafina et al., 2016). On the other hand, Mota et al. (2021) showed different results, with RNAi suppression of *SvAT1*, the *Setaria viridis* BdAT1 ortholog, decreasing *p*CA not FA on GAX. They therefore suggest that BdAT1 and SvAT1 have differing specificities for *p*CA-CoA and FA-CoA donors and conduct some protein structural modelling to support this. The clearest evidence of FAT activity thus far, was obtained through RNAi-based silencing of *SvAT9* (*SvBAHD01*) in Setaria resulting in a 60% decrease in FA-GAX with a significant increase *p*CA-GAX; whereas, downregulation of *OsAT9* ortholog in Brachypodium showed only small effects on FA (de Souza et al., 2018). *AT7* ortholog downregulation in Brachypodium did not yield any significant changes in FA (Buanafina

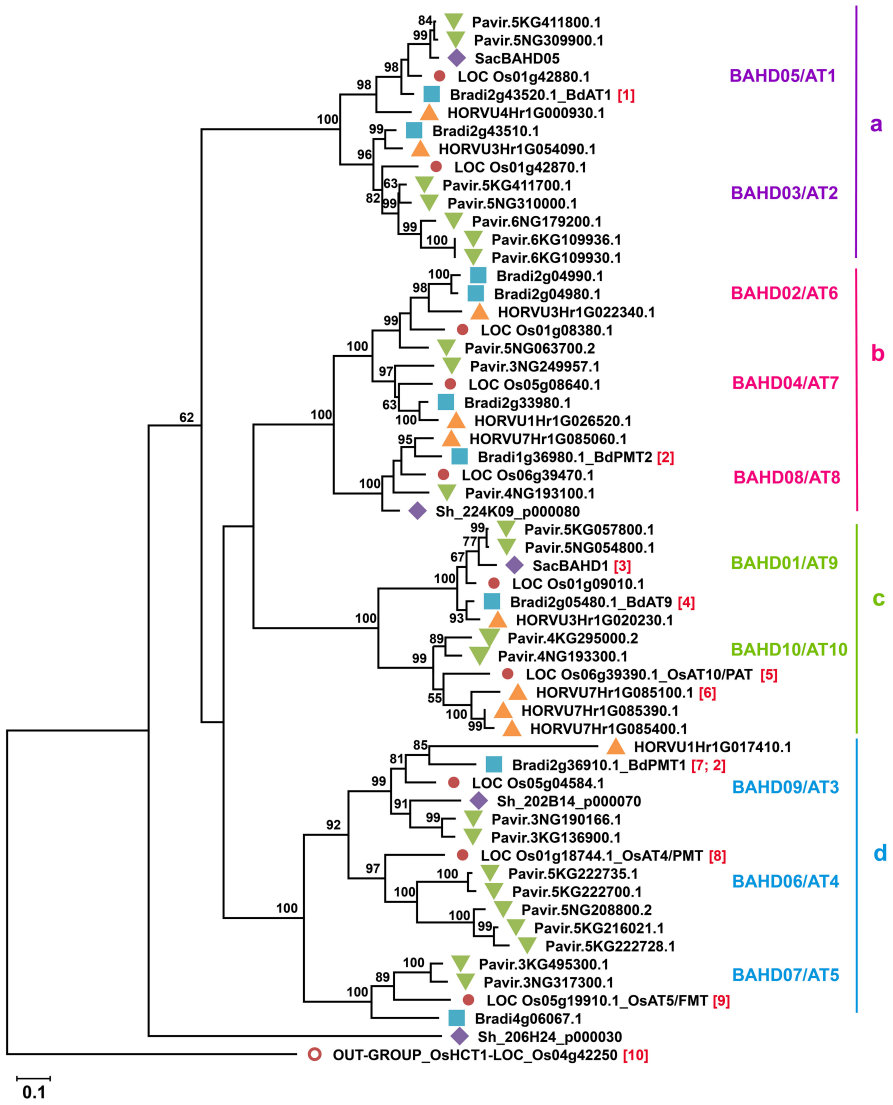


FIGURE 4
Phylogenetic reconstruction of Mitchell subclade of grass BAHD acyltransferases. Phylogenetic tree was generated by maximum likelihood method (Jones et al., 1992) using Mega X software (Kumar et al., 2018) after multiple sequence alignment by MUSCLE (3.8). All sequences of rice genes (Os) originally identified in Mitchell subclade (Bartley et al., 2013; Molinari et al., 2013) and their orthologs in *Brachypodium distachyon* v3.1 (Bradi), *Saccharum* (Sh; sequences from (de Souza et al., 2019), *Hordeum vulgare* (HORVU) and *Panicum virgatum* v5 (Pavir) are included. The rice HCT gene OsHCT1 Kim et al. (2012) [10] was used as an outgroup. The division of proteins into sub-groups (a-d) as in Karlen et al. (2016) is indicated on the right. Proteins that are functionally characterized are as follows: [1] Buanafina et al. (2016), [2] (Sibout et al., 2016), [3] de Souza et al. (2019) [4] de Souza et al. (2018) [5] Bartley et al. (2013) [6] Houston et al. (2020), [7] Petrik et al. (2014), [8] Withers et al. (2012), [9] Karlen et al. (2016). Where assigned, enzyme activities are PMT *p*-Coumaryl CoA Monolignol Transferase, FMT Feruloyl CoA Monolignol Transferase, PAT *p*-Coumaryl CoA Arabinoxylan Transferase. Numbers on tree nodes are percentage bootstrap support; values <50 not shown. Scale bar indicates branch lengths measured in the number of substitutions per site.

et al., 2016), though preliminary evidence suggested that a rice T-DNA insertion line for this gene has less leaf sheath FA (Bartley et al., 2013). In general, whilst genetic manipulation of ATs in grasses has achieved many fold increases and abolition in some tissues of *p*CA-GAX, this has not been reported for FA-GAX where the range of effects is narrower. This may point to a critical functionality of FA-GAX in grass cell walls where abolition would be lethal and large increases are difficult to achieve due to tight regulation.

BAHD enzymes acylate lignin monomers

Independent of the bioinformatics identification of the Mitchell subclade as candidates for GAX feruloylation, other groups established that some members of this subclade add *p*CA to lignin by acylating monolignols. Withers et al. (2012) showed that OsAT4 functions *in vitro* as a *p*CA monolignol acyltransferase (PMT) that transfers *p*CA from *p*CA-CoA onto H and S monolignols. The maize

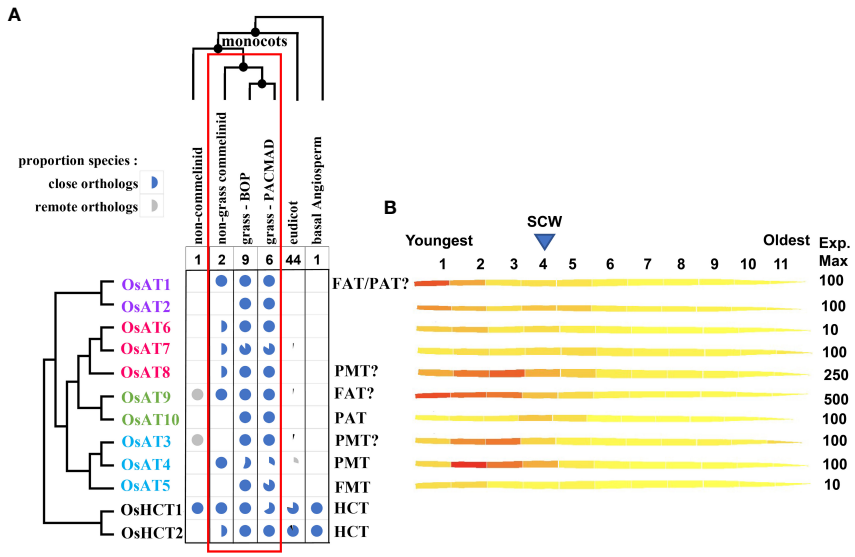


FIGURE 5
Taxonomy and expression profiles of Mitchell subclade acyltransferases. **(A)** Taxonomic distribution of ATs and their known and putative activities. Pie charts indicate the proportion of species with orthologs to the Mitchell subclade OsAT genes and to two BAHD Clade V HCT genes encoding shikimate O-hydroxycinnamoyl transferases, a key enzyme in the phenylpropanoid pathway expected to be present in all plants. All species have close orthologs to HCT1 and/or HCT2 whereas only commelinid monocots have close orthologs to *AT* genes, which matches the taxonomic distribution of FA-GAX indicated by the red rectangle. The species are all angiosperms with fully sequenced genomes present in Ensembl Plants release 54; close and remote orthologs defined as reciprocal blastp top hits with bitscore > 400 and 300 respectively. Text colors indicated the subgroups in Figure 4. FAT indicates likely feruloyl arabinose transferase; PAT indicates likely *p*-coumaryl arabinose acyltransferase; FMT indicates ferulate monolignol transferase; PMT indicates *p*-coumarate monolignol transferase. Question marks indicate that existing evidence is contradictory or relatively weak. **(B)** *AT* expression across the rice leaf gradient generated from the eFP (electronic fluorescent pictograph) browser of the Bio-Analytic Resource for Plant Biology (BAR), the University of Toronto (Sullivan et al., 2019) using Wang et al. (2014a) leaf expression data. The rice leaf gradient includes the youngest tissue in the 1st segment and the oldest in the 11th segment with secondary cell wall (SCW) related expression peaking at segment 4.

ortholog of *OsAT3* also shows PMT activity, and RNAi suppression of the maize led to large decreases in *p*CA ester-linked to lignin (Marita et al., 2014). The strongest line showed a reduction in S lignin. Similarly, a complete knock-out mutant of *BdPMT1*, the Brachypodium ortholog of *OsAT3*, had <0.5% *p*CA on mature lignin; whereas *p*CA on GAX was unaffected. Conversely, overexpression of *BdPMT1* boosted *pCA*-lignin above wild-type levels (Petrik et al., 2014). Heterologous expression of *BdPMT1* and *BdPMT2* (ortholog of *OsAT8*) in *Arabidopsis*, under the control of the *Arabidopsis* cinnamate-4-hydroxylase promoter, introduced *p*CA onto lignin, showing a gain of function since there is no *p*CA on lignin in wild-type Arabidopsis (Sibout et al., 2016). Though not focused on lignin modification, an early study found an enzyme from the commelinid species, *Musa sapientum* (i.e., banana alcohol acyltransferase, BanAAT), to have the highest activity on an aromatic acceptor substrate (Beekwilder et al., 2004). In retrospect, based on phylogenic analysis (Bartley et al., 2013), this enzyme is likely a banana PMT.

BAHD enzymes that acylate monolignols with FA have also been discovered. A feruloyl-monolignol transferase (FMT) from Chinese angelica [*Angelica sinensis* (*As*), a dicotyledonous medicinal plant, was heterologously expressed in hybrid poplar generating monolignol-FAs that were incorporated into lignin polymers (Wilkerson et al., 2014). The use of this activity to facilitate cell wall deconstruction are discussed further below (see Biotechnological Applications). The AsFMT is in Clade III of the BAHD superfamily, which is distant from the Mitchell subclade within Clade V. Surprisingly, overexpression of *OsAT5* in rice increased feruloylated monolignols, suggesting that *OsAT5* also encodes an FMT (Karlen et al., 2016). Thus, AsFMT and OsFMT are the result of convergent evolution, one of several examples in the BAHD family (Luo et al., 2007). A recent discovery on substrate specificity was made by Smith et al. (2022) looking at FMT and PMT enzymes from sorghum (*Sorghum bicolor*) and switchgrass (*Panicum virgatum*) as synthesized with wheat germ extract followed by *in vitro* characterization. The FMT enzymes, including OsAT5, produced both monolignol FA and monolignol *p*CA conjugates; whereas, the PMT enzymes produced exclusively monolignol *p*CA conjugates. A tolerance of differing acyl-CoA donors is another known feature of many BAHD enzymes (D'Auria, 2006).

Taxonomic distribution and patterns of expression of *AT* genes

The taxonomic distribution and expression of *AT* genes provide functional clues and be used to identify other candidate genes involved in the same processes for basic and applied purposes. Phylogenetic analyses from selected grass species here (Figures 4, 5) and elsewhere (Bartley et al. (2013); Karlen et al. (2016); de Souza et al. (2018); Fanelli

et al. (2021)) shows that the Mitchell subclade ATs are highly conserved in grasses. Here, we identified orthologs from a novel set of species to better assess how their distribution compares that with that of FA and *p*CA ester-linked GAX and lignin discussed above. The distribution of orthologs of the Mitchell subclade ATs are shown in Figure 5 and their corresponding demonstrated and putative activities noted. As outlined above, FA-GAX is likely a fundamental feature of grass cell walls, conferring a mode of cross-linking absent in cell walls of plants outside of the commelinid monocots which plausibly represents a trait that contributed to the evolutionary success of the grasses. Therefore, we might expect the enzymes responsible to be highly conserved in all grasses. Consistent with this, in fully sequenced genomes of 15 grasses, all have clear orthologs to *OsAT1*, *OsAT2*, *OsAT3*, *OsAT6*, *OsAT8*, *OsAT9*, and *OsAT10* (Figure 5). Thus, genes demonstrated to be responsible for acylation of GAX with FA and *p*CA and of lignin with *p*CA in some grasses are conserved, suggesting conservation of these functions across the Poaceae. Indeed, biochemical analysis of sorghum and switchgrass orthologs of rice and Brachypodium monolignol ATs, supports the notion that sequence conservation indicates functional conservation, albeit with variation in enzymatic parameters (Smith et al., 2022). However, the absence of conservation of AT5 suggests that either the FMT activity may be dispensable, or another AT may have this activity, either primarily or due to low substrate specificity.

Looking more broadly across monocots, there are also clear orthologs of Mitchell subclade ATs in non-grass commelinids (*Musa acuminata* and *Ananas comosus*) of one member of each enzyme group (a-d), i.e., OsAT1, *OsAT3*/4, OsAT6, OsAT9. The occurrence of close orthologs thus matches the distribution of GAX feruloylation, believed to be confined to commelinid monocots (Harris and Trethewey, 2010). As described above, *p*CA-lignin occurs in all commelinid monocots examined by Karlen et al. (2018) but has also recently been reported in the eudicot mulberry (Hellinger et al., 2022) but mulberry is not within set of eudicot genomes used in Figure 5. None of analyzed species outside the commelinid monocots encode close orthologs to the AT proteins, but there are some remote orthologs to *OsAT3* and *OsAT9* detected in the non-commelinid monocot, *Dioscorea rotundata*, and of *OsAT3*, OsAT4, and OsAT7 and OsAT9 within eudicots. This suggests the origin of Mitchell clade was a gene present in the common ancestor to monocots and eudicots that underwent sequence divergence and gene duplication first in commelinid moncocots and then further in grasses, whereas the genes were lost in most eudicots.

The distribution of cell wall hydroxycinnamates is, however, known to be broader than that of the Mitchell subclade genes. Karlen et al. (2016) showed that whilst feruloylated lignin occurs in all grasses tested, it is also detected in dicots like poplar, balsa, aspen, red maple, Babylon willow, eucalyptus, hibiscus, and *Angelica sinensis*. As those authors discussed, this is likely due to convergent evolution of other unrelated BAHDs such as AsFMT. FA also occurs as a cross-linking moiety on pectin in cell walls of dicots in order *Caryophyllales*, e.g. spinach (Fry, 1986), and FA and *p*CA have both also been reported in primary cell walls of gymnosperms ester-linked to an unknown component (Carnachan and Harris, 2000). Another unrelated BAHD has recently been shown to be responsible for the acylation of lignin with the phenolic acid *p*-hydroxybenzoate in poplar (de Vries et al., 2021; Zhao et al., 2021). While convergent evolution of addition of hydroxycinnamates and similar phenolics to cell wall polymers therefore appears widespread in seed plants, to-date, hydroxycinnamates acylation of GAX appears to be confined to commelinid monocots.

The Mitchell subclade ATs fall into four groups (a-d, Figure 4) due to multiple small differences in sequence. Group 'd' contains the studied hydroxycinnamate monolignol transferases (AT3, AT4, AT5) and 'c' contains the apparent GAX-transferases (AT9 and AT10). Thus, enzymes within groups 'd' and 'c' likely act on common acceptor substrates, but varied CoA donors. The other two groups contain less well- or un-characterized genes and furthermore, the bootstrap support from the phylogenetic analysis does not position group 'a' confidently relative to the others (Figure 4).

Gene expression data can hint at roles of ATs, particularly relative to their function in synthesis of PCW and SCW in grasses. Figure 5B illustrates Mitchell subclade AT gene expression in rice leaf [from (Wang et al., 2014a)] along a developmental gradient, from the intercalary meristem at the leaf base (segment 1, on the left), an elongation zone (segments 2 and 3), to the transition to SCW formation [approximately segment 4, based on peak expression of SCW-inducing transcription factors (Li et al., 2010)]. *OsAT3* and *OsAT4*, encoding the PMTs, possess similar profiles that peak just in segments 2 and 3, respectively. Consistent with its assignment as another PMT (Sibout et al., 2016), this is also the pattern of expression for *OsAT8* (group d), which shows the second highest transcript abundance among the Mitchell subclade. On the other hand, *OsAT9*, which shows the highest transcript abundance in the clade, is highest in segment 1, where mostly PCW synthesis occurs, and continues to be abundant until past the SCW peak. *OsAT1* (group a), which has also been tentatively assigned as a FAT (Buanafina et al., 2016), shows a similar pattern. Potentially with implications for a particular function of the *p*CA-GAX modification later in development, *OsAT10* displays the latest expression peak, at segment 4. By contrast, consistent with a potential role in maintaining lignin flexibility/lability early in development (i.e. in still elongating vascular cells), putative FMT, *OsAT5*, which is among the lowest expressed of the Mitchell subclade in the leaf, peaks early in development. The uncharacterized ATs (*OsAT2*, *OsAT6*, and *OsAT7*) all exhibit a similar double peak of expression, with an initial peak in segments 1 or 2 and a second peak in segments 4 or 5, suggesting these genes might function in both PCW and SCW synthesis (Lin et al., 2016).

Models for the mechanism of incorporation of HCAs into lignin and xylan

FA on GAX may derive from a different pool of phenylpropanoids than *p*CA on lignin

The *p*CA-CoA and FA-CoA molecules that act as donors for the ATs are metabolites within the phenylpropanoid pathway that synthesizes monolignols. In grasses, recent evidence points to the presence of two largely separate *p*CA-CoA pools derived from the two phenylpropanoid pathway precursors phenylalanine and tyrosine

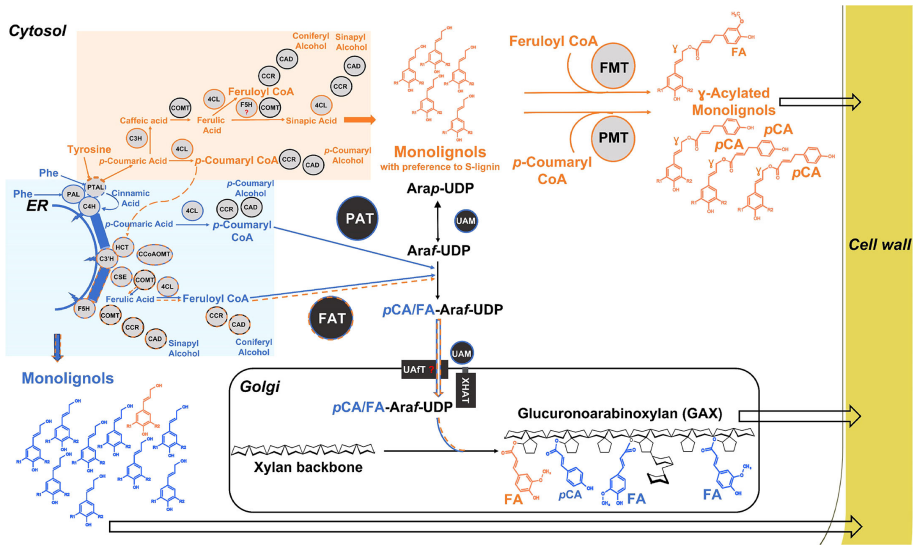


FIGURE 6
Model of hydroxycinnamoylation of lignin and GAX in grasses. The model shows putative pathways for the synthesis of HCAs and monolignols and their routes into the grass cell wall. The phenylalanine precursor-based primary phenylpropanoid pathway with early steps catalyzed by ER-localized enzymes is highlighted in blue and the orange color represents the tyrosine precursor-based pathway; some enzymes (4CL, CAD, CCR, COMT) are shared between these pathways but are shown separately for clarity. Dashed lines are our speculations. Question marks are added for the cytosolic F5H and UAfT as they have yet to be identified. γ-acylated monolignols: S lignin – R1 = R2 = OMe, G lignin – R1 = OMe, R2 = H. Acyltransferases – PMT, *p*-Coumaryl CoA Monolignol Transferase, FMT, Feruloyl CoA Monolignol Transferase; PAT, *p*-Coumaryl CoA Arabinoxylan Transferase; FAT, Feruloyl CoA *Araf* transferase; Lignin biosynthetic enzymes – PAL – monofunctional Phenylalanine Ammonia-Lyase, PTAL, bifunctional Phenylalanine/Tyrosine Ammonia-Lyase; C4H, Cinnamate 4-Hydroxylase; C3'H - 4, Coumaroyl Shikimate/Quinate 3-Hydroxylase, HCT - Hydroxycinnamoyl CoA Shikimate/Quinate Hydroxycinnamoyl Transferase, 4CL - 4-Coumarate : CoA Ligase, F5H1, Ferulate 5-Hydroxylase; C3H - bifunctional 4-Coumarate 3-Hydroxylase/Cytosolic Ascorbate Peroxidase, COMT - Caffeic Acid/5-Hydroxyferulic Acid 3/5-O-Methyltransferase, CSE, caffeoyl shikimate esterase; CCoAOMT, caffeoyl CoA 3-O-methyltransferase; CCR, cinnamoyl CoA reductase; CAD - cinnamyl alcohol dehydrogenase; GAX related enzymes – UAM – UDP-arabinose mutase, UAfT, UDP-*Araf* transporter; XHAT, xylan (hydroxycinnamoyl)-*Araf* transferase; ER, Endoplasmic reticulum.

(Barros et al., 2016; Wang et al., 2018; Simpson et al., 2021; Barros et al., 2022). Current thinking is that endoplasmic reticulum (ER)-associated cytochrome P450 enzymes, C4H, C3′H, and F5H, form a metabolon with the soluble enzymes, PAL, 4CL, HCT, facilitating metabolic channeling (Winkel, 2004; Bassard et al., 2012; Gou et al., 2018; Zhang et al., 2022). (See the legend of Figure 6 for the enzyme acronyms used here.) PTAL, a bifunctional enzyme, might also be involved in metabolon formation with ER-associated enzymes. Initial evidence for multiple *p*CA pools is the observation that PTAL's tyrosine ammonia-lyase activity provides half the total lignin in Brachypodium stems, and wall-bound *p*CA with minimal contribution to wall-bound FA (Barros et al., 2016). Further, downregulation of C3′H and F5H in rice decrease unacylated G/S-lignin but do not alter amounts of *p*-coumaroylated G- or S-lignin; C3′H downregulation also significantly decreases wall-bound FA (Takeda et al., 2017; Takeda et al., 2018; Takeda et al., 2019). Thus, one explanation for the observation that unacylated monolignols and FA on GAX and *p*CA-monoligols appear to require different enzymes, is that there are separate *p*CA pools, though other explanations, such as metabolic compensatation (Vanholme et al., 2012), is possible.

Figure 6 summarize a model, which remains to be tested, of separate tyrosine and phenyalanine-derived hydroxycinnmate pools and partially distinct cell wall products. When tyrosine enters as a substrate to the PTAL, the product, *p*CA, is not utilized by C4H and hence escapes into the cytosol avoiding the initial metabolon channel. The pool of "escaped" *pCA* is utilized by the cytosolic enzymes C3H, COMT, 4CL, CCR, and CAD to produce a part of *p*CA-CoA, FA-CoA, and γ-hydroxycinnamoyl acylated monolignols. [C3H is a recently discovered cytosolic enzyme that directly catalyzes the 3-hydroxylation of 4-coumarate to caffeate, bypassing the previously known shikimate shunt involving C3'H and HCT (Barros et al., 2019)]. A part of the *p*CA-CoA and FA-CoA produced from the "escaped *p*CA" enters back to the monolignol pathway, which is supported both by the results with the PTAL mutant and the observation that heavy atom labeled tyrosine feeding studies in sorghum do result in labeled *p*-coumaryl shikimate (Simpson et al., 2021). Thus, these recaptured hydroxycinnamates can contribute to producing minor amounts of FA-CoA, utilized by FATs to substitute FA on GAX. In contrast, the phenylalanine precursor-based PAL/PTAL-ER-associated enzymes contribute to a major part of cell-wall-associated monolignols, FA-CoA, and *p*-CA-CoA which are utilized by FATs and PATs to decorate GAX. Recently, loss-of-function of two rice 4CL homologs, Os4CL3 and Os4CL4, differentially altered non-acylated and acylated monolignol content (Afifi et al., 2022, indicating divergent roles of 4CL protein isoforms and providing further support for the model. A final step of lignin acylation, it was recently hypothesized that monolignol-FA and monolignol-*p*CA are synthesized in the cytosol and exported into the cell wall by the same simple diffusion mechanism as monolignols (Vermaas et al., 2019).

Mechanisms of hydroxycinnamoyl incorporation onto arabinoxylans

Figure 6 also illustrates a plausible model for how Mitchell subclade ATs can be responsible for FA and *p*CA incorporation into arabinoxylans (AX). Biosynthesis of AX is carried out by glycosytransferases confined to the Golgi lumen where IRX9 and IRX10 proteins participate in a xylan synthase complex (Zeng et al., 2016) and grass XAT proteins mediate Ara*f* decoration (Anders et al., 2012). However, the BAHD ATs are known to be cytosolic, as expected from their sequences which lack transmembrane domains and secretory pathway sequences. In addition, their hydroxycinnamoyl-CoA substrates are cytosolic and not known to occur in the Golgi lumen. Therefore, it seems that F/PATs must acylate a cytosolic precursor to AX synthesis just as P/FMTs acylate cytosolic lignin precursors. This conclusion can also explain the apparently surprising early result that feruloylation activity was found in the cytosolic fraction, not the membrane fraction, of rice cell cultures (Yoshida-Shimokawa et al. (2001); the Ara*f*-Xyl*p*-Xyl*p* acceptor used there is presumably not the natural one but is sufficiently close to be recognized by an endogenous FAT).

The obvious candidate for the natural cytosolic AX precursor is UDP-β-L-arabinofuranose (UDP-Ara*f*) since the UDP-arabinose mutase (UAM) responsible for its generation is localized outside the Golgi lumen, either in the cytsosol or to the Golgi perihpheral region (Konishi et al., 2011; Rautengarten et al., 2011). This is believed to be the last cytosolic step, and UDP-Ara*f* would then be transported by a nucleotide sugar transporter (UAfT) into the Golgi lumen. Therefore, the simplest model is that cytosolic BAHD ATs catalyze the acylation of UDP-Ara*f* to give UDP-Ara*f*-FA/*p*CA as intermediates (Figure 6). However, these putative products have not been identified despite targeted searches in grass tissues. One possibility is that these metabolites are only stable when bound to proteins, being generated by the action of ATs on UDP-Ara*f* whilst this is still bound to UAM before it is transferred to the UAfT transporter (Hatfield et al., 2017). A protein complex involving both UAM and AT localized to Golgi periphery has been postulated (Hatfield et al., 2017) but proteomics from Brachypodium callus suggest that whereas UAM occurs both in peripheral and cytosolic fractions, ATs occur only in cytosol (JF and RACM, unpublished). One possibility is that UAM with bound UDP-Ara*f* shuttles from the Golgi periphery to the cytosol where acylation of UDP-Ara*f* occurs before returning to the periphery to engage with UAfT. This transporter could be similar to known UDP-Ara*f* transporters (Rautengarten et al., 2017) with variation that permits the FA/*p*CA modification, which are small in comparison to UDP. Sharing most of the machinery for generating FA and *p*CA acylated UDP-Ara*f* could also explain the apparent trade off in abundance of FA- and pCA-GAX in many experiments on different grass species when PAT or FAT expression is modified.

This model also necessitates a Golgi-localized GT enzyme to attach FA/*p*CA-Ara*f* to the growing xylan molecule i.e. a xylan (hydroxcinnamoyl)-Ara*f* transferase (XHAT; Figure 6). Addition of non-acylated Ara*f* to xylan is mediated by grass-specific enzymes in GT family 61 (Anders et al., 2012) and it was reported that a closely related GT61 enzyme was responsible for addition of a xylosyl residue to GAX, so this enzyme was named XAX1 (Chiniquy et al., 2012). However more recent LC-MS analysis of sugar products released by mild acid treatment from the rice *xax1* mutant suggests that XAX1 functions in the transfer of hydroxycinnamoyl-Ara*f* to xylan, as all FA-Ara*f* and *p*CA-Ara*f* decorations of GAX were decreased in the mutant compared with the wild type (Feijao et al., 2022). This study therefore provides strong evidence that XAX1 is an XHAT responsible for the incorporation of FA/*p*CA-Ara*f* onto xylan in the Golgi lumen.

Overall, the models in Figure 6 highlight the similarity of action of PMT/FMT and FAT/PAT in acylating hydroxyl groups on, respectively, monolignols and the Ara*f* sugar in the cytosol, consistent with their similar primary sequences. Structural studies of these enzymes are required to understand the factors determining their specificities for acceptor and donor substrates.

Catalytic mechanisms of Mitchell subclade acyltransferases

The major conserved domain shared by BAHD family enzymes contains a HXXXDG motif, located near the center portion of each enzyme, with the second highly conserved region being the DFGWG motif, located near the C-terminus (D'Auria, 2006). The first crystal structure of a BAHD enzyme, vinorine synthase, was obtained by Ma et al. (2005), making a large contribution to understanding the function of conserved domains that are shared among BAHD family members. In general, the proposed catalytic mechanism involves the histidine residue in the HXXXDG motif, which deprotonates the oxygen or nitrogen atom on the corresponding acceptor substrate, allowing a nucleophilic attack on the carbonyl carbon of the CoA thioester donor, which in turn forms a tetrahedral intermediate between the CoA thioester and the acceptor substrate. This intermediate is then reprotonated, giving rise to free CoA and the acylated ester or amide. This general catalytic mechanism has however not yet been confirmed for Mitchell subclade ATs, but generalized forms of bot motifs do occur as HXXXDG and D[FY]GXG motifs in them. Although no experimental structures have been reported for the Mitchell clade ATs, the convergently evolved AsFMT structure has been solved (Liu et al., 2022). The authors showed several unique structural features of AsFMT compared to other BAHD homologs, and molecular docking studies suggest that T375 in AsFMT may function as an oxyanion hole to stabilize the reaction intermediate. These studies also proposed a role of H278 in the binding of the nucleophilic hydroxyl group of monolignols.

Biotechnological applications

Mature plant biomass, composed principally of SCWs and therefore termed lignocellulosic biomass, is a promising feedstock for production of next-generation fuels and chemicals that can replace fossil carbon sources thereby reducing greenhouse gas emissions (Farrell et al., 2006; Fargione et al., 2008; Schmer et al., 2008; Chundawat et al., 2011). To be economically and environmentally viable, this biomass can be non-food residues of crops (sugarcane bagasse, corn stover, paper mill waste and cereal straw) or from dedicated energy crops grown with minimal inputs.

However, the cost and inefficiency of depolymerizing polysaccharides to fermentable sugars, also known as cell wall recalcitrance, are important impediments to large-scale lignocellulosic biofuel production (Lynd et al., 2008).

Manipulation of acylation of SCW polymers with ferulate, in particular, is a promising approach to improve the digestibility of biomass because it combines ester links with capacity for oxidative coupling, with opposite direction of effects depending on the polymer context. Feruloylation of GAX in grasses cross-links xylan strands to each other and to lignin, increasing recalcitrance. Conversely acylation of monolignols with ferulate results in the introduction of alkali-labile ester bonds into the body of the lignin polymer improving ease of saccharification, a technology referred to as "Zip-lignin", by Ralph and colleagues (Wilkerson et al., 2014). Therefore, decreasing feruloylation of GAX in grass biomass and introducing or boosting feruloylation of monolignols in important biomass crops such as poplar are both promising biotechnological approaches.

Decreasing feruloylation in grass biomass

The *AT* genes that modify GAX represent promising targets to improve the suitability of grass lignocellulosic biomass for biofuel production. Since FA on GAX is believed to be the main means by which polysaccharide is cross-linked to lignin, grass SCW FA amounts are therefore a key to recalcitrance. Suppression of the putative FAT-encoding *SvBAHD01*/*SvAT9* in the model grass *Setaria viridis* resulted in a ~40% increases in ease of digestion of cell wall polysaccharides into sugars in the modified plants compared with the wild type (de Souza et al., 2018). Similarly, suppression of the ortholog in sugarcane (*ScBAHD01*/*ScAT9*) improved the digestibility of sugarcane straw by approximately 20% after Organosolv pretreatment, compared to non-transformed plants (de Souza et al., 2019). These results are exciting because sugarcane (*Saccharum* spp.) covers vast areas of land (around 25 million ha worldwide), and its processing is already linked into infrastructure for producing bioethanol in many countries, especially in Brazil. Furthermore, sugarcane straw and bagasses are the main industrial residues after sugarcane processing (Menandro et al., 2017). Also, the Organosolv process involves the use of an organic liquid and water to partially hydrolyze lignin bonds and lignin-carbohydrate bonds, resulting in a solid residue consisting of mainly cellulose and some hemicellulose (Zhao et al., 2009). Thus, the biomass of suppressed *ScBAHD01*/*ScAT9* plants combined with Organosolv pretreatment is an interesting approach to be incorporated in the sugarcane industry for bioethanol production (de Souza et al., 2019). In addition, the reduction in FA-AX that often accompanies increases in *p*CA-AX due to altered expression of PATs in grasses, has been accompanied by a 10 to 40% increase in saccharification depending on the assay conditions. Thus, PAT enzymes like AT10 are also an attractive biotechnological target (Bartley et al., 2013; Li et al., 2018; Mota et al., 2021).

Recently, the world's first CRISPR-edited sugarcane plants, the so-called Flex I and Flex II sugarcane, were reported (Brazilian Agricultural Research Corporation, 2021). Both Flex I and Flex II plants have CRISPR/Cas9-edited *AT* genes, and these sugarcane varieties presented higher cell wall digestibility and higher concentration of sucrose in plant tissues, respectively. The precise acyltransferase genes that were edited in these plants were not revealed, but both varieties have decreased levels of ferulate in the cell wall. Moreover, these CRISPR-edited plants were considered non-transgenic by the Brazilian National Technical Commission on Biosafety, representing an important step towards the use of this modified biomass by the bioethanol industry, as edited plants lacking foreign DNA can bypass the costly process of genetically modified-crop regulation.

Zip-lignin (feruloyl lignin) and other lignin hydroxycinnamates

Feruloyl lignin, i.e., lignin containing feruloyl monolignol conjugates (ML-FAs), facilitates depolymerization of lignin polymers by industrial processes due to the introduction of mild base-labile ester bonds into the lignin polymer. Early work on this technology revealed that incorporation of synthetic coniferyl ferulate into lignin of cell cultures enhanced alkaline delignification and enzymatic hydrolysis (Grabber et al., 2008; Ralph, 2010). Wilkerson et al. (2014) then introduced the Chinese angelica feruloyl-monolignol transferase (AsFMT) to poplar to generate ML-FAs that were incorporated into lignin polymers. The resulting biomass presented improved saccharification after mild base pretreatment (Wilkerson et al., 2014). The generation of the "zip-lignins" can be achieved either through a linear linkage, by extending the polymer chain, or by crosslinking two lignin polymers, as demonstrated elsewhere (Ralph, 2010; Rencoret et al., 2013; Lu et al., 2015; Smith et al., 2015; Kaal et al., 2018). Both biophysical and chemical changes in cell wall accessibility have been observed due to the introduction of ML-FAs in poplar lignin (Shen et al., 2019). As discussed above, FMT activity has also been demonstrated for AT5s in grasses (Karlen et al., 2016). One way to boost the effect of this endogenous enzyme was demonstrated in maize by suppression of the first lignin specific biosynthetic enzyme, cinnamoyl-CoA reductase (CCR) resulting in an increase in the intercellular pool of feruloyl-CoA and in ML-FAs and an overall decrease in lignin content thereby enhancing the digestibility of stem rind tissue (Smith et al., 2017). Similarly, ectopic expression of *PMT* genes increased saccharification yields under some reaction pretreatment conditions both in Brachypodium (Petrik et al., 2014) and *via* heterologous expression in Arabidopsis (Sibout et al., 2016). The mechanism could be due to the tendency of *p*CA-acylated monolignols to end lignin polymerization and not be included within the lignin polymer, consistent with the greater alkali solubility of Arabidopsis lignin esterified with *p*CA (Sibout et al., 2016). These results show that ML-hydroxycinnamate conjugates are a promising means for engineering bioenergy crops and waste streams of mainstream crops by conferring low-cost lignin breakdown and separation for biorefining applications.

Other biotechnological applications

Grains with increased feruloylated arabinoxylans are emerging as a potential multifunctional food and hydroxycinnamates themselves are being used as precursors for material applications. Some have reported that the presence of ferulic acid on the AX can contribute to antioxidant, anticancer and prebiotic properties (Srinivasan et al.,

2007; Snelders et al., 2014). In addition, the cross-linking of FA-AX can form covalently linked gels with potential as drug delivery systems with anticancer or antioxidant properties (Mendez-Encinas et al., 2018). Recently, a study demonstrated the optimization of FA-AX isolation from wheat bran at a pilot scale using subcritical water extraction, demonstrating the feasibility of multifunctional FA-AX-based products for food and material applications in industrial scale (Rudjito et al., 2019). Furthermore, hydroxycinnamates in pure or mixed forms are themselves being used in higher value applications. For example, bacteria have been engineered to use HCAs to synthesize muconic acid (Johnson et al., 2016), a precursor of nylon. Techno-economic analyses indicate the value of engineering biomass to predominantly (>80%) produce only a single hydroxycinnamate for use as a high-value precursor (Karlen et al., 2020). Therefore, the manipulation of *AT* genes in different plant species can improve not only the production of biofuels but may also prove important for food and pharmaceutical applications.

Outstanding questions

This review has highlighted some clear gaps in our knowledge that could be the focus of future research.

For plant science discovery:

- What is the acceptor molecule for ATs responsible for addition of FA and *p*CA to xylan and the pathway for their incorporation?
- Can experimental structural determination of AT enzymes and molecular docking studies explain their acceptor and donor substrate specificities?
- Evidence suggests that GAX-FA plays a key role in cross-linking between xylan chains and from xylan to lignin in grass cell walls; how is this cross-linking controlled?
- What are activities of the uncharacterized Mitchell subclade i and subclade ii ATs?
- What is the function (fitness advantage) of xylan *p*-coumarylation, lignin *p*-coumarylation and lignin feruloylation in commelinids?

For biotechnology applications it seems likely there are limits to manipulation of ATs before negative side effects occur:

- How much lignin FA and lignin *p*CA is too much?
- How much GAX-feruloylation is too little?

Addressing these questions will provide insight into the factors that have driven the evolution of grass cell wall properties, reveal molecular means to incorporate beneficial agronomic features associated with the hydroxycinnamates into food crop species, and potentially lead to the greater utilization of biomass and hydroxycinnamates themselves in the bio-economy.

Author contributions

RM, CZ, WS, LB, and NC wrote the manuscript. NC and RM made the figures. LB, RM, WS, and NC revised the text and figures. All authors contributed to the article and approved the submitted version.

Funding

LB and NC's contributions were supported by US-DOE-BER award DE-SC0021126 and USDA-NIFA Hatch project #1015621. WS contributions were supported by Fapesp award Proc. 2019/04878-7. JF and RM contributions supported by UK BBSRC award BB/K007599/1.

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.

