

Lawrence Berkeley National Laboratory

LBL Publications

Title

Overexpression of REDUCED WALL ACETYLATION C increases xylan acetylation and biomass recalcitrance in Populus.

Permalink

<https://escholarship.org/uc/item/7d68x2sc>

Journal

Plant Physiology, 194(1)

Authors

Zhang, Jin

Wang, Xiaqin

Wang, Hsin-Tzu

et al.

Publication Date

2023-12-30

























DOI

10.1093/plphys/kiad377

Peer reviewed



Overexpression of *REDUCED WALL ACETYLATION C* increases xylan acetylation and biomass recalcitrance in *Populus*

Jin Zhang ^{1,2,3,*} Xiaqin Wang ¹ Hsin-Tzu Wang ⁴ Zhenzhen Qiao ^{2,3} Tao Yao ^{2,3} Meng Xie ⁵ Breeanna R. Urbanowicz ^{4,6} Wei Zeng ¹ Sara S. Jawdy ^{2,3} Lee E. Gunter ^{2,3} Xiaohan Yang ^{2,3} Olaf Czarnecki ² Sharon Regan ⁷ Armand Seguin ⁸ William Rottmann ⁹ Kimberly A. Winkeler ⁹ Robert Sykes ¹⁰ Anna Lipzen ¹¹ Chris Daum ¹¹ Kerrie Barry ¹¹ Meng-Zhu Lu ¹ Gerald A. Tuskan ^{2,3} Wellington Muchero ^{2,3,*} and Jin-Gui Chen ^{2,3,*}

- 1 State Key Laboratory of Subtropical Silviculture, College of Forestry and Biotechnology, Zhejiang A&F University, Hangzhou, Zhejiang 311300, China
- 2 Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
- 3 Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
- 4 Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA
- 5 Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA
- 6 Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA
- 7 Biology Department, Queen's University, Kingston, Ontario K7L 3N6, Canada
- 8 Laurentian Forestry Center, Natural Resources Canada, Québec, Quebec G1V 4C7, Canada
- 9 ArborGen Inc., Ridgeville, SC 29472, USA
- 10 Bioenergy Science and Technology, National Renewable Energy Laboratory, Golden, CO 80401, USA
- 11 Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

*Author for correspondence: zhangj@zafu.edu.cn (J.Z.), chenj@ornl.gov (J.-G.C.), mucherow@ornl.gov (W.M.)

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plphys/pages/General-Instructions>) are Jin Zhang and Jin-Gui Chen.

Abstract

Plant lignocellulosic biomass, i.e. secondary cell walls of plants, is a vital alternative source for bioenergy. However, the acetylation of xylan in secondary cell walls impedes the conversion of biomass to biofuels. Previous studies have shown that *REDUCED WALL ACETYLATION* (RWA) proteins are directly involved in the acetylation of xylan but the regulatory mechanism of RWAs is not fully understood. In this study, we demonstrate that overexpression of a *Populus trichocarpa* *PtRWA-C* gene increases the level of xylan acetylation and increases the lignin content and S/G ratio, ultimately yielding poplar woody biomass with reduced saccharification efficiency. Furthermore, through gene coexpression network and expression quantitative trait loci (eQTL) analysis, we found that *PtRWA-C* was regulated not only by the secondary cell wall hierarchical regulatory network but also by an AP2 family transcription factor *HARDY* (HRD). Specifically, HRD activates *PtRWA-C* expression by directly binding to the *PtRWA-C* promoter, which is also the *cis*-eQTL for *PtRWA-C*. Taken together, our findings provide insights into the functional roles of *PtRWA-C* in xylan acetylation and consequently saccharification and shed light on synthetic biology approaches to manipulate this gene and alter cell wall properties. These findings have substantial implications for genetic engineering of woody species, which could be used as a sustainable source of biofuels, valuable biochemicals, and biomaterials.

Received March 07, 2023. Accepted June 29, 2023. Advance access publication July 3, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of American Society of Plant Biologists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Open Access

Introduction

Plant lignocellulosic biomass is one of the important alternative sources of bioenergy to meet growing energy use demands and alleviate our dependence on fossil fuels (Carroll and Somerville 2009; Chen et al. 2020). The majority of plant lignocellulosic biomass is found in the carbohydrate-rich walls that surround every plant cell. Some cells deposit a thickened secondary wall after elongation ceases, typically consisting of lignin, cellulose, and hemicellulose. Hemicelluloses are a group of diverse polysaccharides that include xylans, xyloglucans, (gluco)mannans, and mixed-linked glucans (Scheller and Ulvskov 2010). Recently, multidimensional analysis of plant biomass using solid-state nuclear magnetic resonance (NMR) spectroscopy has begun to reveal molecular-level information regarding the organization of native cell walls and shedding light on structure-function relationships of polymer classes that impact the 3D molecular architecture of lignocellulose (Kirui et al. 2022).

Almost all noncellulosic polysaccharides in the cell walls of dicots are esterified to varying degrees with *O*-acetyl moieties. Xylan in the secondary cell walls of woody plants can be highly acetylated, which affects its physicochemical properties (Ebringerová 2005). Investigation of ¹³C-labeled poplar (*Populus × canadensis*) stems to study contacts between polymers in spatially packed woody biomass identified key intramolecular interactions between xylan acetyl groups with lignin and cellulose in a xylan conformation-dependent manner (Kirui et al. 2022). Unrestrained molecular dynamics (MD) simulations of acetylated xylan further supports the notion that both the spacing and position of substituents on xylosyl residues play key roles in tuning xylan-cellulose interactions (Gupta et al. 2021). Acetylation of xylans and other polysaccharides has implications for the use of biomass as feedstocks for biofuel production, and the presence of acetyl groups has a negative impact on saccharification and fermentation (Selig et al. 2009; Jönsson et al. 2013). Deacetylation of lignocellulose by chemical or enzymatic methods increases sugar release during enzymatic hydrolysis (Zhang et al. 2011). Taken together, understanding the identity, function, and genomic and biochemical regulation of enzymes involved in polysaccharide *O*-acetylation can give insights into how to manipulate wood architecture for the more facile production of next-generation fuels and materials.

The acetylation and deacetylation of polysaccharides are regulated by complex enzymatic reactions. During polysaccharide biosynthesis in the Golgi apparatus, acetyl groups are transferred from an as yet to be definitively confirmed *O*-acetyl donor to polysaccharides with the participation of the REDUCED WALL ACETYLATION (RWAs), ALTERED XYLOGLUCAN 9 (AXY9), and TRICHOME BIREFRINGENCY-LIKE (TBL) gene families (Gille et al. 2011; Manabe et al. 2013). The *Arabidopsis* (*Arabidopsis thaliana*) *axy9* mutant has reduced acetylation of xylan and xyloglucan (Schultink et al. 2015). Similarly, mutants of *TBL29* (*ESKIMO1*), *TBL32*, and *TBL33* exhibit an irregular xylem phenotype and

decreased xylan acetylation (Xiong et al. 2013; Yuan et al. 2016). In contrast, the deacetylation of polysaccharides can be catalyzed by carbohydrate esterases. An *Arabidopsis* clade IId member of the GDSL esterase/lipase (GELP) family, AtGELP7 (also named as acetyl xylan esterase1, [AtAXE1]), is regulated by secondary cell wall (SCW) master switch AtMYB46 and has acetyl xylan esterase (AXE) activity and functions to reduce xylan acetylation levels (Rastogi et al. 2022). In addition, the use of biotechnology to introduce related enzymes from other species into plants has also become an effective strategy for altering plant acetylation levels. Pawar et al. (2016) expressed AnAXE1 from *Aspergillus niger*, a Carbohydrate Esterase family 1 (CE1) enzyme that deacetylates polymeric xylan but not pectin, to deacetylate xylan in *Arabidopsis* and improved the sugar saccharification efficiency and ethanol yields. Expression of the *Hypocrea jecorina* CE5 family gene *HjAXE* in poplar driven by a wood-specific promoter also effectively reduces wood acetylation levels and improves cellulose accessibility (Wang et al. 2020).

Members of the RWA family were the first class of proteins reported to be involved in plant cell wall acetylation (Manabe et al. 2011). In *Arabidopsis*, 4 RWA genes (*AtRWA1-4*) are regulated by the secondary cell wall master regulator SND1 and are functionally and redundantly associated with secondary wall thickening. The quadruple *rwa* mutant has cell walls with a significant reduction in the acetylation level of xylan and a defect in secondary wall thickening (Lee et al. 2011). The RWA gene family in *Populus* also has 4 genes (*PtRWA-A*, *-B*, *-C*, and *-D*), which are grouped into 2 clades and have been shown to also be functionally and redundantly involved in wood acetylation. In addition, suppression of the 4 members resulted in improved saccharification efficiency (Pawar et al. 2017b).

In this study, a *Populus* activation-tagged mutant with substantial changes in lignin content and S/G ratio was identified and TAIL-PCR confirmed that the *PtRWA-C* gene was activated in this dominant mutant. Analysis of the coexpression network of *PtRWA-C* reveals that several known secondary cell wall-related transcription factors (TFs) such as *NST1*, *SND2*, and *MYB46/83* are coexpressed with *PtRWA-C*, implying a potential role of *PtRWA-C* in secondary cell wall biosynthesis. We discovered that overexpression of *PtRWA-C* does not affect growth and biomass but does affect stem acetylation level and transcriptome. Expression QTL (eQTL) analysis and yeast 1-hybrid (Y1H) assay indicate that *PtRWA-C* is directly regulated by an AP2 family transcription factor HARDY (HRD). These results provided insights into the function of *PtRWA-C* in cell wall modification and its regulatory pathway.

Results

The *rwa-c-d* mutant changes cell wall chemistry phenotypes

To identify the regulators of plant secondary cell wall biosynthesis, a *Populus* activation-tagged population (generated by

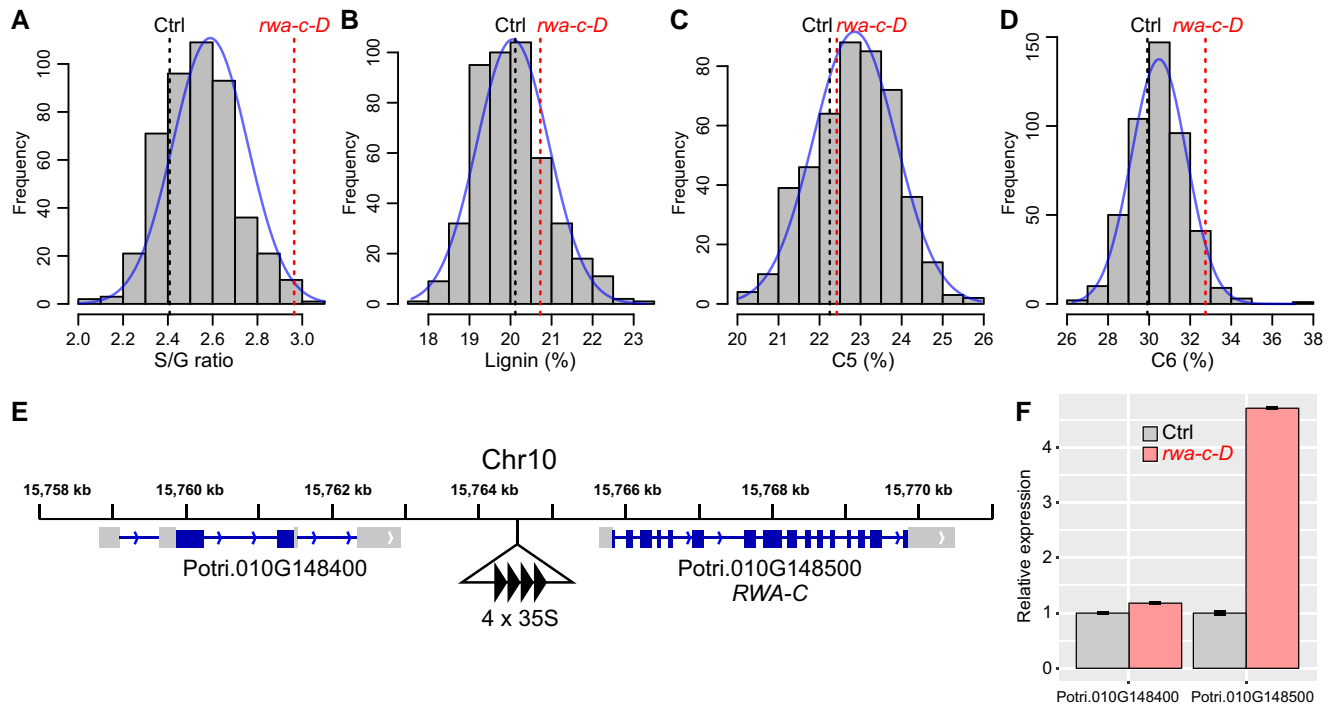


Figure 1. Wood properties of activation-tagged poplar population and the identification of the activation-tagged T-DNA insertion site in the *rwa-c-D* mutant. The distribution of wood properties of the **A**) syringyl/guaiacyl (S/G) lignin ratio, **B**) lignin content, **C**) C5 sugar, and **D**) C6 sugar of the poplar activation-tagged population consisting of 463 independent transgenic lines. The dash lines labelled “Ctrl3” and “*rwa-c-D3*” represent the values of wild type (WT) and “E8-33” (also labelled as “*rwa-c-D*”), respectively. **E**) The T-DNA insertion site in the *rwa-c-D* mutant. The orientation of the 4×35S enhancer repeats in the T-DNA situated 1261 bp upstream of the start codon of the *RWA-C* gene is indicated by arrowheads. **F**) Transcript levels of Potri.010G148400 and Potri.010G148500 in wild-type control (Ctrl) and *rwa-c-D* mutant. Error bars represent means \pm SE ($n = 3$).

transforming *Populus* with the T-DNA insertion containing 35S enhancer tetramer) (Busov et al. 2011) consisting of 463 independent transgenic lines was used to identify transgenic lines with altered cell wall chemistry phenotypes. From the mutagenized population, a gain-of-function mutant “E8-33” was identified with changes in lignin content and syringyl/guaiacyl (S/G) ratio. Specifically, the S/G ratio was increased by 20% in E8-33 compared with that in the wild-type control (Fig. 1, A to D).

In order to identify the location of the T-DNA insertion in the *rwa-c-D* mutant, we used thermal asymmetric interlaced (TAIL) PCR. Our data showed that the T-DNA insertion occurred in chromosome 10 at 1,261-bp upstream of the start codon of Potri.010G148500 (*PtRWA-C*), and 3,079-bp downstream of the stop codon of Potri.010G148400, with the 4 outward-facing 35S enhancers facing *PtRWA-C* (Fig. 1E). RT-qPCR results indicated that expression of *RWA-C*, but not Potri.10G148400, was elevated in the tagged mutant (Fig. 1F), confirming that the *RWA-C* gene was tagged. Thus, the mutant was named *rwa-c-D* (“D” denotes the dominant effect).

Characteristics of *PtRWA-C*

There are 4 *RWA* genes in the *Populus* genome (*PtRWA-A/B/C/D*) (Pawar et al. 2017b). To explore the expression patterns of the 4 *PtRWA* genes during stem development, public gene

expression databases related to poplar stem development were used for query. Among different organs and stem cell types, all the 4 *PtRWA* genes were highly expressed in stem differentiating xylem (SDX) and *PtRWA-B* and *PtRWA-C* were also highly expressed in xylem fiber cells (Fig. 2A). In the SDX protoplast dataset of transient overexpressing *PtSND1-B1*, only *PtRWA-C* was significantly induced at 7 h transfected by *PtSND1-B1* (Fig. 2B). The Aspwood database (<https://plantgenie.org/>) provides a high-resolution transcript profiling of poplar genes during wood formation. The 4 *PtRWA* genes were dramatically induced in secondary cell wall forming xylem. Notably, *PtRWA-C* exhibited a relatively high expression level in phloem, cambium, and expanding xylem compared to the other 3 *PtRWAs* (Fig. 2C).

To explore the potential function of *PtRWA-C*, we then constructed a coexpression network centered on *PtRWA-C* using the PSDX dataset. A total of 466 genes were included in the coexpression network, including 26 TFs, among which 16 are known secondary cell wall (SCW)-related TFs, e.g. *NST1*, *SND2*, *VND1*, *VND4*, *MYB46/83*, and *MYB69*. In addition, 45 SCW-related genes, including 10 *Irregular Xylem* genes (*IRX2/6/8/9/10/14-L*), 9 *Fasciclin-like Arabinogalactan-protein* genes (*FLA7/11/17*), 8 *TBLs* (*TBL3/25/29/31/33/34*), 6 *Cellulose Synthase* genes (*CesA3/4/7/8/9*), 5 *Laccase* genes (*LAC4/17*), 1 *CCoAOMT*, and 1 *PAL*, were included in the network (Fig. 2D and Supplemental Table S1). The coexpression

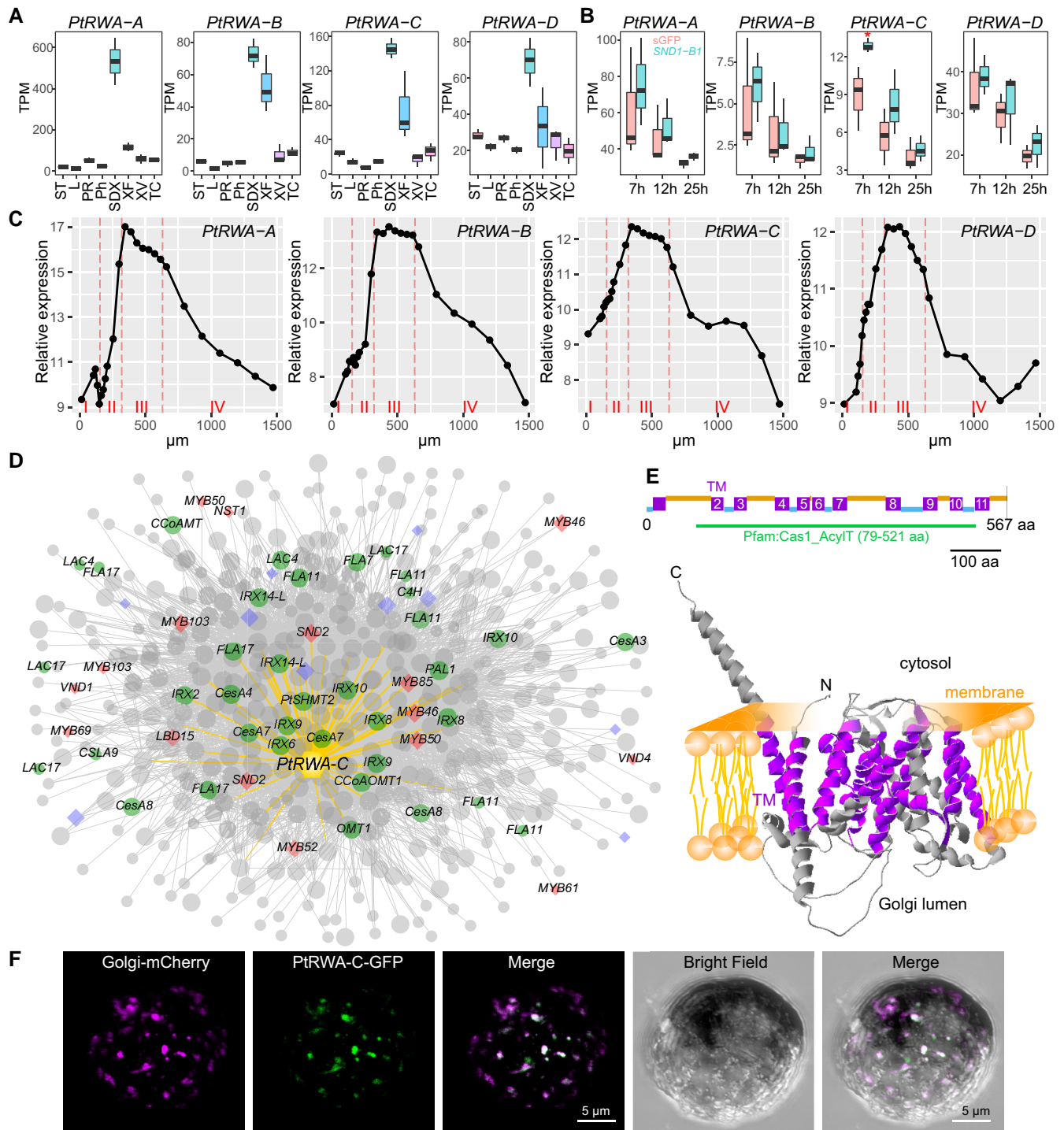


Figure 2. Characterization of the *PtRWA* genes. **A)** Normalized expression of the *PtRWA* genes in different organs and cell types. ST, shoot tips; L, leaves; PR, primary roots; Ph, phloem; SDX, stem differentiating xylem; XF, xylem fiber cells; XV, xylem vessel cells; TC, a combination of fiber, vessel, and ray cells; TPM, transcript per million. The bold line in the center of the boxplots represents the median, the box edges represent the 25th (lower) and 75th (upper) percentiles, and the whiskers extend to the most extreme data points that are no more than 1.5× the length of the upper or lower segment. **B)** Expression of the *PtRWA* genes in poplar SDX protoplasts at 7, 12, and 25 h transfected by overexpressing *PtrSND1-B1*-sGFP and sGFP (as control). * $P < 0.05$. **C)** Expression of the *PtRWA* genes during wood formation. The data were retrieved from the AspWood database (<https://plantgenie.org/>), where relative expression is shown for aspen stem samples, which consist primarily of phloem and cambium zone (I, in the bottom), expanding xylem (II), secondary cell wall forming xylem (III), and programmed cell death zone (IV). **D)** Coexpression network of *PtRWA-C*. Red and purple diamonds represent cell wall-related TFs and other TFs, respectively. Yellow, green, and