Publisher's note

References

Afifi, O. A., Tobimatsu, Y., Lam, P. Y., Martin, A. F., Miyamoto, T., Osakabe, Y., et al. (2022). Genome-edited rice deficient in two 4-COUMARATE:COENZYME a LIGASE genes displays diverse lignin alterations. *Plant Physiol* 190:2155–2172. doi: 10.1093/plphys/kiac450

Akin, D. E. (2008). Plant cell wall aromatics: influence on degradation of biomass. *Biofuels Bioprod. Biorefining* 2 (4), 288–303. doi: 10.1002/bbb.76

Anders, N., Wilkinson, M. D., Lovegrove, A., Freeman, J., Tryfona, T., Pellny, T. K., et al. (2012). Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proc. Natl. Acad. Sci. U.S.A.* 109 (3), 989–993. doi: 10.1073/pnas.1115858109

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 (1), 1994. doi: 10.1038/s41467-019-10082-7

Barros, J., Serrani-Yarce, J. C., Chen, F., Baxter, D., Venables, B. J., and Dixon, R. A. (2016). Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nat. Plants* 2 (6), 16050. doi: 10.1038/nplants.2016.50

Barros, J., Shrestha, H. K., Serrani-Yarce, J. C., Engle, N. L., Abraham, P. E., Tschaplinski, T. J., et al. (2022). Proteomic and metabolic disturbances in lignin-modified brachypodium distachyon. *Plant Cell* 34 (9), 3339–3363. doi: 10.1093/plcell/koac171

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 (4), 1615–1633. doi: 10.1104/pp.112.208694

Bassard, J. E., Richert, L., Geerinck, J., Renault, H., Duval, F., Ullmann, P., et al. (2012). Protein-protein and protein-membrane associations in the lignin pathway. *Plant Cell* 24 (11), 4465–4482. doi: 10.1105/tpc.112.102566

Beekwilder, J., Alvarez-Huerta, M., Neef, E., Verstappen, F. W., Bouwmeester, H. J., and Aharoni, A. (2004). Functional characterization of enzymes forming volatile esters from strawberry and banana. *Plant Physiol.* 135 (4):1865–1878. doi: 10.1104/pp.104.042580

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi: 10.1146/annurev.arplant.54.031902.134938

Brazilian Agricultural Research Corporation (2021) *Embrapa*. Available at: https://www.embrapa.br/en/busca-de-noticias/-/noticia/66969890/brazilian-science-develops-first-non-gm-gene-edited-sugarcane-of-the-world.

Buanafina, M. M. (2009). Feruloylation in grasses: current and future perspectives. *Mol. Plant* 2 (5), 861–872. doi: 10.1093/mp/ssp067

Buanafina, M. M., Fescemyer, H. W., Sharma, M., and Shearer, E. A. (2016). Functional testing of a PF02458 homologue of putative rice arabinoxylan feruloyl transferase genes in brachypodium distachyon. *Planta* 243 (3), 659–674. doi: 10.1007/s00425-015-2430-1

Bunzel, M., Heuermann, B., Kim, H., and Ralph, J. (2008). Peroxidase-catalyzed oligomerization of ferulic acid esters. *J. Agric. Food Chem.* 56 (21), 10368–10375. doi: 10.1021/jf801825z

Bunzel, M., Ralph, J., Kim, H., Lu, F., Ralph, S. A., Marita, J. M., et al. (2003). Sinapate dehydrodimers and sinapate–ferulate heterodimers in cereal dietary fiber. *J. Agric. Food Chem.* 51 (5), 1427–1434. doi: 10.1021/jf020910v

Carnachan, S. M., and Harris, P. J. (2000). Ferulic acid is bound to the primary cell walls of all gymnosperm families. *Biochem. Syst. Ecol.* 28 (9), 865–879. doi: 10.1016/s0305-1978(00)00009-0

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.* 109 (42), 17117–17122. doi: 10.1073/pnas.1202079109

Chundawat, S. P. S., Beckham, G. T., Himmel, M. E., and Dale, B. E. (2011). Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annu. Rev. Chem. Biomol. Eng.* 2 (1):121-145. doi: 10.1146/annurev-chembioeng-061010-114205

D'Auria, J. C. (2006). Acyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol.* 9 (3), 331–340. doi: 10.1016/j.pbi.2006.03.016

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 (9), 1182–1190. doi: 10.1039/B926617H

de Oliveira, D. M., Finger-Teixeira, A., Mota, T. R., Salvador, V. H., Moreira-Vilar, F. C., Molinari, H. B., et al. (2015). Ferulic acid: a key component in grass lignocellulose recalcitrance to hydrolysis. *Plant Biotechnol. J.* 13 (9), 1224–1232. doi: 10.1111/pbi.12292

de Souza, W. R., Martins, P. K., Freeman, J., Pellny, T. K., Michaelson, L. V., Sampaio, B. L., et al. (2018). Suppression of a single BAHD gene in setaria viridis causes large, stable decreases in cell wall feruloylation and increases biomass digestibility. *New Phytol.* 218 (1), 81–93. doi: 10.1111/nph.14970

de Souza, W. R., Pacheco, T. F., Duarte, K. E., Sampaio, B. L., de Oliveira Molinari, P. A., Martins, P. K., et al. (2019). Silencing of a BAHD acyltransferase in sugarcane increases biomass digestibility. *Biotechnol. Biofuels* 12, 111–111. doi: 10.1186/s13068-019-1450-7

de Vries, L., MacKay, H. A., Smith, R. A., Mottiar, Y., Karlen, S. D., Unda, F., et al. (2021). pHBMT1, a BAHD-family monolignol acyltransferase, mediates lignin acylation in poplar. *Plant Physiol.* 188 (2), 1014–1027. doi: 10.1093/plphys/kiab546

Duan, P., Kaser, S. J., Lyczakowski, J. J., Phyo, P., Tryfona, T., Dupree, P., et al. (2021). Xylan structure and dynamics in native brachypodium grass cell walls investigated by solid-state NMR spectroscopy. *ACS Omega* 6 (23), 15460–15471. doi: 10.1021/acsomega.1c01978

Fanelli, A., Rancour, D. M., Sullivan, M., Karlen, S. D., Ralph, J., Riaño-Pachón, D. M., et al. (2021). Overexpression of a sugarcane BAHD acyltransferase alters hydroxycinnamate content in maize cell wall. *Front. Plant Sci.* 12, 626168. doi: 10.3389/fpls.2021.626168

Fargione, J., Hill, J., Tilman, D., Polasky, S., and Hawthorne, P. (2008). Land clearing and the biofuel carbon debt. *Science* 319 (5867), 1235–1238. doi: 10.1126/science.1152747

Farrell, A. E., Plevin, R. J., Turner, B. T., Jones, A. D., O'Hare, M., and Kammen, D. M. (2006). Ethanol can contribute to energy and environmental goals. *Science* 311 (5760), 506–508. doi: 10.1126/science.1121416

Feijao, C., Morreel, K., Anders, N., Tryfona, T., Busse-Wicher, M., Kotake, T., et al. (2022). Hydroxycinnamic acid-modified xylan side chains and their cross-linking products in rice cell walls are reduced in the xylosyl arabinosyl substitution of xylan 1 mutant. *Plant J.* 109 (5), 1152–1167. doi: 10.1111/tpj.15620

Fry, S. C. (1986). Cross-linking of matrix polymers in the growing cell walls of Angiosperms. *Annu. Rev. Plant Biol.* 37, 165–186.

Gao, Y., Lipton, A. S., Wittmer, Y., Murray, D. T., and Mortimer, J. C. (2020). A grass-specific cellulose–xylan interaction dominates in sorghum secondary cell walls. *Nat. Commun.* 11 (1), 6081. doi: 10.1038/s41467-020-19837-z

Gou, M., Ran, X., Martin, D. W., and Liu, C.-J. (2018). The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. *Nat. Plants* 4 (5), 299–310. doi: 10.1038/s41477-018-0142-9

Grabber, J. H., Hatfield, R. D., Lu, F., and Ralph, J. (2008). Coniferyl ferulate incorporation into lignin enhances the alkaline delignification and enzymatic degradation of cell walls. *Biomacromolecules* 9 (9), 2510–2516. doi: 10.1021/bm800528f

Grabber, J. H., Hatfield, R. D., and Ralph, J. (1998). Diferulate cross-links impede the enzymatic degradation of non-lignified maize walls. *J. Sci. Food Agric.* 77 (2), 193–200. doi: 10.1002/(SICI)1097-0010(199806)77:2<193::AID-JSFA25>3.0.CO;2-A

Grabber, J. H., Ralph, J., and Hatfield, R. D. (2002). Model studies of ferulate-coniferyl alcohol cross-product formation in primary maize walls: implications for lignification in grasses. *J. Agric. Food Chem.* 50 (21), 6008–6016. doi: 10.1021/jf0205312

Harris, P. J., and Trethewey, J. A. K. (2010). The distribution of ester-linked ferulic acid in the cell walls of angiosperms. *Phytochem. Rev.* 9 (1):19–33. doi: 10.1007/s11101-009-9146-4

Hatfield, R. D., Ralph, J., and Grabber, J. H. (1999). Cell wall cross-linking by ferulates and diferulates in grasses. *J. Sci. Food Agric.* 79 (3), 403–407. doi: 10.1002/(SICI)1097-0010(19990301)79:3<403::AID-JSFA263>3.0.CO;2-0

Hatfield, R. D., Rancour, D. M., and Marita, J. M. (2017). Grass cell walls: A story of cross-linking. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.02056

Hellinger, J., Kim, H., Ralph, J., and Karlen, S. D. (2022). P-coumaroylation of lignin occurs outside of commelinid monocots in the eudicot genus morus (mulberry). *Plant Physiol.* doi: 10.1093/plphys/kiac485

Houston, K., Learmonth, A., Hassan, A. S., Lahnstein, J., Looseley, M., Little, A., et al. (2020). The p-coumaroyl arabinoxylan transferase HvAT10 underlies natural variation in whole-grain cell wall phenolic acids in cultivated barley. *bioRxiv* 2020, 2012.2021.423816. doi: 10.1101/2020.12.21.423816

Hyde, L. S., Pellny, T. K., Freeman, J., Michaelson, L. V., Simister, R., McQueen-Mason, S. J., et al. (2018). Response of cell-wall composition and RNA-seq transcriptome to methyl-jasmonate in brachypodium distachyon callus. *Planta* 248 (5), 1213–1229. doi: 10.1007/s00425-018-2968-9

Ishii, T. (1997). Structure and functions of feruloylated polysaccharides. *Plant Sci.* 127 (2), 111–127. doi: 10.1016/S0168-9452(97)00130-1

Johnson, K. L., Cassin, A. M., Lonsdale, A., Wong, G. K.-S., Soltis, D. E., Miles, N. W., et al. (2017). Insights into the evolution of hydroxyproline-rich glycoproteins from 1000 plant transcriptomes. *Plant Physiol.* 174 (2), 904–921. doi: 10.1104/pp.17.00295

Johnson, C. W., Salvachúa, D., Khanna, P., Smith, H., Peterson, D. J., and Beckham, G. T. (2016). Enhancing muconic acid production from glucose and lignin-derived aromatic compounds *via* increased protocatechuate decarboxylase activity. *Metab. Eng. Commun.* 3, 111–119. doi: 10.1016/j.meteno.2016.04.002

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8 (3), 275–282. doi: 10.1093/bioinformatics/8.3.275

Kaal, J., Serrano, O., del Rio, J. C., and Rencoret, J. (2018). Radically different lignin composition in posidonia species may link to differences in organic carbon sequestration capacity. *Org. Geochem.* 124, 247–256. doi: 10.1016/j.orggeochem.2018.07.017

Kang, X., Kirui, A., Dickwella Widanage, M. C., Mentink-Vigier, F., Cosgrove, D. J., and Wang, T. (2019). Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* 10 (1), 347. doi: 10.1038/s41467-018-08252-0

Karlen, S. D., Fasahati, P., Mazaheri, M., Serate, J., Smith, R. A., Sirobhushanam, S., et al. (2020). Assessing the viability of recovery of hydroxycinnamic acids from lignocellulosic biorefinery alkaline pretreatment waste streams. *ChemSusChem* 13 (8), 2012–2024. doi: 10.1002/cssc.201903345

Karlen, S. D., Free, H. C. A., Padmakshan, D., Smith, B. G., Ralph, J., and Harris, P. J. (2018). Commelinid monocotyledon lignins are acylated by *p*-coumarate. *Plant Physiol.* 177 (2), 513–521. doi: 10.1104/pp.18.00298

Karlen, S. D., Zhang, C., Peck, M. L., Smith, R. A., Padmakshan, D., Helmich, K. E., et al. (2016). Monolignol ferulate conjugates are naturally incorporated into plant lignins. *Sci. Adv.* 2 (10), e1600393. doi: 10.1126/sciadv.1600393

Kim, I. A., Kim, B. G., Kim, M., and Ahn, J. H. (2012). Characterization of hydroxycinnamoyltransferase from rice and its application for biological synthesis of hydroxycinnamoyl glycerols. *Phytochemistry* 76, 25–31. doi: 10.1016/j.phytochem.2011.12.015

Konishi, T., Aohara, T., Igasaki, T., Hayashi, N., Miyazaki, Y., Takahashi, A., et al. (2011). Down-regulation of UDP-arabinopyranose mutase reduces the proportion of arabinofuranose present in rice cell walls. *Phytochemistry* 72 (16), 1962–1968. doi: 10.1016/j.phytochem.2011.07.012

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35 (6), 1547–1549. doi: 10.1093/molbev/msy096

Lapierre, C., Voxeur, A., Boutet, S., and Ralph, J. (2019). Arabinose conjugates diagnostic of ferulate-ferulate and ferulate-monolignol cross-coupling are released by mild acidolysis of grass cell walls. *J. Agric. Food Chem.* 67 (46), 12962–12971. doi: 10.1021/acs.jafc.9b05840

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 (1), 54. doi: 10.1186/s12896-018-0464-8

Lin, F., Manisseri, C., Fagerstrom, A., Peck, M. L., Vega-Sanchez, M. E., Williams, B., et al. (2016). Cell wall composition and candidate biosynthesis gene expression during rice development. *Plant Cell Physiol.* 57 (10), 2058–2075. doi: 10.1093/pcp/pcw125

Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S. L., et al. (2010). The developmental dynamics of the maize leaf transcriptome. *Nat. Genet.* 42 (12), 1060–1067. doi: 10.1038/ng.703

Liu, X., Dai, S., Zhou, Y., Liu, J., Li, D., Zhang, J., et al. (2022). Crystal structure of the plant feruloyl–coenzyme a monolignol transferase provides insights into the formation of monolignol ferulate conjugates. *Biochem. Biophys. Res. Commun.* 594, 8–14. doi: 10.1016/j.bbrc.2022.01.037

Lu, F., Karlen, S. D., Regner, M., Kim, H., Ralph, S. A., Sun, R.-C., et al. (2015). Naturally p-hydroxybenzoylated lignins in palms. *Bioenergy Res.* 8 (3), 934–952. doi: 10.1007/s12155-015-9583-4

Luo, J., Nishiyama, Y., Fuell, C., Taguchi, G., Elliott, K., Hill, L., et al. (2007). Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from arabidopsis thaliana. *Plant J.* 50 (4), 678–695. doi: 10.1111/j.1365-313X.2007.03079.x

Lynd, L., Laser, M., Bransby, D., Dale, B. E., Davison, B., Hamilton, R., et al. (2008). How biotech can transform biofuels. *Nat Biotechnol* 26, 169–172. doi: 10.1038/nbt0208-169.

Ma, X., Koepke, J., Panjikar, S., Fritzsch, G., and Stockigt, J. (2005). Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. *J. Biol. Chem.* 280 (14), 13576–13583. doi: 10.1074/jbc.M414508200

Marita, J. M., Hatfield, R. D., Rancour, D. M., and Frost, K. E. (2014). Identification and suppression of the p-coumaroyl CoA:hydroxycinnamyl alcohol transferase in zea mays l. *Plant J.* 78 (5), 850–864. doi: 10.1111/tpj.12510

Menandro, L. M. S., Cantarella, H., Franco, H. C. J., Kölln, O. T., Pimenta, M. T. B., Sanches, G. M., et al. (2017). Comprehensive assessment of sugarcane straw: implications for biomass and bioenergy production. *Biofuels Bioprod. Biorefining* 11 (3), 488–504. doi: 10.1002/bbb.1760

Mendez-Encinas, M. A., Carvajal-Millan, E., Rascon-Chu, A., Astiazaran-Garcia, H. F., and Valencia-Rivera, D. E. (2018). Ferulated arabinoxylans and their gels: Functional properties and potential application as antioxidant and anticancer agent. *Oxid. Med. Cell. Longevity* 2018, 2314759–2314759. doi: 10.1155/2018/2314759

Mitchell, R. A. C., Dupree, P., and Shewry, P. R. (2007). A novel bioinformatics approach identifies candidate genes for the synthesis and feruloylation of arabinoxylan. *Plant Physiol.* 144 (1), 43–53. doi: 10.1104/pp.106.094995

Molinari, H., Pellny, T., Freeman, J., Shewry, P., and Mitchell, R. (2013). Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in brachypodium distachyon. *Front. Plant Sci.* 4. doi: 10.3389/fpls.2013.00050

Möller, S. R., Lancefield, C. S., Oates, N. C., Simister, R., Dowle, A., Gomez, L. D., et al. (2022). CRISPR/Cas9 suppression of OsAT10, a rice BAHD acyltransferase, reduces p-coumaric acid incorporation into arabinoxylan without increasing saccharification. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.926300

Mota, T. R., de Souza, W. R., Oliveira, D. M., Martins, P. K., Sampaio, B. L., Vinecky, F., et al. (2021). Suppression of a BAHD acyltransferase decreases p-coumaroyl on arabinoxylan and improves biomass digestibility in the model grass setaria viridis. *Plant J.* 105 (1), 136–150. doi: 10.1111/tpj.15046

Obel, N., Porchia, A. C., and Scheller, H. V. (2002). Dynamic changes in cell wall polysaccharides during wheat seedling development. *Phytochemistry* 60 (6), 603–610. doi: 10.1016/s0031-9422(02)00148-6

Petrik, D. L., Karlen, S. D., Cass, C. L., Padmakshan, D., Lu, F., Liu, S., et al. (2014). P-Coumaroyl-CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in brachypodium distachyon. *Plant J.* 77 (5), 713–726. doi: 10.1111/tpj.12420

Piston, F., Uauy, C., Fu, L., Langston, J., Labavitch, J., and Dubcovsky, J. (2010). Down-regulation of four putative arabinoxylan feruloyl transferase genes from family PF02458 reduces ester-linked ferulate content in rice cell walls. *Planta* 231 (3), 677–691. doi: 10.1007/s00425-009-1077-1

Ralph, J. (2010). Hydroxycinnamates in lignification. *Phytochem. Rev.* 9 (1), 65–83. doi: 10.1007/s11101-009-9141-9

Ralph, J., Bunzel, M., Marita, J. M., Hatfield, R. D., Lu, F., Kim, H., et al. (2004). Peroxidase-dependent cross-linking reactions of p-hydroxycinnamates in plant cell walls. *Phytochem. Rev.* 3 (1), 79–96. doi: 10.1023/B:PHYT.0000047811.13837.fb

Ralph, J., Grabber, J. H., and Hatfield, R. D. (1995). Lignin-ferulate cross-links in grasses: active incorporation of ferulate polysaccharide esters into ryegrass lignins. *Carbohydr. Res.* 275, 167–178. doi: 10.1016/0008-6215(95)00237-N

Ralph, J., Hatfield, R. D., Grabber, J. H., Jung, H.-J. G., Quideau, S., and Helm, R. F. (1998). "Cell wall cross-linking in grasses by ferulates and diferulates," in *Lignin and lignan biosynthesis* (American Chemical Society), 209–236. doi: 10.1021/bk-1998-0697.ch016

Ralph, J., Helm, R. F., Quideau, S., and Hatfield, R. D. (1992). Lignin–feruloyl ester cross-links in grasses. part 1. incorporation of feruloyl esters into coniferyl alcohol dehydrogenation polymers. *J. Chem. Society Perkin Trans.* 21), 2961–2969. doi: 10.1039/P19920002961

Rautengarten, C., Birdseye, D., Pattathil, S., McFarlane, H. E., Saez-Aguayo, S., Orellana, A., et al. (2017). The elaborate route for UDP-arabinose delivery into the golgi of plants. *Proc. Natl. Acad. Sci.* 114 (16), 4261–4266. doi: 10.1073/pnas.1701894114

Rautengarten, C., Ebert, B., Herter, T., Petzold, C. J., Ishii, T., Mukhopadhyay, A., et al. (2011). The interconversion of UDP-arabinopyranose and UDP-arabinofuranose is indispensable for plant development in arabidopsis. *Plant Cell* 23 (4), 1373–1390. doi: 10.1105/tpc.111.083931

Rencoret, J., Ralph, J., Marques, G., Gutiérrez, A., Martínez, Á. T., and del Río, J. C. (2013). Structural characterization of lignin isolated from coconut (Cocos nucifera) coir fibers. *J. Agric. Food Chem.* 61 (10), 2434–2445. doi: 10.1021/jf304686x

Rudjito, R. C., Ruthes, A. C., Jiménez-Quero, A., and Vilaplana, F. (2019). Feruloylated arabinoxylans from wheat bran: Optimization of extraction process and validation at pilot scale. *ACS Sustain. Chem. Eng.* 7 (15), 13167–13177. doi: 10.1021/acssuschemeng.9b02329

Saulnier, L., Vigouroux, J., and Thibault, J.-F. (1995). Isolation and partial characterization of feruloylated oligosaccharides from maize bran. *Carbohydr. Res.* 272 (2), 241–253. doi: 10.1016/0008-6215(95)00053-V

Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61 (1), 263–289. doi: 10.1146/annurev-arplant-042809-112315

Schmer, M. R., Vogel, K. P., Mitchell, R. B., and Perrin, R. K. (2008). Net energy of cellulosic ethanol from switchgrass. *Proc. Natl. Acad. Sci.* 105 (2), 464–469. doi: 10.1073/pnas.0704767105

Shen, W., Collings, C., Li, M., Markovicz, J., Ralph, J., Mansfield, S. D., et al. (2019). Imaging changes in cell walls of engineered poplar by stimulated raman scattering and atomic force microscopy. *ACS Sustain. Chem. Eng.* 7 (12), 10616–10622. doi: 10.1021/acssuschemeng.9b01166

Shen, H., Fu, C., Xiao, X., Ray, T., Tang, Y., Wang, Z.-Y., et al. (2009). Developmental control of lignification in stems of lowland switchgrass variety Alamo and the effects on saccharification efficiency. *Bioenergy Res.* 2, 233–245. doi: 10.1007/s12155-009-9058-6

Sibout, R., Le Bris, P., Legee, F., Cezard, L., Renault, H., and Lapierre, C. (2016). Structural redesigning arabidopsis lignins into alkali-soluble lignins through the expression of p-coumaroyl-CoA: monolignol transferase PMT. *Plant Physiol.* 170 (3), 1358–1366. doi: 10.1104/pp.15.01877

Simmons, T. J., Mortimer, J. C., Bernardinelli, O. D., Pöppler, A.-C., Brown, S. P., deAzevedo, E. R., et al. (2016). Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* 7 (1), 13902. doi: 10.1038/ncomms13902

Simpson, J. P., Olson, J., Dilkes, B., and Chapple, C. (2021). Identification of the tyrosine- and phenylalanine-derived soluble metabolomes of sorghum. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.714164

Smith, R. A., Beebe, E. T., Bingman, C. A., Vander Meulen, K., Eugene, A., Steiner, A. J., et al. (2022). Identification and characterization of a set of monocot BAHD monolignol transferases. *Plant Physiol.* 189 (1), 37–48. doi: 10.1093/plphys/kiac035

Smith, R. A., Cass, C. L., Mazaheri, M., Sekhon, R. S., Heckwolf, M., Kaeppler, H., et al. (2017). Suppression of CINNAMOYL-CoA REDUCTASE increases the level of monolignol ferulates incorporated into maize lignins. *Biotechnol. Biofuels* 10 (1), 109. doi: 10.1186/s13068-017-0793-1

Smith, R. A., Gonzales-Vigil, E., Karlen, S. D., Park, J.-Y., Lu, F., Wilkerson, C. G., et al. (2015). Engineering monolignol *p*-coumarate conjugates into poplar and arabidopsis lignins. *Plant Physiol.* 169 (4), 2992–3001. doi: 10.1104/pp.15.00815

Snelders, J., Dornez, E., Delcour, J. A., and Courtin, C. M. (2014). Impact of wheat bran derived arabinoxylanoligosaccharides and associated ferulic acid on dough and bread properties. *J. Agric. Food Chem.* 62 (29), 7190–7199. doi: 10.1021/jf502315g

Soreng, R. J., Peterson, P. M., Romaschenko, K., Davidse, G., Zuloaga, F. O., Judziewicz, E. J., et al. (2015). A worldwide phylogenetic classification of the poaceae (Gramineae). *J. Systematics Evol.* 53 (2), 117–137. doi: 10.1111/jse.12150

Srinivasan, M., Sudheer, A. R., and Menon, V. P. (2007). Ferulic acid: therapeutic potential through its antioxidant property. *J. Clin. Biochem. Nutr.* 40 (2), 92–100. doi: 10.3164/jcbn.40.92

Sullivan, A., Purohit, P. K., Freese, N. H., Pasha, A., Esteban, E., Waese, J., et al. (2019). An 'eFP-seq browser' for visualizing and exploring RNA sequencing data. *Plant J.* 100 (3), 641–654. doi: 10.1111/tpj.14468

Takahama, U., and Oniki, T. (1994). Effects of ascorbate on the oxidation of derivatives of hydroxycinnamic acid and the mechanism of oxidation of sinapic acid by cell wall-bound peroxidases. *Plant Cell Physiol.* 35 (4), 593–600. doi: 10.1093/oxfordjournals.pcp.a078634

Takeda, Y., Koshiba, T., Tobimatsu, Y., Suzuki, S., Murakami, S., Yamamura, M., et al. (2017). Regulation of CONIFERALDEHYDE 5-HYDROXYLASE expression to modulate cell wall lignin structure in rice. *Planta* 246 (2), 337–349. doi: 10.1007/s00425-017-2692-x

Takeda, Y., Suzuki, S., Tobimatsu, Y., Osakabe, K., Osakabe, Y., Ragamustari, S. K., et al. (2019). Lignin characterization of rice CONIFERALDEHYDE 5-HYDROXYLASE loss-of-function mutants generated with the CRISPR/Cas9 system. *Plant J.* 97 (3), 543–554. doi: 10.1111/tpj.14141

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

Terrett, O., 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

Tian, Y., Lin, C.-Y., Park, J.-H., Wu, C.-Y., Kakumanu, R., Pidatala, V. R., et al. (2021). Overexpression of the rice BAHD acyltransferase AT10 increases xylan-bound p-coumarate and reduces lignin in sorghum bicolor. *Biotechnol. Biofuels* 14 (1), 217. doi: 10.1186/s13068-021-02068-9

Tuominen, L. K., Johnson, V. E., and Tsai, C.-J. (2011). Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among populus paralogues. *BMC Genomics* 12 (1), 236. doi: 10.1186/1471-2164-12-236

Vanholme, R., Storme, V., Vanholme, B., Sundin, L., Christensen, J. H., Goeminne, G., et al. (2012). A systems biology view of responses to lignin biosynthesis perturbations in arabidopsis. *Plant Cell* 24 (9), 3506–3529. doi: 10.1105/tpc.112.102574

Vermaas, J. V., Dixon, R. A., Chen, F., Mansfield, S. D., Boerjan, W., Ralph, J., et al. (2019). Passive membrane transport of lignin-related compounds. *Proc. Natl. Acad. Sci.* 116 (46), 23117–23123. doi: 10.1073/pnas.1904643116

Wakabayashi, K., Hoson, T., and Kamisaka, S. (1997). Osmotic stress suppresses cell wall stiffening and the increase in cell wall-bound ferulic and diferulic acids in wheat coleoptiles. *Plant Physiol.* 113 (3), 967–973. doi: 10.1104/pp.113.3.967

Wang, L., Czedik-Eysenberg, A., Mertz, R. A., Si, Y., Tohge, T., Nunes-Nesi, A., et al. (2014a). Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. *Nat. Biotechnol.* 32, 1158. doi: 10.1038/nbt.3019

Wang, P., Guo, L., Jaini, R., Klempien, A., McCoy, R. M., Morgan, J. A., et al. (2018). A 13C isotope labeling method for the measurement of lignin metabolic flux in arabidopsis stems. *Plant Methods* 14 (1), 51. doi: 10.1186/s13007-018-0318-3

Wang, T., Salazar, A., Zabotina, O. A., and Hong, M. (2014b). Structure and dynamics of brachypodium primary cell wall polysaccharides from two-dimensional (13)C solid-state nuclear magnetic resonance spectroscopy. *Biochemistry* 53 (17), 2840–2854. doi: 10.1021/bi500231b

Wende, G., and Fry, S. C. (1997). 2-O-β-d-xylopyranosyl-(5-O-feruloyl)-l-arabinose, a widespread component of grass cell walls. *Phytochemistry* 44 (6), 1019–1030. doi: 10.1016/S0031-9422(96)00649-8

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

Winkel, B. S. (2004). Metabolic channeling in plants. *Annu. Rev. Plant Biol.* 55, 85–107. doi: 10.1146/annurev.arplant.55.031903.141714

Withers, S., Lu, F., Kim, H., Zhu, Y., Ralph, J., and Wilkerson, C. G. (2012). Identification of grass-specific enzyme that acylates monolignols with *p*-coumarate. *J. Biol. Chem.* 287 (11), 8347–8355. doi: 10.1074/jbc.M111.284497

Yoshida-Shimokawa, T., Yoshida, S., Kakegawa, K., and Ishii, T. (2001). Enzymic feruloylation of arabinoxylan-trisaccharide by feruloyl-CoA:arabinoxylan-trisaccharide O-hydroxycinnamoyl transferase from oryza sativa. *Planta* 212 (3), 470–474. doi: 10.1007/s004250000490

Zeng, W., Lampugnani, E. R., Picard, K. L., Song, L., Wu, A.-M., Farion, I. M., et al. (2016). Asparagus IRX9, IRX10, and IRX14A are components of an active xylan backbone synthase complex that forms in the golgi apparatus. *Plant Physiol.* 171 (1), 93–109. doi: 10.1104/pp.15.01919

Zhang, B., Munske, G. R., Timokhin, V. I., Ralph, J., Davydov, D. R., Vermerris, W., et al. (2022). Functional and structural insight into the flexibility of cytochrome P450 reductases from sorghum bicolor and its implications for lignin composition. *J. Biol. Chem.* 298 (4), 101761. doi: 10.1016/j.jbc.2022.101761

Zhao, X., Cheng, K., and Liu, D. (2009). Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. *Appl. Microbiol. Biotechnol.* 82 (5), 815–827. doi: 10.1007/s00253-009-1883-1

Zhao, Y., Yu, X., Lam, P.-Y., Zhang, K., Tobimatsu, Y., and Liu, C.-J. (2021). Monolignol acyltransferase for lignin p-hydroxybenzoylation in populus. *Nat. Plants* 7 (9), 1288–1300. doi: 10.1038/s41477-021-00975-1

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 10 March 2023
DOI 10.3389/fpls.2023.1118313

OPEN ACCESS

EDITED BY
Wagner Rodrigo De Souza,
Federal University of ABC, Brazil

REVIEWED BY
Jozef Mravec,
University of Copenhagen, Denmark
Leia Colin,
University of Copenhagen, Denmark

*CORRESPONDENCE
Stanislav V. Isayenkov
stan.isayenkov@gmail.com

SPECIALTY SECTION
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

RECEIVED 07 December 2022
ACCEPTED 24 February 2023
PUBLISHED 10 March 2023

CITATION
Dabravolski SA and Isayenkov SV (2023) The regulation of plant cell wall organisation under salt stress. *Front. Plant Sci.* 14:1118313. doi: 10.3389/fpls.2023.1118313

The regulation of plant cell wall organisation under salt stress

Siarhei A. Dabravolski[1] and Stanislav V. Isayenkov[2]*

[1]Department of Biotechnology Engineering, Braude Academic College of Engineering, Karmiel, Israel, [2]Department of Plant Food Products and Biofortification, Institute of Food Biotechnology and Genomics, National Academy of Science (NAS) of Ukraine, Kyiv, Ukraine

Plant cell wall biosynthesis is a complex and tightly regulated process. The composition and the structure of the cell wall should have a certain level of plasticity to ensure dynamic changes upon encountering environmental stresses or to fulfil the demand of the rapidly growing cells. The status of the cell wall is constantly monitored to facilitate optimal growth through the activation of appropriate stress response mechanisms. Salt stress can severely damage plant cell walls and disrupt the normal growth and development of plants, greatly reducing productivity and yield. Plants respond to salt stress and cope with the resulting damage by altering the synthesis and deposition of the main cell wall components to prevent water loss and decrease the transport of surplus ions into the plant. Such cell wall modifications affect biosynthesis and deposition of the main cell wall components: cellulose, pectins, hemicelluloses, lignin, and suberin. In this review, we highlight the roles of cell wall components in salt stress tolerance and the regulatory mechanisms underlying their maintenance under salt stress conditions.

KEYWORDS

salinity stress, stress tolerance, regulatory mechanisms, cell wall composition, transcriptional regulation, root barriers of salt transport

1 Introduction

High salinity threatens more than 20% of irrigated lands worldwide, and this area is dramatically increasing every year, greatly affecting plant growth and yield through sodium accumulation-mediated osmotic and toxicity stresses (Singh, 2021). During their evolution, plants have developed various strategies to cope with the high soil salinity problem, such as osmotic and metabolic adjustment, normalisation of ion levels and Reactive oxygen species (ROS) balance, salts extrusion or safe accumulation and storage, hormonal and epigenetic re-arrangements, life cycle shortening, or re-schedule salt-sensitive developmental stages (Zhao et al., 2020; Rahman et al., 2022).

Despite intensive research, the exact mechanism by which Na^+ and Cl^- enter the roots is still unknown. As was shown in several species, up to 50% of Na^+ and Cl^- of total uptake translocate *via* apoplast, suggesting the importance of salt transport route *via* cell walls during high salinity. The symplastic uptake is mediated *via* various channels and transporters, such as NSCCs (nonselective cation channels) and GLRs (glutamate

receptor–like channels), HKTs (high-affinity K^+ transporters), PIPs (plasma membrane intrinsic proteins), LCTs (low-affinity cation transporters), AKT (Arabidopsis K^+ Transporter) and KAT (K+ channel in Arabidopsis Thaliana), CCCs (cation-coupled chloride cotransporters) and others [reviewed in (Isayenkov and Maathuis, 2019)]. Similarly, several sensors and receptors could recognise osmotic and ionic stresses and initiate signal transduction and adaptation responses in plants. For example, high salt concentration alters the balance of cell wall ions, which is sensed by sensors or receptors [such as RLKs (receptor-like kinases), GIPC (glycosyl inositol phosphorylceramide), and FER (Feronia)] and activates specialised signaling pathways to normalise the ions balance (such as the SOS pathway). The disbalance of outside/inside ions is sensed by osmosensors [such as HPKs (histidine protein kinases), NSCCs, BON (BONZAI1)], which initiate ions uptake and synthesis of osmolytes (such as proline) to normalise the osmotic homeostasis. Further details of the current knowledge of the plant's osmotic and Na^+ sensors could be found in the recent review (Wang et al., 2022). Additionally, many plant hormones [not only ABA (abscisic acid) and cytokinin but also auxin, salicylic acid, gibberellin, ethylene, and others] are involved in the regulation of the defence system and growth adaptation under salt stress (Yu et al., 2020).

Salt stress directly and accompanying ionic, osmotic, and oxidative stresses impair ion homeostasis and damage biomolecules, thus interrupting normal physiological processes and preventing plant normal growth, development, and sometimes even survival (Figure 1A) (van Zelm et al., 2020). The plant cell wall, however, is the first barrier between cell content and external salt. Plant cell walls consist of polysaccharides, various structural proteins, and fatty acid–derived compounds (such as cellulose, hemicelluloses, pectins, lignin, and suberin), which are crucial for plant growth, development and protection from adverse environmental influences (both biotic and abiotic) (Lampugnani et al., 2018). The primary cell wall is a thin layer composed of the polysaccharides cellulose, hemicelluloses (mostly xyloglucan), and pectin; it is permeable to small molecules, flexible, and extensible, thus facilitating the cell's growth. The secondary cell wall is a thick layer formed inside the primary cell wall in some cell types and made of cellulose, hemicellulose (mostly xylan) and lignin, which make it stiff and waterproof (Zhang et al., 2021). The negative charge of the cell wall is physiologically important, because it facilitates the reversible binding of Ca^{2+}, which is used to strengthen the cell wall *via* pectins cross-linking. However, under high salt concentrations, the surplus of Na^+ could replace Ca^{2+}, thus interrupting normal pectins cross-linking and cell elongation (Figure 1B) (Hocq et al., 2017). At the same time, the amount of different negatively charged polymers able to bind Ca^{2+} in the cell wall stronger could be up-regulated, thus preventing various high Na^+-associated toxic effects on the cell wall (Edmond Ghanem et al., 2010).

The transport of water and solutes across plant roots can occur *via* apoplast, symplast, transcellular, or a combination of these three (Figure 2). The endodermis is the first diffusion barrier, preventing radial apoplastic flow. Exodermis is the second barrier developed by some plant species located beneath the epidermis and can consist of one or several layers with increased amounts of suberin, lignin, proteins, and carbohydrates, which would depend on the species, variety, developmental stage, or environmental conditions (Kim et al., 2018). Suberin is the major component, making suberised

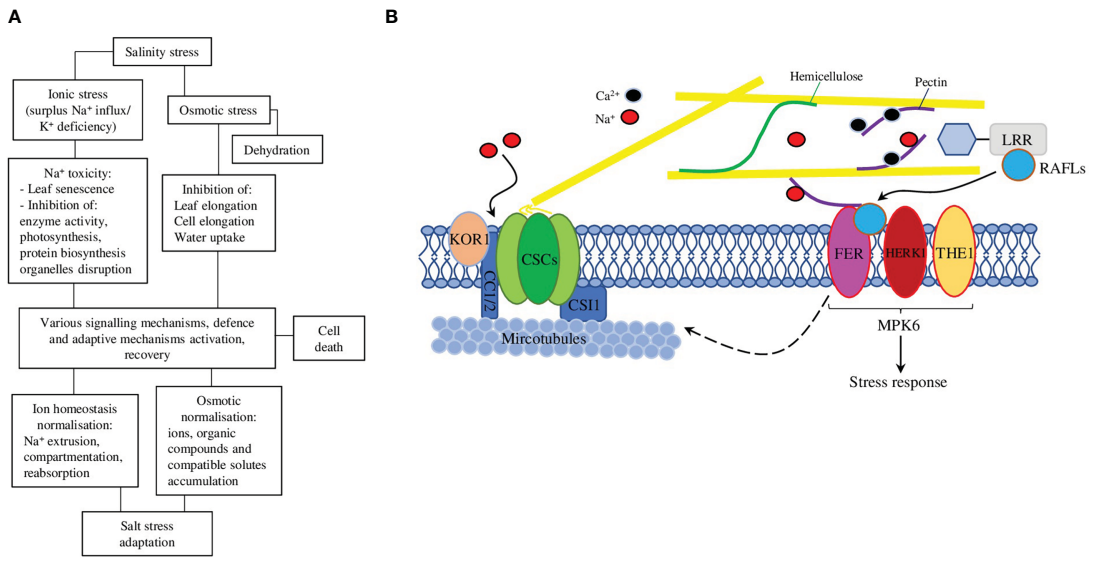


FIGURE 1
(A) A schematic representation of the effects of salt stress on plants and the corresponding responses that plants use to adapt and survive these detrimental effects. **(B)** Salt stress exposure affects cortical microtubule polymerisation and CSC (cellulose synthase complex) assembly. At the recovery stage, FER (FERONIA) regulates cortical microtubule reassembly and the relocation of CSCs to the plasma membrane to synthesise cellulose, thus enhancing plant adaptation to salt stress. The cell wall sensor FER-LRRs-RALF module works in association with HERK1 (HERKULES1) and THE1 (THESEUS1) to perceive salinity through the perturbation in pectins and acts *via* MPK6 (mitogen-activated protein kinase 6) to initiate a salt stress response. The dashed arrow represents in-direct regulation, solid arrows—direct regulations/interactions.

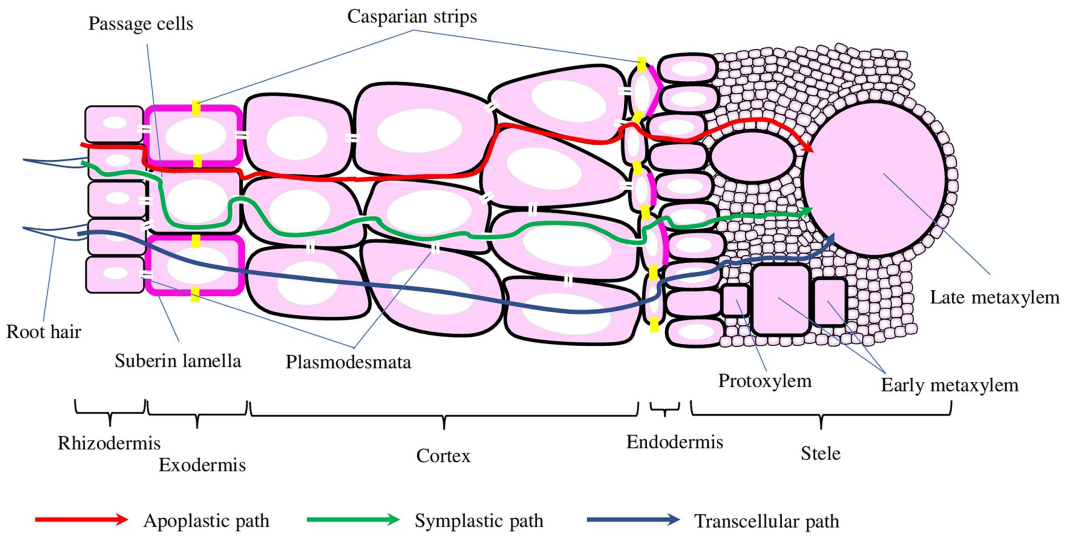


FIGURE 2
Schematic representation of the cross-root transport mechanisms and root cross-section. Water and solutes can be transported into the stele *via* apoplastic, symplastic, and transcellular mechanisms, depicted with red, green and blue arrows, respectively. Casparian strips and suberin lamellae (depicted with yellow and magenta colours, respectively) in the endo- and exodermis interrupt water and solute transport into the stele.

cells almost impermeable to water and providing a critical protective layer preventing water and solutes loss, and the entrance of pathogens. Suberin biosynthesis, assembly, and deposition processes are tightly orchestrated by environmental and developmental stimuli, which involve a wide range of TFs, hormones, and enzymes. The enzymes that are responsible for the synthesis of phenolic and aliphatic monomers and their subsequent esterification, deposition, and assembly (Woolfson et al., 2022) are regulated by TFs (not only of MYB family but also NAC and WRKY) and hormones (mostly by abscisic acid, but jasmonic acid, salicylic acid, and ethylene are also involved) (Nomberg et al., 2022).

The development of endodermis is generally divided into three stages: (1) the deposition of a ring of lignin around the endodermal cells, which resulted in the formation of the Casparian strips (or bands) (CS). Despite CS preventing water and solutes diffusion, endodermal cells on stage 1 still could contribute to both symplastic and trans-cellular flows (Naseer et al., 2012). (2) The deposition of suberin throughout the cell wall greatly slows down nutrient uptake, while preserving symplastic connection with neighbour cells through plasmodesmata (Sager and Lee, 2014). (3) On the final stage, the cell wall is extensively thickened with the deposition of cellulose in a stele-ward direction (Kim et al., 2018). Even though the development of CS and suberin deposition happens at the same time point, the general features of exodermis are similar to the endodermis (Serra and Geldner, 2022).

The exact composition and structure of plant cell wall components are dynamically modulated, thus allowing cells to execute the optimal growth programme and timely respond to different stresses (Rui and Dinneny, 2020). The common salt stress response includes lignin accumulation, disruption of pectins cross-linking, reduction of cellulose content, and enhancement of antioxidant defence. Furthermore, plants defective in cell wall biosynthesis have demonstrated increased sensitivity to salt stress, thus confirming the importance of cell wall integrity maintenance for salt stress adaptation Liu et al., 2021). Further in this review, we focus on the recent discoveries of the molecular mechanisms connecting cell wall integrity and salt tolerance in plants.

2 Role of cell wall component in plant salt tolerance

2.1 Cellulose

The process of cellulose biosynthesis is complex and tightly regulated, and it is interrupted by salt stress on multiple levels and resulted in growth defects/inhibition, anatomical abnormalities, and hypersensitivity to salt stress. Several transcription factors, epigenetic regulatory mechanisms, and hormones, have been shown to affect the cellulose biosynthesis process upon salt stress perception (Table 1). However, salt-mediated modulation of the cellulose biosynthesis process should be always considered in close association with other metabolic and physiological responses caused by salt stress. Further investigation and understanding of the exact molecular mechanisms of every cellulose biosynthesis-associated genes/proteins involved in salt stress tolerance development would promote its practical application in creating salt-tolerant varieties of different plant species.

Cellulose micro-fibrils are β-1,4 linked glucose polymer that is synthesised at the cell surface by cellulose synthase (CesA) complexes (CSCs) not only by higher plants but also algae, some bacteria and even animals, thus, making it the most abundant polymer on the planet (Allen et al., 2021). CesAs are the key part of CSCs, which are active only in the plasma membrane, while actively communicating with the Golgi apparatus and other intracellular compartments *via* exocytosis (trafficking to the plasma membrane)

TABLE 1 List of genes associated with cell wall biosynthesis and salt tolerance.

Species	Gene	Phenotype under salt stress	gene regulation	References
Arabidopsis thaliana	*CESA1* and *CESA6*	*cesA1* and *cesA6* plants are salt sensitive, with severe root tip swelling, reduced cellulose contents and root elongation	–	(Zhang et al., 2016)
	CSI1	*csi1* plants are salt stress hyper-sensitive	–	
	KOR1	*kor1* plants are salt stress-sensitive, with marked cellulose deficiency and retarded growth	–	(Nagashima et al., 2020)
	HSFA7b	*HSFA7b* OE plants had increased proline and soluble sugar levels, and salt-stress tolerance.	Ex: *AtP5CS1* and *AtP5CS2* up-regulated LF: *AtP5CS1* and *AtP5CS2* down-regulated. AtHSFA7b down-regulates 31 cellulose biosynthesis-related genes (including *CESA8*, *CSLG1*, *CSLG2*, and *CSLA9)* and induces other TFs (bHLH, NAC061, NAC036, NAC090, WRKY38, and ZFP2)	(Zang et al., 2019)
	GCN5	*gcn5* plants have reduced cellulose content, and exhibit severe inhibition of growth and cell wall anomaly	GCN5 regulates *CTL1*, *PGX3*, *MYB54*	(Zheng et al., 2019)
	XTH30	*XTH30* Ex plants hypersensitive to salt stress; *xth30* plants are salt tolerant, have lower Na^+ accumulation in shoot and H_2O_2 content		(Yan et al., 2019).
	XTH19 and *XTH23*	*xth23* and *xth19/xth23* plants are additively sensitive to salt stress; *BES1* Ex increased salt tolerance in WT and improved salt tolerance of *xth19/xth23*.	BES1 regulates *XTH19* and *XTH23*	(Xu et al., 2020)
	PME31	*PME31* LF plants are salt-stress hypersensitive to salt stress;	LF: *DREB2A, RD29A* and *RD29B* down-regulated	(Yan et al., 2018)
	MYB3	*myb3* plants have higher salt-stress tolerance	LF: *PAL1*, *C4H*, *COMT*, *4CL3*, *DFR*, *LDOX, TT8* and EGL3 were up-regulated	(Kim et al., 2022)
	SND1	*snd1* plants have a low tolerance to salt stress	SND1 regulates *Myb46* and *ABI4.*	(Jeong et al., 2018)
Brassica oleracea	*CESA1* and *CESA6*	*cesA1* and *cesA6* plants have higher salt-stress tolerance, anatomical and ultrastructural changes in leaves and chloroplasts, dwarf phenotype, reduced content of the cellulose (40%) and pectin (19%), higher soluble sugar content and 3 times higher proline content	LF: *PIP2;2* and *PIP2;3* were 6-7 times up-regulated	(Li et al., 2017)
Nicotiana tabacum	*XTH*	Ex plants have a higher rate of root growth under salt-stress conditions and improved tolerance to frost and heat stresses	–	(Kuluev et al., 2017)
Oryza sativa	*UGE3*	*uge3* plants have diverse growth defects, hypersensitive to osmotic and salt stresses. *UGE3* Ex plants have increased biosynthesis of cellulose and hemicelluloses, increased tolerance to osmotic and salt stresses, improved Na^+ and K^+ homeostasis and higher accumulation of soluble sugars	–	(Tang et al., 2022)
	CSLD4	*nd1* plants are salt-stress sensitive but insensitive to ABA treatments. ABA treatment complemented *nd1* salt-stress sensitive phenotype. *CSLD4* OE plants have increased ABA content and enhanced rice salt-stress tolerance	LF: repressed ABA biosynthesis genes OE: up-regulated ABA biosynthesis genes	(Zhao et al., 2022).
	EIL2	*EIL2* Ex plants have a low tolerance to salt and drought stresses, delayed leaf development and decreased pectin content. *EIL2* LF plants have enhanced tolerance to salt and drought stresses and delayed leaf senescence	*EIL2* binds and regulates BURP14 and BURP16.	(Jin et al., 2020)
	TSD2	*tsd2* plants have increased accumulation of Na^+ and a lower level of K^+ in shoot	LF: KAT1, SOS1 and HKT1 were down-regulated	(Fang et al., 2019)
Panicum virgatum	*NAC1*	*NAC1* Ex plants have enhanced tolerance to salt stress and higher cellulose content, reduced Na^+ and increased K^+ accumulation in roots and shoots. *NAC1* LF plants are salt stress hypersensitive	Ex: antioxidant defence genes and ion-homeostasis-related genes were up-regulated	(Wang et al., 2019b)
Capsicum annuum	*XTH3*	*XTH3 Ex* in tomato plants have increased tolerance to drought and salt stresses		(Choi et al., 2011)

(Continued)

TABLE 1 Continued

Species	Gene	Phenotype under salt stress	gene regulation	References
Populus euphratica	*XTH*	*PeXTH* Ex in tobacco plants have improved water-retaining capacity, and photosynthesis efficiency and increased the number of mesophyll cells		(Han et al., 2013)
Diospyros kaki	*XTH*	*XTH* Ex in Arabidopsis and tomato lead to enhanced tolerance to salt and drought stresses, larger and more irregular cells with a higher density of cell walls and intercellular spaces		(Han et al., 2017)
Nicotiana tabacum	*XTH*	*XTH* Ex plants have a higher rate of root growth under salt-stress conditions and improved tolerance to frost and heat stresses		(Kuluev et al., 2017)
Chorispora bungeana	*PMEI1*	*PMEI1* and *PMEI13* Ex plants have reduced PMEs tissue activity and increased salt tolerance	*CbPMEI1* and *PMEI13* were repressed by ABA and salt stress	(Chen et al., 2018)
Amaranthus hypochondriacus	*DGR2*	*AhDGR2* Ex in Arabidopsis causes general abiotic stress intolerance, hypersensitivity to salt stress and ABA treatment		(Palmeros-Suárez et al., 2017)
Camellia sinensis	*F3H*	*F3H* Ex in tobacco increased tolerance to salt stress, reduced electrolyte leakage and MDA levels, increased the content and activity of antioxidant enzymes		(Mahajan and Yadav, 2014)
Betula platyphylla	*MYB46*	*MYB46 Ex* plants have improved tolerance to salt and osmotic stresses, increased lignin deposition and secondary cell wall thickness; *myb46* plants were hypersensitive to salt and osmotic stresses	Ex: *SOD*, *POD*, *P5CSs* and genes related to the secondary cell wall formation were up-regulated, *P5CDH* and *ProDH* were down-regulated.	(Guo et al., 2017)
	NAC012	*BpNAC012* Ex plants have enhanced lignin accumulation; *nac012* plants have reduced secondary cell wall thickening	Ex: *P5CSs*, *SOD POD, and* lignin biosynthesis genes were up-regulated	(Hu et al., 2019)
Apium graveolens	*NAC1*	*NAC1* Ex in Arabidopsis plants leads to increased salt and drought stresses tolerance, reduced MDA level	OE: *LAC*, *F5H*, *CCoAOMT*, *C3'H*, *COMT*, *CCR*, SOD and POD were up-regulated	(Duan et al., 2020)
Fagopyrum tataricum	NAC16	*NAC16* Ex in Arabidopsis plants leads to salt stress hypersensitivity	Ex: many lignin biosynthesis-related genes were down-regulated; *4CL2* and *CCR2* were up-regulated	(Wang et al., 2021)
Bryum argenteum	*DBL1*	*BaDBL1* Ex in Arabidopsis plants leads to increased salt and osmotic stresses tolerance, and the activity of POD, SOD and CAT was increased	Ex: lignin-biosynthesis-related genes and abiotic stress-related TFs and genes (*LEA*, *AtCOR15A* and *AtRD29A*) were up-regulated	(Liang et al., 2021)
Malus domestica	*SND1*	*SND1* Ex plants have enhanced resistance to salt and osmotic stresses, and higher lignin and antioxidant content; *snd1* plants were sensitive to salt and osmotic stresses	*SND1* binds *MdMYB46/83, MdRD22*, *MdRD29A*, *MdDREB2A*, *MdAREB1B* and *MdAREB1A*	(Chen et al., 2020)
Potentilla atrosanguinea	*SOD*	*SOD* and *APX* Ex in Arabidopsis plants enhanced salt-stress tolerance and increased lignin accumulation, levels of compatible solutes, higher biomass production, better growth rate and increased yield	Double Ex: many lignin biosynthesis-related genes were up-regulated, while *CcAOMT1*, *4CL8*, *CAD1*, *4CL3* and *LAC12* were down-regulated; cell wall-related TFs were up-regulated (VND1, VND2, VND4, VND6, SND1, SND2 and NST1)	(Shafi et al., 2015).

Ex, expression; LF, loss of function.

and endocytosis (removal from the plasma membrane). Microtubules and actin microfilaments serve as tracks along which motor proteins transport cell wall components in vesicles. The direct interactions between microtubules, microtubule-linking proteins, and cellulose synthases made the proper microtubules and microfilaments organisation critical for the positioning and construction of cell wall components (Colin et al., 2023). At the same time, the genes in CesA family are specialised, with CesA1, CesA3, and CesA6 participating in primary cell wall synthesis, while CesA4, CesA7, and CesA8 are involved in secondary cell wall synthesis. Also, some other proteins, such as CSI1 (cellulose synthase–interacting protein 1) and CC1-2 (companion of cellulose synthase 1 and 2), are also required for cellulose biosynthesis (Zhu and McFarlane, 2022). Another recently identified members of the CSC are TTL (tetratricopeptide thioredoxin-like) proteins, which interact with CesA1 and cortical microtubules to promote their polymerisation, thus maintaining cellulose synthesis under salt stress conditions (Kesten et al., 2022).

Experimental data from the model plant *Arabidopsis thaliana* [(L.) Heynh.] demonstrated that *cesA1* and *cesA6* single mutants are salt sensitive, with severe root tip swelling, reduced cellulose contents, and root elongation. Mutant lines of *csi1*, which are known to directly interact with CesA6 and cortical microtubules, are also hyper-sensitive to salt stress (Zhang et al., 2016). Similarly, to Arabidopsis, RNAi knockdowns of *CesA1* and *CesA6* in broccoli (*Brassica oleracea* L.) resulted in anatomical and ultrastructural changes in leaves and chloroplasts, dwarf phenotype, and reduced content of the cellulose and pectins (40% and 19%, respectively). However, mutant plants had higher soluble sugar content and three times higher proline content compared with control plants.

Furthermore, mutants have higher salt-stress resistance, which was associated with six to seven times up-regulated expression of salt-tolerance–related aquaporins PIP2;2 and PIP2;3 (Li et al., 2017). Thus, these genes might play the role of the negative regulator in plant osmoprotection.

CC1 and 2 proteins could interact with CESAs and microtubules and promote microtubule formation and dynamics. Mutations of the CC1 and 2 resulted in salt-sensitive phenotypes with altered microtubule and CSC behaviour *in vivo*. Under salt, stress CSCs are quickly dissociated from the plasma membrane, and functional CC proteins are necessary to maintain CesA migration and microtubule stability to reassemble CSCs during the growth recovery phase after salt treatment (Figure 1B) (Endler et al., 2015). Interestingly, plant CC1 implements an evolutionally conserved microtubule-binding mechanism, similar to that of Tau protein, which is associated with Alzheimer's disease and known to self-aggregate and trigger neurodegeneration. In particular, two tyrosine residues in the N-terminal region of CC1 are responsible for the microtubule binding, both *in vitro* and *in vivo*. Point mutations in these residues interrupted normal microtubule-guided CSC movement and resulted in the generation of a salt stress-hypersensitive microtubule array (Kesten et al., 2019).

Recently, an ubiquitously expressed and salt-stress induced gene *UGE3* (UDP-galactose/glucose epimerase 3) from rice, which provides substrates for polysaccharides polymerization, was demonstrated to improve biomass production, mechanical properties of the cell wall, and increased tolerance to salt and osmotic stresses. *uge3* mutants displayed diverse growth defects and were hypersensitive to osmotic and salt stresses. On the other side, *OsUGE3* overexpressing plants had increased biosynthesis of cellulose and hemicelluloses, which resulted in improved mechanical strength of the cell wall. Furthermore, *OsUGE3* overexpressors showed increased tolerance to osmotic and salt stresses, which was associated with improved Na^+ and K^+ homeostasis and higher accumulation of soluble sugars (Tang et al., 2022).

Additionally, some other cellulose biosynthesis-related proteins have also been reported involved in salt tolerance. For example, KORRIGAN1 (KOR1), a membrane-anchored endo-β-1,4-glucanase, is an integral part of the primary cell wall, physically interacting with CSC and linking cell wall biosynthesis and abiotic stress tolerance. KOR1 was shown to cycle between the trans-Golgi network and the plasma membrane under normal conditions. While under salt stress, this cycling is interrupted, and KOR1 is retained in the plasma membrane or transported to the tonoplast for degradation, thus unable to support cellulose biosynthesis. Not surprisingly, *kor1* mutants have the salt-stress sensitive phenotype, cellulose deficiency, and retarded growth (Nagashima et al., 2020).

Recently, several mechanisms regulating cellulose biosynthesis in relation to salt-stress tolerance have been identified (Figure 3). For example, overexpression of the NAC1 TF (transcription factor)

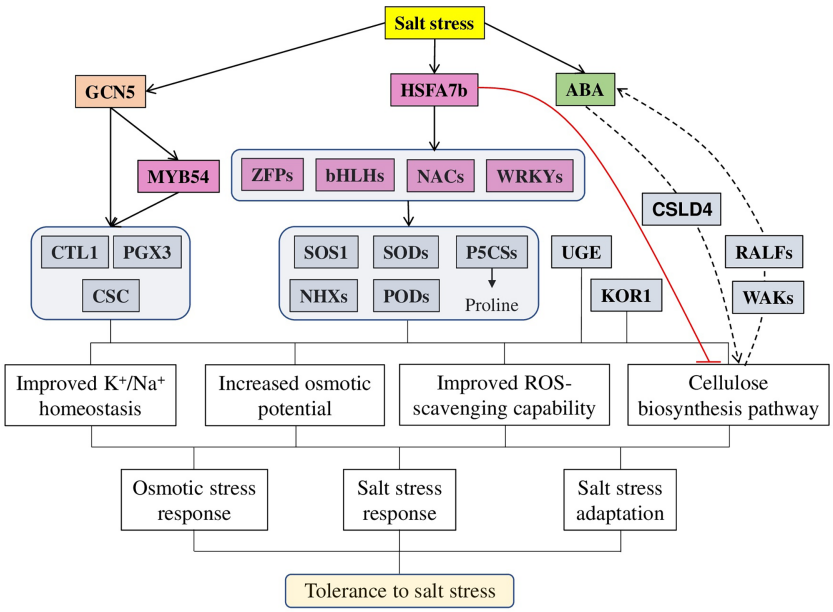


FIGURE 3
A model of different factors regulating cellulose-associated salt stress tolerance in plants. Salt stress induces the expression of a master regulator HSFA7b (heat shock transcription factors), which activates a wide range of genes, such as other TF (from WRKY, NAC, HLH, and ZFP families), ion transporters (SOSs and NHXs), antioxidant defence (SODs, GSTs, and PODs) and osmoprotectors (P5CSs) while inhibiting cellulose synthesis related genes (CesAs, CSLAs, and CSLGs). Similarly, histone acetyltransferase GCN5 (general control non-repressed protein 5) regulates adaptation to salt stress *via* directly targeting *MYB54* and cellulose biosynthesis–related genes [*CTL1* (chitinase-like gene) and *PGX3* (polygalacturonase involved in expansion 3)]. Additionally to other genes, directly affecting the cellulose biosynthesis process [such as *KOR1* (*KORRIGAN1*) and *UGE3* (UDP-galactose/glucose epimerase 3], genes like *CSLD4* modulate the activity and location of cell wall-localised proteins (WAKs and RALFs) to sustain high ABA content to enhance salt and osmotic stress tolerance. In total, all these target genes provide various physiological changes, leading to improved salt stress tolerance: improving ions homeostasis and ROS scavenging capability, increasing osmotic potential and modulating cellulose biosynthesis. The red line represents negative regulation; the dashed line represents other involved components.

in Switchgrass (*Panicum virgatum* L.) leads to higher cellulose content, enhanced tolerance to salt stress, and up-regulated expression of antioxidant defence and ion-homeostasis–related genes, reduced accumulation of Na^+ and increased of K^+ in roots and shoots. At the same time, *NAC1* RNAi plants were hypersensitive to salt stress (Wang et al., 2019b). Arabidopsis AtHSFA7b [heat shock factor (class A)–type transcription factor], known to participate in abiotic stress responses, is also involved in salt tolerance. Under salt conditions, AtHSFA7b improved salt tolerance by regulating genes, such as *SOS1*, *NHXs*, *PODs* (peroxidase), and *SODs* (superoxide dismutase). These gene interactions lead to the reduction of water loss rate, adjustment of osmotic potential, and a decrease in the accumulation of reactive oxygen species. In particular, the content of proline and soluble sugars was increased after NaCl treatment in an AtHSFA7b-dependent way. Under salt-stress conditions, plant lines overexpressing *AtHSFA7b* had increased levels of proline biosynthesis genes (*AtP5CS1* and *AtP5CS2*, encoding Δ1-pyrroline-5-carboxylate synthases), while *hstfa7b* plants had lower *AtP5CS1* and *AtP5CS2* expression. Several other TFs, positive regulators of salt stress, were regulated by AtHSFA7b: bHLH, NAC061, NAC036, NAC090, WRKY38, and ZFP2 (Zinc finger protein). Finally, 31 cellulose biosynthesis-related genes [including *CESA8*, *CSLG1* (cellulose synthase-like glycosyltransferase), *CSLG2*, and *CSLA9*] were down-regulated by AtHSFA7b, suggesting these genes as negative modulators of salt tolerance (Zang et al., 2019).

Epigenetic mechanisms also play an important role in salt tolerance. As it was recently demonstrated, the mRNA levels of histone acetyltransferase GCN5 (general control non-repressed protein 5) were increased after salt treatment, thus suggesting its importance in maintaining the cell wall integrity under salt stress. Several genes, such as *CTL1* (chitinase-like gene), *PGX3* (polygalacturonase involved in expansion 3), and *MYB54* (MYB domain protein 54), involved in cellulose biosynthesis, were identified as direct GCN5 targets. Confirming these results, the *gcn5* mutants exhibited reduced cellulose content, severe growth inhibition, and cell wall anomaly under salt stress. Furthermore, the salt tolerance and cell wall integrity phenotypes of the Arabidopsis *gcn5* mutant were complemented by the expression of the wheat *TaGCN5* gene, suggesting a conserved role of the GCN5-mediated salt tolerance between these two species (Zheng et al., 2019).

Additionally, the plant hormone ABA, known to regulate various aspects of plant growth, development, and stress responses, is also linked to cellulose synthesis-regulated salt stress tolerance. The rice OsCSLD4 (cellulose synthase-like D4 protein) mutant (*nd1*) was sensitive to salt stress but insensitive to ABA treatments. Interestingly, the expression of some ABA synthesis and signaling genes was repressed in *nd1* mutant under both normal and salt stress conditions. Exogenous ABA treatment effectively complemented *nd1* salt-stress sensitive phenotype. Furthermore, *OsCSLD4* overexpression leads to the up-regulated expression of ABA biosynthesis genes, increased ABA content and enhanced rice salt-stress tolerance, suggesting that the role of OsCSLD4 in salt stress tolerance is ABA-mediated (Zhao et al., 2022).

2.2 Hemicelluloses

Hemicelluloses are branched heteropolymer interacting and cross-linking cellulose microfibrils, thus strengthening plant cell walls (Khodayari et al., 2021). A recent study on vascular plants suggested that hemicelluloses could also interact with lignin in some cell walls (Gallina et al., 2018). The abundance and structure of hemicelluloses vary greatly in different plant species, with main examples grouped into xylan, xyloglucan, arabinoxylan, glucuronoxylan, and glucomannan (Scheller and Ulvskov, 2010). Xyloglucan is involved in cell wall strengthening *via* cellulose microfibrils binding during cell elongation, while xylan participates in polysaccharide cross-linking during cell wall architecture establishment. Modification of hemicelluloses by the cell wall remodelling enzymes is an important way to control cell wall extensibility, and XTHs (xyloglucan endotransglucosylase/hydrolases) is the best-studied family facilitating cell wall plasticity (Ishida and Yokoyama, 2022). XTHs can cleave and/or reconnect xyloglucans, which leads to the attachment of the reducing end of the xyloglucans to the non-reducing end of another xyloglucans, thus producing chimeric oligomer or polymer xyloglucan molecules (Park and Cosgrove, 2015). XTHs homologues react to salt stress differently, with some homologues up-regulated, while other – down-regulated and some did not change their level of expression. Therefore, further identification and detailed investigation of every XTHs' (both positive and negative) regulator would help to create new plants with improved tolerance to salt stress.

Early research demonstrated that XTHs are active players in various abiotic stresses, including high-salinity stress. Experiments on *Medicago truncatula* demonstrated that XTHs responded differently to salt stress, with eight up-regulated and 11 down-regulated genes (Xuan et al., 2016). Similar results were reported also for poplar (*Populus* sp.), where 11 differently expressed genes were identified in the roots (five up- and six down-regulated), nine in the stems (four up- and five down-regulated), and seven in the leaves (four up- and three down-regulated) after salt treatment (Cheng et al., 2021). These data were confirmed on other species such as typical halophyte, *Salicornia europaea*, where salt stress up-regulated expression of 27 and 15 SeXTHs (in shoots and roots, respectively) (Tiika et al., 2021), and *Vitis vinifera* L., four VvXTHs were observed (Qiao et al., 2022). Therefore, XTHs' expression is tissue specific, and some members of the family are positive regulators of salinity tolerance, while others are negative.

Thereby, transgenic tomato plants expressing pepper (*Capsicum annuum* L.) *CaXTH3* demonstrated an increased tolerance for drought and salt stresses (Table 1). Transgenic plants had normal phenotype under standard growth conditions and also no significant difference in chlorophyll content and root elongation under salt stress (Choi et al., 2011). Similarly, the expression of poplar (*Populus euphratica* Oliv.) *PeXTH* in tobacco improved water-retaining capacity and photosynthesis efficiency and increased the number of mesophyll cells under salt-stress conditions (Han et al., 2013). Transgenic Arabidopsis and tomato plants expressing persimmon (*Diospyros kaki* Thunb.) *XTH* had

larger and more irregular cells with a higher density of cell walls and intercellular spaces, which resulted in enhanced tolerance for salt and drought stresses (Han et al., 2017).

A negative role in salt tolerance was reported for Arabidopsis *XTH30*, which was strongly up-regulated under salt stress in the stem, root, flower, and hypocotyl. However, plants overexpressing *XTH30* were hyper-sensitive to salt stress, while loss-of-function *XTH30* mutant demonstrated lower Na^+ accumulation in shoot and H_2O_2 content and increased salt tolerance (Yan et al., 2019).

In addition to diverse abiotic stresses, the expression of *XTH* is regulated by plant hormones. For example, *XTH* from tobacco (*Nicotiana tabacum* L.) was up-regulated by both ABA treatment (10 μM) and different stresses (drought, salt, and cadmium), while down-regulated by higher ABA concentration (100 μM) and 0°C stress, thus suggesting its implication in ABA-dependent stress signaling. At the same time, *NtXTH* overexpressing tobacco plants demonstrated a higher rate of root growth under salt-stress conditions, and also greater frost and heat tolerance when compared with control plants (Kuluev et al., 2017). In Arabidopsis, *XTH19* and *XTH23* were induced by salt stress *via* the brassinosteroid signaling TF BES1 (BRI1-EMS-SUPPRESSOR 1). *xth23* single and *xth19/xth23* double mutant were additively sensitive to salt stress, while *BES1* overexpression increased salt tolerance in wild-type and partially complement salt tolerance and phenotype of *xth19/xth23* mutant. Further experiments showed that *XTH19* and *XTH23* expression is directly regulated by BES1, thus postulating a novel role for brassinosteroids in regulating XTH-mediated salt sensitivity (Xu et al., 2020).

2.3 Pectins

Pectins are acidic polysaccharides enriched with α-(1,4)–linked galacturonic acids and accounted for up to 40% of the higher plant cell wall content. Pectins play a crucial role in plant growth and development, immunity, response to stresses, and senescence [reviewed in (Cruz-Valderrama et al., 2019; De Lorenzo et al., 2019; Dauphin et al., 2022)]. The three major types of pectins are HG (homogalacturonan), RG-I (rhamnogalacturonan-I), and RG-II (rhamnogalacturonan-II), which are selectively modified by PMEs (pectin methyl esterases), PAEs (pectin acetylesterases), PGs (polygalacturonases), or PLLs (pectate lyases–like) during cell growth or in response to stresses. The dynamic regulation and interplay between expression and activity of modified enzymes and their inhibitors [such as PMEIs (PME inhibitors) and EGCG (epigallocatechin gallate)] define the extent of cell walls stiffness [reviewed in (Coculo and Lionetti, 2022)] (Du et al., 2022). Different pectin-modified enzymes have been shown to act as positive or negative regulators of salt tolerance in plants.

Thereby, in tobacco (*Nicotiana tabacum* L.) *NtPME005*, *NtPME039*, *NtPME043*, *NtPME047*, *NtPME082*, *NtPME106*, and *NtPME108* were induced and up-regulated, while *NtPME029*, *NtPME056*, *NtPME058*, and *NtPME062* were down-regulated by salt stress. Interestingly, some *NtPME* genes (*NtPME005*, *NtPME039*, and *NtPME106*) were down-regulated by ABA treatments and up-regulated by NaCl stress, suggesting the involvement of ABA pathway in salt stress tolerance (Sun et al., 2022). Expression *PMEI1* from *Chorispora bungeana* (Fisch.) and its Arabidopsis homologue *PMEI13* in Arabidopsis reduced PMEs tissue activity and increased salt tolerance. Interestingly, both *CbPMEI1* and *PMEI13* genes were repressed by salt stress and ABA treatments, thus further supporting the involvement of ABA in PME-mediated salt tolerance regulation (Chen et al., 2018).

Recently, another ABA-mediated mechanism of salt-stress tolerance was demonstrated on rice (*Oryza sativa* L.). In particular, OsEIL2 (EIN3-Like) (ETHYLENE-INSENSITIVE 3), a nuclear-localised TF induced by salt stress and ABA treatment, directly binds and regulates genes encoding β subunit of polygalacturonase subfamilies (OsBURP14 and OsBURP16) (Table 1). *OsEIL2* overexpression resulted in delayed leaf development, decreased (because of the increased PG activity) pectins content and reduced tolerance to salt and drought stresses. On the contrary, *OsEIL2* RNAi knockdown mutant demonstrated delayed leaf senescence and enhanced tolerance to salt and drought stresses (Jin et al., 2020). Another conserved gene, *DGR2* (DUF642 L-GALACTONO-1,4-LACTONE-RESPONSIVE) from *Amaranthus hypochondriacus* (L.), encoding DUF642 domain–containing protein induced by drought and salt stresses, was also associated with cell wall modifications. *AhDGR2* overexpression in Arabidopsis led to hypersensitivity to NaCl and ABA, and general abiotic stress intolerance. Interestingly, the effect of *AhDGR2* overexpression on the PMEs activity was organ-specific, significantly lower in leaves but higher in roots, thus suggesting further insights into the precise regulation of cell wall modifications (Palmeros-Suárez et al., 2017). Similarly, overexpression of tea [*Camellia sinensis* (L.) Kuntze] *CsF3H* gene, encoding flavanone 3-hydroxylase in tobacco, increased tolerance to salt stress *via* decreased PMEs activity in roots and leaves. Also, transgenic plants demonstrated reduced electrolyte leakage and MDA (malondialdehyde) levels, while increasing the content and activity of antioxidant enzymes (Mahajan and Yadav, 2014).

PMEs are novel and often overlooked players in the development of salt stress tolerance. PMEs are closely associated with ABA signaling pathways, a wide range of cell-surface sensors and other abiotic stresses-related genes. However, the exact regulatory circuit and molecular mechanism of their action are unknown and require further investigation.

Arabidopsis *PME31* was significantly up-regulated after salt treatment. It was shown that PME31 acts as a positive regulator of salt stress tolerance. Accordingly, *PME31* knockdown resulted in hypersensitivity to salt stress and down-regulation of other genes related to drought and temperature stresses [such as *DREB2A* (dehydration-responsive element-binding protein), *RD29A* (low-temperature–induced 78 kDa protein) and *RD29B*] (Yan et al., 2018). Similarly, the mutation in rice *PME* (OsTSD2) reduced the expression of several genes maintaining ion homeostasis (*OsKAT1*, *OsSOS1*, and *OsHKT1*), which resulted in increased accumulation of Na^+ and a lower level of K^+ in shoot under salt stress (Fang et al., 2019).

Additionally, cell walls perceive various stresses with cell-surface localised receptors, such as WAKs (wall-associated kinases), LyKs (LysM receptor-like kinases), LRR-RLKs (leucine-rich repeat

receptor-like kinases), LRXs (Leucine-rich repeat extensins) and CrRLK1L (*Catharanthus roseus* receptor-like kinase 1-like) [reviewed in (Bacete et al., 2018)]. FERONIA (FER) is a plasma membrane–localised receptor kinase perceiving salt-stress–mediated damage to the pectins-associated wall. Therefore, root cells of *fer* mutant were shown to explode during growth recovery after salt treatment. Conducted *in vitro* experiments proposed a mechanism where FER directly interacts with pectins, senses salinity-mediated softening of the cell wall and initiates a Ca^{2+}-mediated signaling cascade preventing root cells from exploding during growth under salt stress. The exposure to the salt stress causes the surplus of Na^{+} to replace Ca^{2+} in pectins binding, thus interrupting normal cross-linking of pectins and reducing cell elongation, Ca^{2+} or borate treatment could rescue *fer*-associated phenotype and facilitating pectins cross-linking (Feng et al., 2018). Later studies provide evidence of the close association between FER and LRX3/4/5 (Dünser et al., 2019), and RALF22/23 (RAPID ALKALINIZATION FACTOR) (Yu and Assmann, 2018), thus proposing their functioning as a LRX-RALF-FER module to transduce cell wall signals plant growth in normal condition and tolerance to salt stress (Figure 1B) (Zhao et al., 2018). Furthermore, recent research has identified that LRX-RALF-FER module regulates the homeostasis of jasmonic acid, salicylic acid and abscisic acid, thus controlling plant salt stress response (Zhao et al., 2021). Later, a similar mechanism was described for other sensors, such as THE1 (THESEUS1) and HERK1 (HERKULES1) (Gigli-Bisceglia et al., 2022), WAK2 [*Brachypodium distachyon* (L.) P.Beauv.], and AtWAK2 (Wu et al., 2020). In particular, while *the1* and *herk1* single mutants have wild-type salt stress phenotype, the double mutants have hypersensitive to salt stress phenotype, similar to *fer* mutants. In the proposed model THE-HERK sensors mitigated salt stress through the negative regulation of MAPK6 (mitogen-activated protein kinase 6) (Gigli-Bisceglia et al., 2022), which is known to regulate many cellular processes, including also salt stress response (Zhang and Zhang, 2022).

2.4 Lignin

Lignin is the main structural component of the plant cell wall, composed of the units derived from the polymerisation of sinapyl, coniferyl, *p*-coumaryl, and hydroxycinnamyl alcohols (Figure 4). Accumulation of lignin and cell wall thickening because of the activation of the lignin biosynthesis pathway is an important

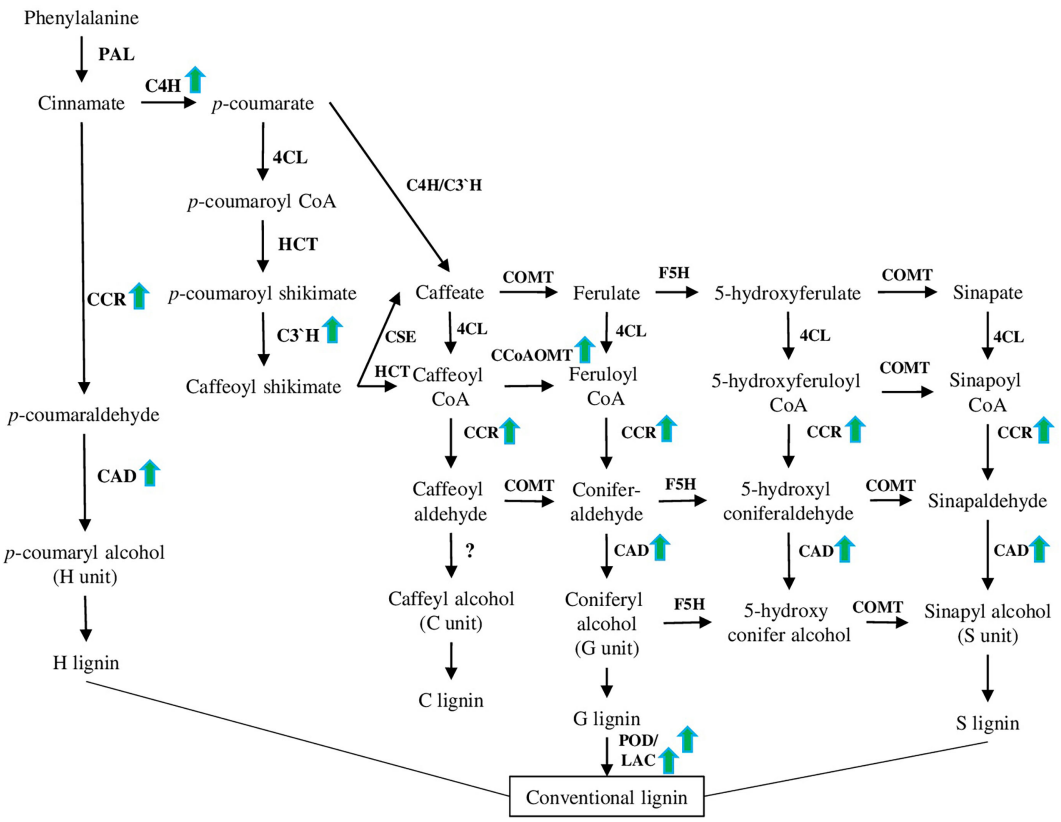


FIGURE 4
Lignin biosynthesis pathway. This pathway has been modified based on recently published papers (please see the main text of section 4.4). PAL (phenylalanine ammonia-lyase), C4H (cinnamate 4-hydroxylase), 4CL (4-coumarate: CoA ligase), CCR (cinnamoyl CoA reductase), CCoAOMT (caffeoyl CoA O-methyltransferase), F5H (ferulate 5-hydroxylase), COMT (caffeate 3-O-methyltransferase), and CAD are involved in the synthesis of monolignols. POD (peroxidase) and LAC (laccase) are involved in the polymerization of monolignols to yield the lignin polymer as a final product. Green arrows represent genes up-regulated in different species under salt stress conditions according to recent papers [(Li et al., 2021), (Isayenkov et al., 2020), (Chun et al., 2019), and (Vasupalli et al., 2021)].

response to various environmental stresses, including high salinity [reviewed in (Vanholme et al., 2019; Coomey et al., 2020)].

2.4.1 Transcriptional regulation of lignin biosynthesis

Salt-mediated activation of lignin biosynthesis and deposition processes are tightly regulated and associated with other mechanisms of abiotic stress tolerance: antioxidant defence, anthocyanin and ABA biosynthesis. At the same time, some TFs play a more important role than others, activating lignin biosynthesis and a wide range of other salt stress-specific and general abiotic stress response mechanisms. Such master regulators TFs (SND1 or NAC83) have great potential as targets for the development of salt-stress tolerant plant species.

A number of transcriptomic studies on different species suggested that up-regulation of the lignin biosynthesis pathway and antioxidant defence are common responses to salt stress (Figure 5). In particular, 28 and 23 differentially expressed genes related to lignin and flavonoid biosynthesis pathways, respectively, have been identified in *Sophora alopecuroides* (L.) after salt treatment (Zhu et al., 2021). Among other cell wall metabolism-related genes, 10-fold up-regulation of laccases was shown on halophytic barley *Hordeum marinum* ssp. *marinum* under salt stress conditions (Isayenkov et al., 2020). Also, many lignin biosynthesis-related genes (such as POD, CAD (cinnamyl alcohol dehydrogenase), C4H (cinnamate 4-hydroxylase), CcoAOMT (caffeoyl CoA O-methyltransferase), CCR (cinnamoyl CoA reductase) and C3′H (*p*-coumarate 3-hydroxylase) (Figure 4), and 36 TFs associated with various families (such as HSF, MYB, NAC, WRKY, bZIP, AP2/ERF-ERF, and C2H2) have been identified in halophytic plant *Eutrema salsugineum* (Pall.) (Li et al., 2021). Moreover, transcriptome profiling of two garlic (*Allium sativum* L.) cultivars (salt sensitive and salt tolerant) revealed that under salt stress, most transcripts of the phenylpropanoid biosynthesis pathway were down-regulated in the salt-sensitive genotype. Additionally, many transcripts related to the brassinosteroid signaling and biosynthesis pathways were down-regulated in the salt-sensitive cultivar (Kong et al., 2021). Interestingly, in a similar experimental setup, genes related to other hormones (such as auxin, ethylene, brassinosteroid, abscisic acid, and jasmonate) demonstrated different time-dependent expression patterns (such as auxin-induced protein 22D, auxin response factor, ethylene-responsive transcription factor, ethylene receptor 2–like, protein phosphatase 2C, BES1/BZR1, BRI1, and ZIM domain–containing protein) (Wang et al., 2019a). Furthermore, we discuss recent results dedicated to the defined signaling pathways regulating lignin accumulation to strengthen cell walls and protect membrane integrity under salt stress.