(continued)

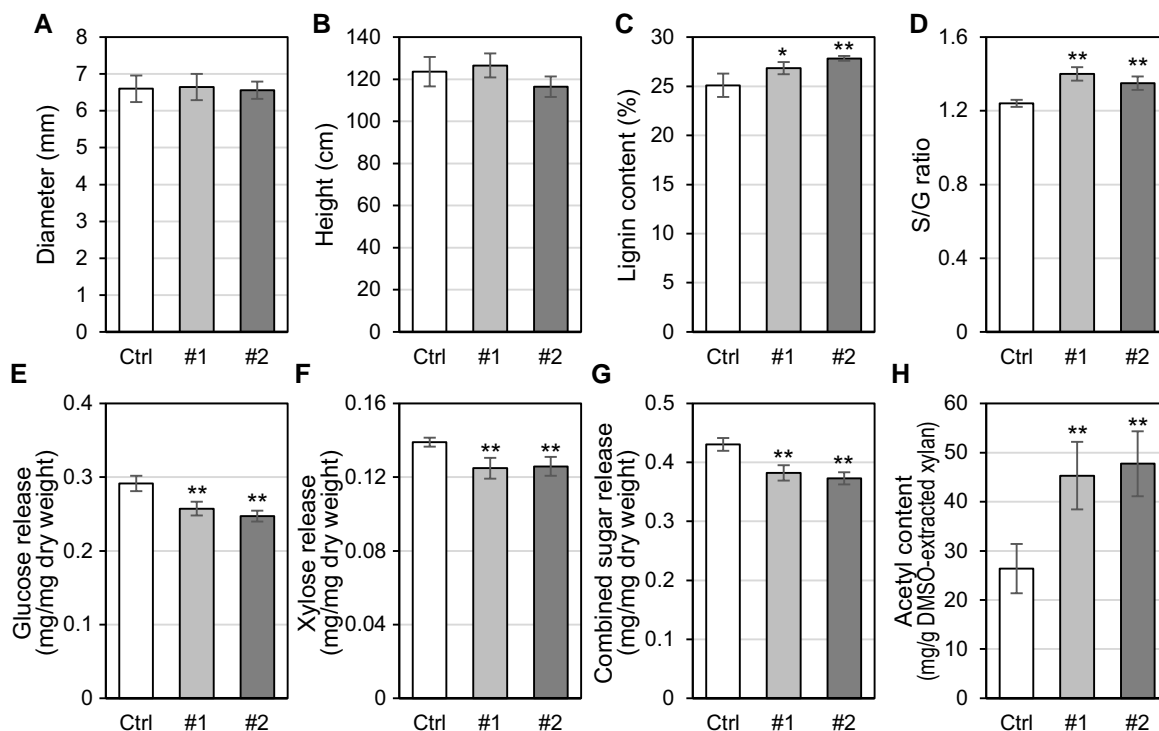


Figure 3. Growth status and cell wall characteristics of *Populus* overexpressing *PtRWA-C*. The diameter of basal stem **A**) and height **B**) were compared between transgenic lines, #1 and #2, and the control line (Ctrl). Lignin content **C**) and syringyl/guaiacyl (S/G) lignin ratio **D**) in *PtRWA-C* overexpression lines were compared to those in the Ctrl plants. Glucose **E**) and xylose **F**) release assay and the combined sugar release with glucose and xylose **G**) of *PtRWA-C* overexpression lines were compared with Ctrl plants. **H**) Acetyl contents in DMSO-extracted xylans of the transgenic and Ctrl plants. Error bars represent means \pm SE ($n = 3$). Asterisks indicate significant differences between transgenic and Ctrl plants based on Student's *t*-test (* $P < 0.05$ and ** $P < 0.01$).

network for *PtRWA-C* suggests that *PtRWA-C* is involved in secondary cell wall biogenesis.

In order to understand the functional organization of *PtRWA-C*, we used AlphaFold 2 to predict protein structure. *PtRWA-C* is 567 aa in length and contains 11 transmembrane (TM) helices based on TMHMM2.0 prediction. 3D structural simulations revealed that the 11 TM helices of *PtRWA-C* are embedded in the membrane lipid bilayer, while the N- and C-termini face the cytoplasm. Although the PFAM database shows that *PtRWA-C* has a Cas1_AcylT motif at 79 to 521 aa, *PtRWA-C* is similar to the *Arabidopsis* RWA proteins—both lack the extended loop containing Gly-Asp-Ser (GDS) and Asp-x-x-His (DxxH) motifs in *Cryptococcus* CnCas1p (Fig. 2E) (Gille and Pauly 2012). Because xylan is biosynthesized in Golgi, to test if the *PtRWA-C* is in fact a membrane protein responsible for xylan acetylation, GFP-fused *PtRWA-C* (*PtRWA-C*-GFP) was coexpressed with Golgi marker (Golgi-mCherry) in poplar protoplasts. The colocalization

of GFP and mCherry signals indicates that *PtRWA-C* is Golgi localized (Fig. 2F). Members of the RWA family have been hypothesized to catalyze translocation of acetyl groups across the Golgi membrane similarly to the mechanism used by the membrane-spanning acyltransferase-3 (AT-3) domain of the peptidoglycan *O*-acetyltransferase A (OatA) from *Staphylococcus* (Jones et al. 2021).

Overexpression of *PtRWA-C* increases the wood acetylation in poplar

To validate and characterize the function of *PtRWA-C* in *Populus*, the full-length CDS of *PtRWA-C* was overexpressed in *Populus deltoides* “WV94”. Two independent lines (#1 and #2) that had high *PtRWA-C* expression levels were selected for further analyses. Compared to control plants, both transgenic lines showed no significant difference in apical nor radial growth (Fig. 3, A and B). Based on pyrolysis

Figure 2. (Continued)

gray nodes represent target gene (*PtRWA-C*), cell wall-related genes, and other genes, respectively. Large, medium, and small nodes represent the target gene, the 1st (direct) coexpressed genes, and the 2nd (indirect) coexpressed genes, respectively. Yellow and gray edges represent the 1st and the 2nd coexpression relationships, respectively. **E**) Protein structure of *PtRWA-C*. Purple and green regions represent the TM domain and Pfam: Cas1_AcylT domain, respectively. **F**) Subcellular localization of *PtRWA-C* in *Populus* protoplast (green). The Golgi marker Golgi-mCherry is shown in purple. Bar = 5 μ m.

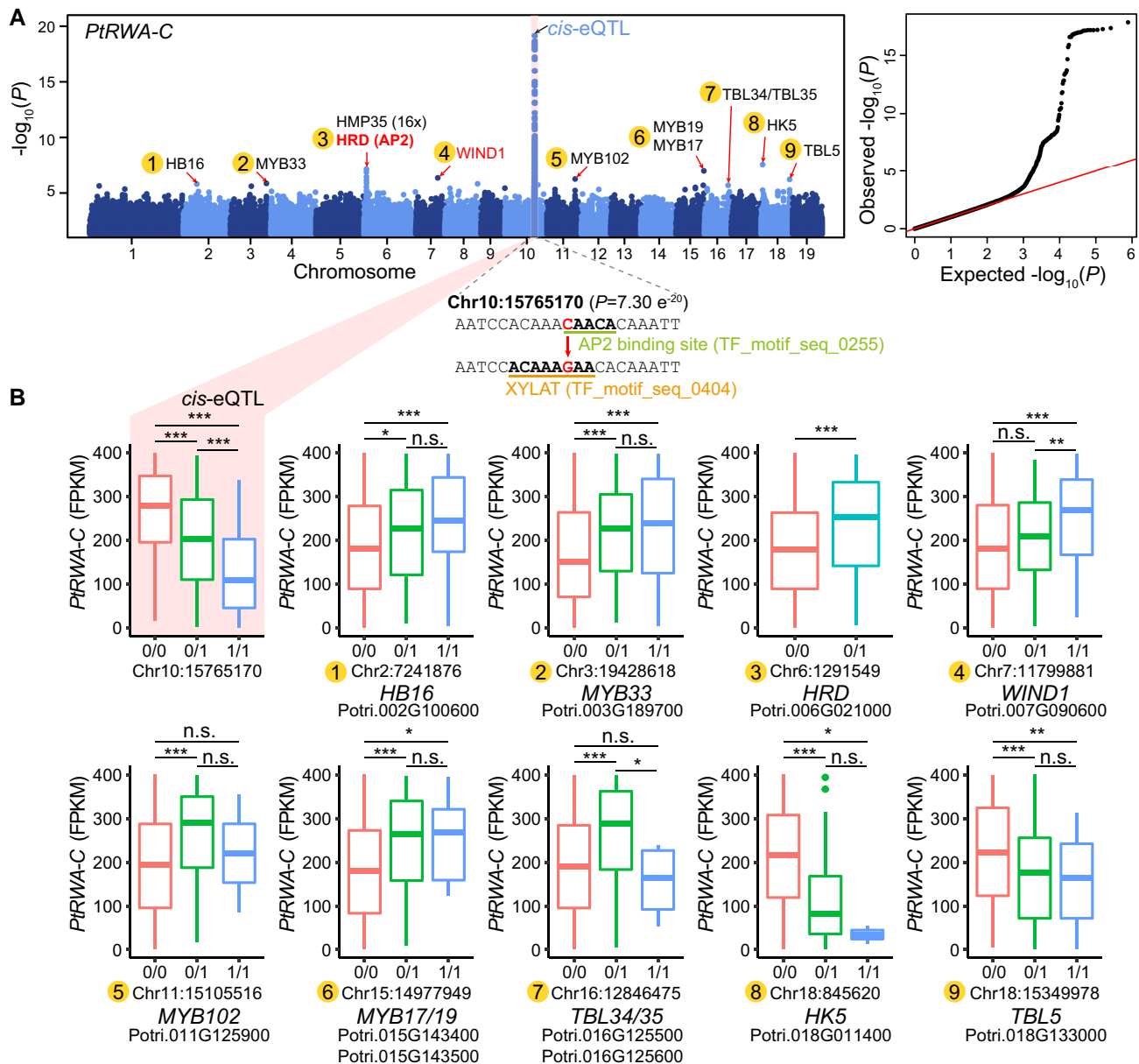


Figure 4. Expression quantitative trait loci (eQTL) mapping of *PtRWA-C* in *Populus* xylem. **A**) eQTL associated with the expression of *PtRWA-C* in xylem. Shadow background and arrows indicate highly associated *cis*-eQTL and *trans*-eQTL, respectively. The right panel represents Quantile-Quantile plot evaluating *P*-value deviation from expectation. **B**) Box plots showing genotypic effects of *PtRWA-C* expression at *cis*-eQTL (shadow background) and 9 *trans*-eQTL (*trans*-eQTL peaks #1-9). Numerical annotations on the X axes denote whether the SNP site is homozygous and consistent (0/0) with the reference genome (*P. trichocarpa* “Nisqually-1”), heterozygous (0/1), or homozygous and inconsistent (1/1) with the reference genome ($n = 444$). The bold line in the center of the boxplots represents the median, the box edges represent the 25th (lower) and 75th (upper) percentiles, and the whiskers extend to the most extreme data points that are no more than 1.5 \times the length of the upper or lower segment. Asterisks indicate significant differences based on Student’s *t*-test (* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$).

molecular beam mass spectrometry (pyMBMS) measurements, the lignin content was slightly increased and the lignin S/G ratio was increased by 12.9% and 8.8% in the 2 overexpression lines, respectively (Fig. 3, C and D), which is consistent with the results in the activation-tagged *rwa-c-D* mutant.