Genome-wide examination of the R2R3-MYB TF family in pigeon pea [*Cajanus cajan*, (L.) Millsp.] identified 30 R2R3-MYB TFs activated by abiotic stresses and involved in lignin and flavonoid biosynthesis. Interestingly, another 122 key enzyme genes from flavonoid and lignin biosynthetic pathways were MeJA (Methyl jasmonate) responsive (Yang et al., 2021).

BpMYB46, the member of the plant MYB transcription family, was characterised in *Betula platyphylla* (Sukaczev.) as associated with abiotic stress tolerance and secondary cell wall biosynthesis (Table 1). Therefore, *BplMYB46* overexpression in plants resulted in

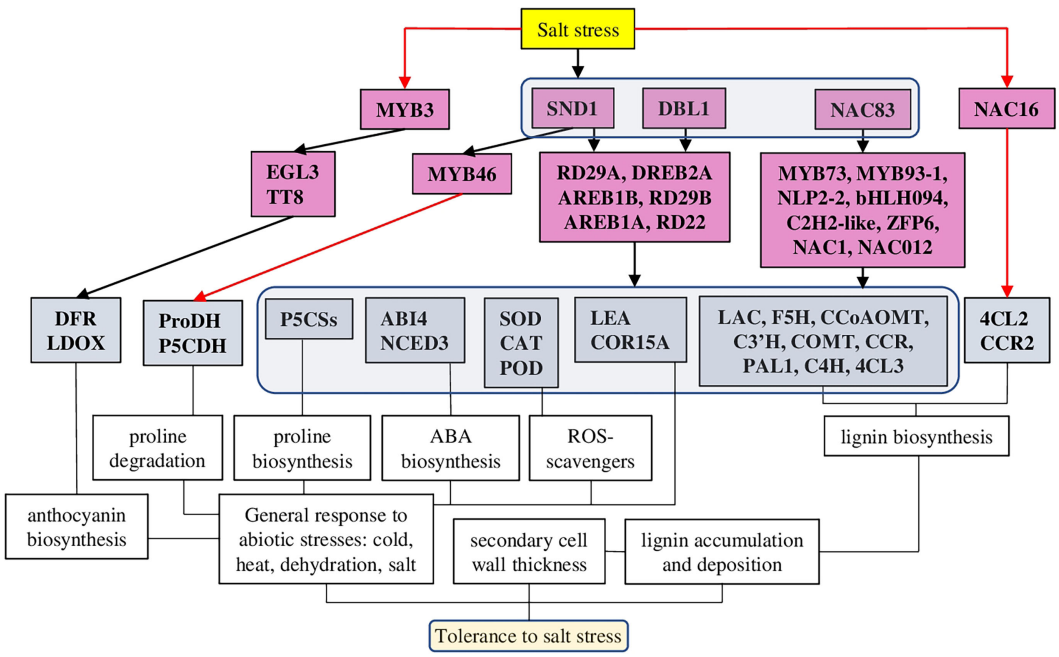


FIGURE 5
A model of different TFs regulating lignin-associated salt stress tolerance in plants. TFs of several families have been identified as salt-mediated activators of lignin biosynthesis and deposition. Among them, several TFs (SND1, DBL1, and NAC83) have been defined as master regulators, which, in addition to the lignin biosynthesis pathway, were activating also other TFs associated with general stress responses, antioxidant defence, ABA, and anthocyanin biosynthesis. The red line represents negative regulation.

increased expression of ROS scavengers (*SOD* and *POD*) and proline biosynthesis genes (*P5CSs*), while proline degrading genes (*P5CDH* and *ProDH*) were down-regulated. Also, *BpMYB46* overexpressing lines showed increased lignin deposition, secondary cell wall thickness, up-regulated genes related to the secondary cell wall formation and reduced stomatal apertures and water loss. In total, mutant lines had improved tolerance to salt and osmotic stresses, while *BpMYB46* knockdown lines were hypersensitive to these stresses (Guo et al., 2017).

Interestingly, Arabidopsis MYB3 TF was identified as a salt-stress–mediated repressor for regulating lignin and anthocyanin biosynthesis. While the expression of many salt response genes was not altered in *myb3* plants, the expression of anthocyanin and lignin biosynthesis genes (such as *PAL1* (phenylalanine ammonia-lyase 1), *C4H*, *COMT* (caffeic acid O-methyltransferase), *4CL3* (4-coumarate–CoA ligase 3), *DFR* (dihydroflavonol 4-reductase), and *LDOX* [leucoanthocyanidin dioxygenase)] and anthocyanin regulators TFs [such as TT8 (Transparent Testa 8) and EGL3 (Enhancer of Glabra 3)] was increased (Kim et al., 2022).

Several members of the NAC transcription family have been characterised as regulators of the expression of lignin-related and salt stress-responsive genes. Indeed, network reconstruction of the lignocellulosic pathway in *Populus davidiana* × *Populus bolleana* under salt stress revealed the involvement of 197 diverse TFs. However, PdbNAC83 was identified as the up-top TF regulating other salt stress- and lignocellulosic synthesis–related genes, including also TFs (*PdbMYB73*, *PdbMYB93-1*, *PdbNLP2-2*, *PdbC2H2-like*, *PdbbHLH094*, and *PdbZFP6*) (Lei et al., 2022).

Analysis of overexpressing and knockdown lines of *NAC012* suggested its function associated with osmotic and salt stress tolerance. Similarly to BpMYB46, BpNAC012 induced the expression of *P5CSs*, *SOD*, and *POD* predominantly in mature stems. Under salt and osmotic stresses, *BpNAC012* overexpressing lines had higher expression levels of lignin biosynthesis genes and enhanced lignin accumulation. On the contrary, knockdown lines greatly reduced the secondary cell wall thickening of stem fibres (Hu et al., 2019). Overexpression of *AgNAC1* (*Apium graveolens* L.) in Arabidopsis resulted in increased tolerance to salt and drought stresses through up-regulation of lignin biosynthesis–related genes [such as *AtLAC* (laccase), *AtF5H* (Ferulate-5-hydroxylase), *AtCCoAOMT*, *AtC3'H*, *AtCOMT*, and *AtCCR*], antioxidants (SOD and POD), and reduced level of MDA (Duan et al., 2020).

Finally, in Arabidopsis, the key NAC family TF SND1 (SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN 1) was characterised as the master regulator, connecting abiotic stress tolerance and secondary cell wall biosynthesis through regulation of abscisic acid levels and activation of other transcription factors. In particular, SND1 was shown to bind *Myb46*, thus activating the lignin biosynthesis pathway, and to directly bind the *ABI4* gene, thus reducing ABA levels. Also, *snd1* knockdown mutants exhibited a low tolerance to salt stress (Jeong et al., 2018). Furthermore, experiments on apple (*Malus* × domestica Borkh.) suggested that *MdSND1* was activated by both ABA and salt stress and bounded directly not only to *MdMYB46/83* but also other stress-related TFs such as *MdRD22*, *MdRD29A*, *MdDREB2A*, *MdAREB1B*, and *MdAREB1A*, thus activating response to a wide range of stresses. Indeed, *MdSND1* overexpressing plants had higher lignin and antioxidant content and enhanced resistance to salt and osmotic stresses, while *MdSND1* knockdown plants were stress sensitive (Chen et al., 2020).

On the contrary, FtNAC16 [*Fagopyrum tataricum* (L.) Gaertn.] was identified as a negative regulator of salt stress tolerance and lignin biosynthesis. *FtNAC16*-expressing Arabidopsis plants demonstrated reduced expression of many lignin biosynthesis–related genes. At the same time, *4CL2* (4-coumarate–CoA ligase 2) and *CCR2* genes were abnormally up-regulated, suggesting a functional compensatory mechanism changing the proportions of the lignin monomer, resulting in hypersensitivity to salt stress (Wang et al., 2021).

DREB family TF isolated from desiccation-tolerant moss *Bryum argenteum* (Hedw.) plays important roles in tolerance to abiotic stresses. *BaDBL1*-expressing Arabidopsis plants demonstrated increased tolerance to salt and osmotic stresses, which was associated with increased activities of antioxidant enzymes [POD, SOD, and CAT (Catalase)] and transcription of lignin-biosynthesis–related genes. Additionally, BaDBL1 up-regulated other abiotic stress-related TFs and genes [such as *AtLEA* (late embryogenesis abundant protein), *AtCOR15A* (Protein COLD-REGULATED) and *AtRD29A*) (Liang et al., 2021).

2.4.2 Lignin biosynthesis genes associated with plant salt tolerance

Activation of lignin biosynthesis and deposition pathways plays an important role in the development of salt stress tolerance. However, limited data available suggested that salt stress affects the biosynthesis of lignin units in a different way, leading to shifts mostly between S and G units. Additionally, the antioxidant defence system, abscisic acid, proline, and anthocyanin biosynthesis are greatly contributing to the development of salt stress tolerance.

Analysis of salt-stress up-regulated genes suggested that *CcoAOMT* (caffeoyl CoA O-methyltransferase) is one of the most highly induced genes (Table 1). The product of CcoAOMT further leads to the guaiacyl and sinapyl lignin formation (G and S units). Interestingly, *ccoaomt1* plants were hypersensitive to salt stress and had lower lignin content, but the amount of only G monomer was reduced, while the amounts of S and H were higher (Figure 4) (Do et al., 2007; Chun et al., 2019). COMT, another crucial enzyme involved in lignin synthesis of S and G monomers, was recently characterised in rice (*Oryza sativa* L.) in relation to salt stress and lignin content. Among 33 identified genes, 5 and 4 *OsCOMTs* were up- and down-regulated under salt stress, respectively. A combined investigation of lignin content and expression analysis of *COMT* genes suggested that *OsCOMT8*, *OsCOMT9*, and *OsCOMT15* as the key players in the synthesis of lignin (Liang et al., 2022).

CAD family is essential for lignin biosynthesis, because it catalyses the final step of all units (H, G, and S) production. As was shown on *Phyllostachys edulis* [(Carrière) J.Houz.], the *PheCAD* family highly co-expressed with biotic and abiotic stress responses, with *PheCAD2*, *PheCAD3*, and *PheCAD5* being the main up-regulated genes after drought and salt stress treatments (Figure 4). Among other *PheCAD* genes highly expressed in many tissues

(*PheCAD1*, *PheCAD2*, *PheCAD6*, *PheCAD8*, and *PheCAD12*), *PheCAD2* demonstrated a positive correlation with most of the lignin biosynthesis enzymes, suggesting that PheCAD2 might be the main enzyme responsible for lignin biosynthesis (Vasupalli et al., 2021).

ROS scavengers also play a crucial role in salt-stress tolerance and are closely connected to lignin biosynthesis. Among other ROS, H_2O_2 is known as a crucial trigger, initiating peroxidase-dependent oxidation of cinnamyl alcohol to lignins (Wen et al., 2020). As it was recently shown, transgenic Arabidopsis overexpressing *SOD* (from *Potentilla atrosanguinea* George Loddiges) and *APX* (L-ascorbate peroxidase) (from *Rheum australe* D.Don) (both individually and combined), had increased lignin accumulation with altered S:G ration and higher levels of compatible solutes, which provided enhanced salt-stress tolerance. In particular, under salt stress conditions, all transgenic plants had twofold higher lignin deposition with a higher S:G ratio. Furthermore, dual transformants had up-regulated expression of many lignin biosynthesis genes under salt stress, while five genes (*CcAOMT1*, *4CL8*, *CAD1*, *4CL3*, and *LAC12)* were down-regulated. Additionally, many cell wall–related TFs were up-regulated [such as VND1 (VASCULAR-RELATED NAC-DOMAIN), VND2, VND4, VND6, SND1, SND2, and NST1 (NAC SECONDARY WALL THICKENING)]. Finally, transgenic plants demonstrated higher biomass production, better growth rate, and increased yield under salt stress when compared to WT plants (Shafi et al., 2015).

3 Root barriers of salt transport – Casparian strips and suberinisation of cell walls

3.1 Regulation of cell wall suberinisation on the transcriptional level

The walls of specialised cells could be modified in response to salt stress, for example, the suberinisation of the endodermis and exodermis (hypodermis) (Shukla and Barberon, 2021). Such modifications affect the transfer of solutes and water, thus preventing the entrance of surplus Na^+ and Cl^- into the plant and the leakage of K^+ and water out (Byrt et al., 2018). TFs members of the MYB and WRKY families are the main regulators of suberin biosynthesis and deposition. Suberin metabolism is closely associated with other fatty acid-based pathways, such as cutin and wax biosynthesis.

Recently, MYB family TFs (MYB41, MYB53, MYB92, and MYB93) have been identified as the main regulators of suberin biosynthesis and deposition, each of which is sufficient to endodermal suberin formation in response to both environmental and developmental signals. Indeed, the quadruple mutant of these four TFs had greatly reduced levels of suberin and affected endodermal deposition of lignin (in CS), probably through a compensatory mechanism. Furthermore, mutant plants had increased sensitivity to salt stress and suberinisation, which was not regulated by ABA and salt treatments. Thus, supporting the central role of suberin in plant adaptation to salt stress (Shukla et al., 2021).

Similarly, MYB family TFs regulate suberin assembly also in the seed coat. As it was recently shown, seeds of *myb9* and *myb107* mutants demonstrated lower germination rates and increased permeability under salt stress due to the reduced amount of suberin monomers and changed levels of other seed coat-associated metabolites (Lashbrooke et al., 2016). Also, MYB49 was shown to contribute to salt-enhanced tolerance in Arabidopsis through the transcriptional activation of cutin deposition and antioxidant defence (Zhang et al., 2020). The plant cuticle, composed of cutin and associated waxes, is usually deposited on the surface of the epidermal cell wall of plant leaves and stems (Kosma et al., 2014). Cutin differs from suberin by lower phenolic content, almost exclusively C16 and C18 monomers, and around 50% of C18:2 monomers (Yang et al., 2017). Therefore, *MYB49* overexpressing plants demonstrated significant up-regulation of the 'cutin, suberin and wax biosyntheses' category of genes under normal and/or salt stress conditions, while in *myb49* mutants these genes were down-regulated. Several biosynthetic genes, such as *ASFT* (omega-hydroxypalmitate O-feruloyl transferase), *FACT* (fatty alcohol:caffeoyl-CoA acyltransferase), *CYP86B1*, which are crucial for incorporating aromatics into suberin, cutin and suberin-associated waxes, and TF MYB41, were defined as the direct targets of MYB49. Additionally, *MYB49* overexpressing lines demonstrate improved antioxidant defence (peroxidases and LEA) and elevated Ca^{2+} levels in leaves, thus enhancing salt tolerance in plants (Zhang et al., 2020).

WRKY family TF 33 was recently identified as a crucial regulator of salt tolerance in both Arabidopsis and rice. Mutants of *wrky33* and its direct target *atcyp94b1 (*involved in suberin biosynthesis) demonstrated reduced suberin content and salt-sensitive phenotype. However, *AtCYP94B1* expressing plants in *wrky33* background and rice wild-type improved both suberin content and salt tolerance, thus confirming the regulatory role of WRKY33 in CYP94B1-mediated salt-tolerance (Krishnamurthy et al., 2020). Similarly, AtWRKY9 was identified as a direct regulator of *AtCYP86B1* and *AtCYP94B3* responsible for salt stress tolerance. *Atwrky9* mutant had reduced roots suberin content and reduced expression of *AtCYP94B3* and *AtCYP86B1*. The salt-sensitive phenotype of *atcyp94b3* and *atcyp86b1* mutants was rescued by the expression of homologues genes from *Avicennia officinalis* (L.) *AoCYP94B3* and *AoCYP86B1* were also associated with reduced Na^+ accumulation in the shoots (Krishnamurthy et al., 2021).

3.2 Suberin biosynthesis genes associated with plant salt tolerance

Suberin homeostasis is closely associated with fatty acid and carbohydrate metabolism. Similarly to other stress-related responses, ABA is involved in the suberin biosynthesis and deposition, CS formation and biosynthesis of compatible solute trehalose. However, the role of CS and CASPs family members requires further investigation.

Many individual genes involved in suberin biosynthesis have been demonstrated to improve the tolerance to salt stress. KCS (β-Ketoacyl-CoA synthase) catalyses the condensation of malonyl-CoA with acyl-CoA; thus, it is a crucial rate-limiting enzyme in the suberin biosynthesis (Rui et al., 2022). Expression of *KCS11* from grape (*Vitis vinifera* L.) in Arabidopsis increased its salt-stress tolerance on germination and seedling stages. This improvement was mediated *via* several mechanisms: accumulation of osmotic regulating substances (increased proline content), membrane stabilisation (reduced MDA content), and normalisation of ions homeostasis (increased expression of genes encoding several ion transporters and channels (such as *AKT1*, *CBL9*, *CIPK23*, *SOS1*, *HKT1* and *CHX14* (Cation/H(+) antiporter)) (Yang et al., 2020). Similarly, expression of *KCS* from quinoa (*Chenopodium quinoa* Willd.) in Arabidopsis increased the occurrence of VLFAs (very long-chain fatty acids) with C22-24 chain lengths and promoting the accumulation of suberin monomers, thus improving salt tolerance (Tariq et al., 2022). These results demonstrated that suberin plays rather a universal role in salt tolerance and suggested *KCS* family members as potential genes to improve the salt-stress tolerance in different species.

Recently, the role of CASPs (Casparian strip membrane domain proteins) in salt tolerance was demonstrated. CASPs (UPF0497) are required for the CS formation at the endodermis and regulation of the selective uptake of mineral elements by roots, in particular Si and Ca^{2+} (Wang et al., 2019). The expression of *CASP4* from sorghum [*Sorghum bicolor* (L.) Mocnch] in Arabidopsis (in both WT and *atcasp5* mutant) increased salt stress tolerance, which was associated with lower level osmotic stress-associated damages and Na^+ accumulation in leaves. Furthermore, transgenic plants had longer roots and higher expression of the genes related to Casparian strip formation (Wei et al., 2021). However, experiments of CS defective mutants suggested that CS formation and suberin deposition processes are interrelated and closely regulated by endogenous factors [small secretory peptides, such as CIF (Casparian strip Integrity Factor) and environmental factors (availability of nutrients, abiotic stresses)] (Doblas et al., 2017). Therefore, further research is required to better understand the complex association between CASPs, lignin, and suberin deposition under the influence of salt stress and the application of different regulatory proteins.

Additionally, TPS (trehalose-6-phosphate synthase), which is required for the biosynthesis of compatible solute trehalose, which is a crucial osmoprotector during salt and drought stresses (Nuccio et al., 2015), also mediated enhanced suberin-associated salt tolerance. In particular, rice *tps8* mutant demonstrated enhanced sensitivity to salt and ABA treatments, reduced content of soluble sugars, CS, and suberin deposition in the roots. In addition, ABA-responsive genes were down-regulated in *tps8* plants. On the contrary, *TPS8* overexpression rescued the salt-sensitivity associated phenotype and enhanced salt stress tolerance. These data suggested that rice TPS8 regulates salt tolerance *via* suberin deposition in an ABA-dependent way (Vishal et al., 2019). ABC family transporters (G type) are known to play a crucial role in suberin biosynthesis through the transport of suberin monomers from the cytoplasm to the apoplastic space (Shanmugarajah et al., 2019).

4 Conclusion and prospects

The cell wall is the first line of defence against environmental stresses. High salinity, as one of the worldwide distributed stress factors, can disrupt cell wall integrity and dysregulate normal signaling and nutritional functions. The severity of caused damage depends on salt concentration, availability of other ions (such as Ca^{2+} and K^+) and associated stresses (drought, heat, and light intensity). Thanks to the great recent progress in our understanding of salt sensing and cell wall maintenance mechanisms under salt stress conditions, several new strategies for the improvement of the plant have been developed. In general, all kinds of plant modifications leading to the enforcement of physical cell wall barriers for salt transport, such as enhancement of suberisation and Casparian strip synthesis in root tissues, might greatly facilitate salt tolerance in plants. The development of gene editing technologies allows to generate of salt-tolerant varieties of cultivated species by target modification of the most relevant genes. First, the TFs—master regulators are responsible for cellulose, lignin, and suberin biosynthesis: HSFA7b, SND1, NAC83, MYB9, MYB41, MYB49, and WRKY33 (Figures 3 and 5). Increased production of ROS is a rather common and unspecific plant response to diverse stresses (including salt stress), thus suggesting antioxidants as an obvious and universal target to improve general plant tolerance to stresses. Members of the PME and XTH families respond to salt stress differently. However, overexpression or knockdown of specific PME or XTH genes may provide desirable outcomes: facilitate growth, yield and biomass accumulation, drought, and salt tolerance. Finally, despite multiple data supporting the role of abscisic acid in the salt stress-mediated modifications of the cell wall, our understanding of the role of plant hormones in the immediate and long-term cell wall modification pathways are rather limited and require further investigation.

In the past decades, extensive research has greatly improved our understanding of the mechanisms of salt stress tolerance in plants. However, the practical application of this vital knowledge to improve plant tolerance to salt stress is a rather slow and tardy process. The primary question in the field of salt stress response in plants is the identification of salt stress sensor/s. In this case it is necessary clearly define Na^+/Cl^- sensors and sensors recognising osmotic, ionic and ROS-related signals. While certain progress has been made in the direction of plasma membrane salt stress sensing, Na^+ perception if other organelles (ER, mitochondria, and chloroplasts) is much less exploited. In the next step, the integrating, coordinating and cross-talk mechanisms of salt-sensing signals from different organelles to provide optimised cellular response should be addressed. Also, it would be interesting to investigate the difference between Na^+ and Cl^- sensors and signaling pathways in shoots and roots.

Moreover, salt-stress sensing and signaling in the cell wall relay and closely associated with Ca^{2+} channels, which need to be identified and characterised. Also, their connection to the cell wall repair mechanisms under salt stress conditions requires further investigation. Also, the circadian clock is known to regulate several crucial proteins (such as SOS1 antiporter) (Cha et al., 2022), thus adding another layer of complexity to the salt stress tolerance regulation. In this regard, it would be interesting to find out if

other transporters and channels are subjected to circadian regulation and how they interact with cell wall biosynthesis processes. For example, how the SOS1-regulated pathway is related to the CesA internalisation, microtubule bundling and stability.

Currently, little is known about how the severity and duration of salt stress exposure correlate with cell wall changes and repair. While modern omics technologies can provide excess information about gene expression, protein, and metabolite abundance, the proper evaluation and establishment of regulatory mechanisms are usually missing. Furthermore, currently used methods of genetic engineering and biotechnology are based on single-gene analysis when the phenotype/genotype of every mutant (overexpressing or loss of function) is analysed independently. Such an approach is well-suited to study a particular pathway; however, it does not allow studying an organism as a system of interconnected and interdependent processes.

The expression of a specific gene in an organ-, tissue-, or cell-dependent context may provide additional complexity to the gene regulatory network at the whole-plant level. Also, while the root is the primary organ for salt sense and response, the molecular mechanisms regulating root-to-shoot signaling and transport need to be further elaborated. Finally, the role of known phytohormones and their cross-talk in the regulation of cell wall properties should be studied on the whole-plant level systemically.

Author contributions

Conceptualization, methodology, formal analysis, SD and SI. Writing—original draft preparation, SD. Supervision, SI. Writing—review and editing, SD and SI. All authors contributed to the article and approved the submitted version.

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.

Publisher's note

References

Allen, H., Wei, D., Gu, Y., and Li, S. (2021). A historical perspective on the regulation of cellulose biosynthesis. *Carbohydr. Polymers* 252, 117022. doi: 10.1016/j.carbpol.2020.117022

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

Byrt, C. S., Munns, R., Burton, R. A., Gilliham, M., and Wege, S. (2018). Root cell wall solutions for crop plants in saline soils. *Plant Sci.* 269, 47–55. doi: 10.1016/j.plantsci.2017.12.012

Cha, J.-Y., Kim, J., Jeong, S. Y., Shin, G.-I., Ji, M. G., Hwang, J.-W., et al. (2022). The Na^+/H^+ antiporter SALT OVERLY SENSITIVE 1 regulates salt compensation of circadian rhythms by stabilizing GIGANTEA in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 119, e2207275119. doi: 10.1073/pnas.2207275119

Chen, J., Chen, X., Zhang, Q., Zhang, Y., Ou, X., An, L., et al. (2018). A cold-induced pectin methyl-esterase inhibitor gene contributes negatively to freezing tolerance but positively to salt tolerance in arabidopsis. *J. Plant Physiol.* 222, 67–78. doi: 10.1016/j.jplph.2018.01.003

Chen, K., Guo, Y., Song, M., Liu, L., Xue, H., Dai, H., et al. (2020). Dual role of MdSND1 in the biosynthesis of lignin and in signal transduction in response to salt and osmotic stress in apple. *Hortic. Res.* 7, 204. doi: 10.1038/s41438-020-00433-7

Cheng, Z., Zhang, X., Yao, W., Gao, Y., Zhao, K., Guo, Q., et al. (2021). Genome-wide identification and expression analysis of the xyloglucan endotransglucosylase/hydrolase gene family in poplar. *BMC Genomics* 22, 804. doi: 10.1186/s12864-021-08134-8

Choi, J. Y., Seo, Y. S., Kim, S. J., Kim, W. T., and Shin, J. S. (2011). Constitutive expression of CaXTH3, a hot pepper xyloglucan endotransglucosylase/hydrolase, enhanced tolerance to salt and drought stresses without phenotypic defects in tomato plants (Solanum lycopersicum cv. dotaerang). *Plant Cell Rep.* 30, 867–877. doi: 10.1007/s00299-010-0989-3

Chun, H. J., Baek, D., Cho, H. M., Lee, S. H., Jin, B. J., Yun, D.-J., et al. (2019). Lignin biosynthesis genes play critical roles in the adaptation of *Arabidopsis* plants to high-salt stress. *Plant Signaling Behav.* 14, 1625697. doi: 10.1080/15592324.2019.1625697

Coculo, D., and Lionetti, V. (2022). The plant Invertase/Pectin methylesterase inhibitor superfamily. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.863892

Colin, L., Ruhnow, F., Zhu, J.-K., Zhao, C., Zhao, Y., and Persson, S. (2023). The cell biology of primary cell walls during salt stress. *Plant Cell* 35, 201–217. doi: 10.1093/plcell/koac292

Coomey, J. H., Sibout, R., and Hazen, S. P. (2020). Grass secondary cell walls, *Brachypodium distachyon* as a model for discovery. *New Phytol.* 227, 1649–1667. doi: 10.1111/nph.16603

Cruz-Valderrama, J. E., Gómez-Maqueo, X., Salazar-Iribe, A., Zúñiga-Sánchez, E., Hernández-Barrera, A., Quezada-Rodríguez, E., et al. (2019). Overview of the role of cell wall DUF642 proteins in plant development. *IJMS* 20, 3333. doi: 10.3390/ijms20133333

Dauphin, B. G., Ranocha, P., Dunand, C., and Burlat, V. (2022). Cell-wall microdomain remodeling controls crucial developmental processes. *Trends Plant Sci.* 27, 1033–1048. doi: 10.1016/j.tplants.2022.05.010

De Lorenzo, G., Ferrari, S., Giovannoni, M., Mattei, B., and Cervone, F. (2019). Cell wall traits that influence plant development, immunity and bioconversion. *Plant J* 97:134–147. doi: 10.1111/tpj.14196

Do, C.-T., Pollet, B., Thévenin, J., Sibout, R., Denoue, D., Barrière, Y., et al. (2007). Both caffeoyl coenzyme a 3-o-methyltransferase 1 and caffeic acid O-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in arabidopsis. *Planta* 226, 1117–1129. doi: 10.1007/s00425-007-0558-3

Doblas, V. G., Geldner, N., and Barberon, M. (2017). The endodermis, a tightly controlled barrier for nutrients. *Curr. Opin. Plant Biol.* 39, 136–143. doi: 10.1016/j.pbi.2017.06.010

Du, J., Anderson, C. T., and Xiao, C. (2022). Dynamics of pectic homogalacturonan in cellular morphogenesis and adhesion, wall integrity sensing and plant development. *Nat. Plants* 8, 332–340. doi: 10.1038/s41477-022-01120-2

Duan, A.-Q., Tao, J.-P., Jia, L.-L., Tan, G.-F., Liu, J.-X., Li, T., et al. (2020). AgNAC1, a celery transcription factor, related to regulation on lignin biosynthesis and salt tolerance. *Genomics* 112, 5254–5264. doi: 10.1016/j.ygeno.2020.09.049

Dünser, K., Gupta, S., Herger, A., Feraru, M. I., Ringli, C., and Kleine-Vehn, J. (2019). Extracellular matrix sensing by FERONIA and leucine-rich repeat extensins controls vacuolar expansion during cellular elongation in arabidopsis thaliana. *EMBO J.* 38, e100353. doi: 10.15252/embj.2018100353

Edmond Ghanem, M., Han, R.-M., Classen, B., Quetin-Leclerq, J., Mahy, G., Ruan, C.-J., et al. (2010). Mucilage and polysaccharides in the halophyte plant species kosteletzkya virginica: Localization and composition in relation to salt stress. *J. Plant Physiol.* 167, 382–392. doi: 10.1016/j.jplph.2009.10.012

Endler, A., Kesten, C., Schneider, R., Zhang, Y., Ivakov, A., Froehlich, A., et al. (2015). A mechanism for sustained cellulose synthesis during salt stress. *Cell* 162, 1353–1364. doi: 10.1016/j.cell.2015.08.028

Fang, C., Li, K., Wu, Y., Wang, D., Zhou, J., Liu, X., et al. (2019). *OsTSD2* -mediated cell wall modification affects ion homeostasis and salt tolerance. *Plant Cell Environ.* 42, 1503–1512. doi: 10.1111/pce.13499

Feng, W., Kita, D., Peaucelle, A., Cartwright, H. N., Doan, V., Duan, Q., et al. (2018). The FERONIA receptor kinase maintains cell-wall integrity during salt stress through Ca2+ signaling. *Curr. Biol.* 28, 666–675.e5. doi: 10.1016/j.cub.2018.01.023

Gallina, G., Cabeza, Á., Grénman, H., Biasi, P., García-Serna, J., and Salmi, T. (2018). Hemicellulose extraction by hot pressurized water pretreatment at 160 °C for 10 different woods: Yield and molecular weight. *J. Supercritical Fluids* 133, 716–725. doi: 10.1016/j.supflu.2017.10.001

Gigli-Bisceglia, N., van Zelm, E., Huo, W., Lamers, J., and Testerink, C. (2022). *Arabidopsis* root responses to salinity depend on pectin modification and cell wall sensing. *Development* 149, dev200363. doi: 10.1242/dev.200363

Guo, H., Wang, Y., Wang, L., Hu, P., Wang, Y., Jia, Y., et al. (2017). Expression of the MYB transcription factor gene *BplMYB46* affects abiotic stress tolerance and secondary cell wall deposition in *Betula platyphylla*. *Plant Biotechnol. J.* 15, 107–121. doi: 10.1111/pbi.12595

Han, Y., Han, S., Ban, Q., He, Y., Jin, M., and Rao, J. (2017). Overexpression of persimmon DkXTH1 enhanced tolerance to abiotic stress and delayed fruit softening in transgenic plants. *Plant Cell Rep.* 36, 583–596. doi: 10.1007/s00299-017-2105-4

Han, Y., Wang, W., Sun, J., Ding, M., Zhao, R., Deng, S., et al. (2013). Populus euphratica XTH overexpression enhances salinity tolerance by the development of leaf succulence in transgenic tobacco plants. *J. Exp. Bot.* 64, 4225–4238. doi: 10.1093/jxb/ert229

Hocq, L., Pelloux, J., and Lefebvre, V. (2017). Connecting homogalacturonan-type pectin remodeling to acid growth. *Trends Plant Sci.* 22, 20–29. doi: 10.1016/j.tplants.2016.10.009

Hu, P., Zhang, K., and Yang, C. (2019). BpNAC012 positively regulates abiotic stress responses and secondary wall biosynthesis. *Plant Physiol.* 179, 700–717. doi: 10.1104/pp.18.01167

Isayenkov, S., Hilo, A., Rizzo, P., Tandron Moya, Y. A., Rolletschek, H. , Borisjuk, L., et al. (2020). Adaptation strategies of halophytic barley hordeum marinum ssp. marinum to high salinity and osmotic stress. *IJMS* 21 9019. doi: 10.3390/ijms21239019

Isayenkov, S. V., and Maathuis, F. J. M. (2019). Plant salinity stress: Many unanswered questions remain. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00080

Ishida, K., and Yokoyama, R. (2022). Reconsidering the function of the xyloglucan endotransglucosylase/hydrolase family. *J. Plant Res.* 135, 145–156. doi: 10.1007/s10265-021-01361-w

Jeong, C. Y., Lee, W. J., Truong, H. A., Trịnh, C. S., Jin, J. Y., Kim, S., et al. (2018). Dual role of SND1 facilitates efficient communication between abiotic stress signalling and normal growth in arabidopsis. *Sci. Rep.* 8, 10114. doi: 10.1038/s41598-018-28413-x

Jin, J., Duan, J., Shan, C., Mei, Z., Chen, H., Feng, H., et al. (2020). Ethylene insensitive3-like2 (OsEIL2) confers stress sensitivity by regulating OsBURP16, the β subunit of polygalacturonase (PG1β-like) subfamily gene in rice. *Plant Sci.* 292, 110353. doi: 10.1016/j.plantsci.2019.110353

Kesten, C., García-Moreno, Á., Amorim-Silva, V., Menna, A., Castillo, A. G., Percio, F., et al. (2022). Peripheral membrane proteins modulate stress tolerance by safeguarding cellulose synthases. *Sci. Adv.* 8, eabq6971. doi: 10.1126/sciadv.abq6971

Kesten, C., Wallmann, A., Schneider, R., McFarlane, H. E., Diehl, A., Khan, G. A., et al. (2019). The companion of cellulose synthase 1 confers salt tolerance through a tau-like mechanism in plants. *Nat. Commun.* 10, 857. doi: 10.1038/s41467-019-08780-3

Khodayari, A., Thielemans, W., Hirn, U., Van Vuure, A. W., and Seveno, D. (2021). Cellulose-hemicellulose interactions - a nanoscale view. *Carbohydr. Polymers* 270, 118364. doi: 10.1016/j.carbpol.2021.118364

Kim, D., Jeon, S. J., Yanders, S., Park, S., Kim, H. S., and Kim, S. (2022). MYB3 plays an important role in lignin and anthocyanin biosynthesis under salt stress condition in arabidopsis. *Plant Cell Rep.* 41, 1549–1560. doi: 10.1007/s00299-022-02878-7

Kim, Y. X., Ranathunge, K., Lee, S., Lee, Y., Lee, D., and Sung, J. (2018). Composite transport model and water and solute transport across plant roots: An update. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00193

Kong, Q., Mostafa, H. H. A., Yang, W., Wang, J., Nuerawuti, M., Wang, Y., et al. (2021). Comparative transcriptome profiling reveals that brassinosteroid-mediated lignification plays an important role in garlic adaption to salt stress. *Plant Physiol. Biochem.* 158, 34–42. doi: 10.1016/j.plaphy.2020.11.033

Kosma, D. K., Murmu, J., Razeq, F. M., Santos, P., Bourgault, R., Molina, I., et al. (2014). At MYB 41 activates ectopic suberin synthesis and assembly in multiple plant species and cell types. *Plant J.* 80, 216–229. doi: 10.1111/tpj.12624

Krishnamurthy, P., Vishal, B., Bhal, A., and Kumar, P. P. (2021). WRKY9 transcription factor regulates cytochrome P450 genes *CYP94B3* and *CYP86B1* , leading to increased root suberin and salt tolerance in arabidopsis. *Physiol. Plantarum* 172, 1673–1687. doi: 10.1111/ppl.13371

Krishnamurthy, P., Vishal, B., Ho, W. J., Lok, F. C. J., Lee, F. S. M., and Kumar, P. P. (2020). Regulation of a cytochrome P450 gene *CYP94B1* by WRKY33 transcription factor controls apoplastic barrier formation in roots to confer salt tolerance. *Plant Physiol.* 184, 2199–2215. doi: 10.1104/pp.20.01054

Kuluev, B., Mikhaylova, E., Berezhneva, Z., Nikonorov, Y., Postrigan, B., Kudoyarova, G., et al. (2017). Expression profiles and hormonal regulation of tobacco NtEXGT gene and its involvement in abiotic stress response. *Plant Physiol. Biochem.* 111, 203–215. doi: 10.1016/j.plaphy.2016.12.005

Lampugnani, E. R., Khan, G. A., Somssich, M., and Persson, S. (2018). Building a plant cell wall at a glance. *J. Cell Sci.* 131, jcs207373. doi: 10.1242/jcs.207373

Lashbrooke, J., Cohen, H., Levy-Samocha, D., Tzfadia, O., Panizel, I., Zeisler, V., et al. (2016). MYB107 and MYB9 homologs regulate suberin deposition in angiosperms. *Plant Cell* 28, 2097–2116. doi: 10.1105/tpc.16.00490

Lei, X., Liu, Z., Xie, Q., Fang, J., Wang, C., Li, J., et al. (2022). Construction of two regulatory networks related to salt stress and lignocellulosic synthesis under salt stress based on a populus davidiana × p. bolleana transcriptome analysis. *Plant Mol. Biol.* 109, 689–702. doi: 10.1007/s11103-022-01267-8

Li, C., Qi, Y., Zhao, C., Wang, X., and Zhang, Q. (2021). Transcriptome profiling of the salt stress response in the leaves and roots of halophytic eutrema salsugineum. *Front. Genet.* 12. doi: 10.3389/fgene.2021.770742

Li, S., Zhang, L., Wang, Y., Xu, F., Liu, M., Lin, P., et al. (2017). Knockdown of a cellulose synthase gene BoiCesA affects the leaf anatomy, cellulose content and salt tolerance in broccoli. *Sci. Rep.* 7, 41397. doi: 10.1038/srep41397

Liang, Y., Li, X., Yang, R., Gao, B., Yao, J., Oliver, M. J., et al. (2021). BaDBL1, a unique DREB gene from desiccation tolerant moss bryum argenteum, confers osmotic and salt stress tolerances in transgenic arabidopsis. *Plant Sci.* 313, 111047. doi: 10.1016/j.plantsci.2021.111047

Liang, S., Xu, S., Qu, D., Yang, L., Wang, J., Liu, H., et al. (2022). Identification and functional analysis of the caffeic acid O-methyltransferase (COMT) gene family in rice (Oryza sativa l.). *IJMS* 23, 8491. doi: 10.3390/ijms23158491

Liu, J., Zhang, W., Long, S., and Zhao, C. (2021). Maintenance of cell wall integrity under high salinity. *IJMS* 22, 3260. doi: 10.3390/ijms22063260

Mahajan, M., and Yadav, S. K. (2014). Overexpression of a tea flavanone 3-hydroxylase gene confers tolerance to salt stress and alternaria solani in transgenic tobacco. *Plant Mol. Biol.* 85, 551–573. doi: 10.1007/s11103-014-0203-z

Nagashima, Y., Ma, Z., Liu, X., Qian, X., Zhang, X., von Schaewen, A., et al. (2020). Multiple quality control mechanisms in the ER and TGN determine subcellular dynamics and salt-stress tolerance function of KORRIGAN1. *Plant Cell* 32, 470–485. doi: 10.1105/tpc.19.00714

Naseer, S., Lee, Y., Lapierre, C., Franke, R., Nawrath, C., and Geldner, N. (2012). Casparian strip diffusion barrier in *Arabidopsis* is made of a lignin polymer without suberin. *Proc. Natl. Acad. Sci. U.S.A.* 109, 10101–10106. doi: 10.1073/pnas.1205726109

Nomberg, G., Marinov, O., Arya, G. C., Manasherova, E., and Cohen, H. (2022). The key enzymes in the suberin biosynthetic pathway in plants: An update. *Plants* 11, 392. doi: 10.3390/plants11030392

Nuccio, M. L., Wu, J., Mowers, R., Zhou, H.-P., Meghji, M., Primavesi, L. F., et al. (2015). Expression of trehalose-6-phosphate phosphatase in maize ears improves yield in well-watered and drought conditions. *Nat. Biotechnol.* 33, 862–869. doi: 10.1038/nbt.3277

Palmeros-Suárez, P. A., Massange-Sánchez, J. A., Sánchez-Segura, L., Martínez-Gallardo, N. A., Espitia Rangel, E., Gómez-Leyva, J. F., et al. (2017). AhDGR2, an amaranth abiotic stress-induced DUF642 protein gene, modifies cell wall structure and composition and causes salt and ABA hyper-sensibility in transgenic arabidopsis. *Planta* 245, 623–640. doi: 10.1007/s00425-016-2635-y

Park, Y. B., and Cosgrove, D. J. (2015). Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol.* 56, 180–194. doi: 10.1093/pcp/pcu204

Qiao, T., Zhang, L., Yu, Y., Pang, Y., Tang, X., Wang, X., et al. (2022). Identification and expression analysis of xyloglucan endotransglucosylase/hydrolase (XTH) family in grapevine (*Vitis vinifera* l.). *PeerJ* 10, e13546. doi: 10.7717/peerj.13546

Rahman, M. A., Woo, J. H., Lee, S.-H., Park, H. S., Kabir, A. H., Raza, A., et al. (2022). Regulation of Na(+)/H(+) exchangers, Na(+)/K(+) transporters, and lignin biosynthesis genes, along with lignin accumulation, sodium extrusion, and antioxidant defense, confers salt tolerance in alfalfa. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1041764

Rui, C., Chen, X., Xu, N., Wang, J., Zhang, H., Li, S., et al. (2022). Identification and structure analysis of KCS family genes suggest their reponding to regulate fiber development in long-staple cotton under salt-alkaline stress. *Front. Genet.* 13. doi: 10.3389/fgene.2022.812449

Rui, Y., and Dinneny, J. R. (2020). A wall with integrity: surveillance and maintenance of the plant cell wall under stress. *New Phytol.* 225, 1428–1439. doi: 10.1111/nph.16166

Sager, R., and Lee, J.-Y. (2014). Plasmodesmata in integrated cell signalling: insights from development and environmental signals and stresses. *J. Exp. Bot.* 65, 6337–6358. doi: 10.1093/jxb/eru365

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

Serra, O., and Geldner, N. (2022). The making of suberin. *New Phytol.* 235, 848–866. doi: 10.1111/nph.18202

Shafi, A., Chauhan, R., Gill, T., Swarnkar, M. K., Sreenivasulu, Y., Kumar, S., et al. (2015). Expression of SOD and APX genes positively regulates secondary cell wall biosynthesis and promotes plant growth and yield in arabidopsis under salt stress. *Plant Mol. Biol.* 87, 615–631. doi: 10.1007/s11103-015-0301-6

Shanmugarajah, K., Linka, N., Gräfe, K., Smits, S. H. J., Weber, A. P. M., Zeier, J., et al. (2019). ABCG1 contributes to suberin formation in arabidopsis thaliana roots. *Sci. Rep.* 9, 11381. doi: 10.1038/s41598-019-47916-9

Shukla, V., and Barberon, M. (2021). Building and breaking of a barrier: Suberin plasticity and function in the endodermis. *Curr. Opin. Plant Biol.* 64, 102153. doi: 10.1016/j.pbi.2021.102153

Shukla, V., Han, J.-P., Cléard, F., Lefebvre-Legendre, L., Gully, K., Flis, P., et al. (2021). Suberin plasticity to developmental and exogenous cues is regulated by a set of MYB transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2101730118. doi: 10.1073/pnas.2101730118

Singh, A. (2021). Soil salinization management for sustainable development: A review. *J. Environ. Manage.* 277, 111383. doi: 10.1016/j.jenvman.2020.111383

Sun, J., Tian, Z., Li, X., Li, S., Li, Z., Wang, J., et al. (2022). Systematic analysis of the pectin methylesterase gene family in nicotiana tabacum and reveal their multiple roles in plant development and abiotic stresses. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.998841

Tang, Y., Wang, M., Cao, L., Dang, Z., Ruan, N., Wang, Y., et al. (2022). *OsUGE3* -mediated cell wall polysaccharides accumulation improves biomass production, mechanical strength, and salt tolerance. *Plant Cell Environ.* 45, 2492–2507. doi: 10.1111/pce.14359

Tariq, F., Zhao, S., Ahmad, N., Wang, P., Shao, Q., Ma, C., et al. (2022). Overexpression of β-ketoacyl CoA synthase 2B.1 from chenopodium quinoa promotes suberin monomers' production and salt tolerance in arabidopsis thaliana. *IJMS* 23, 13204. doi: 10.3390/ijms232113204

Tiika, R. J., Wei, J., Cui, G., Ma, Y., Yang, H., and Duan, H. (2021). Transcriptome-wide characterization and functional analysis of xyloglucan endo-transglycosylase/hydrolase (XTH) gene family of salicornia europaea l. under salinity and drought stress. *BMC Plant Biol.* 21, 491. doi: 10.1186/s12870-021-03269-y

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

van Zelm, E., Zhang, Y., and Testerink, C. (2020). Salt tolerance mechanisms of plants. *Annu. Rev. Plant Biol.* 71, 403–433. doi: 10.1146/annurev-arplant-050718-100005

Vasupalli, N., Hou, D., Singh, R. M., Wei, H., Zou, L.-H., Yrjälä, K., et al. (2021). Homo- and hetero-dimers of CAD enzymes regulate lignification and abiotic stress response in moso bamboo. *IJMS* 22, 12917. doi: 10.3390/ijms222312917

Vishal, B., Krishnamurthy, P., Ramamoorthy, R., and Kumar, P. P. (2019). *Os TPS 8* controls yield-related traits and confers salt stress tolerance in rice by enhancing suberin deposition. *New Phytol.* 221, 1369–1386. doi: 10.1111/nph.15464

Wang, C.-F., Han, G.-L., Yang, Z.-R., Li, Y.-X., and Wang, B.-S. (2022). Plant salinity sensors: Current understanding and future directions. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.859224

Wang, J., Ma, Z., Tang, B., Yu, H., Tang, Z., Bu, T., et al. (2021). Tartary buckwheat (Fagopyrum tataricum) NAC transcription factors FtNAC16 negatively regulates of pod cracking and salinity tolerant in arabidopsis. *IJMS* 22, 3197. doi: 10.3390/ijms22063197

Wang, G.-L., Ren, X.-Q., Liu, J.-X., Yang, F., Wang, Y.-P., and Xiong, A.-S. (2019a). Transcript profiling reveals an important role of cell wall remodeling and hormone signaling under salt stress in garlic. *Plant Physiol. Biochem.* 135, 87–98. doi: 10.1016/j.plaphy.2018.11.033

Wang, Z., Yamaji, N., Huang, S., et al. (2019). OsCASP1 is required for casparian strip formation at endodermal cells of rice roots for selective uptake of mineral elements. *Plant Cell* 31(11):2636–2648. doi: 10.1105/tpc.19.00296

Wang, J., Zhang, L., Wang, X., Liu, L., Lin, X., Wang, W., et al. (2019b). PvNAC1 increases biomass and enhances salt tolerance by decreasing na+ accumulation and promoting ROS scavenging in switchgrass (Panicum virgatum l.). *Plant Sci.* 280, 66–76. doi: 10.1016/j.plantsci.2018.11.007

Wei, X., Liu, L., Lu, C., Yuan, F., Han, G., and Wang, B. (2021). SbCASP4 improves salt exclusion by enhancing the root apoplastic barrier. *Planta* 254, 81. doi: 10.1007/s00425-021-03731-z

Wen, M., Wang, H., Chen, Y., Jiang, Y., Chen, F., and Luo, Z. (2020). Inhibition effect of super atmospheric O2 packaging on H2O2-production and the key enzymes of ligin biosynthesis in fresh-cut Chinese cabbage. *Postharvest Biol. Technol.* 159, 111027. doi: 10.1016/j.postharvbio.2019.111027

Woolfson, K. N., Esfandiari, M., and Bernards, M. A. (2022). Suberin biosynthesis, assembly, and regulation. *Plants* 11, 555. doi: 10.3390/plants11040555

Wu, X., Bacic, A., Johnson, K. L., and Humphries, J. (2020). The role of brachypodium distachyon wall-associated kinases (WAKs) in cell expansion and stress responses. *Cells* 9, 2478. doi: 10.3390/cells9112478

Xu, P., Fang, S., Chen, H., and Cai, W. (2020). The brassinosteroid-responsive xyloglucan endotransglucosylase/hydrolase 19 (*XTH19*) and *XTH23* genes are involved in lateral root development under salt stress in arabidopsis. *Plant J.* 104, 59–75. doi: 10.1111/tpj.14905

Xuan, Y., Zhou, Z. S., Li, H. B., and Yang, Z. M. (2016). Identification of a group of XTHs genes responding to heavy metal mercury, salinity and drought stresses in medicago truncatula. *Ecotoxicol. Environ. Saf.* 132, 153–163. doi: 10.1016/j.ecoenv.2016.06.007

Yan, J., He, H., Fang, L., and Zhang, A. (2018). Pectin methylesterase31 positively regulates salt stress tolerance in arabidopsis. *Biochem. Biophys. Res. Commun.* 496, 497–501. doi: 10.1016/j.bbrc.2018.01.025

Yan, J., Huang, Y., He, H., Han, T., Di, P., Sechet, J., et al. (2019). Xyloglucan endotransglucosylase-hydrolase30 negatively affects salt tolerance in arabidopsis. *J. Exp. Bot.* 70, 5495–5506. doi: 10.1093/jxb/erz311

Yang, Z., Yang, X., Dong, S., Ge, Y., Zhang, X., Zhao, X., et al. (2020). Overexpression of β-Ketoacyl-CoA synthase from vitis vinifera l. improves salt tolerance in arabidopsis thaliana. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.564385

Yang, J., Zhang, S., Li, H., Wang, L., Liu, Y., Niu, L., et al. (2021). Genome-wide analysis and characterization of R2R3-MYB family in pigeon pea (Cajanus cajan) and their functional identification in phenylpropanoids biosynthesis. *Planta* 254, 64. doi: 10.1007/s00425-021-03713-1

Yang, X., Zhao, H., Kosma, D. K., Tomasi, P., Dyer, J. M., Li, R., et al. (2017). The acyl desaturase CER17 is involved in producing wax unsaturated primary alcohols and cutin monomers. *Plant Physiol.* 173, 1109–1124. doi: 10.1104/pp.16.01956

Yu, Y., and Assmann, S. M. (2018). Inter-relationships between the heterotrimeric gβ subunit AGB1, the receptor-like kinase FERONIA, and RALF1 in salinity response. *Plant Cell Environ.* 41, 2475–2489. doi: 10.1111/pce.13370

Yu, Z., Duan, X., Luo, L., Dai, S., Ding, Z., and Xia, G. (2020). How plant hormones mediate salt stress responses. *Trends Plant Sci.* 25, 1117–1130. doi: 10.1016/j.tplants.2020.06.008

Zang, D., Wang, J., Zhang, X., Liu, Z., and Wang, Y. (2019). Arabidopsis heat shock transcription factor HSFA7b positively mediates salt stress tolerance by binding to an e-box-like motif to regulate gene expression. *J. Exp. Bot.* 70, 5355–5374. doi: 10.1093/jxb/erz261

Zhang, B., Gao, Y., Zhang, L., and Zhou, Y. (2021). The plant cell wall: Biosynthesis, construction, and functions. *J. Integr. Plant Biol.* 63, 251–272. doi: 10.1111/jipb.13055

Zhang, S.-S., Sun, L., Dong, X., Lu, S.-J., Tian, W., and Liu, J.-X. (2016). Cellulose synthesis genes *CESA6* and *CSI1* are important for salt stress tolerance in *Arabidopsis*: Cellulose synthesis is important for salt stress tolerance. *J. Integr. Plant Biol.* 58, 623–626. doi: 10.1111/jipb.12442

Zhang, P., Wang, R., Yang, X., Ju, Q., Li, W., Lü, S., et al. (2020). The R2R3-MYB transcription factor ATMYB49 modulates salt tolerance in *Arabidopsis* by modulating the cuticle formation and antioxidant defence. *Plant Cell Environ.* 43, 1925–1943. doi: 10.1111/pce.13784

Zhang, M., and Zhang, S. (2022). Mitogen-activated protein kinase cascades in plant signaling. *J. Integr. Plant Biol* 64:301–341. doi: 10.1111/jipb.13215

Zhao, C., Jiang, W., Zayed, O., Liu, X., Tang, K., Nie, W., et al. (2021). The LRXs-RALFs-FER module controls plant growth and salt stress responses by modulating multiple plant hormones. *Natl. Sci. Rev.* 8, nwaa149. doi: 10.1093/nsr/nwaa149

Zhao, H., Li, Z., Wang, Y., Wang, J., Xiao, M., Liu, H., et al. (2022). Cellulose synthase-like protein OsCSLD4 plays an important role in the response of rice to salt stress by mediating abscisic acid biosynthesis to regulate osmotic stress tolerance. *Plant Biotechnol. J.* 20, 468–484. doi: 10.1111/pbi.13729

Zhao, C., Zayed, O., Yu, Z., Jiang, W., Zhu, P., Hsu, C.-C., et al. (2018). Leucine-rich repeat extensin proteins regulate plant salt tolerance in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 115, 13123–13128. doi: 10.1073/pnas.1816991115

Zhao, C., Zhang, H., Song, C., Zhu, J.-K., and Shabala, S. (2020). Mechanisms of plant responses and adaptation to soil salinity. *Innovation* 1, 100017. doi: 10.1016/j.xinn.2020.100017

Zheng, M., Liu, X., Lin, J., Liu, X., Wang, Z., Xin, M, et al. (2019). Histone acetyltransferase GCN 5 contributes to cell wall integrity and salt stress tolerance by altering the expression of cellulose synthesis genes. *Plant J* 97:587–602. doi: 10.1111/tpj.14144

Zhu, Y., and McFarlane, H. E. (2022). Regulation of cellulose synthesis *via* exocytosis and endocytosis. *Curr. Opin. Plant Biol.* 69, 102273. doi: 10.1016/j.pbi.2022.102273

Zhu, Y., Wang, Q., Wang, Y., Xu, Y., Li, J., Zhao, S., et al. (2021). Combined transcriptomic and metabolomic analysis reveals the role of phenylpropanoid biosynthesis pathway in the salt tolerance process of sophora alopecuroides. *IJMS* 22, 2399. doi: 10.3390/ijms22052399

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 10 May 2023
DOI 10.3389/fpls.2023.1095862

OPEN ACCESS

EDITED BY
Rowan Andrew Craig Mitchell,
Rothamsted Research, United Kingdom

REVIEWED BY
Dyoni M. Oliveira,
Flanders Institute for Biotechnology, Belgium
Laura E. Bartley,
Washington State University, United States

*CORRESPONDENCE
Claire Halpin
c.halpin@dundee.ac.uk
Robbie Waugh
robbie.waugh@hutton.ac.uk
Rachel A. Burton
rachel.burton@adelaide.edu.au

†PRESENT ADDRESS
Mark Looseley,
Xelect Ltd., Scotland, United Kingdom

‡These authors have contributed equally to this work

SPECIALTY SECTION
RECEIVED 11 November 2022
ACCEPTED 06 April 2023
PUBLISHED 10 May 2023

CITATION
Houston K, Learmonth A, Hassan AS, Lahnstein J, Looseley M, Little A, Waugh R, Burton RA and Halpin C (2023) Natural variation in *HvAT10* underlies grain cell wall-esterified phenolic acid content in cultivated barley.
Front. Plant Sci. 14:1095862.
doi: 10.3389/fpls.2023.1095862

Natural variation in *HvAT10* underlies grain cell wall-esterified phenolic acid content in cultivated barley

Kelly Houston[1‡], Amy Learmonth[2‡], Ali Saleh Hassan[3‡], Jelle Lahnstein[3], Mark Looseley[1†], Alan Little[3], Robbie Waugh[1,2,3*], Rachel A. Burton[3*] and Claire Halpin[2*]

[1]Cell and Molecular Sciences, The James Hutton Institute, Scotland, United Kingdom, [2]Division of Plant Sciences, School of Life Sciences, University of Dundee at The James Hutton Institute, Scotland, United Kingdom, [3]School of Agriculture, Food and Wine, University of Adelaide, Urrbrae, SA, Australia

The phenolic acids, ferulic acid and *p*-coumaric acid, are components of plant cell walls in grasses, including many of our major food crops. They have important health-promoting properties in grain, and influence the digestibility of biomass for industrial processing and livestock feed. Both phenolic acids are assumed to be critical to cell wall integrity and ferulic acid, at least, is important for cross-linking cell wall components, but the role of *p*-coumaric acid is unclear. Here we identify alleles of a BAHD *p*-coumaroyl arabinoxylan transferase, *HvAT10*, as responsible for the natural variation in cell wall-esterified phenolic acids in whole grain within a cultivated two-row spring barley panel. We show that *HvAT10* is rendered non-functional by a premature stop codon mutation in half of the genotypes in our mapping panel. This results in a dramatic reduction in grain cell wall-esterifed *p*-coumaric acid, a moderate rise in ferulic acid, and a clear increase in the ferulic acid to *p*-coumaric acid ratio. The mutation is virtually absent in wild and landrace germplasm suggesting an important function for grain arabinoxylan *p*-coumaroylation pre-domestication that is dispensable in modern agriculture. Intriguingly, we detected detrimental impacts of the mutated locus on grain quality traits where it was associated with smaller grain and poorer malting properties. *HvAT10* could be a focus for improving grain quality for malting or phenolic acid content in wholegrain foods.

KEYWORDS

barley, p-coumaric acid, ferulic acid, BAHD, HvAT10, grain, cell wall, malting

1 Introduction

Phenolic acids in the cell walls of cereals limit digestibility when grain or biomass is used for animal feed or processed to biofuels and chemicals. They are also important dietary antioxidant, anti-inflammatory and anti-carcinogenic compounds and contribute to beer flavour and aroma. The hydroxycinnamates, *p*-coumarate and ferulate (*p*CA and

FA respectively), are the major phenolic acids in grasses. Both occur as decorations ester-linked to cell wall arabinoxylan and that can be released by alkali. Lignin also has pendant alkali-releasable esterified *p*CA decorations but FA in lignin can be incorporated directly into the growing polymer by ether linkages (Hatfield et al., 2017) that are not alkali-labile. FA in the cell wall can be involved in radical coupling to cross-link arabinoxylans to each other and to lignin, and it is this cross-linking that may impede digestibility (de Oliveira et al., 2015). There is some evidence for an involvement of *p*CA in cross-linking but, in general, the role of *p*CA in cell walls is less clear (Hatfield et al., 2017). In lignin, it may promote polymerisation of sinapyl alcohol monolignols and act as a termination unit (Ralph et al., 2004; Hatfield et al., 2017) but there are no tested hypotheses about its role when attached to arabinoxylans. However, both phenolic acids have been suggested to contribute to the mechanical strength and digestion-recalcitrance of grain hull cell walls, traits that might influence dormancy or survival, the properties of grain during industrial processing, and the release of nutrients and phytochemicals during digestion. The health-promoting properties of the phenolic acids and other phenylpropanoids, including their antimicrobial, anti-inflammatory, antioxidant, anticancer activities, and their influence on gut health and microbial composition, are all areas of current study (Neelam and Sharma, 2020; Tiozon et al., 2022).

Some but not all of the genes involved in phenolic acid incorporation into grass cell walls have been identified. A specific clade of BAHD acyltransferases, the 'Mitchell' clade (Mitchell et al., 2007), is known to encode enzymes involved in *p*CA and FA addition to an unknown acceptor molecule prior to incorporation into cell wall polymers, although the exact activity of some clade members is still an open question. A complete phylogenetic analysis revealed 20 such genes in the sequenced genome of rice and named them acyltransferases (AT) AT1 to AT20 with only AT1 to AT10 being relatively highly expressed (Bartley et al., 2013). There is currently great interest and research activity into manipulating these Mitchell clade genes in transgenic plants to infer function and to manipulate phenolic acid content to improve the health-promoting properties or industrial uses of cereal crops. A specific motivation of many studies has been to improve digestibility of plant biomass for biofuel and chemical production in future biorefineries. Experimental strategies have ranged from using transgenic activation-tagged mutants ectopically overexpressing BAHD genes (Bartley et al., 2013), to BAHD RNAi suppression or CRISPR/Cas9 knockout (Marita et al., 2014; Petrik et al., 2014; Buanafina et al., 2016; de Souza et al., 2018; Mota et al., 2021; Möller et al., 2022) or overexpression in homologous or heterologous species (Buanafina et al., 2016; Karlen et al., 2016; Sibout et al., 2016; Fanelli et al., 2021; Tian et al., 2021). Results have not always been completely consistent making interpretation difficult, particularly when comparing transgenic experiments using distinct species and genes where enzyme specificity or levels and tissue-localization of gene suppression/overexpression may differ. We therefore pursued an alternative approach, determining the genetic loci underlying the natural variation in cell wall esterified phenolic acids in the grain of elite barley cultivars. This approach should not only help in identifying some of the major determinants of cell wall *p*CA and FA content in grain, but should also provide natural variants that can be exploited more immediately than transgenic approaches for breeding improved crops.

2 Results

2.1 GWAS of cell wall-esterified *p*CA and FA quantified from whole grains

We quantified cell wall-esterified *p*CA and FA in the wholegrain of a replicated GWAS panel of 211 elite 2-row spring barley cultivars grown in a field polytunnel. We observed a 6-fold variation for esterified *p*CA (54 μg/g - 327 μg/g) and a greater than 2-fold variation in esterified FA (277 μg/g -748 μg/g) (Supplementary Figures 1A, B, Supplementary Datas 1, 2). A GWAS of this data using 43,834 SNP markers identified a single highly significant association peak for grain esterified *p*CA on chromosome 7H (Figure 1A, Supplementary Data 3) and a co-locating peak for FA just below statistical significance (Figure 1B, Supplementary Data 3). The peak region above the adjusted false discovery rate (FDR) threshold for *p*CA spanned a 114.4 MB region from 461,023,389bp – 575,489,630bp. The peak for FA showed linkage between markers for a 59.0 MB region from 518,604,829bp – 577,609,887bp.

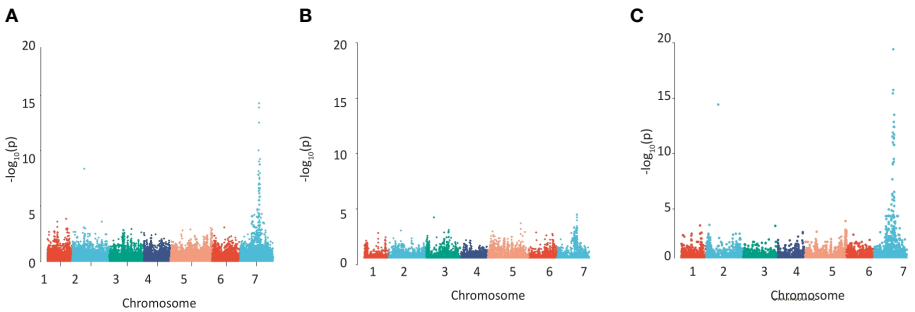


FIGURE 1
Detecting regions of the barley genome associated with grain phenolic acid content using a collection of 211 spring 2-row barleys. Manhattan plots of the GWAS of the phenolic acid content of wholegrain 2-row spring barley indicating regions of the genome associated with grain **(A)** *p*-coumaric acid, **(B)** ferulic acid content, **(C)** using the a ratio of these two phenolic acids calculated by log[FA:p-Coumaric acid]. The –Log 10 (P-value) is shown on the Y axis, and the X axis shows the 7 barley chromosomes. The FDR threshold = –log 10(P)=6.02, plots use numerical order of markers on the physical map.

10.3389/fpls.2023.1095862

Given the closeness of FA and *p*CA on the phenylpropanoid pathway we also conducted a GWAS using FA:*p*CA concentration ratios which provides internal data normalisation and reduces the inherent variability associated with measuring the single compounds (Petersen et al., 2012). Mapping FA:*p*CA ratios as log [FA:*p*CA] values increased both the strength and significance of association with the locus (i.e. the -log10(p) for *p*CA is 13.9 and for FA is 3.9 but for FA:*p*CA it rises to 19.4; Figure 1C, Supplementary Figure 2C, Supplementary Data 3). This increase in the significance of the association when the phenolic acid ratios are used instead of the single compound values confirms a level of dependency between esterified FA and esterified *p*CA concentrations. GWAS on similar data from a semi-independent set of 128 greenhouse-grown barley genotypes identified the same associations (Supplementary Figures 2A–C, Supplementary Data 3).

2.2 Identification of candidate genes on chromosome 7H

The entire peak region above the adjusted false discovery rate (FDR) threshold for the log[FA:*p*CA] values spanned a 65.7MB segment of chromosome 7H, from 459,131,547bp to 524,825,783bp and this region contains 347 high-confidence gene models. We surveyed this region for genes involved in phenolic acid or cell wall biosynthesis. This revealed several candidates including two cinnamyl alcohol dehydrogenases (*CAD*s), a caffeate-O-methyltransferase (*HvCOMT1*; Daly et al., 2019) and three BAHD acyltransferases. Interrogation of an RNA-seq dataset for 16 barley tissues (Colmsee et al., 2015) revealed that five of these six candidates showed only moderate to low levels of expression across all surveyed tissues (Figure 2A). However, the *BAHD* gene HORVU7Hr1G085100 stood out as being highly expressed in the husk lemma and palea where 80% of grain *p*CA is found (Barron et al., 2017) (Figures 2A, B). We then consulted a database of variant calls from a barley RNA-seq dataset that included 118 of our GWAS genotypes (Rapazote-Flores et al., 2019). We observed no SNP variation in two of the candidate genes. Three had one SNP each; *COMT1* (HORVU7Hr1G082280) had a synonymous SNP, one *CAD* (HORVU7Hr1G079380) had a SNP in the 3' UTR and one *BAHD* (HORVU7Hr1G085390) had a non-synonymous but rare SNP. None of these SNPs appeared likely to impair gene function. However, the *BAHD* HORVU7Hr1G085100 had 3 SNPs including one causing a premature stop codon that would lead to loss of a third of the protein sequence. BLASTp of the predicted full-length HORVU7Hr1G085100 protein sequence revealed it was 79% identical to rice *OsAT10* (LOC_Os06g39390.1), a gene previously functionally characterised as a *p*-coumaroyl CoA

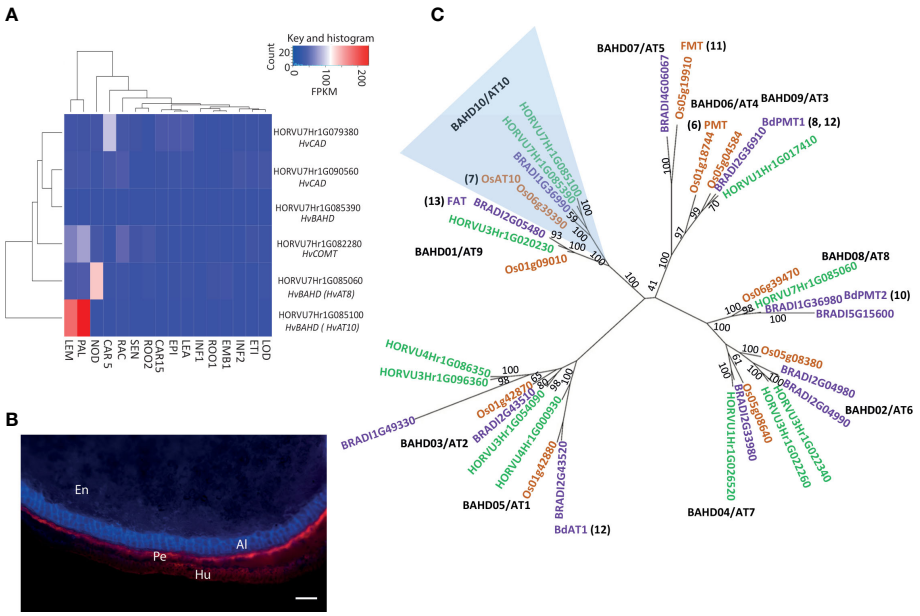


FIGURE 2
Putative candidates contributing to variation in grain *p*-coumaric (*p*CA) and ferulic acid (FA) content from GWAS. **(A)** Expression pattern for candidate genes under GWAS peak on 7H for grain *p*-coumaric and ferulic acid content in 16 different tissues/developmental stages. Tissues are (where DPA is 'days post anthesis' and pa is 'post anthesis'): LEA=Shoot (10cm seedlings), EPI=Epidermis (4 weeks), CAR15=Grain (15 DPA), ROO2=Root (4 week seedling), SEN=Senescing leaf (2 months), LOD=Lodicule (6 weeks pa), ETI=Etiolated (10 day seedlings), INF1=Inflorescence (0.5cm), EMB=Embryo (germinating), ROO1=Root (10cm seedlings), INF2=Inflorescence (1-1.5cm), RAC=Rachis (5 weeks pa), CAR5=Grain (5 DPA), NOD=Tillers (3rd internode), PAL=Palea (6 weeks pa), LEM=Lemma (6 weeks pa). Values are FPKM and a scale bar is provided. This expression data is derived from the publicly available RNAseq dataset BARLEX, https://apex.ipk-gatersleben.de/apex/f?p=284:39. **(B)** Autofluorescence in whole grain sections revealing tissue structure. En: Endosperm, Al: Aleurone, Pe: Pericarp and Hu: Husk. Scale bar = 100µm **(C)** Phylogenetic tree of the *BAHD* acyltransferases. A maximum-likelihood tree of the translation alignment of the coding sequences of group A and B *BAHD* genes from barley, rice and *Brachypodium*. Bootstrap support for branches is provided. Horvu numbers represent the barley gene models in green, BRADI represents *Brachypodium* in purple, and *Os* represents the rice genes in orange. The clade including LOC_OS06g39390 and HORVU7Hr1G085060 is highlighted in blue, *OsAT10* is indicated in green and the closest barley orthologue is marked in red. Where function of a gene model has been assigned the relevant reference is provided. Black text in bold indicates branch names, both BAHD and AT (as in de Souza et al., 2018).

arabinoxylan transferase (Bartley et al., 2013). Critically, ectopic overexpression of *OsAT10* in rice dramatically increases cell wall-esterified *p*CA levels in leaves while concomitantly reducing the levels of esterified FA (Bartley et al., 2013). A maximum likelihood phylogenetic tree of *BAHD* gene sequences confirmed HORVU7Hr1G085100 as the barley ortholog of rice *OsAT10*, i.e. *HvAT10* (Figure 2C). Another of our candidates, HORVU7Hr1G085390, is a possible *HvAT10* paralog but has negligible expression in the tissues surveyed (Figure 2A). The third *BAHD*, HORVU7Hr1G085060, is likely an AT8 (Bartley et al., 2013; de Souza et al., 2018). A gene recently suggested to act as a *p*-coumaroyl CoA arabinoxylan transferase in vegetative tissues in *Setaria viridis* (Mota et al., 2021) has a barley orthologue on chromosome 4H (AT1 clade, Figure 2C), remote from our 7H grain *p*CA/FA locus.

2.3 Discovery of causal mutation – a premature stop codon in *HvAT10*

To more accurately document polymorphisms in *HvAT10*, we PCR-sequenced the gene from 52 genotypes of the GWAS panel (Supplementary Data 1). Two nonsynonymous SNPs, one in each of *HvAT10*'s two exons (Figure 3A), were in complete linkage disequilibrium across the 52 lines. A G/A SNP at 430bp translates to either a valine or isoleucine, substituting one non-polar, neutral amino acid for another, so unlikely to affect function. By contrast, a C/A SNP at 929bp produces either serine in the full length protein, or a premature stop codon that truncates the protein by 124 amino acids, removing the BAHD family conserved DFGWG motif (DVDYG in barley and other grasses) thought to be essential for catalysis (Morales-Quintana et al., 2015) (Figure 3B). The *at10*^{STOP}

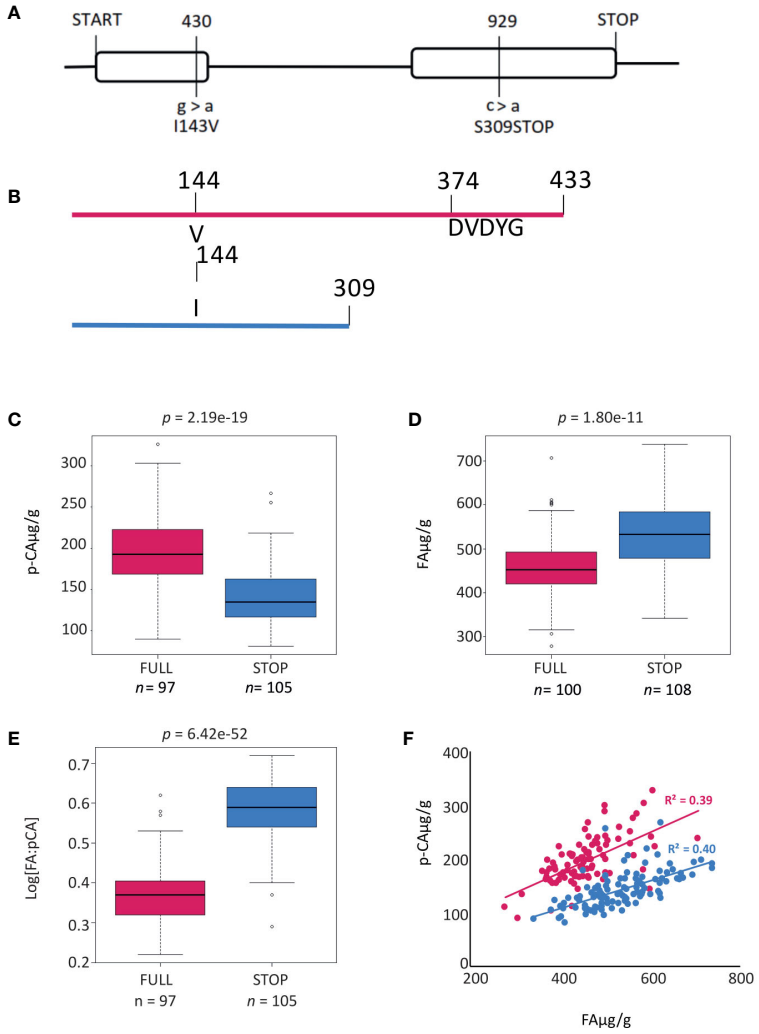


FIGURE 3
Gene and protein models for *HvAT10*. **(A)** Gene model for *HvAT10* including location, and effect of SNPs detected from resequencing this gene in the 211 barley cultivars which have been assayed for *p*-coumaric and ferulic acid. The numbering above the gene model represent locations in the CDS which vary between these cultivars. The SNP, and the resulting change in the particular amino acid are indicated underneath the gene model. The full length of the gene is 2117bp (with a CDS of 1302bp) which translates to a protein of 435 amino acids as indicated. Protein model for translation of HvAT10. **(B)** a Full length protein and **(C)** when the premature stop codon is present this results in a truncated protein. Box plots demonstrate the effect of the SNP at 929bp within *HvAT10* where the grain of the 211 barley cultivars were quantified for **(D)** *p*-coumaric acid levels and **(E)** ferulic acid levels. **(F)** Correlation between *p*CA and FA content based on *HvAT10* allele using 211 lines. The allele which results in full length version of HvAT10 are in pink, and the allele leading to a premature stop codon are coloured blue.

mutation is therefore predicted to knock-out gene function. We designed a diagnostic Kompetitive Allele Specific PCR (KASP) assay to distinguish the wildtype and mutant *HvAT10* alleles and genotyped all 212 cultivars in our GWAS population (Supplementary Table 1). Consistent with the hypothesis that $at10^{STOP}$ is the causal variant mutation underlying the FA:*p*CA ratio GWAS peak, no SNP scored higher than the KASP diagnostic when included in the GWAS although one, JHI-Hv50k-2016-488774, in complete LD, scored equally highly. *HvAT10* had a minor allele frequency of 0.48 and appears to significantly influence levels of both *p*CA (p=4.30e-19) and FA in grain (p=1.80e-11). The median grain esterified-*p*CA content for mutant $at10^{STOP}$ genotypes was 28% lower than it was for genotypes with the wildtype allele (Figure 3C), while median grain esterified-FA content was 14% higher for mutant $at10^{STOP}$ genotypes than it was for wildtype genotypes (Figure 3D). Comparing the median log[FA:*p*CA] for $at10^{STOP}$ cultivars (0.58) to the wildtype cultivar group (0.37) showed an even higher significant difference between the groups (p= 7.56e-50) (Figure 3E, Supplementary Figure 3).

2.4 Consequences of *HvAT10* premature stop codon mutation on *p*CA and FA in whole grain

Splitting the entire population of genotypes into the two allele groups (mutant and wildtype), and then plotting grain esterified *p*CA against FA within each group, revealed positive correlations between *p*CA and FA content in both allele groups (Figure 3F). This suggests that although flux into phenolic acid biosynthesis may differ between cultivars, it co-ordinately affects both phenolic acids irrespective of *Hv*AT10 allele. The $at10^{STOP}$ genotypes as a group show approximately one-third less *p*CA than wildtype genotypes (compare blue and red lines, Figure 3F), reflecting a deficiency of *p*CA on arabinoxylan in cultivars that lack a functional AT10 *p*-coumaroyl CoA arabinoxylan transferase. Nevertheless, two-thirds of cell wall esterified *p*CA remains, likely because most *p*CA is associated with lignin (Ralph, 2010; Lapierre et al., 2018) through the action of other BAHD genes. The influence of $at10^{STOP}$ on FA is evidenced by considering the 27 cultivars with grain esterified FA above 600 µg/g; 23 of these have the $at10^{STOP}$ allele (extreme right hand section of Figure 3F; Supplementary Data 1).

2.5 *HvAT10* mutation is virtually absent in wild and landrace barley

Intrigued by the prevalence of the $at10^{STOP}$ mutation in 50% of our elite barley genepool we were curious about whether this had any ecological, evolutionary, or performance-related significance. To explore this, we PCR-sequenced a collection of 114 georeferenced barley landraces, including accessions from Europe, Africa, and Asia, and 76 wild barley (*Hordeum spontaneum*) genotypes (Russell et al., 2016) across the $at10^{STOP}$ polymorphism (Supplementary Data 1). We found the $at10^{STOP}$ mutation to be extremely rare, present in only three of 114 landraces and absent in all 76 wild genotypes (Supplementary Figure 4A, Supplementary Data 1). The three $at10^{STOP}$ landraces show a clear pattern of identity by descent, clustering in the same clade of the dendrogram (Supplementary Figure 4A). We interpret these data as suggesting strong selection against the premature stop codon in wild germplasm and that $at10^{STOP}$ was a post-domestication mutation that under cultivation has no pronounced negative effects on fitness.

Several possibilities could explain enrichment of the $at10^{STOP}$ mutation in the cultivated genepool. To explore this, we first calculated the genome wide F_{ST} (fixation index) by locus using two groups based on the *HvAT10* allele. HORVU7Hr1G084140 (a serine/threonine-protein kinase not expressed in the lemma or palea) also had an F_{ST} of 1.0, and three other genes had an F_{ST} above 0.875 (Supplementary Figure 5A, B, Supplementary Data 4). Based on their functional annotations and gene expression patterns (Supplementary Data 4, Supplementary Figure 5C) we observed no obvious reason for these to be under strong selection and responsible for enhancing the frequency of the $at10^{STOP}$ mutation *via* extended LD.

2.6 Mutant locus influences grain size and malting properties

Next, due to the exclusive expression of *HvAT10* in the lemma and palea, we measured a series of grain morphometric traits across our panel. We found that, on average, grain from the $at10^{STOP}$ genotypes had significantly reduced grain width compared to cultivars with the wildtype allele (Table 1). While wildtype cultivars had an average grain width of 3.97mm, mutant $at10^{STOP}$ cultivars had a narrower width of 3.92mm and this difference was very highly significant (p=0.0009) suggesting a potential role for arabinoxylan-esterified phenolic acids in determining grain shape.

Prompted by these observations and the prevalence of registered UK barley varieties in our panel, we then explored grain parameters recorded in an extensive historical dataset from the UK's National and Recommended Lists trials 1988-2016 (Looseley et al., 2020). Different grain quality phenotypes were available for up to 107 of our cultivars. Group comparisons of wildtype and $at10^{STOP}$genotypes revealed surprising differences for hot water extract, diastatic power, germinative energy in 4ml, and wort viscosity (Table 1). In all cases, the group of $at10^{STOP}$ cultivars had poorer quality, offering no evidence of positive selection during breeding. The variation associated with the *HvAT10* locus is however highly significant and of potential interest for optimising grain quality traits (Table 1).

2.7 Pedigree and prevalence of the HvAt10 mutation in cultivated germplasm

Finally, to understand more about the origin of the $at10^{STOP}$ mutation in elite germplasm, we investigated its occurrence in the pedigree of our GWAS population. The earliest cultivar with the $at10^{STOP}$ mutation is the cultivar Kenia (a cross between the Swedish landrace Gull and the Danish landrace Binder) released in 1931 and subsequently introduced into north-west

10.3389/fpls.2023.1095862

TABLE 1 T-test results for comparisons between *HvAT10* alleles.