To assess the sugar release performance of the *PtRWA-C* overexpression lines, glucose and xylose releases during the

enzymatic hydrolysis were compared between transgenic and control plants. Both glucose and xylose releases from the 2 transgenic lines were lower than the control plants. The total released sugar from the 2 lines was decreased by 11.3% and 13.4%, respectively, suggesting decreased accessibility of sugar in these transgenic lines (Fig. 3, E to G). To investigate whether *PtRWA-C* is involved in xylan acetylation, we measured the acetyl contents in DMSO-extracted xylans

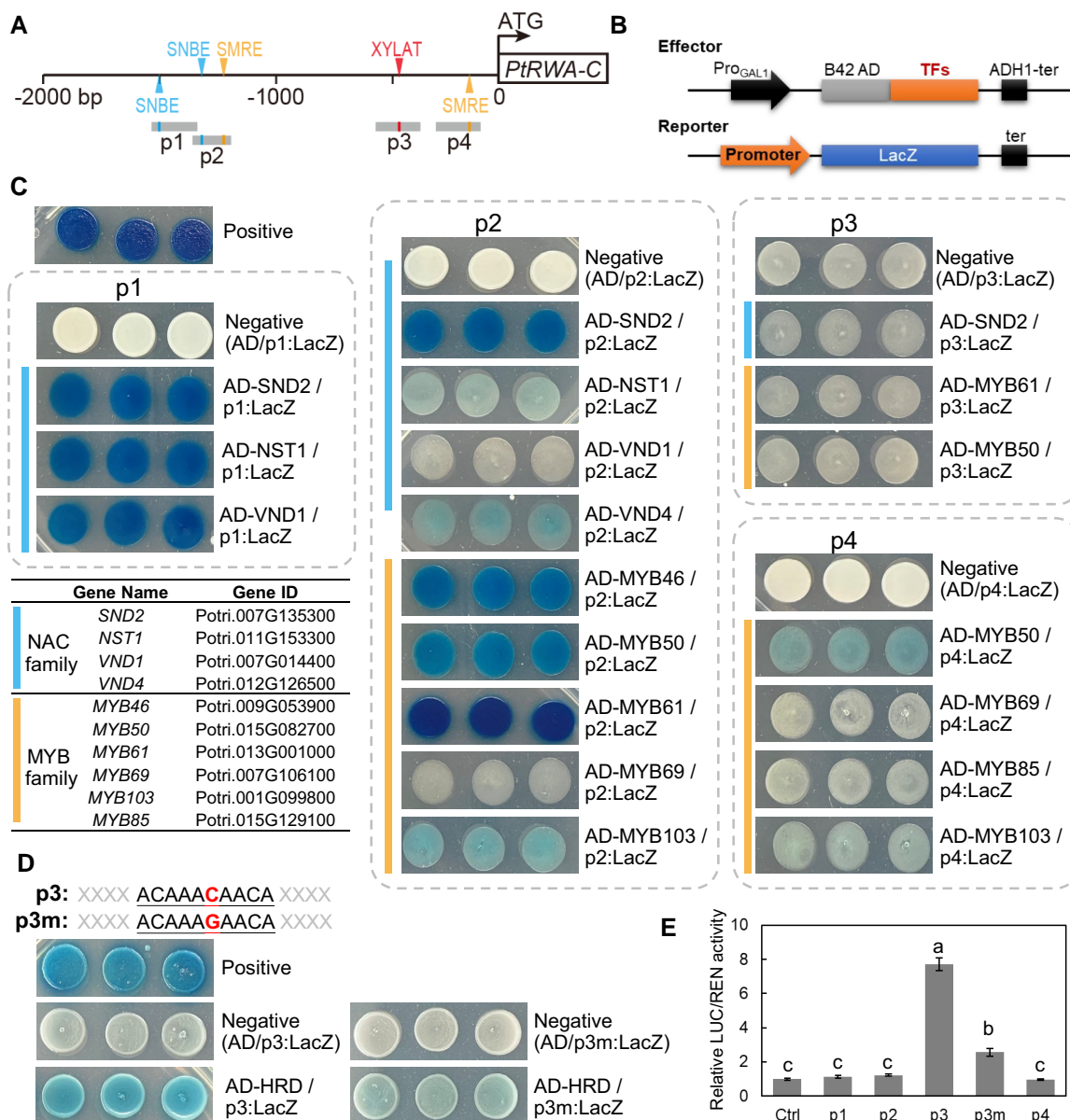


Figure 5. *PtRWA-C* is the direct target gene of secondary cell wall TFs and the *trans*-eQTL TF HRD. **A**) Simplified schematic of the promoter structure indicating the locations of the *cis*-acting elements (SNBE, secondary wall NAC binding elements; SMRE, MYB binding site; XYLAT, core xylem *cis*-element) and amplified promoter fragments (p1–p4). **B**) Schematic diagram of the effector and reporter structures. **C**) Yeast 1-hybrid (Y1H) assay showing the activity of LacZ reporters driven by the promoter fragments and activated by activation domain (AD) fusion effectors. **D**) Binding of the *trans*-eQTL (#3) TF HRD to the *PtRWA-C* promoter fragment p3 and its mutant p3m (C to G). **E**) Binding of HRD to the *PtRWA-C* promoter fragments using the dual-luciferase system. Expression values were determined by calculating the ratio of LUX activity to REN activity. Error bars represent means \pm SE ($n = 4$). Statistical analysis was carried out with 1-way ANOVA followed by Tukey's post hoc test. Statistical significance is indicated by different letters ($P < 0.05$).

of wild-type and the transgenic lines. The acetylation levels of the transgenic lines were approximately 1.7 and 1.8 times of that in the control (Fig. 3H).

Expression QTL analysis of *PtRWA-C*

eQTL mapping is an effective strategy to identify putative *cis*- and *trans*-regulatory elements underlying variation in gene expression (Zhang et al. 2018b). We used the gene expression level of *PtRWA-C* in xylem across a population containing 444

P. trichocarpa accessions as phenotypal data and >8.2 million SNP/InDel from 973 *P. trichocarpa* accessions to perform the eQTL analysis. As shown in Fig. 4A, the expression of *PtRWA-C* in xylem had a strong association with *cis*-eQTL, with the highest association at Chr10:15765170 ($P = 7.30e^{-20}$) (Supplemental Table S2). In addition, at least 9 *trans*-eQTL peaks (red arrows in Fig. 4A) containing potential cell wall-related regulators were associated with the expression of *PtRWA-C*, implying that these *trans*-eQTLs might be

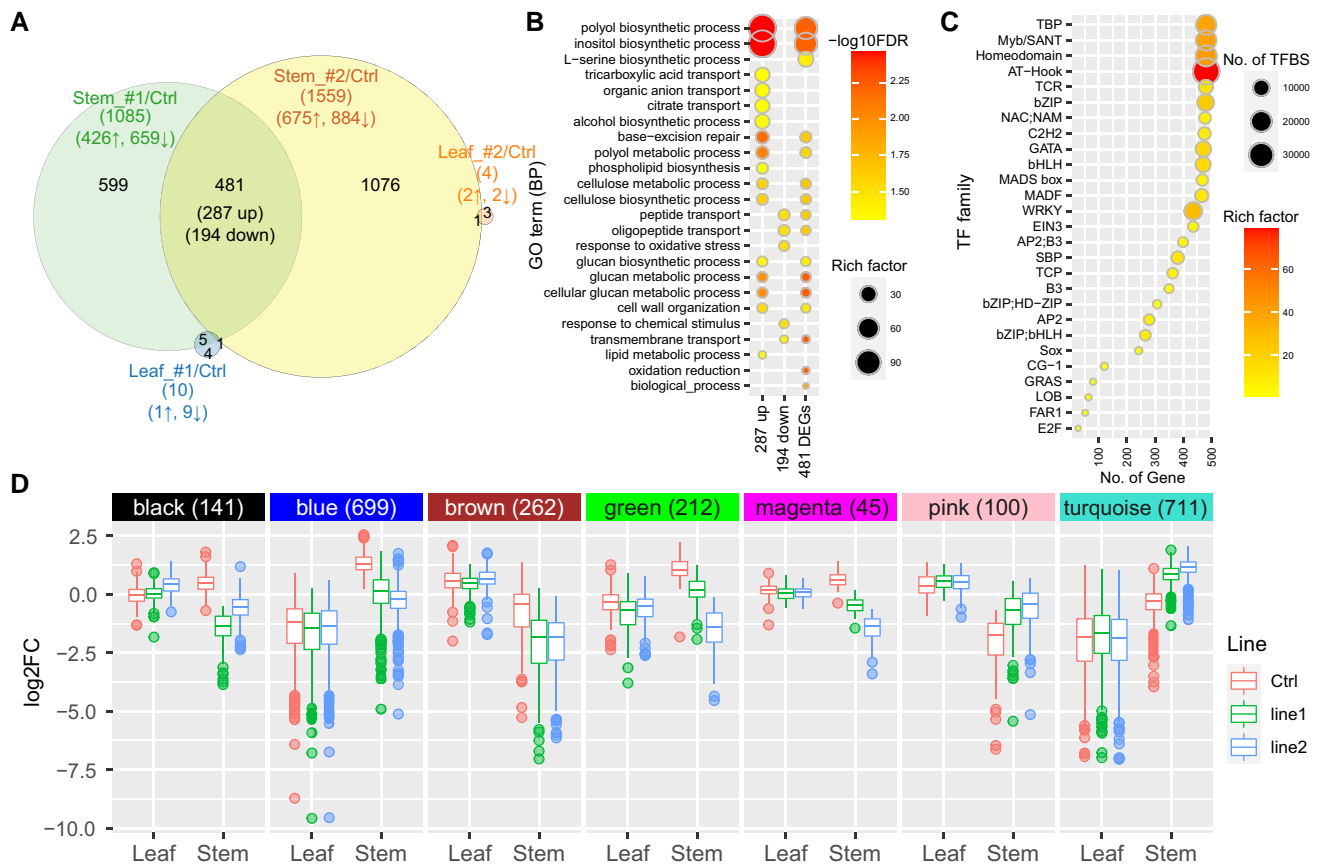


Figure 6. Transcriptomic changes in *Populus* transgenic plants overexpressing *PtRWA-C*. **A**) DEGs overlapped in leaf and stem of the 2 *PtRWA-C* overexpression lines (#1 and #2) compared to control plant (Ctrl). A total of 481 DEGs (286 upregulated and 194 downregulated) overlapped in stem of the 2 transgenic lines were identified as core-DEGs. **B**) Enriched biological process (BP) terms of GO enrichment analysis of the core-DEGs. **C**) Transcription factor binding site (TFBS) enrichment analysis of 2-kb promoter regions of the core-DEGs. **D**) Expression patterns of genes in the 7 coexpression modules. Numbers in parentheses indicate the number of genes (n) in each module. The bold line in the center of the boxplots represents the median, the box edges represent the 25th (lower) and 75th (upper) percentiles, and the whiskers extend to the most extreme data points that are no more than $1.5\times$ the length of the upper or lower segment.

the upstream regulators of *PtRWA-C*. For the *cis*-eQTL, the SNP Chr10:15765170 was located in the promoter region of *PtRWA-C* and its sequence “CAACA” corresponds to an AP2 binding site (TF_motif_seq_0255). When C at this motif is changed to G, it converts the AP2 binding site to another binding site “ACAAAGAA” named XYLAT (TF_motif_seq_0404), which has been reported as a *cis*-element identified among the promoters of the core xylem gene set (Ko et al. 2006). Interestingly, the expression of *PtRWA-C* between the 2 SNP variants in the promoter region showed significant differences in gene expression (Fig. 4B).

The *trans*-eQTL peak#3 (Chr06:1291549) was associated with *HRD* (Potri.006G021000), an AP2 family TF. In addition, 2 *trans*-eQTL harboring 3 TBL genes were highly associated with *PtRWA-C*, including *trans*-eQTL peak#7 (*TBL34* and *TBL35*, $P = 2.03e^{-06}$) and *trans*-eQTL peak#9 (*TBL5*, $P = 5.99e^{-07}$). The TBL family contains a plant-specific DUF231 domain. In *Arabidopsis*, *TBL34* catalyzes xylan 2-*O*- and 3-*O*-monoacetylation and 2,3-di-*O*-acetylation with

differential positional preference and *TBL35* catalyzes 2,3-di-*O*-acetylation (Zhong et al. 2017). The association with TBL family members suggests that *PtRWA-C* may also be involved in the xylan acetylation.

PtRWA-C is a downstream gene regulated by SCW-TFs and *trans*-eQTL TF

To investigate the molecular mechanisms of *PtRWA-C* involvement in SCW biosynthesis, we analyzed the promoter sequence of *PtRWA-C* to explore the potential functional *cis*-acting elements. After scanning the 2-kb promoter region of *PtRWA-C*, we identified 2 secondary wall NAC binding elements (SNBE, located at $-1,502$ and $-1,319$ bp relative to the start codon) (Zhong et al. 2010) and 2 SCW MYB binding sites (SMRE, located at $-1,225$ and -170 bp relative to the start codon) (Zhong and Ye 2012), as well as the XYLAT/AP2 binding site (-471 bp upstream of the start codon) (Ko et al. 2006) (Fig. 5A). This implies the secondary cell wall-related NAC and MYB TFs might directly regulate the expression of *PtRWA-C*. To test this, we cloned 4 promoter

fragments harboring different *cis*-elements (designated p1-p4) and performed a yeast 1-hybrid (Y1H) assay with the known SCW-related NAC and MYB TFs (Zhang et al. 2018a; Han et al. 2022) (Fig. 5B). If a TF binds to a tested promoter region, the synthesis of LacZ will be activated, which is visualized as blue color on the X-Gal containing plate. Y1H assays showed that fragment p1 was specifically bound by the master switches of SCW biosynthesis SND2, NST1, and VND1; fragment p2 was bound by TFs of both the NAC (SND2, NST1, and VND4) and MYB (MYB46/50/61/103) families; p4 was weakly bound by MYB TFs (MYB50 and MYB103), whereas p3 that lacks SNBE and SMRE elements was not bound by NAC or MYB TFs (Fig. 5C). To further verify the regulation of the *trans*-eQTL to *PtRWA-C*, we cloned the AP2 transcription factor *HDR*, which is close to the strong *trans*-eQTL (peak #3) associated with the expression of *PtRWA-C* (Fig. 4A). The strongest *cis*-eQTL located in fragment p3 (XYLAT) of the *PtRWA-C* promoter is also inside the potential binding site of *HDR*. Y1H assays showed that *HRD* strongly bound to the promoter fragment p3. However, when the fragment was mutated (peak *cis*-eQTL was mutated from C to G; p3m), the binding activity was attenuated (Fig. 5D). Next, we utilized a transient dual-luciferase (Dual-LUC) assay in *Nicotiana benthamiana* to further examine whether *HRD* regulates the transcription of *PtRWA-C*. The result of Dual-LUC assay showed that although *HRD* stimulated both fragments p3 and p3m activities, the activation of p3 was 3 times stronger than p3m (Fig. 5E). In summary, we concluded that the SCW-TFs and *HRD* directly regulate the expression of *PtRWA-C* and that mutation in the *cis*-eQTL located in the promoter region of *PtRWA-C* interfered with its transcriptional activation by its upstream transcription regulators.

Overexpression of *PtRWA-C* leads to transcriptomic changes in poplar stem

To explore if gene expression was affected by *PtRWA-C* overexpression, we performed the RNA-Seq analysis of 2 organs (leaf and stem) of the 2 *PtRWA-C* overexpression lines (#1 and #2). Using the cutoff of $\log_2(\text{FC}) > 1$ and FDR corrected $P < 0.05$ for the discovery of significant differentially expressed genes (DEGs), we found that the leaf transcriptome was not generally affected in the *PtRWA-C* overexpression lines, with only 10 and 4 significant DEGs in leaves of #1 and #2 transgenic lines, respectively (Fig. 6A). Noticeably, 1,085 DEGs (426 upregulated and 659 downregulated) and 1,559 DEGs (675 upregulated and 884 downregulated) were detected in stem of line #1 and line #2, respectively. By searching for overlapped DEGs between the 2 overexpression lines, we identified 481 DEGs (287 upregulated and 194 downregulated) (Fig. 6A and Supplemental Table S3), which we defined as “core-DEGs”. Gene Ontology (GO) enrichment analysis revealed that the 481 core-DEGs were enriched in “cell wall organization,” “glucan biosynthetic/metabolic processes,” “polyol/inositol biosynthetic/metabolic processes,”

etc. Comparing the GO terms enriched among the 287 upregulated core-DEGs and 194 downregulated DEGs, we discovered that the more prominent terms such as “cell wall organization,” “glucan biosynthetic/metabolic processes,” and “polyol/inositol biosynthetic/metabolic processes” were mainly contributed by the upregulated core-DEGs, whereas the 194 downregulated core-DEGs were enriched for terms of “peptide transport” and “response to chemical stimulus” (Fig. 6B and Supplemental Table S4).

To further define the functional nature of the core-DEGs, we analyzed the transcription factor binding sites (TFBS) of a 2-kb potential promoter region and classified the TFBS according to their gene family. The top 10 TFBS include targets of several TF families distributed in almost all the core-DEGs: TBP, MYB, Homeodomain, AT-Hook, TCR, bZIP, NAC, C2H2, GATA, and bHLH. In addition, a large number of WRKY binding sites (12,902 TFBS) in the promoters of 435 core-DEGs were detected (Fig. 6C and Supplemental Table S5). We then utilized a weighted gene coexpression network analysis (WGCNA) of the DEGs in both leaf and stem to further our understanding of the core-DEGs. After dynamic gene modules were merged, a total of 7 gene modules were obtained from the 2,170 DEGs (Supplemental Fig. S1 and Table S6). Modules blue and turquoise were clustered together, in which genes were expressed higher in stem than that in leaf whereas genes in module pink were expressed higher in leaf than in stem (Fig. 6D and Supplemental Fig. S1). When comparing the gene expression between transgenic lines and control plants, genes in modules pink and turquoise were expressed higher in stem of transgenic lines than control, whereas genes in the other 5 modules (black, blue, brown, green, and magenta) showed an opposite pattern in stem (Fig. 6D). This suggests that changes in cell wall acetylation levels perturb the expression of a range of genes including TFs and cell wall biosynthesis-related genes, which may be related to the feedback regulation of cell wall biosynthesis.

Discussion

As the major hemicellulose component in cell walls of many woody species, acetylated xylan hinders the conversion of lignocellulosic biomass to biofuels (Carroll and Somerville 2009; Chen et al. 2020). Therefore, understanding the biochemical pathway and key players involved in xylan acetylation and its regulatory mechanism is important not only for our understanding of basic biological processes but also for the directional improvement of bioenergy plants targeting biofuels. During biosynthesis in the Golgi apparatus, polysaccharides in plant cell walls are O-acetylated using acetyl-coenzyme A (CoA) as a donor substrate (Pauly and Scheller 2000). The pattern and degree of acetylation differ not only between plant species but also within cell types and/or developmental stages of plants.

In plants, the process of polysaccharide O-acetylation requires 3 category enzymes: TBLs, AX19, and RWAs. The TBL and AX19 proteins contain a single TM domain that

anchors the proteins to the Golgi membrane and their C-termini are oriented towards the Golgi lumen containing putative catalytic motifs. In contrast, RWA proteins have at least 10 TM helical structures (Manabe et al. 2011). RWA is thought to be responsible for translocation of the acetyl group of acetyl-CoA (or another unidentified donor) from the cytoplasmic pool to the Golgi apparatus in order to provide substrates for the other 2 families of OatAs (i.e. TBLs and AXY9) (Pauly and Scheller 2000).

In *Populus*, there are 4 RWA genes, with the same number as that in the *Arabidopsis* genome (Pawar et al. 2017b), indicating that no gene family expansion has occurred in the *Populus* RWA gene family. Comparative genomics studies have indicated that homologous genes in poplar typically occur in 1.4 to 1.6 times the number of *Arabidopsis* genes, which has been attributed to the Salicoid duplication even (Tuskan et al. 2006). Homologous maintenance of RWA gene numbers in poplar relative to *Arabidopsis* suggests RWA function has been conserved during speciation and evolution of the 2 species. Expression patterns of the 4 *PtRWA* genes indicated that all *PtRWAs* are highly expressed in SDX cells and in the SCW forming stage of xylem development (Fig. 2). Different from the other 3 members, *PtRWA-C* is also highly expressed in the phloem and cambium regions and its expression can be induced by *SND1-B1*, which suggests that *PtRWA-C* may affect the acetylation process of other cell wall matrix components. This is similar to its homolog *AtRWA2* in *Arabidopsis* (amino acid identity 72.4%), which is also expressed in primary cell walls and has a broad acetylation to pectic/nonpectic polysaccharides and xyloglucan (Manabe et al. 2011). The single-mutant *rwa2* in *Arabidopsis* had no obvious growth or developmental phenotypes, but triple and quadruple *rwa* mutants showed severe growth phenotypes. Expression of a single *RWA2* gene in the quadruple mutant compensated for its growth phenotype, suggesting that the RWA homologs are functionally redundant. However, different members have preferences for acetylated substrates such as xylan, (gluco)mannan, and xyloglucan (Manabe et al. 2013). Similarly, suppression of 2 genes in the AB clade or 2 genes in the CD clade in the poplar RWA gene family did not affect its growth but had an effect on its acetylation level of cell wall. When all RWA genes, i.e. A-D, were suppressed in poplar, the acetylation level of xylan in wood was significantly reduced and the saccharification efficiency of biomass was improved (Pawar et al. 2017b).

In our study, *PtRWA-C* contains 11 helical TM structures and its subcellular localization is on the Golgi apparatus (Fig. 2). In addition, the acetyl content in the cell wall of plants overexpressing *PtRWA-C* is significantly increased (Fig. 3). These results indicate that *PtRWA-C* is likely involved in the translocation of the O-acetyl donor, increasing the available substrate pool for OatAs in the Golgi. Furthermore, overexpression of *PtRWA-C* caused changes in the expression levels of a series of genes in poplar stem including a large number of TFs (Fig. 6). It is unclear how the

acetylation level of cell wall components directly affects gene expression, and we propose that changes in xylan structure or substrate utilization trigger a feedback regulation mechanism. Overexpression of a single *PtRWA-C* gene not only led to an increase in xylan acetylation levels but also resulted in a decrease in the release of glucose and fructose (Fig. 3). This is consistent with the negative effect of xylan acetylation on saccharification efficiency. In addition, the lignin content and S/G ratio in the overexpression plants were increased (Fig. 3), indicating that acetylation of xylan has an impact on the biosynthesis of lignin components. Pramod et al. (2021) used a wood-specific promoter to constitutively express the fungal acetyl xylan esterases (*AnAXE1*) in hybrid aspen. The transgenic plants showed lower acetylation levels and a small reduction in lignin S/G ratio, but no significant differences in carbohydrate or lignin contents. Similarly, an earlier study concluded that when the acylation levels of xylan were altered, the composition and solubility of lignin were dramatically modified (Pawar et al. 2017a). However, the direct effect of increase of xylan acetylation level on an increase in lignin content and S/G ratio needs further study.

Our results related to the gene regulatory pathway associated with RWA have bioenergy and biomaterial applications. Previous studies analyzed c. 1-kb promoter sequences of the 4 RWA genes in poplar and found that *RWA-A* and *RWA-B*, but not *RWA-C* and *RWA-D*, were strongly regulated by homologs of SCW synthesis master switches *NST1* and *MYB46* (Pawar et al. 2017b). Our comprehensive gene coexpression network analysis found that *PtRWA-C* was also coexpressed with SCW-TFs such as *NST1* and *MYB46* (Fig. 2), implying that it may also be regulated by the SCW hierarchical regulatory network. Further analysis of the *cis*-acting elements indicates that there are multiple SNBE and SMRE located between 1~2 kb upstream of the *PtRWA-C* gene promoter, which may serve as binding sites for NAC and MYB TFs, respectively (Fig. 5). The Y1H assay also confirmed that NAC and MYB (including the master switches *NST1*, *SND2*, and *MYB46*) directly bind to the promoter of *PtRWA-C* (Fig. 5), indicating that *PtRWA-C* is also directly regulated by the SCW hierarchical regulatory network.

eQTL analysis developed based on population genetics and bioinformatics in recent years has pointed to gene regulatory mechanisms. In this analysis, gene expression levels are considered as quantitative traits and gene expression phenotypes are mapped to specific genomic loci by combining the study of gene expression pattern variation with genome-wide genotyping (Gilad et al. 2008). In forest trees, the potential transcriptional regulatory relationships identified by eQTL mapping have been confirmed experimentally. For example, using GWAS analysis, it was found that a hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase gene *PtHCT2* in poplar is a key gene that controls the biosynthesis of metabolites such as 3-O-caffeoylquinic acid. In parallel, the eQTL analysis discovered that *PtHCT2* expression was regulated by the upstream WRKY-HCT2 regulatory module through *cis*-eQTL (Zhang et al. 2018b). In this

study, we found that the expression of *PtRWA-C* was strongly regulated by *cis*-eQTL through eQTL mapping and the *cis*-element containing the *cis*-eQTL is the binding site of HRD, which is close to another strong *trans*-eQTL associated with *PtRWA-C* expression (Fig. 4). Y1H and Dual-LUC experiments also confirmed the direct regulation of *PtRWA-C* by HRD (Fig. 5). These results indicate that eQTL mapping analysis is an effective platform for exploring regulatory relationships in woody plants, the application of which needs to be further expanded.

In summary, our study provides insights into the functional roles of *PtRWA-C* in modulating xylan acetylation and saccharification efficiency, which has important implications for the genetic engineering of forest trees for bioenergy or biomaterial production. Specifically, we identified upstream regulatory genes and *cis*-eQTL that regulate *PtRWA-C* expression, which could serve as potential targets for gene and promoter editing to alter gene expression levels and acetylation patterns, ultimately affecting bioprocessing recalcitrance. These findings have important implications for the development of engineered wood species as sustainable sources of biofuels and biomaterials.

Materials and methods

Screening of poplar activation tagging population and TAIL-PCR

Binary vector pSKI074 was used for transformation of a hybrid aspen (*Populus tremula* × *Populus alba* clone INRA 717-1B4) through *Agrobacterium tumefaciens*-mediated transformation according to Harrison et al. 2007; Busov et al. 2011. A field study was established in Quebec and Ontario.

The TAIL-PCR and plasmid rescue were performed as previously described (Busov et al. 2011). For details, genomic DNA was extracted using Qiagen DNAeasy kit following the manufacturer's instructions. TAIL-PCR was performed using vector-specific primers (LBr1, 5'-AAGCCCCCA TTTGGACGTGAATGTAGACAC-3'; LBr2, 5'-TTGCTTTCCG CCTATAAATACGACGGATCG-3', LBr3, 5'-TAACGCTGC GGACATCTAC-3') and arbitrary degenerated primers (AD21, 5'-NGTCGASWGANAW GAA-3'; AD22, 5'-NGTC GASWGANAWGTT-3'; AD23, 5'-NGTCGASWGANAWGAC -3'; AD24, 5'-NGTCGASWGANAWCAA-3'; AD25, 5'-NGTC GASWGANAWCTT-3'; AD3, 5'-WCAGNTGWTNGTNC TG-3'; AD4, 5'-NGTAWAASGTNTSCAA-3'). PCR amplification products were purified and sequenced using the LBr3 primer.

Bioinformatics analysis

Normalized gene expression values of the *PtRWA* genes in different organs and cell types and in PtrSND1-B1 transfected SDX protoplasts were obtained from NCBI BioProject (accessions PRJNA320431 and PRJNA215318). For gene expression during wood formation, a nm-scale high-resolution gene

expression database AspWood (<https://plantgenie.org/>) was used, which includes 25 samples from phloem, cambium, and expanding and maturing xylem. A coexpression network of *PtRWA-C* was created according to the method reported by Zhang et al. 2020 using the above gene expression datasets. Cytoscape (Shannon et al. 2003) was used to visualize the resulting network. The protein structure of the *PtRWA-C* was predicted using ColabFold notebook using the default settings, which is powered by AlphaFold2 combined with a fast, multiple-sequence alignment generation stage using MMseqs2 (Mirdita et al. 2022).

Generation of transgenic poplar

For the *PtRWA-C* overexpression lines, the full-length open-reading frame of *PtRWA-C* was amplified from *Populus trichocarpa* "Nisqually-1" and was cloned into the pAGW560 binary vector with the Ubiquitin 3 promoter. The genetic transformation of *P. deltoides* "WV94" through *Agrobacterium*-mediated methods was conducted at ArborGen Inc. (Ridgeville, SC, USA). The transgenic plants were transferred and grown in the greenhouse at Oak Ridge National Laboratory (Oak Ridge, TN, USA) at 25 °C and 16/8 h light/dark photoperiod.

To estimate the stem cylinder volume, plant height and stem base diameter were measured in ~6-month-old plants. We measured primary stem length from stem base to shoot tip for plant height and measured the diameter of the stem base.

Subcellular localization

For the analysis of protein subcellular localization, *PtRWA-C* was cloned into the YFP fusion vector and expressed in the *Populus* leaf mesophyll protoplasts (Zhang et al. 2020).

Quantitative measurement of acetyl groups

The peeled stem segments (~30 cm) above the ground were obtained from 6-month-old plants and used for acetylation analysis. Alcohol-insoluble residue (AIR) of the milled poplar stem tissues was prepared according to Kong et al. 2018. The obtained AIR was then destarched, depectinated, and delignified by using amylase, 50 mM ammonium oxalate (pH 5.0), and 11% (v/v) peracetic acid, respectively. Acetylated xylan from stem cell wall residues was extracted with dimethyl sulfoxide (DMSO) following the procedure described (Gonçalves et al. 2008). Briefly, 100% (v/v) DMSO was added to destarched, depectinated, and delignified stem tissues (20 mg/mL) and the mixture was incubated at 60 °C overnight with shaking. The mixture was then filtered through a nylon mesh, and the flowthrough was mixed with ethanol-methanol-acetic acid (7:2:1) to precipitate xylan after incubation at 4 °C overnight. The xylan-containing precipitant was separated through centrifugation and was further used for the following acetyl group quantification. The amount of acetyl groups in DMSO-extracted xylan was determined by quantifying acetic acid released from xylan upon treatment with NaOH (Teleman et al. 2002). In brief, 20 mg DMSO-extracted xylan was saponified by treating

with 1,100 μL 0.1N NaOH overnight at room temperature. After overnight incubation, the samples were neutralized by adding 100 μL 1 M HCl. The amount of acetic acid was quantitated using an acetic acid assay kit (Megazyme). The data were the average of 5 separate pools of samples.

Saccharification assay

Air-dried stems of the *Populus* control and transgenic plants were ground in a Thomas Scientific Wiley Mini Mill 3383-L10 (Swedesboro, NJ, USA) to pass through a 40-mesh screen. Ground samples with particle size ~ 0.420 mm were used for saccharification assays according to the method described by Zhang et al. (2020). In brief, biomass was extracted with 0.25% (w/v) α -amylase and 1.5% (w/v) α -glucosidase in 0.1 M sodium acetate (24 h, 55 $^{\circ}\text{C}$, pH 5.0) to remove possible starch content, followed by an ethanol (95% v/v) Soxhlet extraction for 24 h. After drying, 5 ± 0.5 mg of biomass was weighed into 1 of 96 wells in a solid Hastelloy microtiter plate and 250 μL of water was added. For pretreatment, the samples were reacted at 180 $^{\circ}\text{C}$ for 17.5 min, and after being cooled, 40 μL of buffer-enzyme stock (8% [v/v] CTec2 [Novozymes, Bagsværd, Denmark] [excess enzyme loading of 70 mg/g biomass] in 1 M sodium citrate buffer) was added. The samples were then gently mixed and left to statically incubate at 50 $^{\circ}\text{C}$ for 70 h. After incubation, an aliquot of the saccharified hydrolysate was diluted and tested using megazymes GOPOD (glucose oxidase/peroxidase) and XDH assays (xylose dehydrogenase).

Expression-based quantitative trait loci (eQTL) analysis

Whole genome resequencing, SNP/InDel calling, and SNPEff analysis for the 917 individuals of this *Populus* GWAS population were previously described by Zhang et al. (2018b). To assess genetic control, we used the EMMA algorithm in the EMMA software with kinship as the correction factor for genetic background effects. The normalized FPKM transcript counts of *PtRWA-C* in xylem from 444 accessions were used as phenotypes.

RNA-Seq and data analysis

For total RNA isolation, approximately 100-mg fully expanded leaves and stem (the 5th internode) from 3 biological replicates were ground in liquid nitrogen and total RNA was extracted using a Spectrum Total Plant RNA extraction kit (Sigma-Aldrich, St. Louis, USA) with the on-column RNase-free DNase I treatment to remove the residual genomic DNA. RNA quality and quantity were determined using a Nanodrop Spectrophotometer (Thermo Fisher Scientific, Hudson, NH). RNA-Seq libraries were generated and quantified using RT-qPCR. Sequencing was performed on an Illumina HiSeq 2,500 (150mer paired-end sequencing). Raw fastq file reads were filtered and trimmed using the JGI QC pipeline. Using BBduk (<https://sourceforge.net/projects/bbmap/>), raw reads were evaluated for sequence artifacts by kmer matching (kmer = 25) allowing 1 mismatch and

detected artifacts were trimmed from the 3' end of the reads. RNA spike-in reads, PhiX reads, and reads containing any Ns were removed. Quality trimming was performed using the phred trimming method set at Q6. Following trimming, reads under the length threshold were removed. Raw reads from each library were aligned to the *P. trichocarpa* reference genome (<https://phytozome.jgi.doe.gov/pz/portal.html#>) using TopHat2 (Kim et al. 2013). Only reads that mapped uniquely to 1 locus were counted. FeatureCounts (Liao et al. 2014) was used to generate raw gene counts, which were used to evaluate the level of correlation between biological replicates. DESeq2 (Love et al. 2014) was subsequently used to determine which genes were differentially expressed between pairs of conditions ($P < 0.05$). GO enrichment was performed using agriGO (Tian et al. 2017). For the promoter analysis, the TF binding sites were identified using PlantPAN (Chow et al. 2016).

RT-qPCR analysis

One microgram of total RNA was used to generate cDNA by means of the Rite Aid reverse transcriptase following manufacturer's instruction (Thermo Fisher Scientific, Hudson, NH). Gene-specific primers were designed using Primer3 software (<http://frodo.wi.mit.edu/primer3/input.htm>) with an annealing temperature of 58 to 60 $^{\circ}\text{C}$ and amplicon size of 150 to 250 bp. RT-qPCR was performed using Maxima SYBR Green/ROX qPCR master mix (Thermo Fisher Scientific) according to the manufacturer's instructions. The relative gene expression was calculated by $2^{-\Delta\Delta\text{Ct}}$ method (Livak and Schmittgen 2001) using *PtUBQ10b* (Potri.001G263000) as the internal control. All the experiments were repeated at least 3 times with similar results. The primers used in this study are listed in Supplemental Table S7.

Yeast 1-hybrid assay (Y1H)

The activation domain-transcription factor fusion constructs (AD-TFs) were cotransformed with the LacZ reporter plasmids (promoter:LacZ) into yeast strain EGY48. Transformants were grown on synthetic dropout plates without tryptophan or uracil containing 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal) for color development.

Dual-luciferase activity assay

The promoter fragments of *PtRWA-C* were inserted into pGreen0800-LUC vector, which expresses Renilla luciferase (REN) driven by the 35S promoter serving as an internal reference and firefly luciferase (LUC) driven by the promoter fragments of *PtRWA-C* serving as a reporter. The Dual-LUC assay was performed using Dual-Luciferase Reporter Assay System (Promega, Durham, USA). Measure the value in the luminometer detection system after mixing. Four biological repeats were measured for each sample.

Statistical analysis

Statistical analysis to determine statistical significance was performed by Student's *t*-tests for paired samples or one-way

ANOVA followed by Tukey's post hoc test for multiple pairwise comparisons. The asterisk or different letters in each figure indicate significant difference compared to control samples ($P < 0.05$).

Accession numbers

The RNA-Seq data were deposited in NCBI SRA with the following accessions: SRP138581, SRP138583, SRP138623-26, SRP138628, SRP138630, SRP138641, SRP183923, SRP183947-49, SRP183954, SRP183990, SRP184008-09, and SRP184021.

Acknowledgments

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

Author Contributions

J.Z., G.A.T., W.M., and J.-G.C. conceived and designed the experiments. J.Z. analyzed the data and wrote the manuscript. J.Z. and X.W. performed the experiments. H.-T.W. and B.R.U. performed acetylation analysis. Z.Q., T.Y., and M.X. carried out subcellular localization analysis. S.S.J. and L.E.G. performed transgenic line phenotyping. O.C. performed activation tagging identification. S.R. and A.S. created the activation-tagged lines and established the field trials for screening. X.Y., W.R. and K.A.W. created transgenic lines. R.S. performed cell wall chemistry analysis. J.Z., A.L., C.D., and K.B. performed RNA-Seq analysis. J.Z., W.Z., M.Z.L., G.A.T., and J.-G.C. interpreted the data and revised the manuscript.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. WGCNA coexpression modules of *Populus* transgenic plants.

Supplemental Table S1. Coexpressed genes of *PtRWA-C*.

Supplemental Table S2. DEGs between pairs of samples ($P_{\text{FDR}} < 0.05$ and $|\log_2(\text{fold-change})| > 1$).

Supplemental Table S3. Gene Ontology enrichment analysis of core-DEGs in the biological process category (P), molecular function category (F), and cellular component category (C).

Supplemental Table S4. Transcription factor binding site (TFBS) enrichment analysis of 2-kb promoter regions of the core-DEGs.

Supplemental Table S5. Coexpression modules identified using weighted gene coexpression network analysis (WGCNA).

Supplemental Table S6. Significant SNPs from eQTL analysis of *PtRWA-C*.

Supplemental Table S7. Primers used in this study.

Funding

J.Z. was funded by the Zhejiang Science and Technology Major Program on Agricultural New Variety Breeding (2021C02070-1) and the National Natural Science Foundation of China (32171814). M.X. was supported by the Office of Science of the U.S. Department of Energy, Office of Biological and Environmental Research, as part of the Quantitative Plant Science Initiative (QPSI) at Brookhaven National Laboratory. This research was supported by the Center for Bioenergy Innovation (CBI). CBI is supported by the Office of Biological and Environmental Research (BER) in the Office of Science of the U.S. Department of Energy. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract Number DE-AC05-00OR22725. The work (proposal: 10.46936/10.25585/60000710) conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. The activation tagging population was created through a Genome Canada grant to S.R. and A.S.

Conflict of interest statement. None declared.

References

- Busov V, Yordanov Y, Gou J, Meilan R, Ma C, Regan S, Strauss S.** Activation tagging is an effective gene tagging system in *Populus*. *Tree Genet Genomes*. 2011;7:91–101. <https://doi.org/10.1007/s11295-010-0317-7>
- Carroll A, Somerville C.** Cellulosic biofuels. *Ann Rev Plant Biol*. 2009;60:165–182. <https://doi.org/10.1146/annurev.arplant.043008.092125>
- Chen YK, Lin CH, Wang WC.** The conversion of biomass into renewable jet fuel. *Energy*. 2020;201:117655. <https://doi.org/10.1016/j.energy.2020.117655>
- Chow CN, Zheng HQ, Wu NY, Chien CH, Huang HD, Lee TY, Chiang-Hsieh YF, Hou PF, Yang TY, Chang WC.** PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Res*. 2016;44(D1):D1154–D1160. <https://doi.org/10.1093/nar/gkv1035>
- Ebringerová A.** Structural diversity and application potential of hemicelluloses. *Macromol Symp*. 2005;232(1):1–12.
- Gilad Y, Rifkin SA, Pritchard JK.** Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet*. 2008;24(8):408–415. <https://doi.org/10.1016/j.tig.2008.06.001>
- Gille S, de Souza A, Xiong G, Benz M, Cheng K, Schultink A, Reca I-B, Pauly M.** O-acetylation of Arabidopsis hemicellulose xyloglucan

- requires *AXY4* or *AXY4L*, proteins with a TBL and DUF231 domain. *Plant Cell*. 2011;**23**(11):4041–4053. <https://doi.org/10.1105/tpc.111.091728>
- Gille S, Pauly M.** O-acetylation of plant cell wall polysaccharides. *Front Plant Sci*. 2012;**3**:12. <https://doi.org/10.3389/fpls.2012.00012>
- Gonçalves VM, Evtuguin DV, Domingues MRM.** Structural characterization of the acetylated heteroxylan from the natural hybrid *Paulownia elongata*/*Paulownia fortunei*. *Carbohydr Res*. 2008;**343**(2): 256–266. <https://doi.org/10.1016/j.carres.2007.11.002>
- Gupta M, Rawal TB, Dupree P, Smith JC, Petridis L.** Spontaneous rearrangement of acetylated xylan on hydrophilic cellulose surfaces. *Cellulose*. 2021;**28**(6):3327–3345. <https://doi.org/10.1007/s10570-021-03706-z>
- Han X, Zhao Y, Chen Y, Xu J, Jiang C, Wang X, Zhuo R, Lu M-Z, Zhang J.** Lignin biosynthesis and accumulation in response to abiotic stresses in woody plants. *For Res*. 2022;**2**:9:1–10. <https://doi.org/10.48130/FR-2022-0009>
- Harrison EJ, Bush M, Plett JM, McPhee DP, Vitez R, O'Malley B, Sharma V, Bosnich W, Seguin A, MacKay J, et al.** Diverse developmental mutants revealed in an activation-tagged population of poplar. *Botany*. 2007;**85**(11):1071–1081. <https://doi.org/10.1139/B07-063>
- Jones CS, Anderson AC, Clarke AJ.** Mechanism of *Staphylococcus aureus* peptidoglycan O-acetyltransferase A as an O-acyltransferase. *Proc Natl Acad Sci U S A*. 2021;**118**(36):e2103602118. <https://doi.org/10.1073/pnas.2103602118>
- Jönsson LJ, Alriksson B, Nilvebrant N-O.** Bioconversion of lignocellulose: inhibitors and detoxification. *Biotechnol Biofuels*. 2013;**6**(1):16. <https://doi.org/10.1186/1754-6834-6-16>
- Kim D, Perteza G, Trapnell C, Pimentel H, Kelley R, Salzberg SL.** Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;**14**(4):R36. <https://doi.org/10.1186/gb-2013-14-4-r36>
- Kirui A, Zhao W, Deligey F, Yang H, Kang X, Mentink-Vigier F, Wang T.** Carbohydrate-aromatic interface and molecular architecture of lignocellulose. *Nat Commun*. 2022;**13**(1):538. <https://doi.org/10.1038/s41467-022-28165-3>
- Ko JH, Beers EP, Han KH.** Global comparative transcriptome analysis identifies gene network regulating secondary xylem development in *Arabidopsis thaliana*. *Mol Genet Genomics*. 2006;**276**(6): 517–531. <https://doi.org/10.1007/s00438-006-0157-1>
- Kong Y, O'Neill M, Zhou G.** Plant cell walls: isolation and monosaccharide composition analysis. In: Guo Y, editor. *Plant senescence. Methods in molecular biology*, vol 1744. New York (NY): Humana Press; 2018. p. 313–319.
- Lee C, Teng Q, Zhong R, Ye ZH.** The four *Arabidopsis* reduced wall acetylation genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. *Plant Cell Physiol*. 2011;**52**(8):1289–1301. <https://doi.org/10.1093/pcp/pcr075>
- Liao Y, Smyth GK, Shi W.** Featurecounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;**30**(7):923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Livak KJ, Schmittgen TD.** Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*. 2001;**25**(4):402–408. <https://doi.org/10.1006/meth.2001.1262>
- Love MI, Huber W, Anders S.** Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;**15**(12):550. <https://doi.org/10.1186/s13059-014-0550-8>
- Manabe Y, Nafisi M, Verherbruggen Y, Orfila C, Gille S, Rautengarten C, Cherk C, Marcus SE, Somerville S, Pauly M, et al.** Loss-of-function mutation of REDUCED WALL ACETYLATION2 in *Arabidopsis* leads to reduced cell wall acetylation and increased resistance to *Botrytis cinerea*. *Plant Physiol*. 2011;**155**(3):1068–1078. <https://doi.org/10.1104/pp.110.168989>
- Manabe Y, Verherbruggen Y, Gille S, Harholt J, Chong S-L, Pawar PM-A, Mellerowicz EJ, Tenkanen M, Cheng K, Pauly M, et al.** Reduced wall acetylation proteins play vital and distinct roles in cell wall O-acetylation in *Arabidopsis*. *Plant Physiol*. 2013;**163**(3): 1107–1117. <https://doi.org/10.1104/pp.113.225193>
- Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M.** Colabfold: making protein folding accessible to all. *Nat Methods*. 2022;**19**(6):679–682. <https://doi.org/10.1038/s41592-022-01488-1>
- Pauly M, Scheller HV.** O-Acetylation of plant cell wall polysaccharides: identification and partial characterization of a rhamnogalacturonan O-acetyl-transferase from potato suspension-cultured cells. *Planta*. 2000;**210**(4):659–667. <https://doi.org/10.1007/s004250050057>
- Pawar PM, Derba-Maceluch M, Chong S-L, Gandla ML, Bashar SS, Sparrman T, Ahvenainen P, Hedenström M, Özparpucu M, Rüggeberg M, et al.** In muro deacetylation of xylan affects lignin properties and improves saccharification of aspen wood. *Biotechnol Biofuels*. 2017a;**10**:98. <https://doi.org/10.1186/s13068-017-0782-4>
- Pawar PM, Derba-Maceluch M, Chong SL, Gómez LD, Miedes E, Banasiak A, Ratke C, Gaertner C, Mouille G, McQueen-Mason SJ, et al.** Expression of fungal acetyl xylan esterase in *Arabidopsis thaliana* improves saccharification of stem lignocellulose. *Plant Biotechnol J*. 2016;**14**(1):387–397. <https://doi.org/10.1111/pbi.12393>
- Pawar PM, Ratke C, Balasubramanian VK, Chong SL, Gandla ML, Adriasola M, Sparrman T, Hedenstrom M, Szwaj K, Derba-Maceluch M, et al.** Downregulation of RWA genes in hybrid aspen affects xylan acetylation and wood saccharification. *New Phytol*. 2017b;**214**(4):1491–1505. <https://doi.org/10.1111/nph.14489>
- Pramod S, Gandla ML, Derba-Maceluch M, Jönsson LJ, Mellerowicz EJ, Winstrand S.** Saccharification potential of transgenic greenhouse-and field-grown aspen engineered for reduced xylan acetylation. *Front Plant Sci*. 2021;**12**:704960. <https://doi.org/10.3389/fpls.2021.704960>
- Rastogi L, Chaudhari AA, Sharma R, Pawar PA-M.** *Arabidopsis* GELP7 functions as a plasma membrane-localized acetyl xylan esterase, and its overexpression improves saccharification efficiency. *Plant Mol Biol*. 2022;**109**(6):781–797. <https://doi.org/10.1007/s11103-022-01275-8>
- Scheller HV, Ulvskov P.** Hemicelluloses. *Annu Rev Plant Biol*. 2010;**61**: 263–289. <https://doi.org/10.1146/annurev-arplant-042809-112315>
- Schultink A, Naylor D, Dama M, Pauly M.** The role of the plant-specific ALTERED XYLOGLUCAN9 protein in *Arabidopsis* cell wall polysaccharide O-acetylation. *Plant Physiol*. 2015;**167**(4):1271–1283. <https://doi.org/10.1104/pp.114.256479>
- Selig MJ, Adney WS, Himmel ME, Decker SR.** The impact of cell wall acetylation on corn stover hydrolysis by cellulolytic and xylanolytic enzymes. *Cellulose*. 2009;**16**(4):711–722. <https://doi.org/10.1007/s10570-009-9322-0>
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T.** Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;**13**(11):2498–2504. <https://doi.org/10.1101/gr.1239303>
- Teleman A, Tenkanen M, Jacobs A, Dahlman O.** Characterization of O-acetyl-(4-O-methylglucurono) xylan isolated from birch and beech. *Carbohydr Res*. 2002;**337**(4):373–377. [https://doi.org/10.1016/S0008-6215\(01\)00327-5](https://doi.org/10.1016/S0008-6215(01)00327-5)
- Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, Xu W, Su Z.** agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res*. 2017;**45**(W1):W122–W129. <https://doi.org/10.1093/nar/gkx382>
- Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al.** The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*. 2006;**313**(5793):1596–1604. <https://doi.org/10.1126/science.1128691>
- Wang Z, Pawar PM-A, Derba-Maceluch M, Hedenström M, Chong S-L, Tenkanen M, Jönsson LJ, Mellerowicz EJ.** Hybrid aspen expressing a carbohydrate esterase family 5 acetyl xylan esterase under control of a wood-specific promoter shows improved saccharification. *Front Plant Sci*. 2020;**11**:380. <https://doi.org/10.3389/fpls.2020.00380>

- Xiong G, Cheng K, Pauly M.** Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the *Arabidopsis* mutant *tbl29*. *Mol Plant*. 2013;**6**(4):1373–1375. <https://doi.org/10.1093/mp/sst014>
- Yuan Y, Teng Q, Zhong R, Haghighat M, Richardson EA, Ye Z-H.** Mutations of *Arabidopsis* *TBL32* and *TBL33* affect xylan acetylation and secondary wall deposition. *PLoS One*. 2016;**11**(1):e0146460. <https://doi.org/10.1371/journal.pone.0146460>
- Zhang J, Siika-Aho M, Tenkanen M, Viikari L.** The role of acetyl xylan esterase in the solubilization of xylan and enzymatic hydrolysis of wheat straw and giant reed. *Biotechnol Biofuels*. 2011;**4**(1):60. <https://doi.org/10.1186/1754-6834-4-60>
- Zhang J, Xie M, Li M, Ding J, Pu Y, Bryan AC, Rottmann W, Winkler KA, Collins CM, Singan V, et al.** Overexpression of a prefoldin β subunit gene reduces biomass recalcitrance in the bioenergy crop *Populus*. *Plant Biotechnol J*. 2020;**18**(3):859–871. <https://doi.org/10.1111/pbi.13254>
- Zhang J, Xie M, Tuskan GA, Muchero W, Chen J-G.** Recent advances in the transcriptional regulation of secondary cell wall biosynthesis in the woody plants. *Front Plant Sci*. 2018a;**9**:1535. <https://doi.org/10.3389/fpls.2018.01535>
- Zhang J, Yang Y, Zheng K, Xie M, Feng K, Jawdy SS, Gunter LE, Ranjan P, Singan VR, Engle N, et al.** Genome-wide association studies and expression-based quantitative trait loci analyses reveal roles of HCT2 in caffeoylquinic acid biosynthesis and its regulation by defense-responsive transcription factors in *Populus*. *New Phytol*. 2018b;**220**(2):502–516. <https://doi.org/10.1111/nph.15297>
- Zhong R, Cui D, Ye Z-H.** Regiospecific acetylation of xylan is mediated by a group of DUF231-containing O-acetyltransferases. *Plant Cell Physiol*. 2017;**58**(12):2126–2138. <https://doi.org/10.1093/pcp/pcx147>
- Zhong R, Lee C, Ye Z-H.** Global analysis of direct targets of secondary wall NAC master switches in *Arabidopsis*. *Mol Plant*. 2010;**3**(6):1087–1103. <https://doi.org/10.1093/mp/ssf062>
- Zhong R, Ye Z-H.** MYB46 And MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol*. 2012;**53**(2):368–380. <https://doi.org/10.1093/pcp/pcr185>