	Grain Area	Grain Length (mm)	Grain Width (mm)	Hot water extract l°/kg	Germinative energy 4ml %	Fermentable extract	Diastatic power (IoB)	Wort viscosity (mPa/s)	Friability %
at10^{STOP}	27,67	9,02	3,92	309,44	97,34	70,72	100,77	1,50	87,97
SD	1,48	0,37	0,10	3,39	0,16	0,38	14,18	0,04	5,23
n=	107	107	107	69	56	51	59	56	59
WT	28,16	9,09	3,97	311,23	97,40	70,88	106,07	1,48	89,95
SD	1,46	0,38	0,10	2,61	0,15	0,36	12,67	0,02	2,93
n=	96	96	96	84	74	72	75	73	73
p value	0,0214*	0,152	0,0009***	0,0005***	0,0206*	0,0203*	0,0282*	0,0009***	0,0123*

For grain area, length and width data are available in Supplementary Data 1 and analysis was carried out using BLUPS derived from 2 – years' worth of samples. Data used for comparison of hot water extract, germinative energy, fermentable extract, diastatic power, wort viscosity and friability between HvAT10 alleles are published (Xu et al., 2018). SD, standard deviation; n, number of cultivars. *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001.

European breeding programmes. Despite smaller grain and slightly poorer malting properties compared to its contemporary UK varieties (i.e. phenotypes consistent with our analyses of *at10^{STOP}* mutation cultivars as a group), it established a long-standing position as a parent for further crop improvement due to its short stiff straw, earliness and high yield (Bell, 1951). Several decades later, *at10^{STOP}* mutation-containing derivatives of Kenia, such as cv. Delta (National list 1959), were still being used as parents in our pedigree chart. We conclude that the continued prevalence of Kenia-derived germplasm may go some way to explaining the frequency of the *at10^{STOP}* mutant allele in our population.

3 Discussion

The phenolic acids FA and *p*CA in the cell walls in grasses are assumed to be critical to cell wall integrity. FA, at least, is important for cross-linking cell wall components, but the function of *p*CA is unknown. Here, we show that the natural variation for *p*CA and FA in cultivated barley grain is due to a very common knock-out mutation in a gene that links *p*CA to cell wall arabinoxylan precursors. The mutation is associated with smaller grain and poorer malting properties. Surprisingly, the mutation is completely absent in wild barley, suggesting an important function for grain arabinoxylan p-coumaroylation in the wild that is dispensable in modern agriculture.

The *HvAt10* gene that underlies this variation is part of the 'Mitchell' clade (Mitchell et al., 2007) and, in barley, its expression appears to be largely restricted to developing grain lemma and palea (Figure 2A). In rice and sugarcane also, AT10s are predominantly or only expressed in inflorescences although Brachypodium and maize show broader expression profiles (Fanelli et al., 2021). Compatible with *HvAt10* expression in barley grain, the group of *at10^{STOP}* cultivars that do not make a functional enzyme had distinct grain phenotypes i.e. narrower grain and reduced grain area compared to wildtype cultivars (Table 1). Xu et al. (2018) previously identified a QTL hotspot on chromosome 7H for traits including grain area, and grain width. The eight 9K iSelect markers defining their QTL can be positioned on the current barley physical map at 482-500MB on 7H, corresponding to the location of *HvAT10*. Wang et al. (2019) also identified a QTL for grain length:width, grain perimeter, and grain roundness at the same location. Bartley et al. (2013) noted changes in total seed mass per plant in the OsAt10-overexpressing rice mutant although Mota et al., 2021 noted no changes in BAHD RNAi-silenced lines of *Setaria viridis*. Collectively, these observations suggest a potential role for arabinoxylan-esterified phenolic acids in influencing grain development and shape, possibly through controlling stiffness and mechanical properties of grain husk. Differences in grain shape and husk mechanical properties could, in turn, explain the differences in malting properties between *at10^{STOP}* and wildtype cultivars. We can only speculate as to why *HvAt10*'s influence on grain width or hull mechanical properties might have been more important to grain survival or dispersal in wild versus cultivated genotypes, although the better germinative energy score of genotypes with the functional *HvAt10* allele (Table 1) may provide some insight.

AT10s share a conserved role in the *p*CA decoration of hemicellulose in different grasses. Initial work on the rice *OsAT10* characterised an activation-tagged mutant that ectopically overexpressed the gene resulting a dramatic 300% increase in cell wall-esterified *p*CA in arabinoxylan in young leaves and concomitant 60% reduction in esterified FA (Bartley et al., 2013). Constitutive overexpression of the rice gene in switchgrass similarly increased *p*CA in green leaves while decreasing FA (Li et al., 2018), although overexpressing the same gene in sorghum increased xylan-bound *p*CA without reducing overall FA (Tian et al., 2021). Recently, CRISPR/Cas9 knock-out of Os*AT10* in rice resulted in mutant plants with approximately 40% decrease in cell wall esterified *p*CA and essentially devoid of *p*CA associated with arabinoxylan (no change in *p*CA associated with lignin) but with some compensatory increase in grain husk arabinoxylan FA (Möller et al., 2022). Overexpression of the orthologous sugarcane Sc*AT10* in maize resulted in up to 75% increase in total *p*CA content with the increase restricted to hemicellulose, while total FA was reduced (Fanelli et al., 2021).

While interpretation of experiments where genes are massively over- and miss-expressed can be complicated, it is notable that an inverse interaction between levels of *p*CA and FA was also seen in our elite barley populations of many distinct cultivars when

genotypes with either the full length wildtype or *at10*^{STOP} mutant alleles of *HvAt10* are compared. Collectively, genotypes with the *at10*^{STOP} allele had a median *p*CA level 28% lower than those with the wildtype allele, but a median FA level 14% higher than wildtype genotypes (Figure 3C). This effect on FA might occur in several ways: *p*CA that cannot be esterified onto arabinoxylan could be methoxylated to produce FA thereby increasing FA pools for transfer onto arabinoxylan, or alternatively, *p*CA and FA may compete for transfer onto a shared acceptor (likely UDP-arabinose) before incorporation into arabinoxylan such that loss of *p*CA transfer by *HvAT10* leaves more free acceptor for FA transfer. Either mechanism could explain how *at10*^{STOP} can indirectly increase grain cell wall esterified FA.

Since both phenolic acids are altered by *HvAt10* knock-out, we cannot be conclusive about which cell wall changes underly the grain shape and malting phenotypes – these could be due to changes in either *p*CA or FA content or to the ratio between them. It is notable in this context that FA dimerization in the cell wall has been proposed to increase crosslinks between arabinoxylan and lignin, thereby increasing wall rigidity and potentially limiting growth (Ishii, 1997; Hatfield et al., 2017).

Taken together, we conclude that the prevalence of the *at10*^{STOP} mutant allele in our elite barley populations may simply be a straightforward genetic legacy of historical barley breeding and use of Kenia-derived germplasm. However, our data suggests that purging the *at10*^{STOP} allele could assist the development of superior quality malting barley varieties. Conversely, much research has focussed on the beneficial bioactivity of FA in the diet, and use of the *at10*^{STOP} allele could facilitate breeding for increased FA in wholegrain products.

4 Methods

4.1 Plant material and growth conditions

Two populations of 2-row spring type barley were used to carry out the GWAS. The first population includes 211 elite lines grown in a polytunnel under field conditions in Dundee, Scotland in 2011 (Oakey et al., 2013). For each line, 5 whole grains were ground to a fine powder using a ball mill (Mixer Mill MM400; Retsch Haan Germany) and stored in dry conditions until the HPLC analysis. The second population which was used for verification of the results of the analysis of the first subpopulation includes 128 elite lines grown in a glasshouse compartment in a mix of clay-loam and cocopeat (50:50 v/v) at daytime and night time temperatures of 22° C and 15°C respectively in The Plant Accelerator, Adelaide, Australia (Hassan et al., 2017). The collection of germplasm these populations are sampled from has minimum population structure while maintaining as much genetic diversity as possible. Mature grains were stored until phenolic acid content analysis.

4.2 Genotyping with SNP markers

All lines were genotyped using the 50K iSelect SNP genotyping platform described previously (Bayer et al., 2017). Prior to marker-trait association analysis, all markers with a minimum allele frequency of <5% and markers with missing data >5% were excluded from the analysis.

4.3 Phenotyping for cell wall-bound phenolic acids

A ~ 20 mg amount of wholegrain barley (i.e. including husk) was used per sample. *Trans*-ferulic and *trans-p*-coumaric acid standards were purchased from SIGMA Aldrich (Castle Hill NSW, Australia). Standards were prepared at 62.5 µm, 250 µm and 1000 µm by dissolving the appropriate amount of powder in 50% methanol. Extraction of cell wall esterified phenolic acids was carried out following a previously described method (Irakli et al., 2012) with the following modifications. Samples were washed twice with 500 µl 80% ethyl alcohol, with shaking for 10 minutes at room temperature to remove free phenolic acids. To release total cell wall esterified phenolic acids, alkaline treatment was carried out by adding 600 µl 2M NaOH to the pellet. Samples were incubated on a rotary rack under nitrogen for 20 h in the dark at room temperature. Samples were centrifuged at 15000 x g for 15 minutes at room temperature, after which the supernatant was collected, acidified by adding 110 µl concentrated HCL and extracted three times with 1 mL ethyl acetate. Following each extraction, samples were centrifuged at 5000 x g for 7 minutes and the organic solution was collected. Extracts were combined, evaporated to dryness in a rotary evaporator and dissolved in 100 µl of 50% methanol prior to injecting 40 µl into the HPLC column. For each sample two technical replicates were applied.

4.4 HPLC conditions

An Agilent Technologies 1260 Infinity HPLC equipped with a Diode Array detector was used. Samples were analysed on an Agilent Poroshell 120 SB-C18 3.0x100mm 2.7- micron column kept at 30 C°. Eluents were A (0.5mM trifluoroacetic acid) and B (0.5mM trifluoroacetic acid, 40% methanol, 40% acetonitrile, 10% water). Starting conditions were 85% A and 15% B. Flow rate was 0. 7 mL/min. Eluting gradients were as follow; min 0-10: 15% to 55% B, min 11-12: column washed with 100% B, min 13 back to the starting condition (85% A and 15% B). Detection was carried out at 280 nm and spectral data was collected from 200 to 400 nm when required. Ferulic and *p*-coumaric acid peaks were identified by comparing retention times and spectra to their corresponding standards. The area under the peaks was quantified at 280 nm for *trans* forms.

4.5 GWAS analysis of grain alkaline extractable *p*CA and FA and FA:*p*CA ratio

Marker- trait association analysis was carried out using R 2.15.3 (www.R-project.org) and performed with a compressed mixed linear model (Zhang et al., 2010) implemented in the GAPIT R package. For phenotype values, the mean values of the barley wholegrain total alkaline extractable *trans*-ferulic and *trans- p-*

coumaric acid (w/w) were used. To identify genes within intervals associated with our trait we used BARLEYMAP (Cantalapiedra et al., 2015). We also used the ratio of FA:*p*CA as a trait in our GWAS analysis. The ratio between the two compounds was log transformed i.e. log(FA:*p*CA) to provide a more normally distributed dataset. When using ratios in GWAS, a significant increase in the *p*-gain statistic (a comparison between the lowest -log10(p) values of the individual compounds and the -log10(p) value of the ratio; Petersen et al., 2012) indicates that ratios carry more information than the corresponding metabolite concentrations alone. A significant p-gain identifies a biologically meaningful association between the individual compounds. We used B/(2*α) to derive a critical value of $3.42x10^5$ for the FDR-adjusted p-gain, where α is the level of significance (0.05) and B the number of tested metabolite pairs (Petersen et al., 2012). Therefore, as we tested two traits our threshold was 2×10^1 and our p-gain was above this threshold.

To identify local blocks of LD, facilitating a more precise delimitation of QTL regions Linkage disequilibrium (LD) was calculated across the genome between pairs of markers using a sliding window of 500 markers and a threshold of $R^2<0.2$ using Tassel v 5. We anchored markers that passed FDR and represented initial borders of the QTL on 7H to the physical map and then expanded this region using local LD derived from genome wide LD analysis as described above. When the GWAS had not resulted in an association that passed the FDR we used the arbitrary threshold of -LOG10(P) to define the initial border. The SNP with the highest LOD score was used to represent the QTL. After identification and Sanger sequencing of the candidate gene *HvAT10* the GWAS was repeated including the allele present at the S309Stop as an additional marker. The physical map positions and gene model names used in this analysis are from Morex v1 (Mayer et al., 2012, Mascher et al., 2017).

4.6 Bioinformatics and gene identification

We used BARLEX (Colmsee et al., 2015) to identify gene models present with the QTL defined by our analysis and their expression profile based on RNAseq data in 16 different tissues/developmental stages.

4.7 Phylogenetic analysis of barley BAHD acyltransferases

Coding sequences of all BAHD acyltransferases with the PFAM domain PF02458 from rice, barley and *Brachypodium* were downloaded from the Ensembl Plants database (http://plants.ensembl.org/). Sequences were aligned using the MUSCLE alignment function available in the Geneious 9.1.4 (https://www.geneious.com). The translation alignment option was used. A neighbour-joining tree was produced from the alignment. Barley genes within group A and B clades were identified, realigned with their rice and *Brachypodium* orthologs and a maximum likelihood tree was produced from the translation alignment of the sequences. The following settings were applied: substitution model: General-Time-Reversible (GTR), branch support: bootstrap, number of bootstrap: 1000.

4.8 Resequencing and genotyping of *HvAT10* in the main and supplemental set

Aligning the translation of AK376450 to Os06g39390 allowed the identification of the putative genomic sequence of *HvAT10*. We designed four pairs of primers, details of sequences and reaction conditions are in Supplementary Data 5, to amplify the full length CDS. To facilitate quick and efficient genotyping of large numbers of cultivars we subsequently designed a KASP genotyping assay to a SNP at 430bp in *HvAT10* (Supplementary Data 5) (Semagn et al., 2014). Reactions were performed in an 8.1 µL reaction volume, with 3 µL H2O, 1 µL DNA (20ng/µl), 4 µL KASP genotyping master mix, and 0.11 µL of the KASP assay.

Box plots to demonstrate the contribution of the SNP at 436bp in *HvAT10* to variation in grain *p*CA and FA content were produced using R 2.15.3 (www.R-project.org). To test for identity by descent of the *HvAT10* allele within the set of accessions using for the GWAS a dendrogram was constructed using maximum likelihood using the genotypic data from the 9k-select array in MEGA7 with default settings except for including bootstrapping and visualised in FigTree (v.1.4.4) http://tree.bio.ed.ac.uk/software/figtree/.

4.9 Characterization of diversity of *HvAT10* in *Hordeum spontaneum* from the fertile crescent and barley landraces

DNA was extracted as described above from 76 *H. spontaneum* and 114 barley landraces from (Russell et al., 2016). The S309Stop SNP was PCR amplified and Sanger sequenced with primer pair 5 using conditions described above. A dendrogram was constructed using maximum likelihood using 4000 exome capture derived SNPs from Russell et al. (2016) in MEGA7 with default settings except for including bootstrapping and visualised in FigTree (v.1.4.4) http://tree.bio.ed.ac.uk/software/figtree/.

4.10 Genome wide F_{ST} analysis

The fixation index (FST) is a measure of genetic differentiation between groups of individuals. Genome wide FST was calculated by locus using GenAlEx 6.502 after dividing the accessions into two populations based on their HvAT10 allele using all informative 50K iSelect markers.

4.11 Phenotypic analysis of cultivars with wildtype vs $at10^{STOP}$ allele

We characterised mature grain morphology using from plants grown in a polytunnel under field conditions in Dundee, Scotland as described above, over two years (2010 and 2011). Grain area, width and length were quantified using the MARVIN Seed Analyzer (GTA Sensorik GmbH, 2013). BLUPs (Best Linear Unbiased Predictions) calculated from this data using R 2.15.3 (www.R-project.org) were used in subsequent comparisons between allelic groups.

Data availability statement

The data presented in the study are deposited in the NCBI repository, accession numbers OQ320054 - OQ320323..

Author contributions

RW, KH, RB, CH, ALi, designed experiments. KH, AH, JL, ALe, carried out experiments. KH, AH, ALe, ML, ALi, JL analyzed data. The manuscript was written by CH, KH, RW, RB, ALe, AH with contributions from all other authors. All authors contributed to the article and approved the submitted version.

Funding

KH and RW acknowledge support from the Rural & Environment Science & Analytical Services Division of the Scottish Government and BBSRC (BB/J014869/1, BB/L026317/1). CH and RW are grateful for funding from BBSRC (BB/N023455/1, BB/G016232/1 and BB/P025951/1). ALe was supported by BBSRC grant BB/J01446X/1. ML and KH acknowledge funding from BB/K008188/1. RB, AH, and ALi acknowledge funding provided by the Australian Research Council Centre of Excellence in Plant Cell Walls CE110001007.

Acknowledgments

We would also like to acknowledge technical support from Malcolm Macaulay and advice on analysis from Helen Oakey.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1095862/full#supplementary-material

SUPPLEMENTARY FIGURE 1

Phenolic acid content of wholegrain flour from 211 2-row spring barleys linea. **(A)** *p*CA and **(B)** ferulic acid content. Values represent the mean for FA and pCA expressed as w/w. Error bars represent standard deviation of the replicates.

SUPPLEMENTARY FIGURE 2

Manhattan plots of the GWAS of the phenolic acid content of wholegrain flour from 128 2-row spring barley lines indicating regions of the genome associated with grain phenolic acid content. Manhattan plots of the GWAS of the phenolic acid content of wholegrain 2-row spring barley indicating regions of the genome associated with grain a. *p*-coumaric acid, b. ferulic acid content and c. log[FA:p-Coumaric acid]. The −Log 10 (P-value) is shown on the Y axis, and the X axis shows the 7 barley chromosomes. FDR threshold = −log 10(P)=6.02, plots use numerical order of markers on the physical map.

SUPPLEMENTARY FIGURE 3

Distribution of ratio between two phenolic acids quantified in the grain of 211 spring 2 row barleys lines and used to carry out GWAS. The ratio was calculated as log[FA:p-Coumaric acid]. Accessions containing the allele which results in a full length version of *HvAT10* are in pink, and accessions containing the allele leading to a premature stop codon are coloured blue.

SUPPLEMENTARY FIGURE 4

Distribution of the *HvAT10* premature stop codon in *H. vulgare* landraces and cultivated barley lines. **(A)** A dendrogram of 114 *H. vulgare* landraces constructed using a selection of SNPs with a genome-wide distribution with maximum likelihood methods. **(B)** A dendrogram of cultivated barley germplasm using a selection of SNPs with a genome-wide distribution using maximum likelihood methods. Accessions containing the allele which results in full length version of HvAT10 are in pink, and accessions containing the allele leading to a premature stop codon are coloured blue.

SUPPLEMENTARY FIGURE 5

F_{ST} analysis based on *HvAT10*. **(A)** Plot displaying genome wide F_{ST} with F_{ST} index provided on the Y axis, an F_{ST} of 1 indicating a complete fixation of each allele within the two subpopulations determined by their allele of *HvAT10*. **(B)** Just F_{ST} of markers at 7H. Red box indicates location of the centromere. Two SNPs whose location overlap on this plot, including one in *HvAT10*, have an F_{ST} of 1.0. Note shape of peak appears different in **(A, B)** due to the difference in scale of the plots. **(C)** RNAseq data for genes with F_{ST}>0.875 from 16 different tissues/developmental stages. Values are FPKM and a scale bar is provided. This expression data is derived from the publicly available RNAseq dataset BARLEX, https://apex.ipk-gatersleben.de/apex/f?p=284:39. Tissue abbreviations as in main paper text.

SUPPLEMENTARY DATA SHEET 1

Phenolic acid and genetic data for all cultivars included in this study. *p*-coumaric and ferulic acid content, KASP data and NCBI number for those lines that where sequenced for *HvAT10* is included.

SUPPLEMENTARY DATA SHEET 2

Summary of number of accessions used for each GWAS. A total of 211 accessions were included in the main dataset but data for both phenolic acids is not available for all lines, therefore the number of individuals included in different analysis varies. Includes number of accessions for the GWAS presented in the main and supplementary analysis for both individual trait and the ratio analysis. Number of individuals with each allele of *HvAT10* based on genotyping of A430G is also included.

SUPPLEMENTARY DATA SHEET 3

Details of QTL identified on 7H for all analysis carried out. Physical location, LOD score, and 50k iSelect marker with the highest LOD score are provided. * indicates that analysis passed the FDR threshold of -log10(p)=6.1.

SUPPLEMENTARY DATA SHEET 4

Gene models containing SNPs that have an F_{ST}>0.875 when F_{ST} analysis carried out based on *HvAT10* allele. This table includes 50k iSelect marker name, the chromosome the marker is located on, gene model and annotation based on Morex v1 Gene Models (2016).

SUPPLEMENTARY DATA SHEET 5

Details of primers and genotyping assays used in this study. This includes details of primers for Sanger sequencing and KASP genotyping assay sequence for *HvAT10*.

References

Barron, C., Holopainen-Mantila, U., Sahlstrom, S., Hotekjolen, A. K., and Lullien-Pellerin, V. (2017). Assessment of biochemical markers identified in wheat for monitoring barley grain tissue. *J. Cereal Sci.* 74, 11–18. doi: 10.1016/j.jcs.2017.01.004

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

Bayer, M. M., Rapazote-Flores, P., Ganal, M., Hedley, P. E., Macaulay, M., Plieske, J., et al. (2017). Development and evaluation of a barley 50k iSelect SNP array. *Front. Plant Sci.* 8, 1792. doi: 10.3389/fpls.2017.01792

Bell, G. D. H. (1951). Barley breeding and related researches. *J. Institute Brewing* 57, 247–260. doi: 10.1002/j.2050-0416.1951.tb01625.x

Buanafina, M. M. D., Fescemyer, H. W., Sharma, M., and Shearer, E. A. (2016). Functional testing of a PF02458 homologue of putative rice arabinoxylan feruloyl transferase genes in brachypodium distachyon. *Planta* 243, 659–674. doi: 10.1007/s00425-015-2430-1

Cantalapiedra, C. P., Boudiar, R., Boudiar, R., Casas, A. M., Igartua, E., Contreras-Moreira, B., et al. (2015). BARLEYMAP: physical and genetic mapping of nucleotide sequences and annotation of surrounding loci in barley. *Mol. Breed.* 35, 13. doi: 10.1007/s11032-015-0253-1

Colmsee, C., Beier, S., Himmelbach, A., Schmutzer, T., Stein, N., Scholz, U., et al. (2015). BARLEX - the barley draft genome explorer. *Mol. Plant* 8 (6), 964–966. doi: 10.1016/j.molp.2015.03.009

Daly, P., McClellan, C., Maluk, M., Oakey, H., Lapierre, C., Waugh, R., et al. (2019). RNAi-suppression of barley caffeic acid O-methyltransferase modifies lignin despite redundancy in the gene family. *Plant Biotechnol. J.* 17, 594–607. doi: 10.1111/pbi.13001

de Oliveira, D. M., Finger-Teixeira, A., Mota, T. R., Salvador, V. H., Moreira-Vilar, F. C., Molinari, H. B. C., et al. (2015). Ferulic acid: a key component in grass lignocellulose recalcitrance to hydrolysis. *Plant Biotechnol. J.* 13, 1224–1232. doi: 10.1111/pbi.12292

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

Fanelli, A., Rancour, D. M., Sullivan, M., Karlen, S. D., Ralph, J., Riano-Pachon, D. M., et al. (2021). Overexpression of a sugarcane BAHD acyltransferase alters hydroxycinnamate content in maize cell wall. *Front. Plant Sci.* 12, 626168. doi: 10.3389/fpls.2021.626168

Hassan, A. S., Houston, K., Lahnstein, J., Shirley, N., Schwerdt, J. G., Gidley, M. J., et al. (2017). A genome wide association study of arabinoxylan content in 2-row spring barley grain. *PloS One* 12, e0182537. doi: 10.1371/journal.pone.0182537

Hatfield, R. D., Rancour, D. M., and Marita, J. M. (2017). Grass cell walls: a story of cross-linking. *Front. Plant Sci.* 7, 2056. doi: 10.3389/fpls.2016.02056

Irakli, M. N., Samanidou, V. F., Biliaderis, C. G., and Papadoyannis, I. N. (2012). Development and validation of an HPLC-method for determination of free and bound phenolic acids in cereals after solid-phase extraction. *Food Chem.* 134, 1624–1632. doi: 10.1016/j.foodchem.2012.03.046

Ishii, T. (1997). Structure and functions of feruloylated polysaccharides. *Plant Sci.* 127, 111–127. doi: 10.1016/S0168-9452(97)00130-1

Karlen, S. D., Zhang, C. C., Peck, M. L., Smith, R. A., Padmakshan, D., Helmich, K. E., et al. (2016). Monolignol ferulate conjugates are naturally incorporated into plant lignins. *Sci. Adv.* 2, 1–9. doi: 10.1126/sciadv.1600393

Lapierre, C., Voxeur, A., Karlen, S. D., Helm, R. F., and Ralph, J. (2018). Evaluation of feruloylated and p-Coumaroylated arabinosyl units in grass arabinoxylans by acidolysis in Dioxane/Methanol. *J. Agric. Food Chem.* 66, 5418–5424. doi: 10.1021/acs.jafc.8b01618

Li, G. T., 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

Looseley, M., Ramsay, L., Bull, H., Swanston, J. S., Shaw, P. D., Macaulay, M., et al. (2020). Association mapping of malting quality traits in UK spring and winter barley cultivar collections. *Theor. Appl. Genet.* 133, 2567–2582. doi: 10.1007/s00122-020-03618-9

Marita, J. M., Hatfield, R. D., Rancour, D. M., and Frost, K. E. (2014). Identification and suppression of the p-coumaroyl CoA: hydroxycinnamyl alcohol transferase in zea mays l. *Plant J.* 78, 850–864. doi: 10.1111/tpj.12510

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., et al. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature* 544, 426–433. doi: 10.1038/nature22043

Mayer, K. F. X., Waugh, R., Langridge, P., Close, T. J., Wise, R. P., Graner, A., et al. (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491, 711–716. doi: 10.1038/nature11543

Mitchell, R. A. C., Dupree, P., and Shewry, P. R. (2007). A novel bioinformatics approach identifies candidate genes for the synthesis and feruloylation of arabinoxylan. *Plant Physiol.* 144, 43–53. doi: 10.1104/pp.106.094995

Möller, S. R., Lancefield, C. S., Oates, N. C., Simister, R., Dowle, A., Gomez, L. D., et al. (2022). CRISPR/Cas9 suppression of OsAT10, a rice BAHD acyltransferase, reduces p-coumaric acid incorporation into arabinoxylan without increasing saccharification. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.926300

Morales-Quintana, L., Alejandra Moya-Leon, M., and Herrera, R. (2015). Computational study enlightens the structural role of the alcohol acyltransferase DFGWG motif. *J. Mol. Model.* 21, 216. doi: 10.1007/s00894-015-2762-6

Mota, T. R., de Souza, W. R., Oliveira, D. M., Martins, P. K., Sampaio, B. L., Vinecky, F., et al. (2021). Suppression of a BAHD acyltransferase decreases p-coumaroyl on arabinoxylan and improves biomass digestibility in the model grass setaria viridis. *Plant J.* 105, 136–150. doi: 10.1111/tpj.15046

Neelam, K. A., and Sharma, K. K. (2020). Phenylpropanoids and its derivatives: biological activities and its role in food, pharmaceutical and cosmetic industries. *Crit. Rev. Food Sci. Nutr.* 60, 2655–2675. doi: 10.1080/10408398.2019.1653822

Oakey, H., Shafiei, R., Comadran, J., Uzrek, N., Cullis, B., Gomez, L. D., et al. (2013). Identification of crop cultivars with consistently high lignocellulosic sugar release requires the use of appropriate statistical design and modelling. *Biotechnol. Biofuels* 6, 185. doi: 10.1186/1754-6834-6-185

Petersen, A. K., Krumsiek, J., Wagele, B., Theis, F. J., Wichmann, H. E., Gieger, C., et al. (2012). On the hypothesis-free testing of metabolite ratios in genome-wide and metabolome-wide association studies. *BMC Bioinf.* 13, 120. doi: 10.1186/1471-2105-13-120

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

Ralph, J. (2010). Hydroxycinnamates in lignification. *Phytochem. Rev.* 9, 65–83. doi: 10.1007/s11101-009-9141-9

Ralph, J., Bunzel, M., Marita, J. M., Hatfield, R. D., Lu, F., Kim, H., et al. (2004). Peroxidase-dependent cross-linking reactions of *p*-hydroxycinnamates in plant cell walls. *Phytochem. Rev.* 3, 79–96. doi: 10.1023/B:PHYT.0000047811.13837.fb

Rapazote-Flores, P., Bayer, M., Milne, L., Mayer, C. D., Fuller, J., Guo, W. B., et al. (2019). BaRTv1.0: an improved barley reference transcript dataset to determine accurate changes in the barley transcriptome using RNA-seq. *BMC Genomics* 20, 968. doi: 10.1186/s12864-019-6243-7

Russell, J., Mascher, M., Dawson, I. K., Kyriakidis, S., Calixto, C., Freund, F., et al. (2016). Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat. Genet.* 48, 1024–1030. doi: 10.1038/ng.3612

Semagn, K., Babu, R., Hearne, S., and Olsen, M. (2014). Single nucleotide polymorphism genotyping using kompetitive allele specific PCR (KASP): overview of the technology and its application in crop improvement. *Mol. Breed* 33, 1–14. doi: 10.1007/s11032-013-9917-x

Sibout, R., Le Bris, P., Legee, F., Cezard, L., Renault, H., and Lapierre, C. (2016). Structural redesigning arabidopsis lignins into alkali-soluble lignins through the expression of p-coumaroyl-coA: monolignol transferase PMT. *Plant Physiol.* 170, 1358–1366. doi: 10.1104/pp.15.01877

Tian, Y., CY, L., Park, J. H., Wu, C. Y., Kakumanu, R., Pidatala, V. R., et al. (2021). Overexpression of the rice BAHD acyltransferase AT10 increases xylan-bound p-coumarate and reduces lignin in sorghum bicolor. *Biotechnol. Biofuel* 14, 217. doi: 10.1186/s13068-021-02068-9

Tiozon, R. J. N., Sartagoda, K. J. D., Serrano, L. M. N., Fernie, A. R., and Sreenivasulu, N. (2022). Metabolomics based inferences to unravel phenolic compound diversity in cereals and its implications for human gut health. *Trends Food Sci. Technol.* 127, 14–25. doi: 10.1016/j.tifs.2022.06.011

Wang, Q., Sun, G. L., Ren, X. F., Du, B. B., Cheng, Y., Wang, Y. X., et al. (2019). Dissecting the genetic basis of grain size and weight in barley (Hordeum vulgare l.) by QTL and comparative genetic analyses. *Front. Plant Sci.* 10, 469. doi: 10.3389/fpls.2019.00469

Xu, X., Sharma, R., Tondelli, A., Russell, J., Comadran, J., Schnaithmann, F., et al. (2018). Genome-wide association analysis of grain yield-associated traits in a pan-European barley cultivar collection. *Plant Genome* 11, 170073. doi: 10.3835/plantgenome2017.08.0073

Zhang, Z. W., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 06 June 2023
DOI 10.3389/fpls.2023.1155188

OPEN ACCESS

EDITED BY
Wagner Rodrigo De Souza,
Federal University of ABC, Brazil

REVIEWED BY
Jozef Mravec,
University of Copenhagen, Denmark
Jacek Zebrowski,
University of Rzeszow, Poland

*CORRESPONDENCE
Maurice Bosch
mub@aber.ac.uk

RECEIVED 31 January 2023
ACCEPTED 05 May 2023
PUBLISHED 06 June 2023

CITATION
Iacono R, Slavov GT, Davey CL, Clifton-Brown J, Allison G and Bosch M (2023) Variability of cell wall recalcitrance and composition in genotypes of *Miscanthus* from different genetic groups and geographical origin. *Front. Plant Sci.* 14:1155188. doi: 10.3389/fpls.2023.1155188

Variability of cell wall recalcitrance and composition in genotypes of *Miscanthus* from different genetic groups and geographical origin

Rosario Iacono[1], Gancho T. Slavov[1,2], Christopher L. Davey[1], John Clifton-Brown[1,3], Gordon Allison[1] and Maurice Bosch[1]*

[1]Institute of Biological Environmental and Rural Sciences (IBERS), Aberystwyth University, Gogerddan, Aberystwyth, United Kingdom, [2]Radiata Pine Breeding Company, Rotorua, New Zealand, [3]Department of Agronomy and Plant Breeding, Justus Liebig University Giessen, Giessen, Germany

Miscanthus is a promising crop for bioenergy and biorefining in Europe. The improvement of *Miscanthus* as a crop relies on the creation of new varieties through the hybridization of germplasm collected in the wild with genetic variation and suitable characteristics in terms of resilience, yield and quality of the biomass. Local adaptation has likely shaped genetic variation for these characteristics and is therefore important to quantify. A key biomass quality parameter for biorefining is the ease of conversion of cell wall polysaccharides to monomeric sugars. Thus far, the variability of cell wall related traits in *Miscanthus* has mostly been explored in accessions from limited genetic backgrounds. Here we analysed the soil and climatic conditions of the original collection sites of 592 *Miscanthus* genotypes, which form eight distinct genetic groups based on discriminant analysis of principal components of 25,014 single-nucleotide polymorphisms. Our results show that species of the genus *Miscanthus* grow naturally across a range of soil and climate conditions. Based on a detailed analysis of 49 representative genotypes, we report generally minor differences in cell wall characteristics between different genetic groups and high levels of genetic variation within groups, with less investigated species like *M. floridulus* showing lower recalcitrance compared to the other genetic groups. The results emphasize that both inter- and intra- specific variation in cell wall characteristics and biomass recalcitrance can be used effectively in *Miscanthus* breeding programmes, while also reinforcing the importance of considering biomass yield when quantifying overall conversion efficiency. Thus, in addition to reflecting the complexity of the interactions between compositional and structural cell wall features and cell wall recalcitrance to sugar release, our results point to traits that could potentially require attention in breeding programmes targeted at improving the *Miscanthus* biomass crop.

KEYWORDS

biomass, cell wall, genetic diversity, geographical distribution, germplasm, *Miscanthus*, pedoclimatic conditions, recalcitrance (saccharification)

1 Introduction

Mitigating the effects of climate change while coping with shortages of fossil fuels is the challenge that defines our time and calls for shifting to renewable sources of fuel and chemicals. Plant biomass is an abundant source of building blocks for fuel and chemicals, thus having the potential to replace fossil feedstock for their production. Facilities dedicated to processing biomass to produce fuel, electricity, and chemicals are defined as biorefineries. However, to become a realistic alternative to petrol refineries, biorefineries require a constant supply of feedstock with optimal characteristics for the specific conversion processes (Hoang et al., 2015; Torres et al., 2019; Oyedeji et al., 2021; Brancourt-Hulmel et al., 2022).

Perennial biomass crops, ideally grown on lands not used for food production (e.g. marginal lands), represent attractive feedstocks for biorefining (Kang et al., 2013; Von Cossel et al., 2019; Yang et al., 2019; Pancaldi and Trindade, 2020). Perennial rhizomatous grass species of the genus *Miscanthus*, exhibiting C4 photosynthesis and widely distributed in East Asia with 16 species described in the wild (The Plant List, 2013), are one of the best candidates for the implementation of these cropping systems in Europe (Clifton-Brown et al., 2019a; Clifton-Brown et al., 2019b).

The cell wall, rich in sugar and aromatic molecules is the main energy and carbon sink in nature incorporating 45% of the fixed carbon (Barnes and Anderson, 2018; Tang et al., 2018). It is the main component of plant biomass and the one with the highest potential for producing chemicals and fuel (Albersheim et al., 2011). Its structure and composition depend on the organ considered (da Costa et al., 2017), the phenological stage (Rancour et al., 2012) as well as on the genetic background of the plant (da Costa et al., 2017). Importantly, being the dynamic interface between the plant and the external environment, the composition and structure of the cell wall can change in response to a plant's exposure to external stressors (Moura et al., 2010; Domon et al., 2013; Le Gall et al., 2015; Vaahtera et al., 2019; Gladala-Kostarz et al., 2020). The interplay between the maintenance of cell wall integrity and reaction to environmental stresses has been reviewed recently by Baez et al. (2022).

A key biomass quality parameter for biorefining is the ease by which cellulosic and hemicellulosic cell wall polysaccharides can be converted to monomeric sugars (DeMartini et al., 2013). The complex and dynamic structure of the cell wall has evolved to maintain its functional integrity in response to developmental and environmental cues. This has resulted in the natural recalcitrance of the cell wall to deconstruction (McCann and Carpita, 2015), which hinders the profitable use of biomass for the production of fuel and chemicals (DeMartini et al., 2013).

Miscanthus species can intercross and produce sterile hybrids (Tamura et al., 2016) and the use of hybridization between selected lines has been proposed as a suitable breeding technique to develop *Miscanthus* into a specialized bioenergy and biorefining crop (Lewandowski et al., 2016). In the last decade, this approach led to the development of new hybrids (Kalinina et al., 2017; Clifton-Brown et al., 2019a; Clifton-Brown et al., 2019b) commercialized by companies such as Terravesta (Lincoln, UK). The hybridization technique to create new hybrids relies on the existence of a sufficient range of intra- and inter-specific variability in the target traits and on the ability to identify the most promising parental lines expressing them. However, the extent of variability for cell wall traits in the genus *Miscanthus* has mostly been explored in the natural hybrid *M.* × *giganteus* and its parent species *M. sacchariflorus* and *M. sinensis* (Lygin et al., 2011; van der Weijde et al., 2017b; da Costa et al., 2019; Bilska-Kos et al., 2022). There is little information on the variation in cell wall related traits between *Miscanthus* genotypes from a wider range of different genetic groups.

The process of selecting parental lines for the production of hybrids starts with the collection of promising germplasm in the wild. This phase is time-consuming and expensive. Notably, the structure and composition of the cell wall are presumably adapted to the environmental conditions where the genotype was originally collected. Thus, geographical information plays an important role in the process of selection of germplasm for the creation of varieties with a superior biomass quality for conversion. For instance, Li et al. (2016) found that there is a relation between the geographical area where a *Miscanthus* accession was sampled and the amount of glucose released with enzymatic digestion after acid or alkali pretreatment.

In this context, spatial analysis using geographic information system (GIS) technology (Hyman et al., 2013) can assist plant breeding by uncovering the environmental associations of germplasm across a wide range of collection sites (Hijmans and Spooner, 2001). Breeders routinely use spatial analysis (i.e., either explicitly or implicitly) to inform the decision about where to test and disseminate crop varieties (Oshunsanya and Aliku, 2016) and to identify interesting starting materials such as for the resilience to abiotic stress in wild accessions (Hijmans and Spooner, 2001). For example, Malinowska et al. (2017) found that a model based on the precipitation and temperature conditions in the area of origin of *Miscanthus* genotypes can predict their resistance to drought.

Here, we focused on *Miscanthus* genotypes grown in a spaced field trial and used analysis of single-nucleotide polymorphism (SNP) data and the pedo-climatic conditions of their original collection sites to select representatives belonging to broadly diverse and distinct genetic groups. We demonstrate the value of using environmental data to identify sites where germplasm of the genus *Miscanthus* with suitable characteristics can be found, with the potential to assist breeding programmes. We report differences in cell wall characteristics between and within the different genetic groups of *Miscanthus*, including for recalcitrance to enzymatic sugar release, identify compositional and structural features that correlate with cell wall recalcitrance and discuss the practical implications of our findings.

2 Materials and methods

2.1 Plant material, genetic groups and sample preparation

2.1.1 Experimental field and origin of *Miscanthus* genotypes

Genotypes of *Miscanthus* were selected from the ABR33 replicated field trial established between 2012 and 2014 near

Aberystwyth (Wales, UK; 52° 25' 57.7" N - 4° 01' 33.2" W). Briefly, 953 *Miscanthus* accessions (i.e., genotypes), which had previously been brought into and grown across Europe, were propagated vegetatively using rhizome division and planted in a Randomised Complete Block Design, with one replicate per genotype in each of three blocks planted at 1.5 × 1.5 m spacing. In addition to previously established trials near Aberystwyth, rhizome propagules had been collected from trials near Catania (Sicily, Italy) and Braunschweig (Lower Saxony, Germany). The original source geographic coordinates were available for 592 out of the 953 genotypes in ABR33 and were downloaded from the IBERS MScan Database (Huang et al., 2019). These 592 Miscanthus genotypes were originally collected from an area between 18° 30' 0" N - 109° 18' 18" E and 45° 12' 25.92" N - 144° 26' 42" E and included South-East China, South Korea, Taiwan, and Japan. In terms of altitude, the genotypes were collected from sea level to 3000 m above sea level.

2.1.2 Identification of genetic groups

Leaf samples were collected from all 953 genotypes that survived through the spring of 2015 and DNA was extracted using previously described protocols (Slavov et al., 2013a). RAD-Seq genotyping was then performed by Floragenex as described by Slavov et al. (2014). Because several species of *Miscanthus* were sampled (Figure 1), and reference genome data was only available for some of these, RAD-Seq reads were aligned and single nucleotide polymorphism (SNP) genotyping data generated using the reference-free UNEAK pipeline (Lu et al., 2013), which is particularly suitable for species with highly repetitive and complex genomes, such as *Miscanthus* (Mitros et al., 2020; De Vega et al., 2021; Zhang et al., 2021). The default parameters of UNEAK were used and data was exported for 25,014 SNPs that had minor allele frequencies of at least 0.0025 (i.e., at least 5 minor allele copies) and call rates of at least 80%. To define genetic groups objectively, the resulting SNP data were subjected to discriminant analysis of principal components (DAPC) using the *adegenet* R package (Jombart, 2008; Jombart et al., 2010), and the 'optimal' number of groups was selected using the *find.clusters* function within the *adegenet* R package.

2.1.3 Sample collection and processing

Sample collection and processing was performed as described by Huang et al. (2019) with some minor modifications. Briefly, above-ground biomass was collected in early spring 2016 from 49 completely senesced *Miscanthus* genotypes across different genetic groups, with three biological replicates for each genotype. At the time of sampling, the plants were 4 years old. The cut biomass (comprised of stem and leaf material) was weighed to determine the total fresh weight. A sub-sample of approx. 200 g was removed and its relative moisture content was determined after drying at 60°C to constant weight. The percentage of moisture was used to calculate the approximate dry weight per harvest per plant. Dried samples were ground to a 1.5 mm mesh and stored at room temperature until further processing. Yield data for the biomass harvested for each genotype in 2016 was obtained from the MScan database at IBERS.

2.1.4 Cell wall preparation

All compositional analyses were carried out on purified cell walls. Approximately 70 mg of the 1.5 mm mesh biomass samples were weighed in 2 mL microtubes (Sarstedt, Cat. 72.609.001) along with two stainless steel balls and positioned in a Plant Grinding and Preparation System from Labman® similar to the one described by Santoro et al. (2010).

Cell wall material was extracted as described by da Costa et al. (2014) with minor modifications. The efficacy of the extraction at several steps was tested, as suggested by Fry (1988). Cell wall purification was performed using an alcohol-insoluble residue (AIR) preparation followed by starch removal. Preparation of AIR started with the addition of 1.5 mL of 70% v/v ethanol, followed by thorough vortexing (2400 rpm x 30 sec) and incubation for 16 h in a

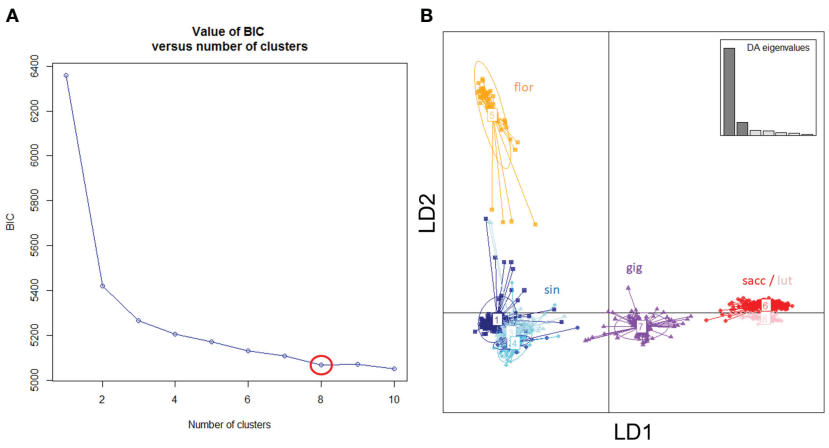


FIGURE 1
Genetic groups of *Miscanthus* in the ABR33 field trial based on Discriminant Analysis of Principal Components (DAPC). **(A)** Inference of 'optimal' number of groups based on the rate of change of model Bayesian Information Criterion (BIC). **(B)** Scatter plot of the first two linear discriminants (LD1 and LD2), with groups labelled using *a priori* species designation: sin = *Miscanthus sinensis*; flor = *M. floridulus*; sacc = *M. sacchariflorus*; lut = *M. lutarioriparius*; gig = *M.* × *giganteus* and inset showing the eigenvalues of discriminant analysis (DA).

shaking incubator set at 40°C and 150 rpm. Samples were then centrifuged at 10,000 rpm to pellet the residue. The pellet was repeatedly washed with 70% v/v ethanol until a stable A280 reading of the supernatant. Subsequently, the pellet was washed twice by adding 1.5 mL of a chloroform/methanol (1:1) solution, vortexed to resuspend the pellet, and incubated for 30 min at 25°C and 150 rpm. After each wash, samples were centrifuged at 10,000 rpm for 10 min and the supernatant was discarded. Finally, the pellet was washed three times with 1.5 mL of 100% acetone with resuspension by vortexing, incubation, centrifugation, and discarding of supernatant performed as in the previous step. The samples were then dried under airflow at room temperature. After this step, dry samples were stored at room temperature until further processing.

Starch was removed using α-amylase from the porcine pancreas (Novozyme E-PANAA-3G), following the method described by Foster et al. (2010) as adapted by da Costa et al. (2014), with minor additional modifications. Briefly, starch removal was initiated by re-suspending the AIR pellet in 1.5 mL of 0.1 M sodium acetate buffer pH 5. Tubes were capped and incubated at 80°C for 20 min to gelatinize the starch. Afterwards, samples were cooled on ice, centrifuged at 10,000 rpm, and the supernatant was discarded. The resulting pellet was washed three times with 1.5 mL of H_2O. Next, 1.5 mL of a mixture containing 0.3 µL of 1% w/v sodium azide, 15.6 µL of porcine α-amylase (3000 U mL^{-1}) and 1.4841 mL of H_2O were added to each sample. Samples were incubated in a shaking incubator for 48 h at 35°C and 110 rpm. The digestion was terminated by heating the samples to 95°C for 15 min, and the samples were subsequently cooled on ice. Finally, samples were centrifuged (10,000 rpm for 10 min), and the supernatant was discarded. Pellets were washed three times with water and two times with 100% acetone, as described previously, and dried at room temperature under a gentle flow of air. The absence of starch was confirmed using Lugol's staining test as described by Barnes and Anderson (2017).

2.2 Acid hydrolysis for total monosaccharide release

The total monosaccharide composition of cell wall material was determined by double hydrolysis (Saeman, 1945; Sluiter et al., 2008; Pettolino et al., 2012; Petit et al., 2019). Approximately 10 mg of the previously prepared cell wall material was weighed into 10 mL Pyrex glass tubes fitted with polypropylene caps. Subsequently, 0.100 mL of 72% (w/w) H_2SO_4 was added, and the tubes were capped and placed on a heating block set at 30°C for 60 min, during which time the samples were vortexed every 5 to 10 minutes. A set of sugar recovery standards (SRS) was prepared and taken through the remaining hydrolysis to correct for losses due to the destruction of sugars during dilute acid hydrolysis. Subsequently, deionized water was added to obtain a 4% (w/w) H_2SO_4 solution, and samples were mixed to eliminate phase separation. The sealed tubes were placed in an autoclave at 121°C for 1 h. Once at room temperature, the tubes were centrifuged to produce a particulate-free supernatant, and the samples were diluted ten-fold by taking 0.100 mL of each sample and mixing it with 0.900 mL of deionized water. Samples were stored at −20°C until analysis. Just before analysis, samples were diluted 1 to 200 in ultrapure water and enzymatically released amounts of glucose (Glc), xylose (Xyl), and arabinose (Ara) were quantified using the high-performance anion exchange chromatography system as described below in section 2.4.

2.3 Saccharification analysis

Enzymatic release of monosaccharides was performed as described by Resch et al. (2015) and modified by da Costa et al. (2015) and Petit et al. (2019). Briefly, approximately 10 mg of cell wall material was manually weighed out in 2 mL polypropylene microtubes with a screw cap (Sarstedt, Cat. 72.609.001). The exact sample weight was recorded. Then, 0.3 mL of 100% acetone was added to each sample to collect the material at the bottom of the tube. Acetone was left to evaporate under a stream of air overnight. Next, 1 mL of saccharification mixture was added having the following composition: 0.957 mL of 0.025 M potassium acetate buffer (pH5.6), 0.0024 mL of cellulase from *Trichoderma reesei* (Cellulase, Sigma Aldrich, code C2730), 0.0006 mL of *β*-glucosidase from *Aspergillus niger* (Novozyme 188, Novozyme, discontinued), and 0.040 mL of 1% sodium azide to repress bacterial growth. The mixture was prepared in a single batch and kept at 4°C until use. Samples were incubated for 48 h at 50°C in a shaking incubator set at 150 rpm. Subsequently, samples were centrifuged at 10,000 rpm for 5 min, and 0.9 mL of the supernatant was transferred to a new 2 mL polypropylene tube and stored for 2-3 days at −20°C until monosaccharide quantification. Just before analysis, samples were diluted 1:50 in ultrapure water and enzymatically released amounts of glucose, xylose, and arabinose were quantified using the high-performance anion exchange chromatography system as described below in section 2.4.

2.4 Monosaccharide quantification

Monosaccharides in the solutions were separated and quantified by high-performance anion exchange chromatography (Thermo Fisher Scientific Inc. ICS-5000) coupled with pulsed amperometric detection (HPAEC-PAD) operated at 45°C using a CarboPac SA10 (4×250 mm) column with a CarboPac SA10G (4×50 mm) guard column. An eluent generator coupled to the system continuously prepared a KOH solution at 0.001 M for isocratic elution at a flow rate of 1.5 mL/min for 14 min. A volume of 0.025 mL of the sample was injected into the column and detected by PAD using a gold working electrode and an Ag/AgCl reference electrode. A set of calibration standards was prepared. The calibration curve was validated between 5 µg/mL and 40 µg/mL. Immediately before the HPAEC-PAD analysis, 0.080 mL of each 1:10 diluted sample was neutralized by adding 0.320 mL of 0.02 M KOH. Samples were further diluted 1:4 with deionised water, to a final dilution of 1:200 (da Costa et al., 2015). Aliquots of 0.400 mL of the diluted samples were then filtered through 0.45 µm nylon filter vials (Thomson SINGLE STEP; Thomson Instrument Company, Oceanside, California, USA). The Chromeleon software (v. 7.1; Thermo

Fisher) was used for data processing. External calibration standards were used to identify and quantify the three most prominent monosaccharides detected in the chromatograms: glucose (Glc), xylose (Xyl), and arabinose (Ara).

2.5 Near infrared spectroscopy data

For each genotype, the content of cellulose, hemicellulose and lignin predicted by near-infrared spectroscopy (NIR) was determined by Analytical Chemistry (IBERS, Aberystwyth University) following procedures described by Allison et al. (2011).

2.6 Geo-environmental data sourcing

2.6.1 Genotype geographical origin details

Based on data from MScan, a complete list of the accessions included in the ABR33 trial was compiled. For accessions that were originally collected in the wild, source geographic coordinates, as determined using the Global Positioning System (GPS), were downloaded in a comma-separated file format.

2.6.2 Species distribution in the wild

The distribution of *Miscanthus* species in the wild was mapped using data from the Global Biodiversity Information Facility (GBIF) (Gbif.Org, 2019). The GBIF database contains data on species observations, including exact coordinates, from different sources. Data were standardized between various sources using the Darwin Core Standard (Wieczorek et al., 2012). The number of species observations per geographic point was counted and mapped using the R programming language (R Core Team, 2018).

2.6.3 Climatic data

Temperature-related and precipitation-related bio-climatic variables (Table S1) were retrieved from the WorldClim database v.1.4 (Hijmans et al., 2005) for all georeferenced genotypes using R as described by Hijmans and Elith (2013). The resulting dataset was downloaded in comma-separated format.

2.6.4 Soil data

Chemical and physical properties of the topsoil (TS, 0 – 100 cm from ground level) and the subsoil (SS, >100 cm from ground level) for each sampled location were downloaded from the Harmonized World Soil Database (HWSD) (Fao/Iiasa/Isric/Isscas/Jrc, 2009) in Microsoft Access format, then converted to a SQL 3 format using the MS Access to Sqlite3 Converter software (https://github.com/sanandrea/mdb2sq3). The resulting data was downloaded into a comma-separated file. Table S2 shows an overview of the soil variables.

2.6.5 Geo-environmental data analysis

Principal Component Analysis (PCA) was used to identify the traits with the largest contribution to the overall variability and consequently select genotypes (Maji and Shaibu, 2012) using the *get_pca_var* function in the factoextra package of R (R Core Team, 2018).

2.7 Statistical analysis of cell wall characteristics

Linear mixed-effect models (LMEMs) were used to test the effect of genetic group on the cell wall characteristics quantified in this study. We chose to use LMEMs as opposed to conventional analyses of variance (ANOVA) because of the unbalanced design of our experiment (i.e., different numbers of genotypes in different genetic groups), heteroscedasticity revealed by exploratory plots, and potential lack of independence of genotypes within genetic groups (i.e., because of coancestry). In LMEMs aimed at identifying contrasts among groups, genetic group was treated as a fixed effect and genotype and biological replicate as random effects. We also used LMEMs with all effects treated as random to partition the overall trait variance into group, genotype, and environmental sources (i.e., biological replicate and residual). All LMEMs were fitted using the *lmer* function in the lmerTest R package (Kuznetsova et al., 2017), which is an extended version of the same function in the widely used lme4 package (Bates et al., 2015). To identify significant contrasts, LMEMs with genetic group treated as a fixed effect were then passed on to the *emmeans* function from the emmeans R package (Lenth, 2022). Finally, results from these contrasts were visualized using a compact letter display obtained using the *cld* function from the R package multcomp (Hothorn et al., 2008). The datasets and the code used for the statistical analysis presented in this paper are available at: https://github.com/RosarioIacono/Iaconoetal2023Data.

3 Results

3.1 Genetic groups and geospatial distribution of *Miscanthus* genotypes

Based on results from DAPC, the 'optimal' number of groups was set at eight (i.e., there was no further reduction of the Bayesian Information Criterion with higher numbers of groups, Figure 1). *M. sinensis* was the only species represented by multiple groups ($n = 4$, Table 1). As expected, *M.* × *giganteus* formed a separate group with intermediate clustering between its parental species (*M. sinensis* and *M. sacchariflorus*).

The four species of *Miscanthus* included in this study (*M. sacchariflorus*, *M. sinensis*, *M. floridulus*, and *M. lutarioriparius)* are distributed across a broad geographic area (blue shading in Figure 2). The distribution of the 592 *Miscanthus* genotypes included in the ABR33 trial (red dots in Figure 2) fell within the distribution of their corresponding genetic group reported in the GBIF database. Despite the extensive distribution overlap between *M. sacchariflorus* and *M. sinensis*, collection points of both *M. sacchariflorus* and *M. sinensis* in mainland China were recorded in areas where there was no GBIF report of the two species in the wild (red dots outside of the blue area), corroborated by collection points

TABLE 1 List of genetic groups and their ID as used in this paper.

Group ID	Genetic group
M1	*M. sinensis* South Japan
M2	*M. sinensis* (EMI/PRI)
M3	*M. sinensis* North Japan
M4	*M. sinensis* Taiwan
M5	*M. floridulus*
M6	*M. sacchariflorus/robustus*
M7	*M.* × *giganteus*
M8	*M. lutarioriparius*

of inter-specific *M.* × *giganteus* hybrids in close proximity (Figure 2). In addition, *M. floridulus* and *M. lutarioriparius* genotypes were collected from a relatively small part of their respective distribution areas (Figure 2). Taken together, these observations point to the opportunity offered by databases like the one from GBIF to inform future germplasm collection campaigns and to the presence of vast unexplored areas where interesting new accessions could be found. Moreover, they highlight the possibility that more inter-specific hybrids, similar to commercially grown *M.* × *giganteus*, could perhaps be found with targeted collection campaigns in areas of coexistence of different species.

3.2 Pedo-climatic conditions

Using the coordinates for each collection point of the 592 *Miscanthus* genotypes, it was possible to retrieve a complete series of climatic and soil variables from the WorldClim database (Fick and Hijmans, 2017) and the Harmonised World Soil Database (Fao/Iiasa/Isric/Isscas/Jrc, 2009), respectively. The data describing climatic conditions over the last 50 years and soil conditions of the sites of origin of the 592 genotypes resulted in a multivariate dataset comprising 51 variables. The list and description of the variables are available in Tables S1, S2. The high number of variables and the well-known interdependence between precipitation, temperature, and pedologic conditions in specific environments suggested that the dataset may not be of full rank and some variables were partly redundant. To visualize the relations

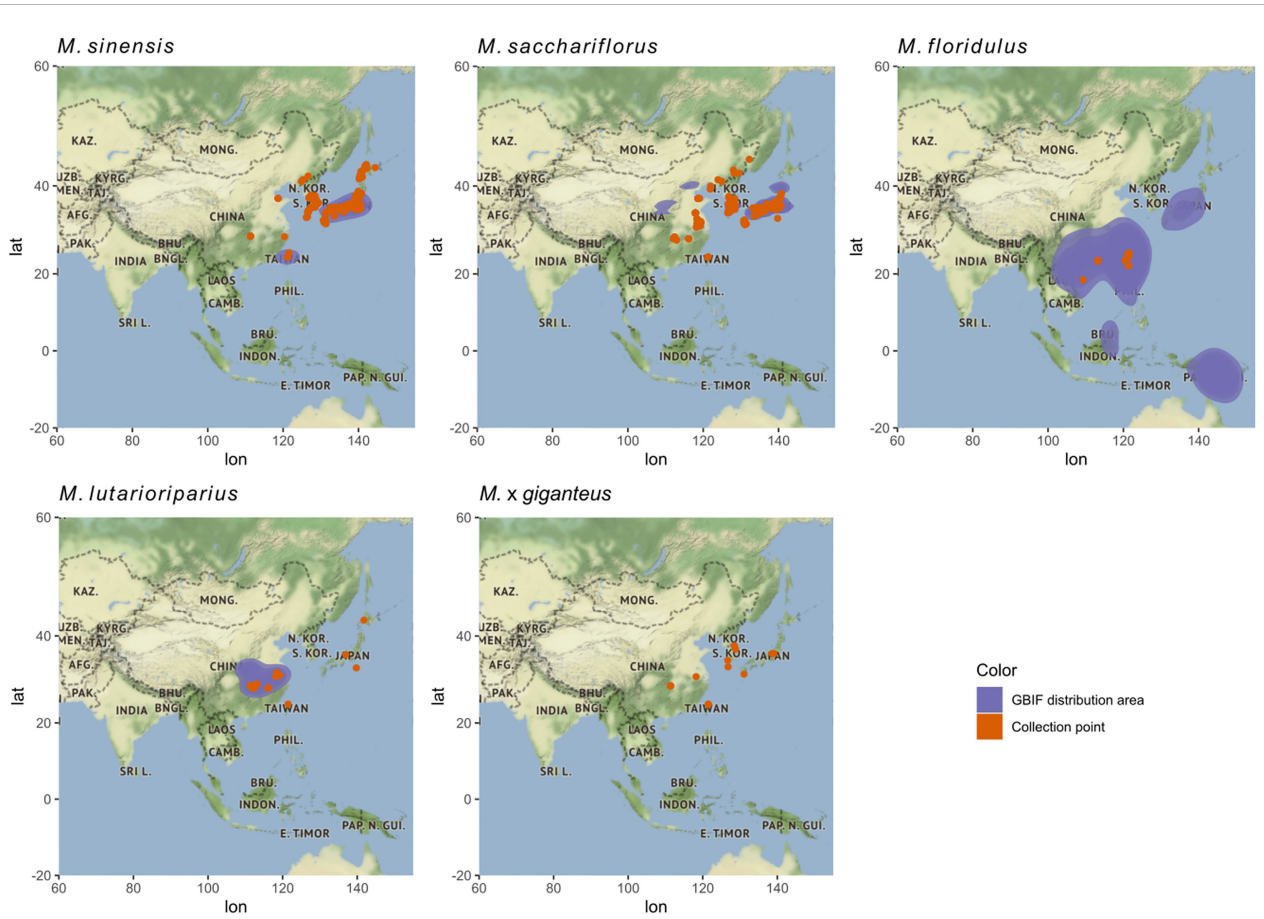


FIGURE 2
Area of distribution of *Miscanthus* species in the wild and collection points for genotypes in ABR33. The distribution of *Miscanthus* species is presented using data from the Global Biodiversity Information Facility (GBIF). The x-axis and y-axis indicate longitude and latitude, respectively. Blue shading represents the areas where each species has been reported in the wild according to the GBIF database, while the red dots represent the collection sites of genotypes in the ABR33 trial.

between the environmental variables we calculated the correlation between the variables in the soil-climate dataset (Figure S1). These correlations indicated that *Miscanthus* species can grow across a wide range of conditions in terms of salinity, seasonality of precipitation and soil composition.

3.3 Selection of 49 genotypes

We selected a number of representative genotypes to examine the cell wall properties of *Miscanthus* across different source locations and genetic groups. The pedo-climatic PCA score plot is presented in Figure 3 and the plot of PCA loadings is shown in Figure S2. The first two components of the PCA model explained 40% of the total variance in the dataset. The first principal component (explaining 25.4% of the model variability) separated the *M. floridulus* group (mostly positive scores of green dots) from the other groups, particularly *M. sacchariflorus* (purple dots; Figure 3).

Analysis of the loadings shows that the inverse relation between the amount of precipitation (PTA, positive) and soil pH (T_PH and S_PH, negative) at *Miscanthus* collection sites are major contributors to the ordination along PC1 (Figure S2). Accessions in the *M. floridulus* group tended to be found in environments with high precipitation and low pH, while accessions of the *M. sacchariflorus* group were collected from environments characterized by a low amount of precipitation and higher values of pH in the soil. It is worth noting that pH is one of the main factors affecting soil nutrient availability. Soils with low pH generated under high precipitation conditions tend to have lower nutrient availability.

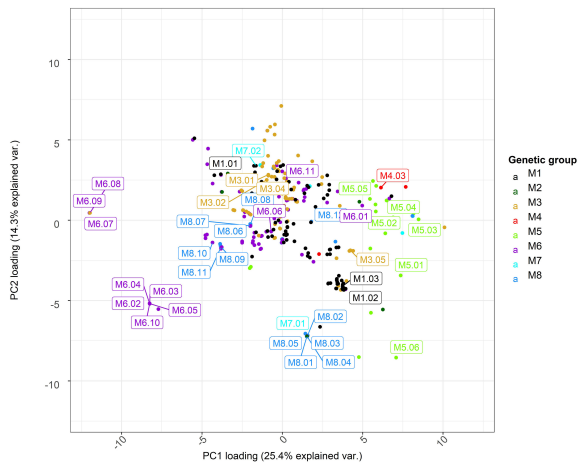


FIGURE 3
Principal component analysis (PCA) of climate and soil variables for sampling locations of ABR33 genotypes. Scores for the first and second principal components (PC1 and PC2) are shown along the x- and y-axis, respectively, with proportions of variance explained shown in parentheses. Labels indicate genotypes selected for biomass characterization. The loadings for PC1 and PC2 are shown in Figure S2. M1 = *M. sinensis* from South Japan, M2 = *M. sinensis* EMI/PRI, M3 = *M. sinensis* from North Japan, M4 = *M. sinensis* from Taiwan, M5 = *M. floridulus*, M6 = *M. sacchariflorus/robustus*, M7 = *M.* × *giganteus*, M8 = *M. lutarioriparius*.

The selection of 49 representative *Miscanthus* genotypes was performed taking into account the PCA and genetic grouping results described above. We selected 35 genotypes from 7 genetic groups (i.e., maximizing environmental variation captured within each group) occurring in the wild and completed the selection with 14 genotypes which included naturally occurring and artificially generated (i.e., at IBERS) *M.* × *giganteus* hybrids (Hodkinson and Renvoize, 2001). Details of the 49 selected genotypes are shown in Table S3.

3.4 Differences in cell wall composition and characteristics among genetic groups

The main cell wall monosaccharides were quantified following acid hydrolysis of purified cell wall material of the 49 *Miscanthus* genotypes. The average content of glucose (Glc), xylose (Xyl) and arabinose (Ara) across the 49 genotypes was 39%, 28% and 2.5% (cell wall material (CWM); Figure S3), respectively, and in accordance with values previously reported for *Miscanthus* (e.g.: da Costa et al., 2017; van der Weijde et al., 2017b). Based on LMEM analyses, there were significant genetic group effects on the total content of Glc, Xyl and Ara in the cell wall (Tables 2, S5). However, group differences only accounted for 9-22% of the total variance and were mostly driven by contrasts with *M. floridulus* (Figure S3). Furthermore, variation among genotypes within groups was generally low and was only significant for Glc (14% of total variance), with most of the variation (56-79%) explained by the residual LMEM terms (Table 2).

A series of variables describing the composition and structural features of the cell wall were collected (Table S4). Measurements of acid detergent fibre (ADF), neutral detergent fibre (NDF), acid detergent lignin (ADL), ash, dry matter (DM), and Klason Lignin were determined by near-infrared (NIR) spectroscopy. The amounts of cellulose and hemicellulose were derived as described by Van Soest (1963) and Van Soest et al. (1991). Ratios between the content of certain monosaccharides and cellulose/hemicellulose were also used as indicators of structural cell wall features.

The general trend for these composition variables was that genetic group effects were substantial (i.e., as high as 50% of the total variance for cellulose and 49% for ADF, Table 2) and once again mainly driven by contrasts involving *M. floridulus* (Figure S4). In contrast to monosaccharides, variation among genotypes within groups also tended to be high (i.e., up to 53% for Klason Lignin) and was only not statistically significant for DM (Table 2). Monosaccharide ratios followed a similar trend, though both genetic group and genotype-within-group effects were substantially weaker (Table 2; Figure S5). As expected, variance components for the cellulose/hemicellulose ratio were roughly intermediate to those for cellulose and hemicellulose (Table 2; Figure S5).

In summary, genetic group effects for cell wall composition were typically strong but driven almost exclusively by differences between *M. floridulus* (M5) and the other genetic groups. However,

10.3389/fpls.2023.1155188

TABLE 2 Variance components associated with genetic group (σ^2_{Group}), genotypes within groups (σ^2_{Geno}), biological replicates (σ^2_{Rep}), and residual error (σ^2_{Err}).

	Random Effects percentage			
	σ^2_{Group}	σ^2_{Geno}	σ^2_{Rep}	σ^2_{Err}
Glc	9.6	13.6	21.3	55.6
Xyl	21.5	6.7$^{n.s.}$	10.3	61.4
Ara	8.7	3.3$^{n.s.}$	9.1	78.9
NDF	45.8	35.0	0.3$^{n.s.}$	18.9
ADF	49.0	32.0	2.4	16.6
ADL	36.4	38.7	1.9	23.0
Ash	48.0	25.0	1.6	25.4
DM	19.9	4.2$^{n.s.}$	3.7$^{n.s.}$	72.2
K Lignin	18.7	53.2	1.8	26.3
Cellulose	49.7	32.3	2.3	15.6
Hemicellulose	37.4	27.8	7.3	27.4
Cel Hem	45.7	28.6	5.2	20.6
Ara Xyl	22.0	0.3$^{n.s.}$	13.8	63.9
Xyl Glc	16.4	16.1	18.8	48.7
Glc E	41.8	12.4	4.1	41.7
Xyl E	33.0	16.5	1.2$^{n.s.}$	49.2
Ara E	5.1$^{n.s.}$	25.9	9.62E-11$^{n.s.}$	69.0
DW	24.9	45.5	0.7%	28.9
Glc E DW	17.9	60.2	0.5$^{n.s.}$	21.5
Xyl E DW	10.4$^{n.s.}$	54.7	0.1$^{n.s.}$	34.8
Ara E DW	2.0$^{n.s.}$	25.9	0.00E+00$^{n.s.}$	72.1

Variance components are expressed as percentage of the total variance and were estimated using LMEMs in which the effects of genetic groups, genotypes and biological replicates were all treated as random (see Materials and Methods). Superscripts "n.s." indicate LMEM terms that were not statistically significant at $\alpha = 0.05$.

there was a comparably high amount of genetic variation among genotypes within groups, indicating that breeding could be effective both within and among groups.

3.5 Variation in cell wall recalcitrance

Next, we wanted to investigate if there are differences in cell wall recalcitrance to sugar release between *Miscanthus* genotypes from different genetic groups. To measure recalcitrance, the cell wall material from aboveground senesced biomass was digested for 48 hours using an enzymatic mixture containing various cellobiohydrolases, and endo-(1 → 4)-β-glucanases with collateral xylanase activity. To increase the ability to detect genetically related differences in recalcitrance, the cell wall material was not pre-treated. The relative recalcitrance was quantified using the amount of glucose (GlcE), xylose (XylE) and arabinose (AraE) released in solution after the digestion and was expressed as a percentage of the total amount of the respective monosaccharides present in the cell wall (Figure 4).

Unlike monosaccharide quantities, genetic group effects for GlcE and XylE were strong to moderate (i.e., 42% and 33% of the total variance, respectively, Tables 2, S5) and once again driven by contrasts involving *M. floridulus*, which was consistently less recalcitrant. This was not the case for AraE, but there was significant variation among genotypes within groups (i.e., 12-26% of the total variance) for all three measures (GlcE, XylE, and AraE) of recalcitrance.

3.6 Cell wall recalcitrance in perspective to biomass yield

The amount of monosaccharides that can be enzymatically released from biomass depends not only on cell wall recalcitrance but also on the annual biomass yield. Therefore, we normalized our enzymatic sugar release data (GlcE, XylE and AraE) against the amount of biomass that can be harvested annually from each genotype [i.e., yield in kg of dry weight (DW)].

As expected, genetic groups differed significantly for DW (Tables 2, S5), with the *M.* × *giganteus* group having the highest biomass yields (Figure S6). However, these differences accounted for only 25% of the total variance, whereas variation among genotypes within genetic groups was considerably more pronounced (i.e., 46% of the total variance). Consequently, DW-normalized recalcitrance varied strongly among genotypes within groups (i.e., up to 60% of the total variance for DW-normalized GlcE), whereas group effects were weak and only significant for GlcE (Tables 2, S5), though no specific group contrasts were statistically significant (Figure 5). Thus, recalcitrance and biomass yield tended to offset each other at the genetic group level, but with very high levels of genetic variation within groups.

3.7 Structural bases of the differences in recalcitrance

We next wanted to determine the compositional and structural cell wall features contributing to the observed differences in recalcitrance to enzymatic sugar release. To explore the structural basis of recalcitrance across the 49 selected *Miscanthus* genotypes, independent from their genetic groups, correlation analysis with Bonferroni correction was performed between enzymatic sugar release (GlcE, XylE, and AraE) and measures for cell wall composition and structure (Figure 6). GlcE and XylE were negatively correlated with the total amount of glucose and xylose in the cell wall ($r < -0.75$), while only GlcE was negatively correlated with lignin content (K lignin and ADL). Furthermore, there was a positive correlation between the ash content and the amount of glucose released. This was also the case for the arabinose-to-xylose ratio. The cellulose-to-hemicellulose ratio negatively correlated with GlcE (Figure 6). Finally, there was a positive correlation between the leaf-to-stem ratio (LSR) and GlcE, possibly due to the lower recalcitrance of leaf material when compared to stem material (da

10.3389/fpls.2023.1155188

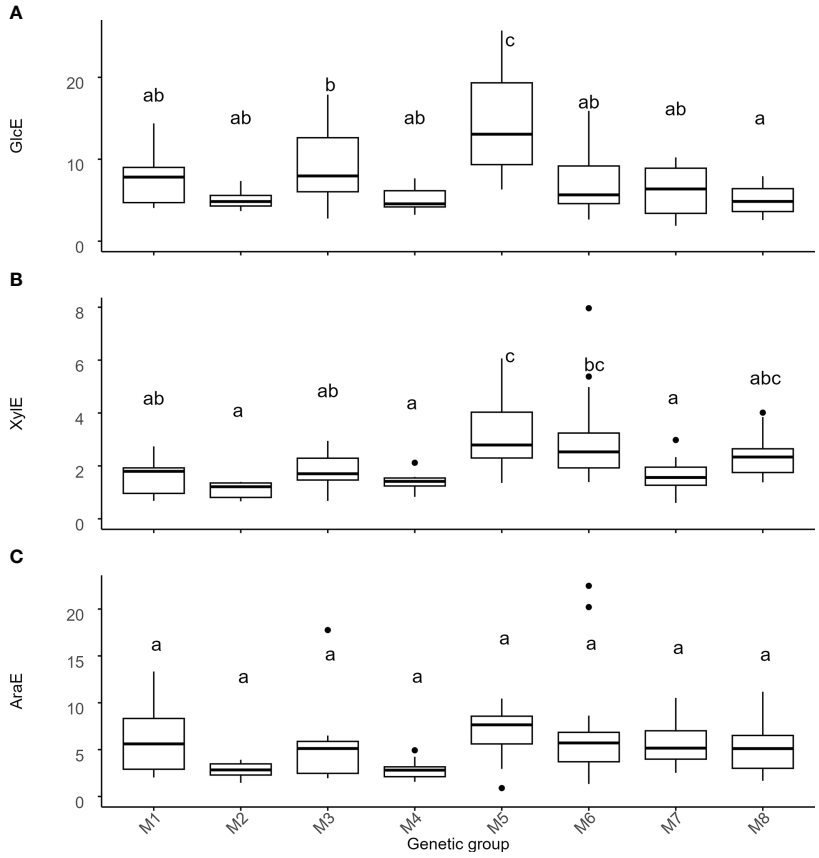


FIGURE 4
Variation of recalcitrance across genetic groups of *Miscanthus*. Recalcitrance was expressed as the percentage of the total amount of the respective monosaccharides present in the cell wall. Values on the y-axis are the amount of **(A)** glucose (GlcE), **(B)** xylose (XylE), and **(C)** arabinose (AraE) enzymatically released. The thick line in the box represents the median value. The box itself indicates the interquartile range, where 75% of measurements fall. Letters represent significant differences as detected by estimation of marginal means after a LMEM with genetic group treated as a fixed effect and with $p < 0.05$. Labels on the x-axis are the 8 genetic groups delineated using single-nucleotide polymorphism data (Table 1, Figure 1). M1 = *M. sinensis* from South Japan, M2 = *M. sinensis* EMI/PRI, M3 = *M. sinensis* from North Japan, M4 = *M. sinensis* from Taiwan, M5 = *M. floridulus*, M6 = *M. sacchariflorus/robustus*, M7 = *M.* × *giganteus*, M8 = *M. lutarioriparius*.

Costa et al., 2017). Besides pointing to the complexity of the interactions between compositional and structural cell wall features and cell wall recalcitrance to sugar release, these results also provide some indications of the traits that require attention in breeding programs that target specific end uses of biomass.

3.8 Variation in cell wall related variables within and among genetic groups

Having established the overall cell wall features that correlate with enzymatic sugar release across the 49 genotypes, we next wanted to examine if these variables resulted in distinctive correlations with *Miscanthus* genotypes and the genetic group they belong to. Measures for the various cell wall features and enzymatic sugar release data were scaled and two trees were generated using the correlation coefficients between cell wall features and genotypes, respectively (Figure 7). Cell wall and phenotypic variables clustered in four distinct groups (I-IV), indicating that the variables within each group tend to follow a similar pattern for a given genotype. For example, cluster I comprised the variables hemicellulose (Hem), arabinose (Ara), arabinose to xylose ratio (AraXyl), leaf to stem ratio (LSR), and ash content (Figure 7). These variables (except Ara) all correlated positively with GlcE (Figure 6).

Based on the cell wall characteristics (excluding recalcitrance), the genotypes clustered in 5 groups (A-E). The most distinguishing feature of the genotypic tree is that all the *M. floridulus* genotypes (M5) are included in two small clusters (cluster A and cluster D). For both of these clusters, measures of cell wall and phenotypic variables belonging to cluster I tended to be higher for *M. floridulus* genotypes when compared to most other genotypes (Figure 7). In addition, both groups A and D contained the genotypes with the lowest recalcitrance (high GlcE values). Most of the genotypes with the highest glucose content (Glc) group together in cluster C, dominated by *M. sacchariflorus/robustus* (M6), *M.* × *giganteus* (M7), and *M. lutarioriparius* (M8); interestingly this is also the cluster with the highest recalcitrance (low GlcE values).

10.3389/fpls.2023.1155188

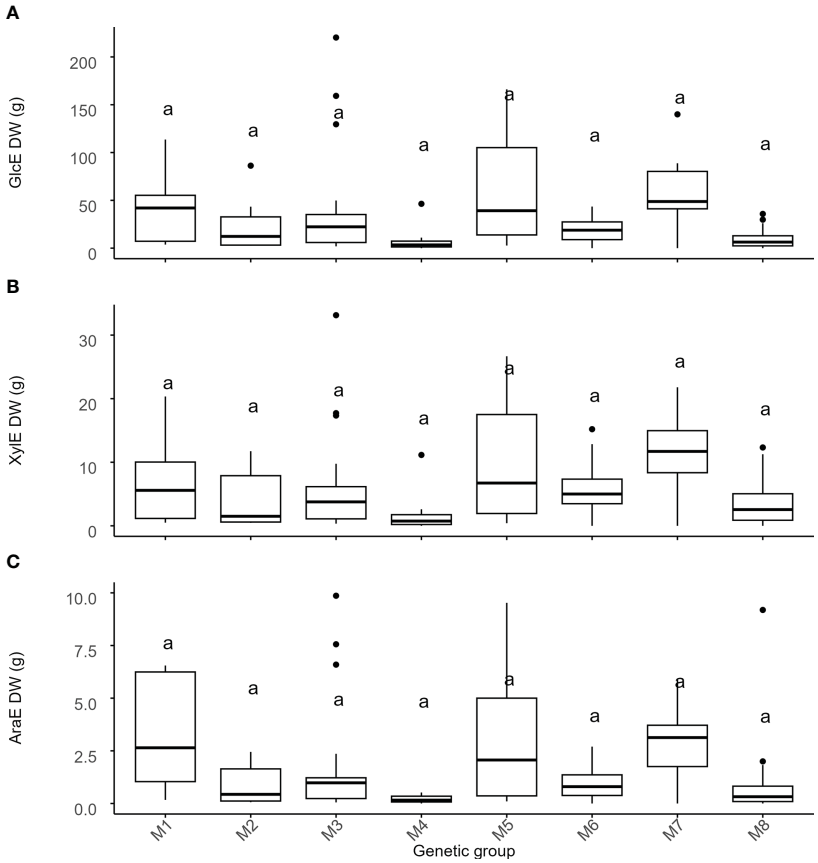


FIGURE 5
Variation of recalcitrance across genetic groups of *Miscanthus*. *Miscanthus* biomass to enzymatic sugar release normalized for the amount of biomass produced. Values on the y-axis are the total amount of **(A)** glucose (GlcE), **(B)** xylose (XylE) and **(C)** arabinose (AraE) enzymatically released normalized for the amount of biomass (dry weight) produced. The thick line in the box represents the median value. The box itself indicates the interquartile range, where 75% of measurements fall. Letters represent significant differences as detected by estimation of marginal means after a LMEM with genetic group treated as a fixed effect and with $p < 0.05$. Labels on the x-axis are the 8 genetic groups identified in the *Miscanthus* population. M1 = *M. sinensis* from South Japan, M2 = *M. sinensis* EMI/PRI, M3 = *M. sinensis* from North Japan, M4 = *M. sinensis* from Taiwan, M5 = *M. floridulus*, M6 = *M. sacchariflorus/robustus*, M7 = *M.* × *giganteus*, M8 = *M. lutarioriparius*.

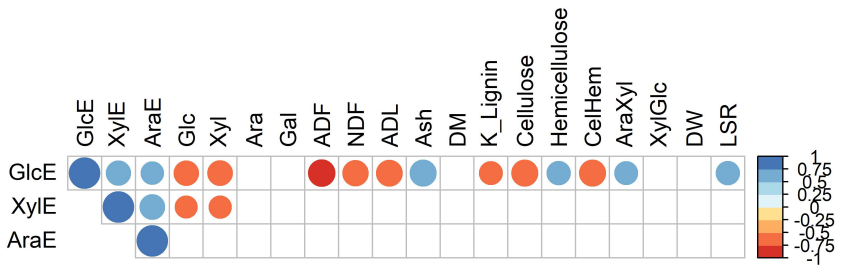


FIGURE 6
Pearson correlation coefficients between glucose (GlcE), xylose (XylE) and arabinose (AraE) enzymatically released from the cell wall and measures of cell wall structure and composition. Circle sizes are proportional to the significance of the correlation. Only correlations with $p < 0.05$ (i.e., after Bonferroni corrections) are shown.

10.3389/fpls.2023.1155188

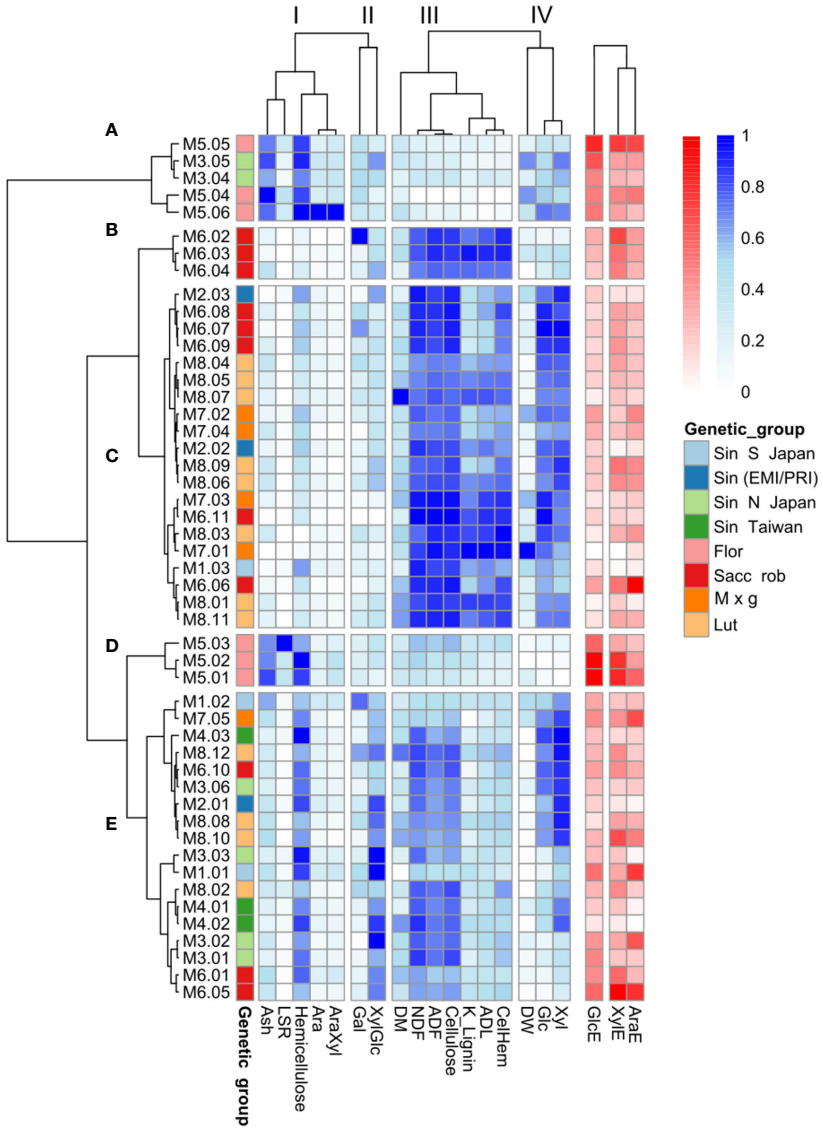


FIGURE 7
Clustering of the 49 *Miscanthus* genotypes based on cell wall traits. Relative abundances of cell wall compositional features are represented in shades of blue. Shades of red represent saccharification efficiency indices: the percentages of total glucose (GlcE), xylose (XylE) and arabinose (AraE) released upon enzymatic saccharification. Only cell wall composition data (blue), and not the three saccharification-related variables (red), were considered for genotype classification. **(A–E)** Represent the correlation clusters between genotypes, based on the cell wall variables and I-IV indicate the correlation clusters between cell wall features across genotypes. M1 = *M. sinensis* from South Japan, M2 = *M. sinensis* EMI/PRI, M3 = *M. sinensis* from North Japan, M4 = *M. sinensis* from Taiwan, M5 = *M. floridulus*, M6 = *M. sacchariflorus/robustus*, M7 = *M.* × *giganteus*, M8 = *M. lutarioriparius*.

4 Discussion

Despite the potential of *Miscanthus* as a biomass feedstock for the production of fuels and chemicals, few studies have evaluated if biomass recalcitrance to sugar release is governed by the same cell wall features when considering different genetic backgrounds. In this study, we evaluated *Miscanthus* genotypes from eight distinct genetic groups and identified the compositional and structural properties that correlate with the enzymatic release of cell wall sugars. Results suggest that the cell wall properties that correlate with recalcitrance are mostly similar when comparing different genetic *Miscanthus* groups and that differences in biomass allocation to leaf and stem between these groups may contribute to differences in cell wall properties and recalcitrance. Of immediate practical interest, there was significant genetic variation within groups for almost all traits, suggesting that breeding for cell wall characteristics can be effective both within and among groups.

4.1 Geospatial analysis to inform the collection of wild accessions

Our analysis of the geographical information of the area of collection of each genotype in the wild showed the opportunity offered by the GBIF database as a source of information for future germplasm collection campaigns. We showed how there are vast

areas where the presence of *Miscanthus* has been reported, but have not yet been targeted for the collection of genotypes for breeding programmes. The idea that geographical and environmental information should drive germplasm selection was proposed already in 1972 (Simmonds, 1972). More recently, it was argued that the importance of geographical information for breeders resides in the direct influence of the environmental conditions of an area on the phenotype and genotype of the organisms through adaptation (Parra-Quijano et al., 2012).

In this study, we used a PCA approach to study the contribution of climatic and soil variables to the species' geographical distribution. A similar approach has been used to study the geographic distribution of tomato accessions (Ramírez-Ojeda et al., 2021) and barley (Teklemariam et al., 2022). The first component of the PCA showed a strong association with the soil pH and total annual precipitation and genotypes of *Miscanthus* under investigation separated along this dimension, with the genotypes of *M. floridulus* forming a separate group. A study using the *Miscanthus* Genomic DataBase (MGDB) and characterizing 485 genotypes of *Miscanthus* (Xiang et al., 2020), showed that 10% of elite germplasm of *M. floridulus* covers most areas in southern China and authors inferred that soil characteristics could play a role. Notably, soil pH is one of the principal determinants of the chemical availability of minerals in the soil (Roem and Berendse, 2000; Stark et al., 2012; Penn and Camberato, 2019) and adaptation to different soil pH conditions could point to differences in nutrient use efficiency and mineral mobilization among *Miscanthus* species and/or genetic groups. It has been shown that the pH of the rhizosphere of *M. floridulus* decreases when grown in heavy metal-contaminated soils while the content of soil organic matter increases. Furthermore, the level of heavy metals was low in the rhizosphere compared with the non-rhizosphere, due to the increased uptake of heavy metals (Qin et al., 2022). Knowledge of the level of adaptation to soil pH could inform the choice of breeding specifically for phytoremediation.

4.2 Variation in cell wall composition between genetic groups

M. floridulus genotypes clearly showed the most distinctive cell wall compositional features. Accessions belonging to this group showed a lower glucose, xylose and ADL content and a higher Ara/Xyl ratio, when compared with some of the other groups. It is known that the composition of leaf and stem cell walls significantly differs in *Miscanthus* (da Costa et al., 2014; da Costa et al., 2017). Glucose, which is mainly derived from cellulose, and lignin content tend to be higher in stems while arabinose and the Ara/Xyl ratio tend to be higher in leaves (da Costa et al., 2017). Thus, the relative proportions of leaf versus stem biomass in the total aboveground biomass influence cell wall compositional features. For example, the leaf-to-stem ratio (LSR) for *M. floridulus* (M5) was significantly higher (LSR = 3.51) compared with those of the other genetic groups (LSR ranging from 0.12 in M8 to 0.70 in M4). Therefore, some of the observed differences in cell wall composition for *M. floridulus* are likely caused by the fact that most of their biomass is comprised of leaf tissue. This is corroborated by the high levels of ash found in *M. floridulus* as it is well established that leaves have a higher ash content than stems (Monti et al., 2008). High levels of ash negatively affect biomass quality for combustion. Another aspect that distinguishes *M. floridulus* from the other genetic groups is that it does not senesce over winter, exhibiting a stay-green phenotype. It has been shown before that genotypes of *M. sinensis* with a higher percentage of the leaf over the total amount of biomass and the ability to keep the leaves over-winter, are characterised by better performance for the production of methane through anaerobic digestion compared to *M.* × *giganteus* (Mangold et al., 2019). From an agronomic point of view, *M. floridulus* could help to increase the harvesting window for *Miscanthus*, when considering their biomass utilization in biorefining and anaerobic digestion.

Effects of genotype and genetic group on cell wall composition have been reported previously. For example, Allison et al. (2011) studied the cell wall composition of 244 genotypes belonging to *M.* × *giganteus*, *M. sacchariflorus*, and *M. sinensis*, and found that *M.* × *giganteus* biomass contained significantly more NDF and cellulose, and less hemicellulose compared with *M. sacchariflorus* or *M. sinensis*. Clustering of eight *Miscanthus* genotypes following a detailed cell wall analysis resulted in two distinctive groupings, a cluster comprised of the *M. sinensis* genotypes and a cluster comprised of hybrids and *M. sacchariflorus* (da Costa et al., 2019). Analysis of 15 accessions of *M. sinensis*, *M. sacchariflorus*, and *M.* × *giganteus* grown at different locations found that differences in cell wall composition between genotypes were mainly caused by genotype-by-environment (G x E) interactions (van der Weijde et al., 2017a).

We found that there were high levels of genetic variation within groups for nearly all cell wall composition traits. This result underlies the existence of a high percentage of intra-specific variation in species of *Miscanthus*. The existence of intra-specific variability within species of Miscanthus had been pointed out by Slavov et al. (2013b). In addition, our results agree with the one presented by Xu et al. (2020) that reported the existence of intra-specific variability in cell wall composition in *M. sinensis*, *M. floridulus*, *M. nudipes*, *M. sacchariflorus*, *M. lutarioriparius*, and their hybrids. Intra-specific genetic variability has been reported for *M. sinensis* (Shimono et al., 2013), *M. lutarioriparius* (Yang et al., 2019), and *M.* × *giganteus* (Głowacka et al., 2015). Thus, breeding programmes aimed at improving cell wall traits do not necessarily need to rely on inter-specific hybrids.

Further research, using growth under controlled environmental stress conditions and multiple genetic groups, is required to determine the full extent of variability in cell wall composition and its plasticity in relation to the environment within the genus *Miscanthus*.

4.3 Differences in cell wall recalcitrance between genetic groups

Cell wall recalcitrance is a trait hindering our ability to profitably deconstruct the cell wall to obtain molecules for the

biorefining process. In our study, we used the amount of the main cell wall sugars that can be released enzymatically from purified cell walls as a measure of recalcitrance. Effects of genotype on cell wall recalcitrance have been reported before between hybrids and *M. sinensis* (Belmokhtar et al., 2017). Our results show significant differences in cell wall recalcitrance between genetic groups in the genus *Miscanthus*.

It is striking that while genotypes of *M.* × *giganteus* show the highest content of glucose in the cell wall and *M. floridulus* the lowest, when it comes to recalcitrance the situation is reversed, with *M. floridulus* being the genetic group with the lowest recalcitrance and the highest levels of glucose released. A similar trend can be observed for xylose. However, when the biomass yield is taken into account, the higher recalcitrance of *M.* × *giganteus* genotypes is offset by their higher yield when compared with *M. floridulus*, resulting in similar amounts of sugars that can be released from *Miscanthus* plants belonging to these two genetic groups. *M. lutarioriparius* showed similar characteristics to *M.* × *giganteus* for the content and enzymatic release of cell wall sugars, but their lower biomass yield meant that the amount of sugars released on a plant biomass basis remained lower when compared to *M.* × *giganteus*. Although the abundance of the main cell wall sugars is similar across the four *M. sinensis* groups, genotypes from Japan are less recalcitrant to enzymatic glucose release compared with the other two similar yielding *M. sinensis* groups.

In addition, our results point to the existence of intra-specific variability in cell wall recalcitrance in the *Miscanthus* species under study. Intra-specific variability in recalcitrance has been reported before in biomass crops. For example, Ohlsson et al. (2019) described intra-specific variability in recalcitrance between 286 natural *Salix viminalis* clones. Ostos Garrido et al. (2018), working with different genotypes of three *Poaceae* species, found that the intra-specific variability in recalcitrance is a trait that depends on the species investigated. Our findings emphasize that both biomass yield and biomass recalcitrance need to be taken into account when developing new *Miscanthus* cultivars for biorefining purposes. Moreover, the intra-specific variability available for this trait could be a resource for breeding programs targeting specific biomass end uses.

4.4 Structural bases of the differences in cell wall recalcitrance

It has been shown that recalcitrance is a complex trait depending both on the cell wall composition and the structural organization of its components (De Souza et al., 2015). The main components of recalcitrance vary according to the species considered (DeMartini et al., 2013). We investigated the correlation between the cell wall composition and structural features and the amount of glucose, xylose and arabinose enzymatically released to determine which cell wall traits affected recalcitrance. Although da Costa et al. (2019) previously carried out such a correlation analysis, here the group of genotypes used included a wider genetic background as demonstrated by the number of species represented with a large geographical distribution. In addition, the set of cell wall and structural variables used in the current study was significantly different. Not surprisingly, our correlation test confirmed the negative effect of lignin content on the amount of glucose that can be enzymatically released from the cell wall. Lignin represents between 15 and 25% of *Miscanthus* biomass (da Costa et al., 2014), and its content has been shown to play a significant role in recalcitrance (Chen and Dixon, 2007). Indeed, a previous study showed that lignin content correlated negatively with cell wall sugar release in *Miscanthus* (da Costa et al., 2019). However, it has been observed that the monomeric composition of lignin, in particular the syringyl (S) to guaiacyl (G) lignin monomer ratio also has a major role in cell wall recalcitrance (Yoo et al., 2018). Here only the lignin content was used for the correlation and we cannot exclude that a full characterization of the lignin could have provided a better understanding of the relationship observed between the traits. Several studies have shown that lignin is not the only contributor to cell wall recalcitrance (Studer et al., 2011; DeMartini et al., 2013; da Costa et al., 2017; da Costa et al., 2019). DeMartini et al. (2013) compared the cell wall components affecting the recalcitrance of switchgrass and poplar biomass. They found that while lignin removal reduces poplar biomass recalcitrance, xylose removal is more effective in reducing switchgrass biomass recalcitrance. Similarly, Mangold et al. (2019) found that hemicellulose rather than lignin content had a higher effect on methane production from Miscanthus biomass. Indeed, our results show that hemicellulose content and the arabinose to xylose (AraXyl) ratio have a positive correlation with enzymatic glucose release. This is possibly linked to the positive effect of LSR on glucose release as the hemicellulose content and AraXyl ratio tends to be higher in leaves than in stems. Indeed, da Costa et al. (2019) found that the AraXyl ratio correlated positively with glucose release in *Miscanthus* leaves. Similarly, the positive correlation between ash content and sugar release is possibly also linked to LSR. More detailed analysis, evaluating leaf and stem organs separately and expanding the cell wall variables to include for instance lignin composition, measures for pectins and hydroxycinnamic acids could provide further information to explain the differences in cell wall recalcitrance observed.

In summary, we identified a number of cell wall related variables important for biomass quality related to using *Miscanthus* as a biomass crop for biorefining. Although our study identified significant variation in cell wall related features across the 49 selected *Miscanthus* genotypes belonging to eight different genetic groups, with the exception of *M. floridulus*, this variation was generally not distinctive enough to separate the genotypes according to their genetic background. The results emphasize the inter- and intra- specific variation in cell wall characteristics and biomass recalcitrance in the genus *Miscanthus* and the importance of also considering yield and organ related parameters when analyzing cell wall properties and biomass recalcitrance aimed at improving *Miscanthus* as a biomass crop.

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 author.

Author contributions

RI, GA and MB contributed to conception and design of the study. CD and GS performed the experimental work and data analysis aimed at defining genetic groups. GS provided assistance with the LMEM statistical analyses of cell wall characteristics. JC-B instigated the collection of the NIRS data. All remaining experimental work and data analysis was performed by RI. RI and MB wrote the initial manuscript draft with feedback from GA and GS. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Biotechnology and Biological Sciences Research Council (BB/M017389/1; BBS/E/W/0012843A, and BB/X011062/1).

Acknowledgments

We thank analytical chemistry staff, in particular Sue Lister (IBERS), for providing the NIRS data. We also acknowledge the many members of IBERS who were involved in setting up and maintenance of the ABR33 field trail, harvesting and sample preparation for NIRS.

Conflict of interest

The authors wish to declare that DRAX Power Limited UK was a project partner for the PhD studentship of RI that was supervised by GA and MB. 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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1155188/full#supplementary-material

References

Albersheim, P., Darvil, A., Roberts, K., Sederof, R., and Staehelin, A. (2011). *Plant cell walls: from chemistry to biology* (New York: Garland Science (Taylor & Francis Group). doi: 10.1086/662480

Allison, G. G., Morris, C., Clifton-Brown, J., Lister, S. J., and Donnison, I. S. (2011). Genotypic variation in cell wall composition in a diverse set of 244 accessions of miscanthus. *Biomass Bioenergy* 35, 4740–4747. doi: 10.1016/j.biombioe.2011.10.008

Baez, L. A., Tichá, T., and Hamann, T. (2022). Cell wall integrity regulation across plant species. *Plant Mol. Biol.* 109, 483–504. doi: 10.1007/s11103-022-01284-7

Barnes, W. J., and Anderson, C. T. (2017). Acetyl bromide soluble lignin (ABSL) assay for total lignin quantification from plant biomass. *Bio-Protocol* 7, e2149. doi: 10.21769/BioProtoc.2149

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

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Soft.* 67, 1–48. doi: 10.18637/jss.v067.i01

Belmokhtar, N., Arnoult, S., Chabbert, B., Charpentier, J.-P., and Brancourt-Hulmel, M. (2017). Saccharification performances of miscanthus at the pilot and miniaturized assay scales: genotype and year variabilities according to the biomass composition. *Front. Plant Sci.* 8, 740. doi: 10.3389/fpls.2017.00740

Bilska-Kos, A., Pietrusińska, A., Suski, S., Niedziela, A., Linkiewicz, A. M., Majtkowski, W., et al. (2022). Cell wall properties determine genotype-specific response to cold in *Miscanthus* × *giganteus* plants. *Cells* 11, 547. doi: 10.3390/cells11030547

Brancourt-Hulmel, M., Arnoult, S., Cézard, L., El Hage, F., Gineau, E., Girones, J., et al. (2022). A comparative study of maize and miscanthus regarding cell-wall composition and stem anatomy for conversion into bioethanol and polymer composites. *Bioenerg. Res.* 15, 777–791. doi: 10.1007/s12155-020-10239-z

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

Clifton-Brown, J., Harfouche, A., Casler, M. D., Dylan Jones, H., Macalpine, W. J., Murphy-Bokern, D., et al. (2019a). 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

Clifton-Brown, J., Schwarz, K. U., Awty-Carroll, D., Iurato, A., Meyer, H., Greef, J., et al. (2019b). Breeding strategies to improve miscanthus as a sustainable source of biomass for bioenergy and biorenewable products. *Agronomy* 9, 673. doi: 10.3390/agronomy9110673

da Costa, R., Allison, G., and Bosch, M. (2015). Cell wall biomass preparation and Fourier transform mid-infrared (FTIR) spectroscopy to study cell wall composition. *Bio-Protocol* 5 (11), e1494. doi: 10.21769/BioProtoc.1494

da Costa, R. M. F., Lee, S. J., Allison, G. G., Hazen, S. P., Winters, A., and Bosch, M. (2014). Genotype, development and tissue-derived variation of cell-wall properties in the lignocellulosic energy crop miscanthus. *Ann. Bot.* 114, 1265–1277. doi: 10.1093/aob/mcu054

da Costa, R. M. F., Pattathil, S., Avci, U., Lee, S. J., Hazen, S. P., Winters, A., et al. (2017). A cell wall reference profile for miscanthus bioenergy crops highlights compositional and structural variations associated with development and organ origin. *New Phytol.* 213, 1710–1725. doi: 10.1111/nph.14306

da Costa, R. M. F., Pattathil, S., Avci, U., Winters, A., Hahn, M. G., and Bosch, M. (2019). Desirable plant cell wall traits for higher-quality miscanthus lignocellulosic biomass. *Biotechnol. Biofuels* 12, 1–18. doi: 10.1186/s13068-019-1426-7

DeMartini, J. D., Pattathil, S., Miller, J. S., Li, H., Hahn, M. G., and Wyman, C. E. (2013). Investigating plant cell wall components that affect biomass recalcitrance in poplar and switchgrass. *Energy Environ. Sci.* 6, 898–909. doi: 10.1039/C3EE23801F

De Souza, A. P., Kamei, C. L. A. A., Torres, A. F., Pattathil, S., Hahn, M. G., Trindade, L. M., et al. (2015). How cell wall complexity influences saccharification efficiency in *Miscanthus sinensis*. *J. Exp. Bot.* 66, 4351–4365. doi: 10.1093/jxb/erv183

De Vega, J., Donnison, I., Dyer, S., and Farrar, K. (2021). Draft genome assembly of the biofuel grass crop *Miscanthus sacchariflorus*. *F1000Res* 10, 29. doi: 10.12688/f1000research.44714.1

Domon, J. M., Baldwin, L., Acket, S., Caudeville, E., Arnoult, S., Zub, H., et al. (2013). Cell wall compositional modifications of miscanthus ecotypes in response to cold acclimation. *Phytochemistry* 85, 51–61. doi: 10.1016/j.phytochem.2012.09.001

Fao/Iiasa/Isric/Isscas/Jrc (2009). *Harmonized world soil database (version 1.1)* (Rome, Italy and IIASA, Laxenburg, Austria: FAO).

Fick, S. E., and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatology* 37, 4302–4315. doi: 10.1002/joc.5086

Foster, C. E., Martin, T. M., and Pauly, M. (2010). Comprehensive compositional analysis of plant cell walls (lignocellulosic biomass) part I: lignin. *J. Visualized Experiments* 37, e1745. doi: 10.3791/1837

Fry, S. C. (1988). "The growing plant cell wall: chemical and metabolic analysis," in *Longman scientific & technical* (Caldwell: The Blackburn Press). Available at: https://books.google.co.uk/books?id=WITwAAAAMAAJ.

Gbif.Org (2019). *Occurrence download*. 454461. doi: 10.15468/DL.KTFBBB

Gladala-Kostarz, A., Doonan, J. H., and Bosch, M. (2020). Mechanical stimulation in *Brachypodium distachyon*: implications for fitness, productivity, and cell wall properties. *Plant Cell Environ.* 43, 1314–1330. doi: 10.1111/pce.13724

Głowacka, K., Clark, L. V., Adhikari, S., Peng, J., Stewart, J. R., Nishiwaki, A., et al. (2015). Genetic variation in *Miscanthus* × *giganteus* and the importance of estimating genetic distance thresholds for differentiating clones. *GCB Bioenergy* 7, 386–404. doi: 10.1111/gcbb.12166

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatology* 25, 1965–1978. doi: 10.1002/joc.1276

Hijmans, R. J., and Elith, J. (2013). *Species distribution modeling with R. R Cran Project*.

Hijmans, R. J., and Spooner, D. M. (2001). Geographic distribution of wild potato species. *Am. J. Bot.* 88, 2101–2112. doi: 10.2307/3558435

Hoang, N. V., Furtado, A., Botha, F. C., Simmons, B. A., and Henry, R. J. (2015). Potential for genetic improvement of sugarcane as a source of biomass for biofuels. *Front. Bioeng. Biotechnol.* 3, 182. doi: 10.3389/fbioe.2015.00182

Hodkinson, T. R., and Renvoize, S. (2001). Nomenclature of *Miscanthus* x *giganteus* (*Poaceae*). *Kew Bull.* 56, 759. doi: 10.2307/4117709

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical J.* 50, 346–363. doi: 10.1002/bimj.200810425

Huang, L. S., Flavell, R., Donnison, I. S., Chiang, Y.-C., Hastings, A., Hayes, C., et al. (2019). Collecting wild miscanthus germplasm in Asia for crop improvement and conservation in Europe whilst adhering to the guidelines of the united nations' convention on biological diversity. *Ann. Bot.* 124, 591–604. doi: 10.1093/aob/mcy231

Hyman, G., Hodson, D., and Jones, P. (2013). Spatial analysis to support geographic targeting of genotypes to environments. *Front. Physiol.* 4, 40. doi: 10.3389/fphys.2013.00040

Jombart, T. (2008). Adegenet: a r package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., Balloux, F., Falush, D., Stephens, M., Pritchard, J., et al. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11, 94. doi: 10.1186/1471-2156-11-94

Kalinina, O., Nunn, C., Sanderson, R., Hastings, A. F. S., van der Weijde, T., Özgüven, M., et al. (2017). Extending miscanthus cultivation with novel germplasm at six contrasting sites. *Front. Plant Sci.* 8, 563. doi: 10.3389/fpls.2017.00563

Kang, S., Post, W. M., Nichols, J. A., Wang, D., West, T. O., Bandaru, V., et al. (2013). Marginal lands: concept, assessment and management. *J. Agric. Sci.* 5, 129–139. doi: 10.5539/jas.v5n5p129

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Software* 82, 1–26. doi: 10.18637/jss.v082.i13

Le Gall, H., Philippe, F., Domon, J.-M., Gillet, F., Pelloux, J., and Rayon, C. (2015). Cell wall metabolism in response to abiotic stress. *Plants* 4, 112–166. doi: 10.3390/plants4010112

Lenth, R. V. (2022) *Emmeans: estimated marginal means, aka least-squares means*. Available at: https://CRAN.R-project.org/package=emmeans.

Lewandowski, I., Clifton-Brown, J., Trindade, L. M., van der Linden, G. C., Schwarz, K.-U., Müller-Sämann, K., et al. (2016). Progress on optimizing miscanthus biomass production for the European bioeconomy: results of the EU FP7 project OPTIMISC. *Front. Plant Sci.* 7, 1620. doi: 10.3389/fpls.2016.01620

Li, X., Liao, H., Fan, C., Hu, H., Li, Y., Li, J., et al. (2016). Distinct geographical distribution of the miscanthus accessions with varied biomass enzymatic saccharification. *PLoS One* 11, e0160026. doi: 10.1371/journal.pone.0160026

Lu, F., Lipka, A. E., Glaubitz, J., Elshire, R., Cherney, J. H., Casler, M. D., et al. (2013). Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genet.* 9 (1), e1003215. doi: 10.1371/journal.pgen.1003215

Lygin, A. V., Upton, J., Dohleman, F. G., Juvik, J., Zabotina, O. A., Widholm, J. M., et al. (2011). Composition of cell wall phenolics and polysaccharides of the potential bioenergy crop miscanthus. *GCB Bioenergy* 3, 333–345. doi: 10.1111/j.1757-1707.2011.01091.x

Maji, A. T., and Shaibu, A. A. (2012). Application of principal component analysis for rice germplasm characterization and evaluation. *J. Plant Breed. Crop Sci.* 4, 87–93. doi: 10.5897/JPBCS11.093

Malinowska, M., Donnison, I. S., and Robson, P. R. H. (2017). Phenomics analysis of drought responses in miscanthus collected from different geographical locations. *GCB Bioenergy* 9, 78–91. doi: 10.1111/gcbb.12350

Mangold, A., Lewandowski, I., Möhring, J., Clifton-Brown, J., Krzyżak, J., Mos, M., et al. (2019). Harvest date and leaf:stem ratio determine methane hectare yield of miscanthus biomass. *GCB Bioenergy* 11, 21–33. doi: 10.1111/gcbb.12549

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

Mitros, T., Session, A. M., James, B. T., Wu, G. A., Belaffif, M. B., Clark, L. V., et al. (2020). Genome biology of the paleotetraploid perennial biomass crop miscanthus. *Nat. Commun.* 11, 5442. doi: 10.1038/s41467-020-18923-6

Monti, A., Di Virgilio, N., and Venturi, G. (2008). Mineral composition and ash content of six major energy crops. *Biomass Bioenergy* 32, 216–223. doi: 10.1016/j.biombioe.2007.09.012

Moura, J. C. M. S., Bonine, C. A. V., de Oliveira Fernandes Viana, J., Dornelas, M. C., and Mazzafera, P. (2010). Abiotic and biotic stresses and changes in the lignin content and composition in plants. *J. Integr. Plant Biol.* 52, 360–376. doi: 10.1111/j.1744-7909.2010.00892.x

Ohlsson, J. A., Hallingbäck, H. R., Jebrane, M., Harman-Ware, A. E., Shollenberger, T., Decker, S. R., et al. (2019). Genetic variation of biomass recalcitrance in a natural salix viminalis (L.) population. *Biotechnol. Biofuels* 12, 135. doi: 10.1186/s13068-019-1479-7

Oshunsanya, S. O., and Aliku, O. (2016). "GIS applications in agronomy," in *Geospatial technology - environmental and social applications*. Eds. P. Imperatore and A. Pepe (London, UK: InTech). doi: 10.5772/64528

Ostos Garrido, F. J., Pistón, F., Gómez, L. D., and McQueen-Mason, S. J. (2018). Biomass recalcitrance in barley, wheat and triticale straw: correlation of biomass quality with classic agronomical traits. *PLoS One* 13, e0205880. doi: 10.1371/journal.pone.0205880

Oyedeji, O., Langholtz, M., Hellwinckel, C., and Webb, E. (2021). Supply analysis of preferential market incentive for energy crops. *Biofuels Bioprod. Bioref.* 15, 736–748. doi: 10.1002/bbb.2184

Pancaldi, F., and Trindade, L. M. (2020). Marginal lands to grow novel bio-based crops: a plant breeding perspective *Front. Plant Sci.* 11, 227. doi: 10.3389/fpls.2020.00227

Parra-Quijano, M., Iriondo, J. M., and Torres, E. (2012). Review. applications of ecogeography and geographic information systems in conservation and utilization of plant genetic resources. *Span J. Agric. Res.* 10, 419. doi: 10.5424/sjar/2012102-303-11

Penn, C., and Camberato, J. (2019). A critical review on soil chemical processes that control how soil pH affects phosphorus availability to plants. *Agriculture* 9, 120. doi: 10.3390/agriculture9060120

Petit, J., Gulisano, A., Dechesne, A., and Trindade, L. M. (2019). Phenotypic variation of cell wall composition and stem morphology in hemp (*Cannabis sativa* l.): optimization of methods. *Front. Plant Sci.* 10, 959. doi: 10.3389/fpls.2019.00959

Pettolino, F. A., Walsh, C., Fincher, G. B., and Bacic, A. (2012). Determining the polysaccharide composition of plant cell walls. *Nat. Protoc.* 7, 1590–1606. doi: 10.1038/nprot.2012.081

Qin, J., Zhao, H., Dai, M., Zhao, P., Chen, X., Liu, H., et al. (2022). Speciation distribution and influencing factors of heavy metals in rhizosphere soil of *Miscanthus floridulus* in the tailing reservoir area of dabaoshan iron polymetallic mine in northern guangdong. *Processes* 10, 1217. doi: 10.3390/pr10061217

Ramírez-Ojeda, G., Peralta, I. E., Rodríguez-Guzmán, E., Chávez-Servia, J. L., Sahagún-Castellanos, J., and Rodríguez-Pérez, J. E. (2021). Climatic diversity and ecological descriptors of wild tomato species (*Solanum* sect. *Lycopersicon*) and close related species (*Solanum* sect. *Juglandifolia* and sect. *Lycopersicoides*) in Latin America. *Plants* 10, 855. doi: 10.3390/plants10050855

Rancour, D. M., Marita, J. M., and Hatfield, R. D. (2012). Cell wall composition throughout development for the model grass *Brachypodium distachyon*. *Front. Plant Sci.* 3, 266. doi: 10.3389/fpls.2012.00266

R Core Team. (2018) *R: a language and environment for statistical computing*. Available at: https://www.r-project.org/.

Resch, M. G., Baker, J. O., and Decker, S. R. (2015). Low solids enzymatic saccharification of lignocellulosic biomass. *Tech. Rep. NREL/TP-5100-63351 Lab. Analytical Procedure (LAP)*, 1–9.

Roem, W. J., and Berendse, F. (2000). Soil acidity and nutrient supply ratio as possible factors determining changes in plant species diversity in grassland and

heathland communities. *Biol. Conserv.* 92, 151–161. doi: 10.1016/S0006-3207(99)00049-X

Saeman, J. F. (1945). Kinetics of wood saccharification - hydrolysis of cellulose and decomposition of sugars in dilute acid at high temperature. *Ind. Eng. Chem.* 37, 43–52. doi: 10.1021/ie50421a009

Santoro, N., Cantu, S. L., Tornqvist, C. E., Falbel, T. G., Bolivar, J. L., Patterson, S. E., et al. (2010). A high-throughput platform for screening milligram quantities of plant biomass for lignocellulose digestibility. *Bioenergy Res.* 3, 93–102. doi: 10.1007/s12155-009-9074-6

Shimono, Y., Kurokawa, S., Nishida, T., Ikeda, H., and Futagami, N. (2013). Phylogeography based on intraspecific sequence variation in chloroplast DNA of *Miscanthus sinensis* (Poaceae), a native pioneer grass in Japan. *Botany* 91, 449–456. doi: 10.1139/cjb-2012-0212

Simmonds, N. W. (1972). Genetic resources in plants - their exploration and conservation. edited by o. h. Frankel and e. Bennett Oxford: Blackwell scientific publications (1970). *Ex. Agric.* 8, 87–87. doi: 10.1017/S0014479700023553

Slavov, G., Allison, G., and Bosch, M. (2013b). Advances in the genetic dissection of plant cell walls: tools and resources available in miscanthus. *Front. Plant Sci.* 4, 217. doi: 10.3389/fpls.2013.00217

Slavov, G. T., Nipper, R., Robson, P., Farrar, K., Allison, G. G., Bosch, M., et al. (2014). Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass *Miscanthus sinensis*. *New Phytol.* 201, 1227–1239. doi: 10.1111/nph.12621

Slavov, G., Robson, P., Jensen, E., Hodgson, E., Farrar, K., Allison, G., et al. (2013a). Contrasting geographic patterns of genetic variation for molecular markers vs. phenotypic traits in the energy grass *Miscanthus sinensis*. *GCB Bioenergy* 5, 562–571. doi: 10.1111/gcbb.12025

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., et al. (2008). Determination of structural carbohydrates and lignin in biomass. *Lab. Anal. Proced.* 1617 (1), 1–16.

Stark, S., Eskelinen, A., and Männistö, M. K. (2012). Regulation of microbial community composition and activity by soil nutrient availability, soil pH, and herbivory in the tundra. *Ecosystems* 15, 18–33. doi: 10.1007/s10021-011-9491-1

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

Tamura, K., Uwatoko, N., Yamashita, H., Fujimori, M., Akiyama, Y., Shoji, A., et al. (2016). Discovery of natural interspecific hybrids between *Miscanthus sacchariflorus* and *Miscanthus sinensis* in southern Japan: morphological characterization, genetic structure, and origin. *Bioenergy Res.* 9, 315–325. doi: 10.1007/s12155-015-9683-1

Tang, C., Yang, X., Chen, X., Ameen, A., and Xie, G. (2018). Sorghum biomass and quality and soil nitrogen balance response to nitrogen rate on semiarid marginal land. *Field Crops Res.* 215, 12–22. doi: 10.1016/j.fcr.2017.09.031

Teklemariam, S. S., Bayissa, K. N., Matros, A., Pillen, K., Ordon, F., and Wehner, G. (2022). The genetic diversity of Ethiopian barley genotypes in relation to their geographical origin. *PLoS One* 17, e0260422. doi: 10.1371/journal.pone.0260422

The Plant List. (2013). *The plant list. version 1.1.* Available at: http://www.theplantlist.org/.

Torres, A. F., Xu, X., Nikiforidis, C., Bitter, J. H., and Trinidade, L. M. (2019). Exploring the treasure of plant molecules with integrated biorefineries. *Front. Plant. Sci.* 10, 478. doi: 10.3389/fpls.2019.00478

Vaahtera, L., Schulz, J., and Hamann, T. (2019). Cell wall integrity maintenance during plant development and interaction with the environment. *Nat. Plants* 5, 924–932. doi: 10.1038/s41477-019-0502-0

van der Weijde, T., Dolstra, O., Visser, R. G. F., and Trindade, L. M. (2017a). Stability of cell wall composition and saccharification efficiency in miscanthus across diverse environments. *Front. Plant Sci.* 7, 2004. doi: 10.3389/fpls.2016.02004

van der Weijde, T., Kamei, C. L. A., Severing, E. I., Torres, A. F., Gomez, L. D., Dolstra, O., et al. (2017b). Genetic complexity of miscanthus cell wall composition and biomass quality for biofuels. *BMC Genomics* 18, 1–15. doi: 10.1186/s12864-017-3802-7

Van Soest, P. J. (1963). Use of detergents in the analysis of fibrous feeds. 2. a rapid method for the determination of fiber and lignin. *J. Assoc. Off. Agric. Chem.* 46, 829–835. doi: 10.1093/jaoac/46.5.829

Van Soest, P. J., Robertson, J. B., and Lewis, B. A. (1991). Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *J. Dairy Sci.* 74, 3583–3597. doi: 10.3168/jds.S0022-0302(91)78551-2

Von Cossel, M., Lewandowski, I., Elbersen, B., Staritsky, I., Van Eupen, M., Iqbal, Y., et al. (2019). Marginal agricultural land low-input systems for biomass production. *Energies* 12, 3123. doi: 10.3390/en12163123

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., et al. (2012). Darwin Core: an evolving community-developed biodiversity data standard. *PLoS One* 7 (1), e29715. doi: 10.1371/journal.pone.0029715

Xiang, W., Xue, S., Liu, F., Qin, S., Xiao, L., and Yi, Z. (2020). MGDB: a database for evaluating *Miscanthus* spp. to screen elite germplasm. *Biomass Bioenergy* 138, 105599. doi: 10.1016/j.biombioe.2020.105599

Xu, P., Cheng, S., Han, Y., Zhao, D., Li, H., Wang, Y., et al. (2020). Natural variation of lignocellulosic components in miscanthus biomass in China. *Front. Chem.* 8, 595143. doi: 10.3389/fchem.2020.595143

Yang, Y., Reilly, E. C., Jungers, J. M., Chen, J., and Smith, T. M. (2019). Climate benefits of increasing plant diversity in perennial bioenergy crops. *One Earth* 1, 434–445. doi: 10.1016/j.oneear.2019.11.011

Yang, S., Xue, S., Kang, W., Qian, Z., and Yi, Z. (2019). Genetic diversity and population structure of miscanthus lutarioriparius, an endemic plant of China. *PLoS One* 14, e0211471. doi: 10.1371/journal.pone.0211471

Yoo, C. G., Dumitrache, A., Muchero, W., Natzke, J., Akinosho, H., Li, M., et al. (2018). Significance of lignin S/G ratio in biomass recalcitrance of *Populus trichocarpa* variants for bioethanol production. *ACS Sustain. Chem. Eng.* 6, 2162–2168. doi: 10.1021/acssuschemeng.7b03586

Zhang, G., Ge, C., Xu, P., Wang, S., Cheng, S., Han, Y., et al. (2021). The reference genome of *Miscanthus floridulus* illuminates the evolution of *Saccharinae*. *Nat. Plants* 7, 608–618. doi: 10.1038/s41477-021-00908-y

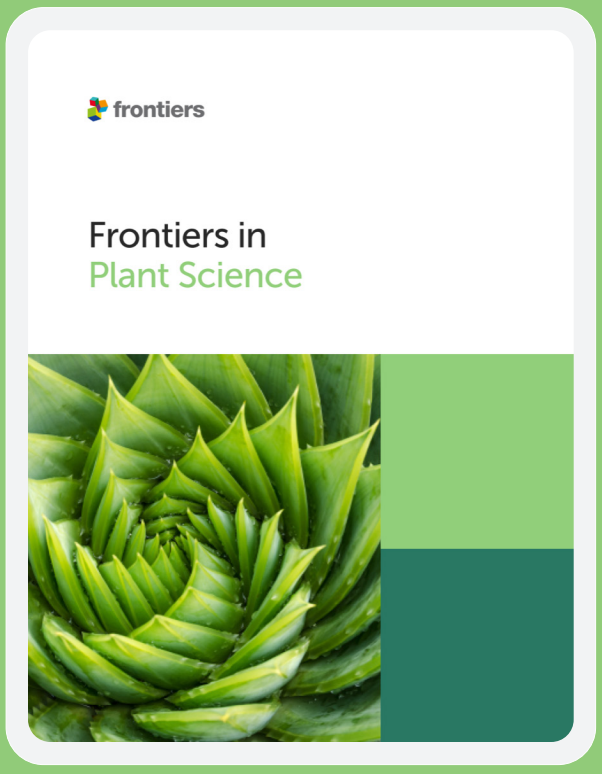
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact