Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Step out of the ERF comfort zone: in silico hunting for novel regulators of primary cell wall formation in celery collenchyma

Natalya Syrchina

ngsyrchina@gmail.com

FRC Kazan Scientific Center of RAS

Natalia Mokshina

FRC Kazan Scientific Center of RAS

Research Article

Keywords: celery (Apium graveolens L.), collenchyma, primary cell wall, transcription factor, ERF (ethylene response factor), coexpression network

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11511214/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Step out of the ERF comfort zone: in silico hunting for novel regulators of primary cell wall formation in celery collenchyma

Natalya Syrchina[1]* and Natalia Mokshina[1]

[1]Kazan Institute of Biochemistry and Biophysics, FRC Kazan Scientific Center of RAS, Lobachevsky Str., 2/31, Kazan, 420111, Russia

*https://orcid.org/0009-0002-0606-7268
ngsyrchina@gmail.com
+79046616101

Main conclusion

An integrative in silico approach identified eight non-ERF transcriptional regulators and proposed an interconnected regulatory network governing primary cell wall formation in celery collenchyma.

Abstract

The regulation of primary cell wall (PCW) formation remains considerably less understood than that of the secondary cell wall. Using RNA-seq data from celery (*Apium graveolens* L.) petiole tissues, we analyzed an *AgrERF* (ethylene response factor)-centered coexpression network associated with collenchyma, a tissue characterized by thickened yet extensible PCWs. The network contained eight transcriptional regulator genes from the Aux/IAA, KNOX, MADS-box, ZF-HD, THX, C2H2, atypical bHLH, and bZIP families. These genes showed strong collenchyma-associated expression and were coexpressed with genes involved in the collenchyma-specific biosynthesis and modification of the major PCW components, cellulose, hemicellulose, and pectin. Cis-element and pseudotime analyses of gene expression patterns suggested a highly interconnected regulatory system potentially involving both ERF- and IAA-associated pathways. Although the proposed relationships require experimental validation, the identified candidates provide a useful framework for future studies of plant PCW formation, particularly using celery collenchyma as a model system.

Key words: celery (*Apium graveolens* L.), collenchyma, primary cell wall, transcription factor, ERF (ethylene response factor), coexpression network

Author contribution

SN: conceptualization, methodology, software, data analysis, visualization, writing – original draft preparation; MN: project administration, supervision, data curation, writing – review and editing.

Funding

This work was performed with financial support from the government assignment for FRC Kazan Scientific Center of RAS, Kazan, Russia.

Data Availability Statement

Data are contained within the article and Supplementary Information. The RNA-seq data have been deposited at the European Nucleotide Archive (http://www.ebi.ac.uk/ena/) under the accession number PRJNA884180.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Introduction

Plant cell development is critically dependent on the cell wall, a dynamic supramolecular structure surrounding every plant cell. Depending on their biochemical composition, mechanical properties and physiological context, three types of plant cell wall can be distinguished: primary wall (PCW), secondary (SCW) (Alberts et al. 2002; Harris 2006), and tertiary wall (TCW, also known as the gelatinous layer or G-layer) (Gorshkova et al. 2018). While SCWs provide mechanical strength to vascular and fiber cells and TCWs enable specialized mechanical functions in fibers, PCWs are unique in being present in every plant cell and in remaining extensible during growth.

Despite their universal importance for plant development and morphogenesis, the transcriptional mechanisms coordinating PCW formation remain far less understood than those governing SCW biosynthesis. Consequently, elucidating the transcriptional regulation of PCW formation represents one of the major unresolved questions in plant cell wall biology.

PCWs are typically around 0.1 μm thick and are able to stretch during cell growth. PCWs consist almost entirely of polysaccharides synthesized by different glycosyltransferases (GTs). Approximately 30% of the PCW is cellulose, a β-1,4-glucan produced by cellulose synthase complexes containing a combination of six trimers of cellulose synthase A (CESA) subunits: CESA1, CESA3, and CESA6 (the latter can be replaced by CESA2, CESA5, or CESA9) (Persson et al. 2007; Nixon et al. 2016). Long cellulose molecules assemble into microfibrils that form the structural scaffold of the cell walls (Wolters-Arts et al. 1993; Crow and Murphy 2000). Between these microfibrils lies hemicellulose (≈25%), represented mainly by xyloglucan in dicots or mixed-linkage glucan (MLG) in grasses (*Poaceae*) synthesized by cellulose synthase-like (CSL) C and F proteins, respectively (Davis et al. 2010; Vega-Sánchez et al. 2012). Xyloglucan can be modified by xyloglucan endohydrolases (XEHs) and xyloglucan endotransglycosylases (XETs), both collectively referred to as XTHs and belonging to the GH16 protein family (Rose et al. 2002; Baumann et al. 2007). Hemicelluloses also include heteromannans, which are synthesized by CSLA proteins and are present as minor components of PCWs (Liepman et al. 2012). Heteromannans may influence cellulose organization and cell wall mechanics, and may contribute to tissue-specific wall properties. The space between cellulose microfibrils and hemicellulose molecules is filled with pectins (≈30%), mainly homogalacturonan (HG) and rhamnogalacturonan I (RG-I). The (1,4)-α-galacturonan backbone of HG is synthesized by GT8 family members (Galacturonosyltransferases, GAUTs) (Amos and Mohnen 2019), whereas RG-I rhamnosyltransferases (RRTs) from the GT106 family and RG-I galacturonosyltransferases (RGGATs) from the GT116 family synthesize the backbone of RG-I (Wachananawat et al. 2020; Amos et al. 2023). Because RG-I is one of the major pectic polysaccharides of celery collenchyma (Mokshina et al. 2025), regulation of its biosynthesis is likely to play an important role during thickened PCW formation. Additionally, pectin methyl esterases (PMEs) and their inhibitors (PMEIs) regulate HG methylesterification, thereby influencing cell wall mechanical properties (Peaucelle et al. 2015; Coculo and Lionetti 2022). Beyond methylesterification, acetylation of cell wall polymers (e.g., pectins and xyloglucans) is also widespread and can modulate wall properties (Gille and Pauly 2012; Zhang et al. 2024). In addition, expansins (EXPs) facilitate cell wall loosening by disrupting non-covalent interactions between cellulose and matrix polysaccharides without exhibiting enzymatic activity (Cosgrove et al. 1997; Park and Cosgrove 2012). Ultimately, coordinated deposition and remodeling of PCW polysaccharides depend not only on the activities of biosynthetic enzymes but also on precise transcriptional regulation of their encoding genes.

Unlike PCWs, transcriptional regulation of SCW formation has been investigated extensively during the past two decades. These studies established a transcriptional regulatory network (TRN) centered on NAC master regulators, including VND (VASCULAR-RELATED NAC DOMAINS, vessel-specific) and NST/SND (NAC SECONDARY WALL THICKENING PROMOTING FACTORS/SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEINS, fiber-specific) proteins, which activate downstream MYB transcription factors controlling the expression of genes required for cellulose, hemicellulose, and lignin biosynthesis (Nakano et al. 2015). Additional TF families further refine this network in a tissue- and developmental stage-specific manner. For SCWs, a broad hierarchical TRN has been described in *Arabidopsis thaliana* and subsequently partially validated in several other species. Among them, the class II KNOX transcription factor AtKNAT7 has attracted particular attention because its regulator activity is highly context-dependent: depending on its interacting partners, it can function either as a transcriptional activator or as part of transcriptional repressor complexes controlling SCW deposition (Yu 2019). This complexity illustrates that even well-characterized transcriptional networks rely on context-dependent interactions among multiple TF families rather than on linear regulatory cascades alone.

This remarkable progress stands in sharp contrast to the still limited understanding of PCW regulation. Despite the universal importance of PCWs, no transcriptional regulatory network comparable to the NAC-MYB hierarchy has yet been identified for PCW biosynthesis. Evidence accumulated for SCW strongly suggests that transcription factors rarely function in isolation but instead operate as components of interconnected regulatory hierarchies. Whether a similar organization exists for thickened PCWs remains completely unknown.

The trihelix (THX) family is currently the only TF family implicated in PCW regulation. This regulatory role has been described for the Type II PCW, characteristic of grasses and other commelinoid monocots and distinguished by the presence of MLG (Vogel 2008). In *Brachypodium distachyon*, *BdTHX1* is strongly coexpressed with *BdCSLF6*. BdTHX1 and BdTHX2 were also shown to bind different regions of both *BdCSLF6* and *BdXTH8* promoters (Fan et al. 2018).

Most TFs identified as influencing PCW formation in dicots are, in fact, negative regulators of SCW biosynthesis, preventing SCW deposition in tissues where it is unnecessary. These include AtWRKY12, AtCAN/AtHB15, and AtVNI2, which repress the activity of NAC-dependent SCW regulators in tissues that normally retain only PCWs (Wang et al. 2010; Yamaguchi et al. 2010; Du et al. 2015). In terms of positive regulators, only the APETALA2/ETHYLENE RESPONSE FACTOR (AP2/ERF) family has been implicated in dicot PCW regulation. Studies in *A. thaliana*, cotton (*Gossypium hirsutum*), and tomato (*Solanum lycopersicum*) consistently demonstrated that ERFs (AtERF034-041, GhERF41, SlERF.H5 and SlERF.H7 (both are homologs of AtERF040)) activate genes involved in cellulose biosynthesis, promote PCW deposition, and suppress lignification, indicating a conserved role in maintaining the PCW developmental program (Sakamoto et al. 2018; Saelim et al. 2019; Nakata et al. 2021; Gao et al. 2023; Pei et al. 2024). Notably, all previously described ERF proteins involved in PCW regulation belong to the IIId and IIIe phylogenetic subgroups (Nakano et al. 2006). In our recent work, we proposed that other ERF subgroups may also participate in this process. In celery (*Apium graveolens* L.), three ERFs — AgrERF002b, AgrERF041b, and AgrERF079 — belonging to the Vb, IIIe, and VIIIa subgroups, respectively, were identified as potential regulators of thickened PCW biosynthesis in collenchyma tissue, since genes encoding these TFs showed collenchyma-enriched expression and were coexpressed with putative PCW-related *AgrCESA* genes. Another candidate, AgrERF001b (Va), also showed elevated expression in collenchyma, but its expression was not collenchyma-specific, as it was also detected in leaves (Syrchina et al. 2025).

Celery collenchyma is a mechanical tissue characterized by thickened PCWs. Its walls, like those of celery parenchyma, are enriched in pectins compared with those of many other dicots and resemble thin PCWs in overall polysaccharide composition, with collenchyma containing higher proportions of xyloglucan (Leroux 2012; Zujovic et al. 2016; Chen et al. 2017; Chen et al. 2019). Importantly, collenchyma cells are capable of elongation while progressively thickening their walls, with wall deposition ceasing once elongation stops (Leroux 2012; Chen et al. 2019). This indicates that their thickened walls are exclusively primary in nature. Unlike the thin (~0.1 μm) PCWs typical of other celery petiole tissues, collenchyma walls can reach up to 4 μm without loss of expansion ability (Chen et al. 2019), making collenchyma a promising “non-classical” model for studying PCW formation. Because thickening requires increased deposition and modification of PCW components, genes involved in these processes might be expected to show enhanced expression in collenchyma. Based on this assumption, previous in silico studies generated a list of potential TF targets involved in thickened PCW formation among genes encoding synthases of major PCW components in collenchyma: *AgrCESAs*, *AgrCSLs*, *AgrGAUTs* and *AgrRRTs* (Mokshina et al. 2025; Mokshina 2025a,b). However, despite their broadly similar polysaccharide composition, celery parenchyma and collenchyma PCWs still differ in several biochemical and biophysical properties (Mokshina et al. 2023). Consequently, the TRN governing PCW formation is likely partially tissue-specific — analogous to SCW-associated TRNs in vessels and fibers (Nakano et al. 2015) — while still maintaining a conserved core.

In this work, we used transcriptome data from different celery petiole tissues characterized by distinct cell wall types — parenchyma with thin PCWs, collenchyma with thickened PCWs, and vascular bundles containing both thin PCWs and SCWs (Mokshina et al. 2025) — to construct a coexpression network based on previously identified *AgrERFs* potentially associated with collenchyma PCW formation: *AgrERF001b* (*Ag2G02839*), *AgrERF002b* (*Ag4G00940*), *AgrERF041b* (*Ag7G00319*), and *AgrERF079* (*AgUnG00085*) (Syrchina et al. 2025). TFs captured within these networks were considered additional candidate regulators of thickened PCW formation. Multiple in silico analyses of the transcriptome data were then performed to ensure that these TFs were not included in the network by chance, but instead were genuinely associated with regulation of PCW-related genes. A putative hierarchical TRN was reconstructed based on cis-element interactions, diffusion mapping, and a force-directed network layout derived from TF expression patterns. This work makes it possible to step off the beaten ERF path and opens new avenues for investigating other TF families in the context of plant PCW formation.

Materials and methods

Plant material and qPCR analysis

Celery plants (*Apium graveolens* L. var. secalinum 'Yablochny') were grown under previously described conditions (Mokshina et al. 2023). Three biological replicates of celery leaves, petioles, and roots were harvested at the 30-day developmental stage, immediately frozen in liquid nitrogen, and stored at −80 °C until analysis.

Total RNA was extracted using the FastPure Plant Total RNA Isolation Kit (Vazyme, China). RNA concentration and purity were assessed using a NanoPhotometer NP80 (Implen, Germany). Total RNA (1 μg) was reverse-transcribed into cDNA in a total reaction volume of 20 μL using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, USA).

Gene-specific primers were designed using the NCBI Primer-BLAST web tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Complete primer parameters (sequence, amplicon size, melting temperature, and amplification efficiency) and targeted gene details are provided in Table

S1. Primer specificity was confirmed by a single peak in the melting curve analysis. To construct standard curves, cDNA was diluted 5-fold, 25-fold, and 125-fold. For relative expression analysis, cDNA was used at a 10-fold dilution. Quantitative PCR was conducted in an 18-μL reaction volume containing 7.2 μL of 2.5× RT-PCR reaction mix with EvaGreen (Syntol, Russia), 3.8 μL of deionized water, 3.0 μL of diluted cDNA, and 4.0 μL of forward and reverse primer mixture (0.4 μM final concentration each). Thermal cycling conditions were as follows: initial denaturation at 95°C for 2 min, followed by 40 cycles of 94°C for 10 s (denaturation), 60°C for 15 s (annealing), and 72°C for 30 s (extension). All reactions were performed in technical duplicates across three biological replicates, alongside no-template controls (cDNA substituted with nuclease-free water).

Amplification efficiency (E) for each primer pair was calculated from standard curve slope values (Cq < 40) using the formula: E% = (10[1⁄slope]-1) × 100% (Radonic et al. 2004). Relative gene expression was normalized using the reference genes *TIP41* and *CACTIN* (Syrchina et al. 2025). To achieve normality and variance homogeneity, normalized values were log2-transformed prior to statistical testing. Group differences were evaluated using one-way ANOVA (Zar 2010) followed by Tukey's HSD post-hoc test for pairwise comparisons (Tukey 1949).

Transcriptome data analysis

RNA-Seq datasets previously deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena/) under accession number PRJNA884180 were used (Mokshina et al. 2025). Briefly, the dataset comprises transcriptomic profiles of celery petiole tissues, including collenchyma (Col), parenchyma (Par), vascular bundles (Vas), and collenchyma-free (CF) tissue fractions collected at two developmental stages: young (30-40 days, stage 1) and mature (80-90 days, stage 2). A total of 29 samples, including biological replicates for each tissue type and developmental stage, were analyzed: Par1 (4 replicates), Par2 (4 replicates), Col1 (3 replicates), Col2 (3 replicates), CF1 (4 replicates), CF2 (3 replicates), Vas1 (4 replicates), and Vas2 (4 replicates). Raw expression values were represented as total gene reads (TGRs), defined as the sum of reads mapped to coding sequences of individual genes. Differential expression analysis was performed in R using the DESeq2 package (Love et al. 2014), which applies internal size-factor normalization based on the median-of-ratios approach. Significantly regulated genes were identified using false discovery rate-adjusted p-values (padj). SRplot (Tang et al. 2023) was used to calculate gene-gene correlation coefficients.

Gene coexpression network based on *AgrERF* genes

Gene coexpression network analysis was performed in R using celery petiole transcriptome data with the WGCNA (Langfelder and Horvath 2008) and igraph (Csardi and Nepusz 2006) packages. Expression matrices were imported as semicolon-separated files and converted to numeric format, and genes with zero variance across samples were removed prior to analysis. A predefined set of *AgrERF* genes previously described as potential regulators of primary collenchyma cell wall formation was used as the reference gene set: *AgrERF001b* (*Ag2G02839*), *AgrERF002b* (*Ag4G00940*), *AgrERF041b* (*Ag7G00319*), and *AgrERF079* (*AgUnG00085*) (Syrchina et al. 2025). Pairwise Pearson correlation coefficients between genes were calculated using pairwise complete observations for missing values. Genes showing strong positive correlation with the *AgrERF* reference genes ($r \geq 0.7$) were retained, and for each *AgrERF* gene, up to 50 top positively correlated partners were selected to facilitate network visualization and interpretation. Correlation values were converted into an adjacency matrix by setting negative correlations to zero and applying a soft-thresholding power of 6 (selected based on the scale-free topology criterion) to emphasize stronger coexpression relationships. Edges with adjacency values <0.01 and isolated nodes were removed. The resulting undirected weighted network was visualized

using a Fruchterman-Reingold force-directed layout implemented in the igraph package (Fruchterman and Reingold 1991).

Gene annotation and functional characterization

All corresponding amino acid sequences for genes identified in the *AgrERF*-based network were analyzed using InterProScan version 5.48-83.0 (https://www.ebi.ac.uk/interpro/search/sequence/) to determine their protein domain organization. Amino acid sequences corresponding to TFs and proteins associated with PCW formation were selected for further analysis. These sequences were additionally subjected to BLAST searches in the Phytozome database (Goodstein et al. 2012; https://phytozome-next.jgi.doe.gov/blast-search) to identify the closest homologs in *A. thaliana*. Celery gene names were assigned based on the identified homologs.

Functional annotations of celery genes were performed using blast+ 2.2.28 (https://bioweb.pasteur.fr/packages/pack@blast+@2.2.28) against *A. thaliana* gene sequences and the online protein annotation tool Mercator4 (https://www.plabipd.de/mercator_main.html). Additionally, protein subcellular localization was predicted using DeepLoc 2.0 (Almagro Armenteros et al. 2017; https://services.healthtech.dtu.dk/services/DeepLoc-2.0/), and gene functional annotation was further refined using Swiss-Prot (Bairoch and Apweiler 2000), UniProt (The UniProt Consortium 2023; https://www.uniprot.org/), and TAIR (Huala et al. 2001; http://www.arabidopsis.org) databases.

To functionally characterize the transcriptional network, Gene Ontology (GO) enrichment analysis was performed on the set of 164 *A. thaliana* orthologs corresponding to the 184 celery genes from the *AgrERF*-based coexpression network. GO enrichment was conducted using the ShinyGO 0.85.1 web tool (Ge et al. 2020) with a false discovery rate (FDR) cutoff of 0.05. GO terms from the Biological Process, Cellular Component, and Molecular Function ontologies were analyzed. For cases where multiple celery genes corresponded to a single *A. thaliana* gene, duplicates were removed to avoid overrepresentation.

Phylogenetic analysis

Knotted-like homeobox (KNOX) TFs in celery were identified using HMMER-based domain searches (Eddy 2011). Protein sequences containing KNOX1 (PF03790), KNOX2 (PF03791), ELK (PF03789), and Homeobox KN (PF05920) domains were selected based on Pfam annotation (Mistry et al. 2021). KNOX protein sequences from *A. thaliana* were additionally retrieved from the TAIR10 database (Huala et al. 2001; http://www.arabidopsis.org) for comparative phylogenetic analysis. The AtBEL1 (AT5G41410) sequence was included as an outgroup to improve tree stability and clade resolution. Multiple sequence alignment was performed using MAFFT L-INS-i (Katoh and Standley 2013). The resulting alignment was used for phylogenetic reconstruction using IQ-TREE version 1.6.9 with a maximum-likelihood approach (Nguyen et al. 2015). The best-fit substitution model (JTT+F+G4) was selected automatically using ModelFinder under the Bayesian Information Criterion (Kalyaanamoorthy et al. 2017). Branch support was estimated using 1,000 ultrafast bootstrap replicates (Minh et al. 2013). The tree with the Robinson-Foulds distance (Robinson and Foulds 1981) equal to 0 was accepted as reliable. The phylogenetic tree was visualized using the iTOL web server (version 6.5.2; Letunic and Bork 2024; https://itol.embl.de/). Orthologous relationships between *A. thaliana* and celery protein sequences were further assessed using InParanoid-DIAMOND (Persson and Sonnhammer 2022) to validate the phylogenetic analysis.

Potential target genes selection

Eleven genes were selected as potential TF target genes based on the following criteria: (i) they showed preferential expression in young collenchyma (Col1), quantified using an expression ratio calculated as the mean relative expression in Col1 divided by the sum of the mean relative expression values in Col2, Par1, Par2, Vas1, and Vas2, with a ratio >30; and (ii) their *A. thaliana* homologs encode proteins associated with PCW formation. Additionally, *AgrCESA* (Mokshina 2025a), *AgrCSL* (Mokshina 2025b), and pectin synthase genes (Mokshina et al. 2025), previously described as potentially involved in collenchyma PCW thickening, were also included as candidate target genes. The complete list of potential target genes and their associated information is presented in Table S2.

Cis-element analysis

The 2,000-bp upstream regions of potential target genes and selected TF genes (including the *AgrERF* genes analyzed in this study) were extracted from the celery genome sequence using the TBtools-II toolbox (Chen et al. 2023). Cis-elements within these sequences were predicted using PlantPAN 4.0 (Chow et al. 2024; https://plantpan.itps.ncku.edu.tw/plantpan4/index.html).

Diffusion map analysis

Diffusion map analysis was performed on a selected set of *AgrERF* and associated TF genes identified from the previously described coexpression analysis (Coifman and Lafon 2006). The analysis was implemented in R and visualized using the ggplot2 (Wickham 2016) and ggrepel (Slowikowski 2024) packages. A Pearson correlation matrix for the selected genes was used as the input similarity matrix. Correlation values were converted into distances according to D = 1 - R, where R represents the Pearson correlation coefficient. A Gaussian kernel was constructed using a sigma parameter defined as the median of all nonzero distances. The kernel matrix was subsequently normalized by its row sums to generate a transition matrix, which was then subjected to eigendecomposition. The first and second diffusion components (DC1 and DC2) were used for low-dimensional embedding and visualization. A pseudotime-like score was calculated by min–max normalization of DC1 values to describe the relative positions of genes along the first diffusion component.

Force-directed gene network analysis

Force-directed network analysis was performed on the same set of genes identified from the previously described coexpression analysis. The analysis was implemented in R using the igraph package. A Pearson correlation matrix was used as the input for network construction. Only correlations with $r \geq 0.7$ were retained , and self-connections were removed prior to construction of an undirected weighted network. Isolated nodes without remaining connections were subsequently excluded from the analysis. Network visualization was performed using the Fruchterman-Reingold force-directed layout algorithm with 2,000 iterations. Degree, betweenness, and eigenvector centrality measures were calculated to characterize node importance. Node size was scaled according to degree centrality, while edge width was proportional to edge weight. Community structure within the network was identified using the Louvain clustering algorithm (Blondel et al. 2008), and the network was visualized with nodes colored according to their module assignment.

Results

An AgrERF-centered network identifies additional candidate regulators of thickened PCW formation

The coexpression network constructed around four *AgrERF* genes contained 184 additional genes and formed a single connected component. Two partially separated groups were identified: group I around *AgrERF001b* and *AgrERF079*, and group II around *AgrERF002b* and *AgrERF041b* (Fig. 1). InterProScan analysis of amino acid sequences corresponding to genes from the network revealed eight genes encoding TFs from different protein families: MADS-box, zinc-finger homeodomain (ZF-HD), KNOX, THX, C2H2-type zinc finger (C2H2 ZF), basic helix-loop-helix (bHLH), Aux/IAA (not a TF itself, but a direct negative regulator of auxin response TFs), and basic leucine zipper (bZIP). BLAST analysis against the closest homologs in *A. thaliana* allowed assignment of the following names: *AgrAGL62* (*AGAMOUS-LIKE 62*, *Ag10G01798*), *AgrHB21* (*HOMEOBOX DOMAIN PROTEIN 21*, *Ag10G02150*), *AgrKNAT7* (*Ag2G01615*), *AgrPTL* (*PETAL LOSS*, *Ag3G00597*), *AgrZAT4* (*ZINC FINGER OF ARABIDOPSIS THALIANA 4*, *Ag5G00091*), *AgrPRE5* (*PACLOBUTRAZOL RESISTANCE 5*, *Ag5G00597*), *AgrIAA4* (*Ag9G02083*), and *AgrbZIP76* (*AgUnG00415*). *AgrHB21*, *AgrZAT4*, *AgrPRE5*, *AgrIAA4*, and *AgrbZIP76* were assigned to group I; *AgrAGL62*, *AgrKNAT7*, and *AgrPTL* were assigned to group II; notably, *AgrKNAT7*, *AgrAGL62*, and *AgrPRE5* also occupied intermediate positions between the two network parts.

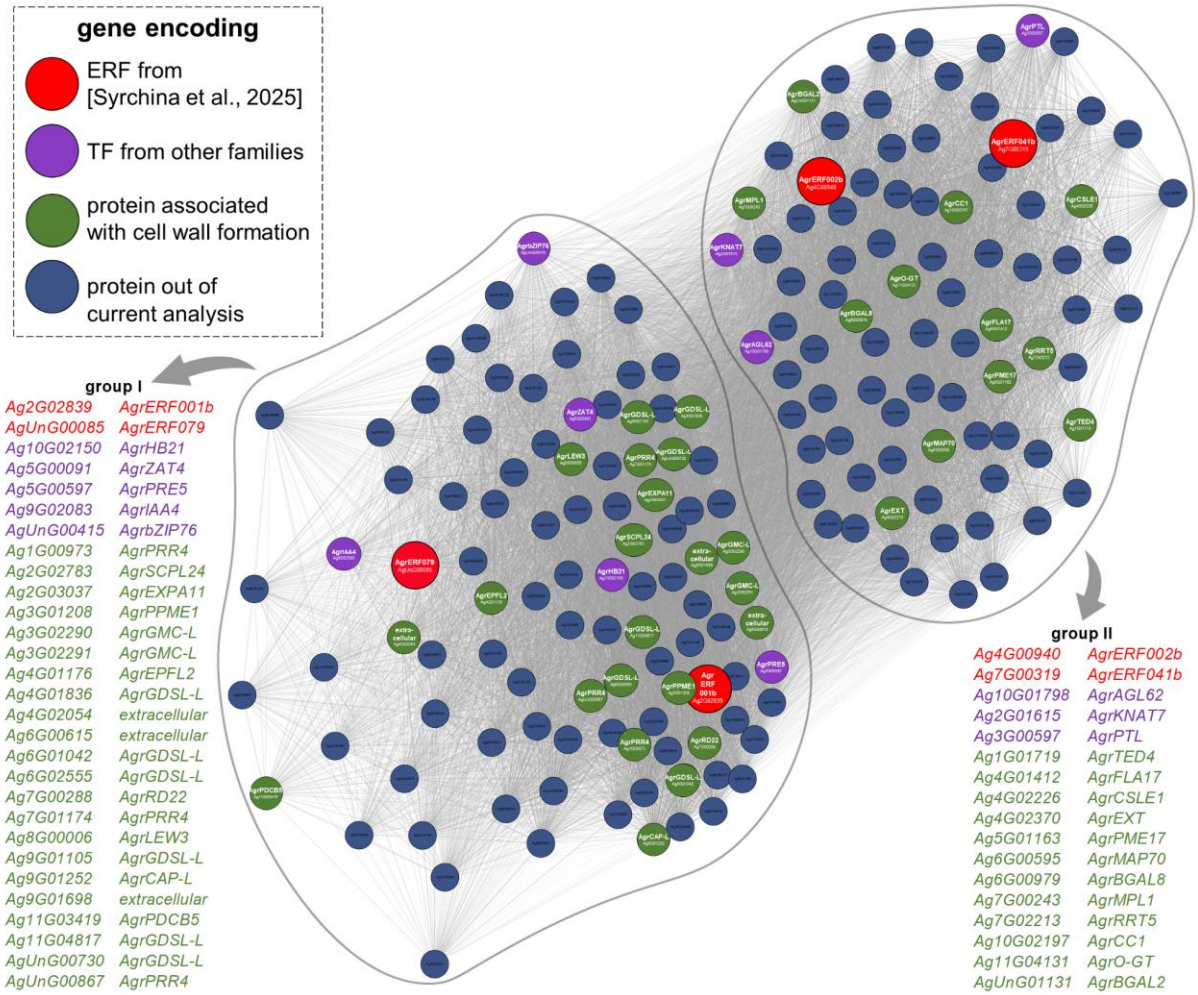


Fig. 1. *AgrERF*-based interconnected coexpression network built using celery petiole RNA-seq data (Mokshina et al. 2025). TF — transcription factor.

In addition to TF genes, approximately 18.5% of the network genes encode proteins involved in plant cell wall biosynthesis and modification. Their identification and annotation were based on homology to *A. thaliana* genes encoding cell wall-related proteins, protein subcellular localization predicted to the cell wall, plasma membrane, or Golgi apparatus using DeepLoc 2.0 (Almagro Armenteros et al. 2017), and literature evidence supporting their involvement in cell wall metabolism. Based on these criteria,

the following genes were selected: *AgrCC1* (*COMPANION OF CELLULOSE SYNTHASE 1*, *Ag10G02197*), *AgrO-GT* (*O-GLYCOSYL TRANSFERASE WITH DUF821*, *Ag11G04131*), *AgrEXPA11* (α-expansin, *Ag2G03037*), *AgrPPME1* (*POLLEN PECTIN METHYLESTERASE 1*, *Ag3G01208*), *AgrFLA17* (*FASCICLIN-LIKE ARABINOGALACTAN PROTEIN 17*, *Ag4G01412*), *AgrCSLE1* (*Ag4G02226*), *AgrPME17* (*Ag5G01163*), *AgrBGAL2* (*β-GALACTOSIDASE 2*, *AgUnG01131*), *AgrLEW3* (*LEAF WILTING 3*, *Ag8G00006*), and *AgrMAP70* (*MICROTUBULE-ASSOCIATED PROTEIN 70*, *Ag6G00595*). Additionally, *AgrRRT5* (*Ag6G00233*), involved in RG-I backbone biosynthesis (Wachananawat et al. 2020), and several GDSL esterases/lipases with predicted extracellular localization, including *AgrGDSL1* (*Ag1G00554*), *AgrGDSL2* (*AgUnG00730*), and *AgrGDSL3* (*Ag8G01854*), which may function in polysaccharide deacetylation (Zhang et al. 2017; Mokshina et al. 2025), were also identified as cell wall-related candidates.

The network also contained genes associated with hormone signaling pathways. Auxin signaling-related genes, including the auxin-responsive SAUR family gene *AgrSAUR67* (*Ag7G01777*), the auxin influx carrier *AgrAUX1-like* (*Ag8G01325*), the auxin-responsive BIG GRAIN 1-LIKE E protein *AgrBIG* (*Ag8G02095*), and the auxin canalization-related protein *AgrVAN3* (*Ag8G01568*) were revealed. One gene *AgrSCP24* (*Ag2G02783*), encoding a serine carboxypeptidase involved in brassinosteroid signaling, was also present. No genes with clear ethylene response annotation were identified in the network.

To characterize the biological functions associated with the network, GO enrichment analysis was performed on its 164 corresponding *A. thaliana* orthologs. The analysis revealed significant enrichment of genes involved in processes closely linked to cell wall regulation and plant development. The most compelling enrichments included auxin transport (Fold Enrichment ~8.5, FDR = 3.1E-02), regulation of auxin polar transport (Fold Enrichment ~21.9, FDR = 5.0E-03), and hormone transport (Fold Enrichment ~7.7, FDR = 1.5E-02) (Fig. S1). Lipid catabolic and metabolic processes were also significantly overrepresented (Fold Enrichment ~4.4-6.5), along with cutin biosynthetic processes and cuticle development (Fold Enrichment ~28.3, FDR = 2.2E-03). Cellular Component enrichment highlighted a strong association with the endomembrane system, particularly the endoplasmic reticulum (Fold Enrichment ~2.6, FDR = 1.2E-02), while Molecular Function analysis pointed to hydrolase activity acting on ester bonds (Fold Enrichment ~2.6, FDR = 3.4E-02). These results, especially the enrichment of auxin-related and lipid metabolic pathways, strongly support the network's association with cell wall-related processes and its proposed regulatory complexity involving auxin signaling. Complete details for all network genes are provided in Table S3.

Selected transcription factor genes are expressed significantly higher in collenchyma than in other celery petiole tissues

All selected TF genes showed significantly higher expression in young collenchyma than in the other young petiole tissues (Fig. 2). However, only *AgrHB21*, *AgrPTL*, *AgrPRE5*, and *AgrIAA4* reached expression levels above 1,000 TGR in young collenchyma. *AgrAGL62* showed the lowest expression level among all analyzed TF genes, whereas *AgrIAA4* was the most highly expressed gene across all samples. Only *AgrPTL* and *AgrIAA4* showed significant differences between mature collenchyma and other mature petiole tissues. In mature samples, *AgrPTL* displayed the highest expression level in collenchyma, whereas *AgrIAA4* showed the lowest expression level in mature collenchyma compared with the other mature tissues.

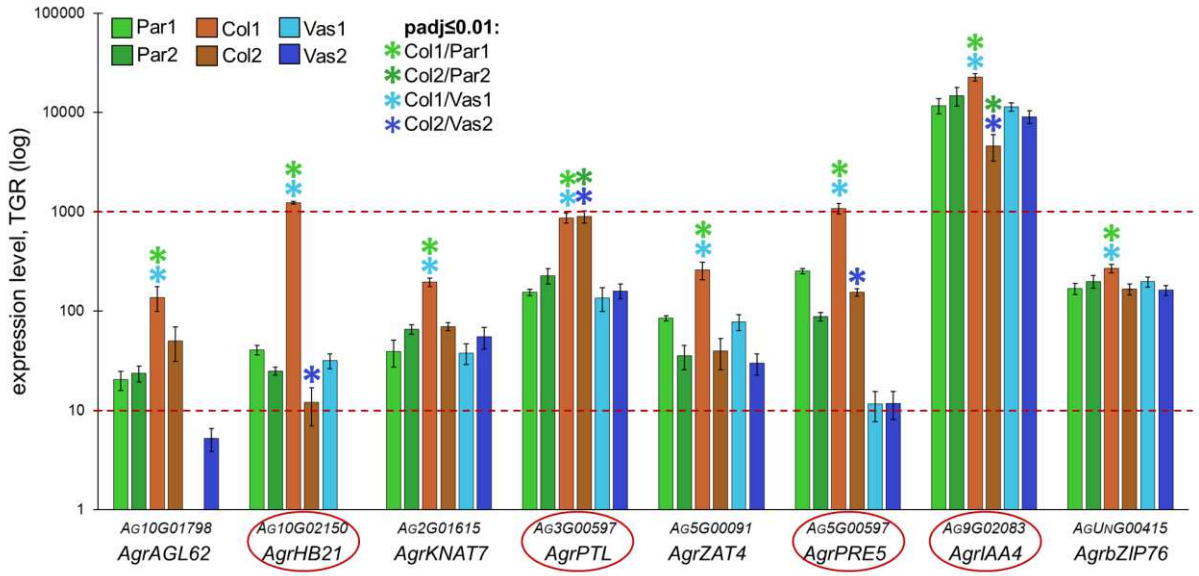


Fig. 2. Expression analysis of genes encoding transcription factors from the *AgrERF*-based coexpression network using celery petiole RNA-seq data (Mokshina et al. 2025). Green and blue asterisks indicate significant differences between collenchyma (Col) and parenchyma (Par) or between collenchyma and vascular bundles (Vas), respectively, at padj ≤ 0.01 (within the same developmental stage). Red circles highlight genes whose expression in young collenchyma exceeds 1,000 total gene reads (TGR). 1 and 2 represent young (30 days) and mature (80-90 days) plant developmental stages, respectively.

All genes except *AgrAGL62* and *AgrHB21* were expressed in all samples and showed peak expression in young collenchyma. *AgrAGL62* was not expressed in the young vascular bundle, whereas *AgrHB21* was not expressed in the mature vascular bundle sample. The differences in expression between young and mature stages in parenchyma and vascular bundle tissues varied among genes, and no apparent common pattern was observed.

Selected transcription factor genes are expressed significantly higher in petioles than in leaves and roots

Quantitative PCR analysis of 30-day-old celery plants revealed that most of selected TF genes were expressed at significantly higher levels in petioles — which contain collenchyma — than in collenchyma-lacking organs (leaves and roots) (Fig. 3).

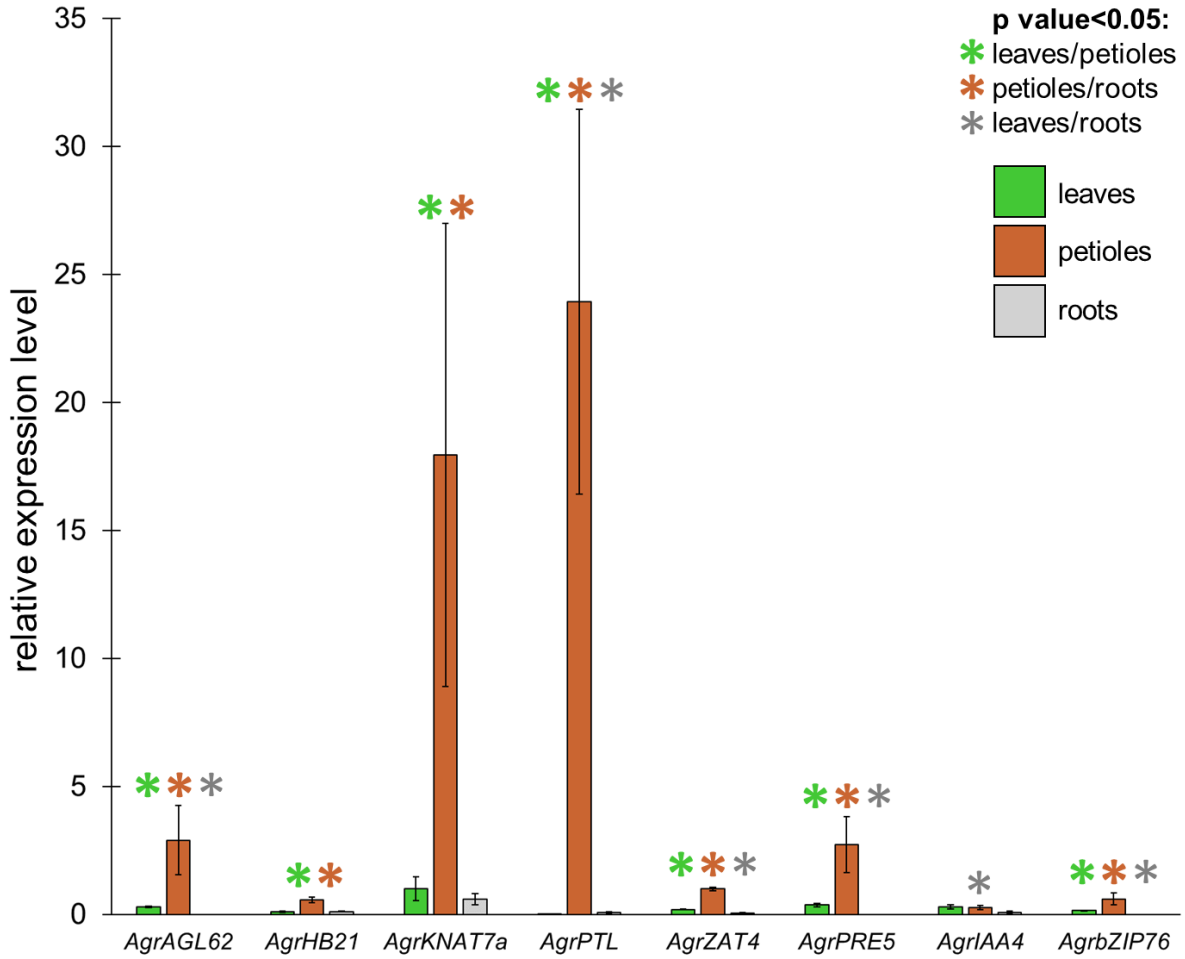


Fig. 3 Relative expression of genes encoding transcription factors from the AgrERF-based coexpression network in leaves, petioles and roots of 30-day-old celery plants determined by qPCR. Expression levels are shown relative to the reference genes TIP41 and CACTIN. Green, brown, and grey asterisks indicate significant differences between leaves and petioles, petioles and roots, or leaves and roots, respectively, at p value < 0.05.

The relative ranking of the transcription factor genes was only partially consistent between the tissue-specific RNA-seq and the organ-level qPCR datasets. *AgrAGL62*, *AgrKNAT7a*, *AgrPTL*, *AgrZAT4* and *AgrbZIP76* showed broadly similar relative positions in both analyses, whereas *AgrHB21*, *AgrPRE5,* and *AgrIAA4*, which were among the most highly expressed genes in the RNA-seq data, displayed low expression in whole petioles. Notably, *AgrIAA4* was the only gene that showed no significant difference in expression between leaves and petioles. Conversely, *AgrPTL* was the most highly expressed gene in petioles, with expression levels over 5-fold higher than those of the other genes, while remaining relatively low in leaves and roots. *AgrAGL62*, *AgrPRE5* and *AgrbZIP76* showed no detectable expression in the roots.

Celery orthologs of KNAT7 have different expression patterns

Due to the role of AtKNAT7 in SCW formation (Nookaraju et al. 2022), a comprehensive analysis of celery orthologs was performed. Identification of KNOX family members in the celery genome revealed 17 genes, compared with 9 in the *A. thaliana* genome. Phylogenetic analysis of corresponding protein sequences combined with InParanoid-DIAMOND homolog searching (Fig. 4a, Table S4) showed that AtKNATM has no celery orthologs, whereas AtSTM (SHOOT MERISTEMLESS) has three celery orthologs (AgrSTMa-c), AtKNAT1 has two (AgrKNAT1a-b), AtKNAT2/AtKNAT6 have three (AgrKNAT2/6a-c), AtKNAT3/AtKNAT4 have five (AgrKNAT3/4a-e), AtKNAT5 has no orthologs, and

AtKNAT7 has four orthologs (AgrKNAT7a-d) encoded in the celery genome. The TF gene identified in the coexpression network (Fig. 1) encoding one of these four AgrKNAT7s (Ag2G01615) was therefore designated *AgrKNAT7a*.

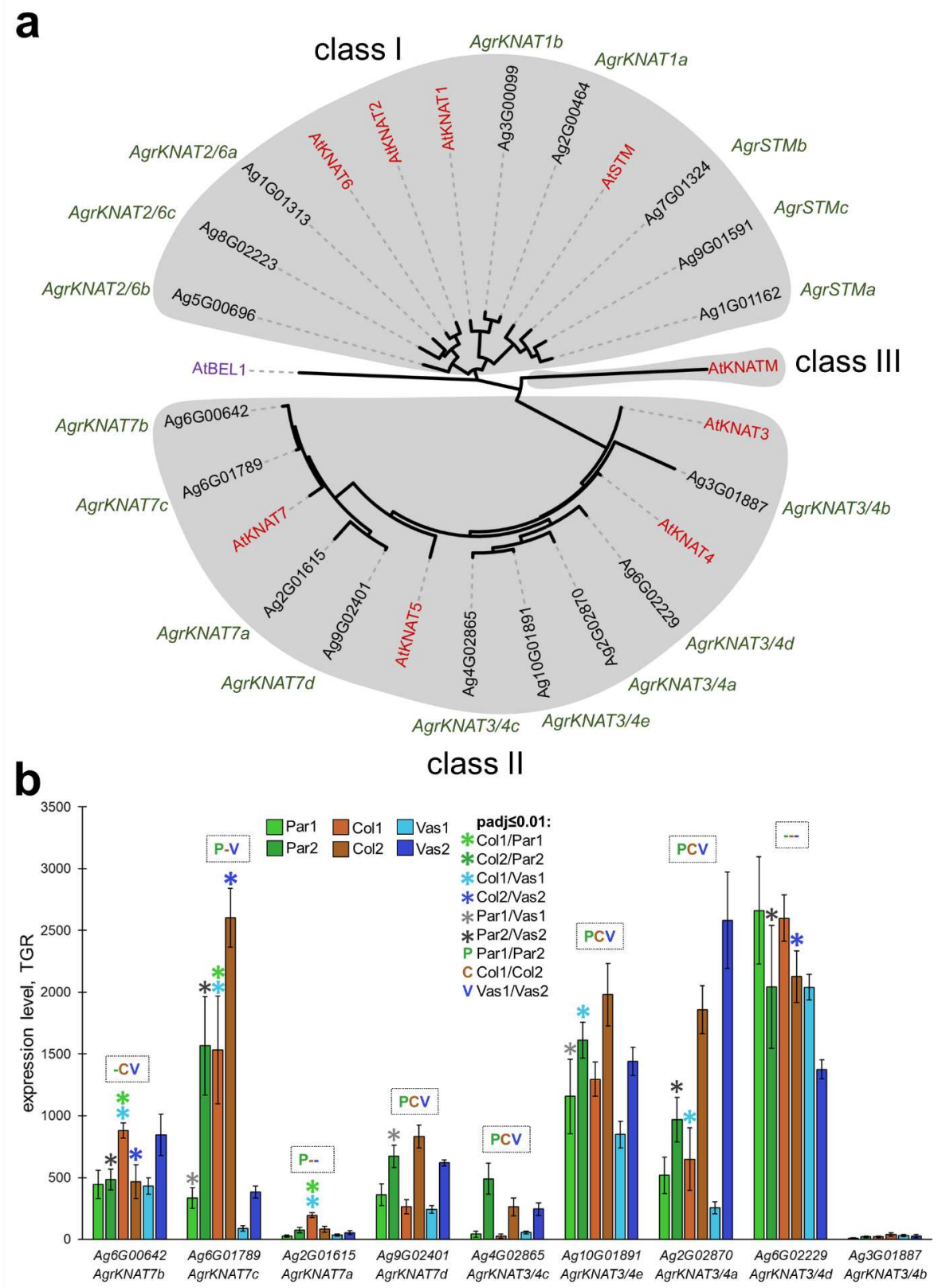


Fig. 4 Analysis of KNOX family members in celery. (a) Comparative phylogenetic analysis of KNOX family members encoded in the celery genome (black) and *A. thaliana* genome (red); AtBEL1 is shown in purple and was used as an outgroup to improve tree stability and clade resolution; celery short protein names assigned based on the closest *A. thaliana* homologs are shown in green. (b) Gene expression analysis of celery KNOX class II proteins using celery petiole RNA-seq data (Mokshina et al. 2025):

green, blue, and grey asterisks indicate significant differences between collenchyma (Col) and parenchyma (Par), between collenchyma and vascular bundles (Vas), or between parenchyma and vascular bundles, respectively, at padj ≤ 0.01 (within the same developmental stage); letters P, C, and V indicate significant differences between young (1; 30-40-day-old) and mature (2; 80-90-day-old) stages of parenchyma, collenchyma, and vascular bundles, respectively, at padj ≤ 0.01. TGR — total gene reads.

Given the established role of AtKNAT7 in cooperation with AtKNAT3 during SCW deposition (Wang et al. 2020; Qin et al. 2020), together with the presence of four celery AtKNAT7 homologs and five celery AtKNAT3/4 homologs, we examined their expression patterns across celery petiole tissues in greater detail. This comparative analysis revealed distinct expression patterns (Fig. 4b). Among the four *AtKNAT7* orthologs, three genes — *AgrKNAT7a*, *AgrKNAT7b* (*Ag6G00642*), and *AgrKNAT7c* (*Ag6G01789*) — showed significantly higher expression levels in young collenchyma compared with other young tissues; however, only *AgrKNAT7a* and *AgrKNAT7b* exhibited peak expression in this tissue among all samples. *AgrKNAT7a* appeared to be more collenchyma-specific, as *AgrKNAT7b* also showed high expression levels in mature vascular bundles. *AgrKNAT7c*, *AgrKNAT7d* (*Ag9G02401*), *AgrKNAT3/4a* (*Ag2G02870*), *AgrKNAT3/4c* (*Ag4G02865*), and *AgrKNAT3/4e* (*Ag10G01891*) showed significantly higher expression levels in mature tissues than in young tissues, although *AgrKNAT7c* did not show significant differences in collenchyma between stages despite following the same overall trend. Meanwhile, *AgrKNAT3/4d* (*Ag6G02229*) showed higher expression in young tissues, with peaks in parenchyma and collenchyma, although these differences were not significant. Only *AgrKNAT3/4b* (*Ag3G01887*) showed no apparent differences in expression among the samples, with its expression level being extremely low.

Selected transcription factor genes and their candidate targets contain predicted cis-regulatory motifs

Analysis of cis-elements in the upstream sequences of the selected TF genes (including the *AgrERFs*) and their possible target genes showed that all analyzed sequences contained numerous cis-elements predicted to be strongly bound by TFs from the ZF-HD, C2H2 ZF, bHLH, and bZIP protein families (Fig. 5). In addition, all sequences contained cis-elements associated with TF families not included in the present analysis, including NAC, MYB, GATA, WRKY, and nuclear factor Y (NF-Y) (data not shown).

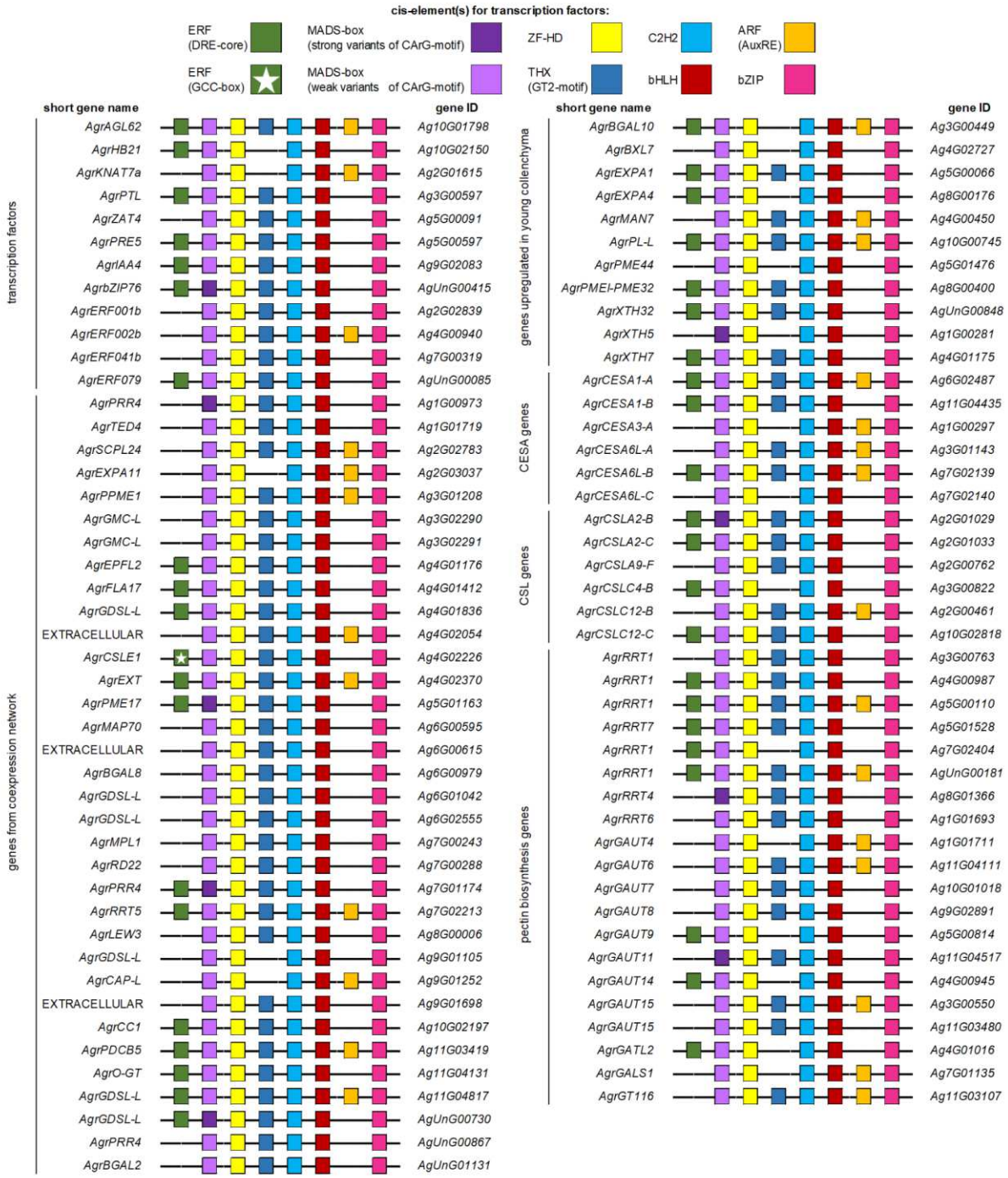


Fig. 5 Predicted cis-regulatory elements in the upstream sequences of transcription factor genes and their potential target genes. The figure indicates the presence or absence of these cis-elements, but does not represent their actual positions, abundance, or order within the sequences.

In the upstream sequences of the selected TF genes, DRE-core motifs for ERF binding were identified in *AgrERF079*, *AgrAGL62*, *AgrHB21*, *AgrPTL*, *AgrPRE5*, *AgrIAA4*, and *AgrbZIP76*. The CArG-box motif, recognized by MADS-box TFs, was strongly represented only in the *AgrbZIP76* upstream sequence, whereas weaker variants were present in all analyzed genes. The GT2-motif, a cis-element recognized by THX, was identified in all sequences except *AgrHB21* and *AgrKNAT7a*. AuxRE motifs, recognized by auxin response factors (ARFs), which are regulated by Aux/IAA, were detected only in *AgrAGL62*, *AgrKNAT7a*, and *AgrERF002b* (Fig. 5).

The upstream regions of all possible target genes also contained numerous cis-elements predicted to bind TFs from the ZF-HD, C2H2 ZF, bHLH, and bZIP families. Weak variants of the CArG-box motif were detected in all analyzed sequences. In addition, most genes contained DRE-core motifs (and GCC-box instead of this in a single case of *AgrCSLE1*) and/or GT2-motifs. In some cases, AuxRE motifs or strong CArG-box motifs were additionally identified. Among them, *AgrXTH5* contained only a strong CArG-box motif, whereas *AgrCESA3-A*, *AgrGAUT4*, *AgrGALS2*, *AgrEXPA11*, and *AgrCAP-L* contained only AuxRE motifs. *AgrBXL7*, *AgrPME44*, *AgrCESA6-C*, and one of the *AgrGDSL-Ls* (*Ag9G01105*) lacked DRE-core, GT2-motif, AuxRE, and strong CArG-box motif (Fig. 5).

Selected transcription factor genes show positive expression correlations with their possible target genes

All selected TF genes showed positive expression correlations with their possible target genes (Fig. S2). Previously, *AgrERF* genes were selected based on positive correlations with genes highly expressed in collenchyma compared with other celery petiole tissues and with PCW-associated *AgrCESA* genes, together with negative correlations with SCW-associated *AgrCESA* genes (*AgrCESA4*, *AgrCESA7*, and *AgrCESA8*) (Syrchina et al. 2025). In the present analysis, we additionally evaluated the coexpression of *AgrERFs* with *AgrCSLA* and *AgrCSLC* genes, as well as with pectin biosynthesis genes (*AgrRRTs* and *AgrGAUTs*). The results showed strong positive correlations ($r > 0.67$) with *AgrCSLA and AgrCSLC* genes, with *AgrERF001b* showing the highest correlations and *AgrERF079* the lowest. A similar pattern was observed for pectin biosynthesis genes, including *AgrRRTs* and *AgrGAUTs* ($r > 0.52$).

For the other TF genes, we analyzed their coexpression with genes highly expressed in collenchyma, PCW- and SCW-associated *AgrCESAs*, *AgrCSLAs* and *AgrCSLCs*, and pectin biosynthesis genes. All of these TF genes showed strong positive correlations with genes highly expressed in collenchyma. *AgrZAT4*, *AgrHB21*, and *AgrAGL62* had the highest correlation coefficients (approximately $r = 0.90$), whereas *AgrPTL* showed the lowest correlations ($r = 0.50$-0.67). All TF genes showed negative correlations with SCW-associated *AgrCESA* genes, with *AgrKNAT7a* displaying the most negative coefficients, and positive correlations with PCW-associated *AgrCESA* genes. *AgrCESA1-A*, *AgrCESA1-B, AgrCESA3-A*, and *AgrCESA6L-A* showed high correlation coefficients ($r = 0.62$-0.93) with all TF genes except *AgrIAA4*. In contrast, *AgrCESA6L-B* and *AgrCESA6L-C* overall showed lower correlations ($r = 0.02$-0.81; $r = 0.27$-0.76), even close to zero in the case of *AgrCESA6L-B and AgrIAA4*. *AgrPTL* showed the highest correlation with *AgrCESA6L-A* ($r = 0.83$) and the lowest correlation with *AgrCESA6L-C* ($r = 0.27$) among all TF genes.

AgrCSL, *AgrRRT*, and *AgrGAUT* genes showed strong positive correlations with all selected TF genes, with the highest values observed for *AgrAGL62*, *AgrHB21*, *AgrZAT4*, and *AgrPRE5*. However, *AgrGAUT15* showed the lowest correlation coefficients among the analyzed *AgrGTs* across all TF genes ($r = 0.45$-0.75), with the highest value observed for *AgrHB21* ($r = 0.80$). All data are presented in Fig. S2.

Selected transcription factor gene coexpression patterns form the basis for pseudotime analysis and force-directed gene network construction

The correlation matrix revealed predominantly strong positive correlations between most genes, with Pearson correlation coefficients ranging from weak to very high values (Fig. 6a). The strongest associations were observed between *AgrHB21* and *AgrPRE5* ($r = 0.97$), *AgrPRE5* and *AgrERF001b* ($r = 0.96$), and *AgrHB21* and *AgrERF001b* ($r = 0.95$). In contrast, *AgrIAA4* and *AgrPTL* exhibited the weakest correlation ($r = 0.16$).

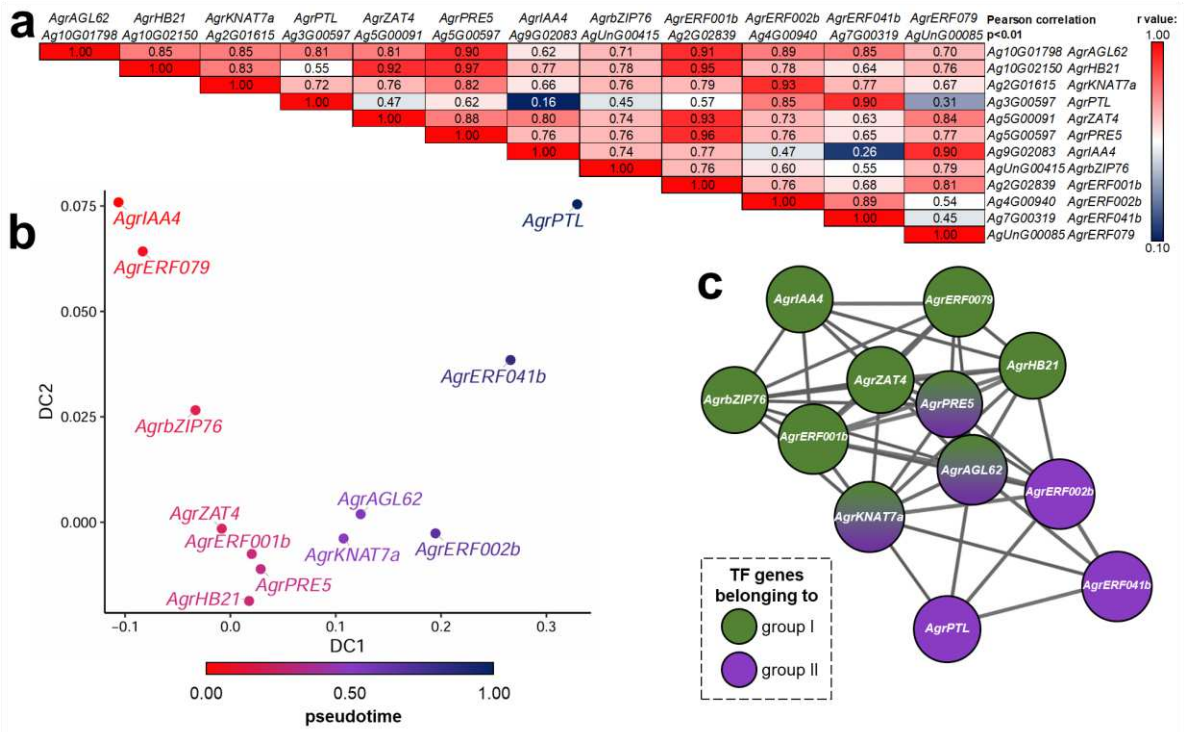


Fig. 6 Coexpression relationships among transcription factor (TF) genes from the *AgrERF*-based coexpression network, including *AgrERFs*, using celery petiole RNA-seq data (Mokshina et al. 2025). (a) Pearson correlation coefficient values for TF genes ($p < 0.01$): the highest and lowest r values are shown in red and blue, respectively; the correlation matrix was built based on total gene read values. (b) Diffusion map of TF gene coexpression based on diffusion components 1 and 2 (DC1 and DC2) with pseudotime calculation (red font indicates earlier-expressed genes, whereas blue font indicates later-expressed genes). (c) Force-directed coexpression network of TF genes; green circles correspond to group I from the *AgrERF*-based coexpression network, and purple circles correspond to group II.

Diffusion embedding analysis distributed the genes along two major diffusion components (DC1 and DC2) (Fig. 6b). Several genes clustered closely within the diffusion space: (i) the pair *AgrIAA4* and *AgrERF079*; and (ii) a larger cluster containing *AgrAGL62*, *AgrHB21*, *AgrKNAT7a*, *AgrZAT4*, *AgrPRE5*, *AgrERF001b*, and *AgrERF002b*. In contrast, *AgrbZIP76*, *AgrERF041b*, and *AgrPTL* occupied more isolated positions. Distinct positioning along the primary diffusion axis was also observed, with *AgrIAA4* and *AgrERF079* located at the beginning of the pseudotime trajectory and *AgrERF041b* and *AgrPTL* at the end. These differences indicate distinct positions within the diffusion-based organization of the TF coexpression patterns but do not, by themselves, establish the temporal order of gene expression during collenchyma differentiation.

The force-directed coexpression network (Fig. 6c) revealed a densely interconnected transcriptional structure composed of two major clusters corresponding to the previously identified coexpression groups. Group I included *AgrHB21*, *AgrZAT4*, *AgrIAA4*, *AgrbZIP76*, *AgrERF001b*, and *AgrERF079*, while group II comprised *AgrPTL*, *AgrERF002b*, and *AgrERF041b*. Despite the separation into two major modules, numerous intergroup connections were retained within the network, indicating extensive coexpression relationships between genes from different clusters. *AgrPRE5*, *AgrAGL62*, and *AgrKNAT7a* occupied intermediate positions between the two groups and formed multiple connections with genes from both modules. In contrast, *AgrIAA4* and *AgrERF079* occupied more external positions within the green group I cluster, whereas *AgrPTL* and *AgrERF041b* were localized at the periphery of the purple group II cluster. Notably, genes from group I were positioned closer to each other than genes from group II, while the intermediate genes, together with *AgrERF002b*, were positioned closer to group

I than as isolated nodes. The observed intermodule interactions complemented the diffusion-based pseudotime ordering by revealing direct coexpression links between genes distributed across different regions of the transcriptional gradient, indicating interconnected regulatory systems rather than fully isolated coexpression modules.

Discussion

The TF-based molecular regulation of PCW formation remains poorly understood, and only a few TFs beyond the ERF family have been implicated in this process (Mokshina and Syrchina 2025). In our previous work, we identified several AgrERFs potentially associated with thickened PCW formation in celery collenchyma (Syrchina et al. 2025). However, the limited diversity of ERFs captured in that analysis prompted us to explore whether additional TF families might be involved in this regulatory network. Using an *AgrERF*-centered coexpression network, we identified eight TF genes from distinct families that showed strong collenchyma-specific expression and connectivity to PCW-related metabolic genes. In this study, we aimed to move beyond the ERF family and explore TFs outside this group. Below, we discuss each of these TFs in the context of literature data and additional data generated in this study, including qPCR analysis showing that most selected TF genes were expressed significantly higher in collenchyma-containing petioles than in leaves and roots.

AgrERFs: central connectors between PCW biosynthesis and regulatory networks

We begin with additional observations regarding *AgrERFs*. All of them showed strong coexpression with genes encoding PCW-associated hemicellulose and pectin synthases, in addition to the previously reported strong coexpression with PCW-associated *AgrCESA* genes and the “top-10” genes highly expressed in young collenchyma (Syrchina et al. 2025). Almost all potential target genes contained the DRE-core motif for ERF binding, so AgrERFs appear to be strong regulators of collenchyma PCW formation. In the upstream sequences of the analyzed TFs, half of them contained DRE-core motifs, and only *AgrERF079* contained such a motif among the *AgrERFs* themselves. Due to *AgrERF079*’s early position in the pseudotime trajectory, it is possible that an additional upstream AgrERF regulator exists above AgrERF079. In contrast, the *AgrERF002b* upstream gene sequence contains an AuxRE motif for ARF binding, suggesting that it may be part of a (semi-)separated regulatory system from the AgrERF079-governed pathway (see explanation below). According to ePlant data (https://bar.utoronto.ca/eplant/), AtERF79 has a potential protein-protein interaction with AtNST1, one of the master switches of SCW formation in fibers (Mitsuda et al. 2007; Zhong et al. 2007). AgrERF079 may therefore be associated with regulatory processes linked to suppression or modulation of SCW-associated programs when it is not required.

AgrIAA4: a potential repressor of PCW thickening

In our analysis, we also included the AgrIAA4 (Ag9G02083). AgrIAA4 encodes an Aux/IAA protein, a member of the Aux/IAA family, that does not itself bind DNA but negatively regulates ARF activity, thereby modulating auxin-dependent processes associated with plant growth, cell expansion, and cell wall formation (Zhuang et al. 2025). However, some studies have shown non-auxin-dependent roles of IAA4 homologs in cell wall formation and modification. Overexpression of a stabilized form of *EgrIAA4* (*Eucalyptus*) in *A. thaliana* caused severe growth defects, auxin insensitivity, and strongly delayed lignification of interfascicular fibers, while SCWs in xylary fibers were nearly absent, indicating a role in fiber differentiation and SCW deposition (Yu et al. 2015). Similarly, in *Camellia sinensis*, CsIAA4 interacts with the R2R3-MYB TF CsMYB4a, which stabilizes CsIAA4 and represses auxin-responsive genes involved in cell elongation (Ma et al. 2024). MYB4 homologs are known regulators of cell wall formation through modulation of lignin, cellulose, and SCW biosynthesis genes in different species. For example, AtMYB4 represses phenylpropanoid and lignin-related genes (Jin et al. 2000; Ko et al. 2009),

while homologs in cotton, birch (*Betula pendula*), and *Pinus massoniana* regulate SCW deposition and lignification in a species- and context-dependent manner (Qiao et al. 2025; Yu et al. 2021; Yao et al. 2021).

ARFs downstream of Aux/IAA proteins can also act as protectors of PCW integrity, as shown in grape (*Vitis vinifera*), where VvARF19 represses PCW loosening and fruit softening via inhibition of VvLBD13-mediated activation of cell wall-remodeling genes (Li et al. 2025). Although ARFs were not explicitly detected in our current network, they may still participate in the proposed regulatory scheme of collenchyma PCW formation, potentially controlled by upstream factors such as AgrIAA4. Moreover, despite being constructed around *AgrERF* genes, the coexpression network contained genes encoding proteins associated with auxin signaling, including *AgrSAUR67*, *AgrAUX1-like*, *AgrBIG*, and *AgrVAN3*. This suggests that the network may integrate hormonal signals with cell wall biosynthetic programs during collenchyma development, with ERFs potentially acting as hubs connecting these processes.

In the present study, *AgrIAA4* showed the highest expression among all TFs, despite relatively low correlation with downstream targets, particularly with *AgrCESA* and *AgrCSL* genes. AuxRE motifs were identified only in *AgrAGL62*, *AgrKNAT7a*, *AgrERF002b* upstream sequences, and a subset of target genes, suggesting partial auxin responsiveness in the network. Interestingly, the *AgrIAA4* upstream sequence also contains a DRE-core motif. The fact that *AgrIAA4* appears earlier than *AgrERF079* in the pseudotime trajectory and is unlikely to be its downstream target raises the possibility that an unknown AgrERF may act as an upstream regulator, bridging two (semi-)separated “tails” of the collenchyma PCW-associated TRN governed by AgrERF079 and AgrIAA4.

In contrast to the other analyzed TFs, *AgrIAA4* exhibited one of the lowest overall expression levels in the qPCR analysis and was the only gene that showed no significant difference between leaves and petioles. The low and relatively uniform expression of *AgrIAA4* in leaves and petioles — in contrast to its high expression in the tissue-specific RNA-seq data — may result from the distinct sampling strategies (tissue-specific vs whole-organ), the use of independent plant batches, the different normalization approaches applied in RNA-seq and qPCR analyses, and the possible misassignment of multi-mapping reads among closely related paralogs in the RNA-seq data. Its lack of clear petiole enrichment in the qPCR experiment suggests a broader role outside strict collenchyma differentiation, which is consistent with the known homeostasis-related functions of certain Aux/IAA family members (Zhuang et al. 2025).

AgrKNAT7a: a multifaceted regulator between PCW and SCW programs

AgrKNAT7a (Ag2G01615) is a homolog of AtKNAT7, a well-studied regulator of SCW formation involved in a complex transcriptional network controlling cellulose, xylan, and lignin biosynthesis (Zhong et al. 2008; Li et al. 2012; He et al. 2018; Wang et al. 2020). AtKNAT7 functions as both a positive and negative regulator depending on tissue context, promoting xylan biosynthesis and xylem vessel SCW formation while repressing lignin deposition and SCW thickening in interfascicular fibers (Li et al. 2012; He et al. 2018; Wang et al. 2020). KNAT7 is directly regulated by NAC and MYB master regulators and represents a central hub coordinating SCW transcriptional programs across plant species (Zhong et al. 2006 2007 2008; Ko et al. 2009; Zhou et al. 2009; McCarthy et al. 2009; Nakano et al. 2010). This TF interacts with multiple transcriptional regulators, including AtMYB75, AtBLH6, AtOFP proteins, and AtMYB6, forming complexes involved in feedback control of SCW biosynthesis (Bhargava et al. 2013; Liu et al. 2014; Liu and Douglas 2015; Wang et al. 2020). KNAT3 also contributes to SCW formation in *Arabidopsis*, cooperating with KNAT7 in xylem vessels; it promotes xylan biosynthesis and can interact with NST1/NST2 to regulate SCW formation in fibers (Wang et al. 2020; Qin et al. 2020).

The celery genome contains four KNAT7 homologs and five KNAT3/4 homologs, whose expression patterns are consistent with subfunctionalization or neofunctionalization following gene duplication events. Interestingly, closely related homologs display divergent expression patterns, with no clear correlation between phylogeny and tissue or temporal specificity (Fig. 4).

Among these genes, *AgrKNAT7a* is of particular interest because its expression is associated with PCW-related genes in collenchyma and negatively correlated with SCW-associated *AgrCESAs*. These patterns are consistent with several non-mutually exclusive hypotheses: *AgrKNAT7a* may contribute to restricting ectopic SCW deposition in PCW-dominated tissues, modulate PCW deposition, or participate more directly in PCW-related processes. However, neither coexpression nor negative correlation establishes a direct regulatory relationship or its direction. In particular, the negative correlation with SCW-associated *AgrCESAs* could reflect SCW repression or the contrasting developmental programs associated with PCW- and SCW-dominated cells. A direct role in promoting PCW biosynthesis therefore remains speculative and requires functional validation.

The expression profiles of other *AgrKNAT7* and *AgrKNAT3/4* homologs likewise suggest diverse, tissue- and stage-dependent functions. *AgrKNAT7c,d* and *AgrKNAT3/4a,c,e* show relatively high expression in mature tissues and, in some cases, in PCW-dominated tissues compared with vascular bundles. This pattern may be consistent with roles in cell wall maturation or the maintenance of tissue-specific wall properties, potentially including the restriction of unnecessary SCW deposition. However, these functions cannot be inferred from expression data alone. *AgrKNAT7b* displays expression peaks in both young collenchyma and mature vascular bundles, suggesting a possible context-dependent role. By contrast, the higher expression of *AgrKNAT3/4d* in young tissues was not statistically significant and therefore provides limited evidence for developmental specificity. *AgrKNAT3/4b* showed the lowest expression among the analyzed genes and no significant differences among the examined petiole tissues, preventing a clear functional interpretation.

The diversification of *AgrKNAT7* and *AgrKNAT3/4* may be particularly relevant to celery, whose collenchyma combines substantial PCW deposition with flexibility and continued growth. Different homologs may contribute to distinct regulatory combinations across tissues and developmental stages, potentially through interactions with one another or with additional transcriptional regulators. Determining whether individual AgrKNAT proteins promote PCW-related processes, repress SCW formation, or coordinate both programs will require functional analyses.

AgrAGL62 and AgrHB21: central candidates with strong connectivity

AgrAGL62 (Ag10G01798) belongs to the Type I subgroup of MADS-box TFs recognizing CArG-box motifs (Parenicová et al. 2003). AtAGL62 is required for suppression of cellularization during the syncytial phase of endosperm development, with expression declining immediately before cellularization (Kang et al. 2008). It also participates in auxin biosynthesis activation in the endosperm in both *A. thaliana* and *Fragaria vesca*, indicating a role in auxin-mediated developmental regulation (Guo et al. 2022). In the present study, *AgrAGL62* showed strong correlations with nearly all candidate target genes and one of the highest correlations with TF genes, making it a potential central regulator in the PCW-associated network. Cis-element analysis revealed putative ERF-, ARF-, and THX-associated motifs in the *AgrAGL62* upstream region, whereas only a few target genes, including *AgrbZIP76*, contained strong CArG-box motifs or their variants. This suggests that *AgrAGL62* may be regulated by both ERF- and IAA-dependent pathways. However, the presence of cis-elements alone does not establish direct regulation, and the limited occurrence of CArG-box motifs among candidate targets suggests that some potential effects of AgrAGL62 may be mediated indirectly. MADS-box proteins can form multicomponent regulatory complexes (Shore and Sharrocks 1995), and ePlant data for AtAGL62 (AT5G60440) show multiple interactions with other AtAGLs. This supports the possibility

that AgrAGL62, like its Arabidopsis homolog, acts in a protein complex and may influence downstream genes through additional TFs not included in the current analysis.

AgrHB21 (Ag10G02150) belongs to the ZF-HD TF I family, containing both a zinc-finger dimerization domain and a homeodomain DNA-binding region (Tan and Irish 2006). Although AtHB15 has been reported to negatively regulate SCW formation in pith cells, it belongs to the distinct HD-ZIP III family (Du et al. 2015); therefore, its function does not provide direct evidence for a similar role of AgrHB21. AtHB21 and its homologs participate in developmental and stress-responsive pathways, particularly ABA signaling, which appears to be conserved across dicots and monocots (González-Grandío et al. 2017; Sánchez-Gerschon et al. 2024). The homolog LlZFHD4 in *Lilium lancifolium* is also involved in ABA-mediated stress responses linked to cell wall remodeling (Yong et al. 2021). AtHB21 interacts with AtHB22, suggesting that ZF-HD proteins function in regulatory complexes (Tan and Irish 2006). Importantly, AtHB21 represses genes associated with anther and pollen wall development, including *AtNST2*, a SCW "master-switch" (Zhang et al. 2024). In the current study, *AgrHB21* showed one of the highest expression levels in collenchyma and was the only gene with no expression in mature vascular bundles, suggesting a preferential association with PCW-dominated tissues. It showed high coexpression with all candidate target genes except *AgrCESA6s*. Its network proximity to *AgrERF079* and the presence of DRE-core motifs suggest ERF-dependent upstream regulation.

AgrPTL, AgrZAT4, AgrPRE5, and AgrbZIP76: diverse regulators with distinct network roles

AgrPTL (*Ag3G00597*) encodes a GT2-clade THX TF recognizing GT2-motifs (Smalle et al. 1998). Although THXs have been implicated in PCW formation in grasses, PTL homologs are absent from grasses but are present in other angiosperm lineages, including some monocots (Kaplan-Levy et al. 2014; Zhao et al. 2021). In *A. thaliana*, AtPTL suppresses cell proliferation in inter-sepal zones, maintaining floral organ separation, while both loss- and gain-of-function mutants disrupt flower development through altered boundary growth and auxin homeostasis (Griffith et al. 1999; Brewer et al. 2004; Lampugnani et al. 2012; Li et al. 2008). AtPTL also regulates auxin-dependent developmental processes and interacts with auxin transport components, suggesting an indirect role in cell wall formation via cell expansion and remodeling (Lampugnani et al. 2013; Kaplan-Levy et al. 2014). Studies in *Hedychium coronarium* suggest functional diversification: unlike *AtPTL*, *HcPTL* is expressed in floral meristems and organ primordia, where it may promote growth and differentiation (Zhao et al. 2021). In the present study, *AgrPTL* showed high expression in young collenchyma and uniquely high expression in mature collenchyma, indicating strong collenchyma specificity. Organ-level qPCR further identified *AgrPTL* as the most highly expressed TF in petioles, with expression levels more than fivefold higher than those of the other analyzed genes, while expression in leaves and roots remained low. *AgrPTL* was strongly coexpressed with most candidate target genes, except for genes associated with pectin biosynthesis. Notably, its expression showed the strongest correlation with *AgrCESA6L-B* and the weakest correlation with *AgrCESA6L-C*. Most analyzed TF and target gene upstream regions contained GT2-related motifs, whereas the upstream region of *AgrCESA6L-C* did not. This difference raises the possibility of distinct regulatory relationships, although motif occurrence and coexpression alone cannot establish direct or selective regulation by AgrPTL. The presence of DRE-core motifs, together with the absence of detected AuxRE motifs in the *AgrPTL* upstream region, suggests potential input from DRE-associated regulatory pathways but does not establish ERF-dependent regulation or exclude ARF-related control through other mechanisms.

AgrZAT4 (Ag5G00091) belongs to the C2H2 zinc-finger TF family. In *A. thaliana*, overexpression of *AtZAT4* (together with *AtZAT1* and *AtZAT9*) strongly inhibits root and organ growth, disrupts root meristem organization, and induces programmed cell death marker genes (Song et al. 2020). AtZAT4 also regulates genes involved in pollen, embryo, and seed development (Puentes-Romero et al. 2022).

In *Brassica napus*, *BnaA04G.ZAT4* expression is negatively correlated with seed oil content and positively correlated with seed coat content (Yu et al. 2023). In celery, *AgrZAT4* showed higher expression in young tissues than in mature tissues and high coexpression with TF genes and PCW biosynthetic genes, particularly *AgrCSLCs* (xyloglucan biosynthesis) and pectin-related genes. Cis-elements for C2H2 TF binding are present in all upstream sequences of analyzed genes, making its regulatory specificity unclear. The *AgrZAT4* upstream region lacked detected ERF- and ARF-associated motifs but contained a THX-related motif. These findings do not identify a definitive upstream regulator, and the relationship between AgrZAT4 and PCW-associated genes may be indirect. Thus, although the expression and coexpression patterns nominate AgrZAT4 as a candidate involved in developmental regulation of cell wall formation, its specific contribution to collenchyma PCW deposition remains unclear.

AgrPRE5 (Ag5G00597) differs from typical bHLH TFs in lacking a functional DNA-binding basic domain and regulating gene expression indirectly through protein-protein interactions with other HLH/bHLH proteins (Xiaoxing et al. 2006). AtPRE5 promotes cell expansion by antagonizing IBH1, thereby releasing HBI1, which activates *AtEXPA* genes involved in cell wall loosening (Oh et al. 2012). *AtPRE5* is a direct target of the BZR1-PIF4 module, which links PRE-mediated growth regulation to brassinosteroid and light signaling. Simultaneous suppression of several PRE-family genes leads to dwarfism and reduced growth responses (Bai et al. 2012). *AgrPRE5* showed high expression in young collenchyma and relatively high expression in young parenchyma, with low levels in vascular bundles. It showed high coexpression with all candidate target genes and TFs, placing it at the center of the putative TRN alongside AgrAGL62. Due to its non-DNA-binding nature, it likely acts as a partner protein, and the widespread occurrence of bHLH-associated cis-elements in the analyzed sequences suggests the involvement of additional bHLH factors still hidden under molecular shadows.

AgrbZIP76 (AgUnG00415) has as its closest *A. thaliana* homolog AT1G58110, which belongs to the poorly studied group E of bZIP TFs (Dröge-Laser et al. 2018). Evidence from phylogenetically related proteins suggests possible links between group E bZIPs and stress-responsive pathways that may indirectly affect cell wall-related processes. For example, the sweet potato (*Ipomoea batatas*) E-group factor IbbZIP1, whose closest *A. thaliana* homolog is also AT1G58110, is involved in ABA-mediated stress responses and may indirectly affect cell wall remodeling via ROS homeostasis (Kang et al. 2019). E-group bZIPs such as AtbZIP34 and AtbZIP61 can heterodimerize with members of bZIP groups I and S, which include proteins associated with vascular development and cell wall organization (Shen et al. 2007). Group I factors such as bZIP29 and SWIZ regulate genes involved in cell wall organization in *A. thaliana* and *B. distachyon*, respectively (Van Leene et al. 2016; Coomey et al. 2024). Therefore, AgrbZIP76 may participate in the regulation of cell wall formation despite the limited direct evidence currently available. Its gene expression is relatively stable across tissues but significantly higher in young collenchyma, suggesting a role in PCW formation, possibly with activity peaks during wall thickening. Its upstream region contains a DRE-core motif, indicating ERF-dependent regulation. Coexpression with candidate target genes and the widespread presence of relevant cis-elements further support its potential involvement in PCW regulation in general, rather than collenchyma-specific regulation.

Network organization and functional specialization

Altogether, these TFs represent potential candidates for regulators of PCW formation in celery collenchyma based on network connectivity, association with PCW-related gene modules, and supporting evidence from literature and cis-element analyses. However, within this set, some TFs appear less likely to play central roles in this process. In particular, AgrbZIP76, AgrZAT4, and AgrERF001b, the latter of which was additionally included in the present study but was not part of the

collenchyma-specific AgrERF set defined in Syrchina et al. (2025), show weaker integration into both AgrIAA4- and AgrERF079-centered network modules, lack clear supporting cis-regulatory elements, and are not strongly supported by functional evidence from homologous genes. This suggests that their involvement in collenchyma PCW formation, if any, is likely indirect or context-dependent.

In addition, all analyzed TF genes and their potential targets contain cis-elements for NAC, MYB, GATA, WRKY, and NF-Y TFs within their 2,000 bp upstream regions. The NAC-MYB system is a well-established regulatory module controlling SCW formation (Nakano et al. 2015). GATA factors are involved in light signaling, chloroplast development, nitrogen metabolism, and developmental regulation (Schwechheimer et al. 2022), whereas WRKY proteins are key regulators of stress responses, immunity, senescence, and hormone signaling (Bakshi and Oelmüller 2014). Some WRKY TFs have also been implicated in cell wall regulation. For example, AtWRKY12 represses the SCW regulators *AtNST2* and *AtC3H14*, thereby inhibiting SCW deposition in pith cells (Wang et al. 2010), while in birch, BpWRKY29/40-binding motifs were identified in the promoter of the PCW-associated xylan biosynthesis gene *BpGUX3* (Ilievska et al. 2025). NF-Y TFs participate in embryogenesis, flowering, root development, stress responses, and carbon/nitrogen metabolism through heterotrimeric complexes (Siriwardana 2024). Because these TF families are large and multifunctional, the presence of their cis-elements in PCW-related genes is not unexpected. Nevertheless, some members of these families may participate in PCW regulation in celery collenchyma and warrant further investigation.

In addition to TF genes, the AgrERF-based coexpression network contained genes encoding proteins potentially involved in cell wall formation and modification. These candidates represented diverse aspects of cell wall metabolism, including cellulose and hemicellulose biosynthesis and pectin modification, cell expansion, cell wall structural organization, and microtubule-mediated regulation of wall deposition. Among them were a cellulose synthase companion protein (AgrCC1), the AgrCSLE1, AgrPME17 and AgrPPME1, AgrEXPA11, a fasciclin-like arabinogalactan protein (AgrFLA17), and a microtubule-associated protein (AgrMAP70). CC1 is a microtubule-associated protein that regulates cortical microtubule organization and cellulose biosynthesis, particularly under salt stress (Kesten et al. 2019; Gurumoorthy et al. 2025). CSLE proteins are thought to participate in hemicellulose biosynthesis and remodeling during PCW formation. Although their substrate specificity remains uncharacterized, expression patterns in celery and other plant species suggest that they may instead be involved in protein or metabolite glycosylation rather than in the biosynthesis of major cell wall polysaccharides (Mokshina 2025b). AtPME17 contributes to pectin demethylesterification and cell wall integrity (Del Corpo et al. 2017), whereas AtPPME1 regulates pollen tube elongation through pectin modification (Tian et al. 2006). AtEXPA11 promotes localized cell expansion during hyponastic growth and is regulated by ethylene (Polko et al. 2011). FLA17 is associated with SCW formation and cellulose deposition in fibers (Wan et al. 2014; MacMillan et al. 2017), while FLA16 and FLA18 have been implicated in cell wall mechanics and anisotropic expansion (Liu et al. 2020; Ashagre et al. 2021; Cao et al. 2025). AtMAP70-1 regulates microtubule organization and the spatial patterning of SCW deposition (Korolev et al. 2005; Pesquet et al. 2010), thereby contributing to proper wall architecture (Korolev et al. 2007; Derbyshire et al. 2015).

Beyond these candidates, the network also contained genes associated with other aspects of cell wall metabolism, including *AgrBGAL2*, *AgrO-GT*, and *AgrLEW3*. BGAL2 participates in pectin remodeling in *A. thaliana* and *Malus×domestica* (Gantulga et al. 2008; Collins et al. 2019), while O-glycosyltransferases regulate glycoprotein modification within the cell wall matrix (Nguema-Ona et al. 2014; Zhang et al. 2025). LEW3 is an ER-localized α-1,2-mannosyltransferase required for N-glycosylation, and mutations in LEW3 affect cellulose synthesis and cell wall integrity (Zhang et al. 2009).

Another group of candidates potentially relevant to cell wall remodeling comprised several *AgrGDSL* genes encoding GDSL esterases/lipases. GDSL proteins constitute a large family with diverse enzymatic activities, including acyl group transfer in hydrolase reactions involving both lipid and non-lipid substrates (Akoh et al. 2004). Although they are best known for their roles in cutin deposition and lipid metabolism (Girard et al. 2012), emerging evidence implicates some members in the deacetylation of cell wall polysaccharides. In rice, DARX1 (Deacetylase on the ARabinosyl sidechain of Xylan 1) was shown to control SCW patterning by modulating xylan acetylation (Zhang et al. 2017). Several *AgrGDSL* genes were coexpressed with *AgrCSLs*, *AgrRRTs*, and *AgrGAUTs*, suggesting their potential involvement in the deacetylation of hemicelluloses and pectins during PCW formation (Mokshina et al. 2025). Their presence in the current network, together with their predicted extracellular localization, further supports the hypothesis that some GDSL proteins may function as polysaccharide deacetylases in the apoplast and contribute to the dynamic remodeling of cell wall polymers during collenchyma development. However, substrate specificity remains uncharacterized for most GDSL proteins, and further biochemical studies are required to confirm their enzymatic activities and roles in cell wall metabolism.

The distribution of cell wall-related genes across the two coexpression modules is consistent with their functional specialization. Group I, which includes TF genes such as *AgrHB21*, *AgrZAT4*, *AgrPRE5*, *AgrIAA4*, and *AgrbZIP76*, is enriched in genes encoding proteins associated with cell wall integrity, extensibility, and active growth processes. In contrast, group II, containing fewer TF genes (*AgrAGL62*, *AgrKNAT7*, and *AgrPTL*), is enriched in genes involved in pectin remodeling and glycoprotein modification, aligning with PCW biosynthesis and modification. Intergroup connections and intermediate nodes, such as *AgrPRE5*, *AgrAGL62*, and *AgrKNAT7a*, indicate that these groups are not independent pathways but rather interconnected regulatory modules that may contribute to different aspects of PCW development.

A key observation is the consistent negative correlation between all analyzed TFs and SCW-associated *AgrCESA* genes, particularly *AgrKNAT7a*. Together with the generally higher expression of most TFs in young collenchyma, this pattern supports an association of the network with PCW formation rather than with cell wall biosynthesis in general. Collenchyma occupies a unique position among supporting tissues because its walls remain extensible despite substantial thickening. Accordingly, its regulatory system may require coordination of biosynthetic activity while limiting activation of SCW-related programs. Strong correlations between TF genes and *AgrCSL*, *AgrRRT*, and *AgrGAUT* genes further support an association with the biosynthesis and modification of non-cellulosic PCW polysaccharides.

Pseudotime and regulatory dynamics

In addition to the structural organization of the network, pseudotime and diffusion analyses suggest temporal organization within this transcriptional system. As noted above, *AgrIAA4* and *AgrERF079* occupy early pseudotime positions, whereas *AgrERF041b* and *AgrPTL* are positioned later. Although pseudotime inferred from bulk transcriptomic data should be interpreted with caution, this pattern is consistent with sequential transcriptional events during collenchyma differentiation. Early-expressed genes may participate in the initiation of transcriptional reprogramming, whereas later-expressed genes may be associated with subsequent stages of cell wall differentiation. Aux/IAA proteins generally function as upstream repressors of ARFs, and DNA-binding activity has been demonstrated for ERF041 homologs (Nakata et al. 2021; Gao et al. 2023); therefore, the observed pseudotime arrangement is biologically plausible.

Importantly, the force-directed network indicates that genes occupying distant pseudotime positions remain highly connected, suggesting substantial integration across the inferred developmental

trajectory. However, the greater the distance between genes on a pseudotime image, the more likely it is that additional TFs act between them. Genes encoding such TFs may not have been identified within the *AgrERF*-based collenchyma-specific coexpression network, possibly because the corresponding TFs participate in multiple biological processes, including PCW formation in other tissues. This possibility is supported by the presence of AgrIAA4 in the network, as Aux/IAA proteins are generally expected to regulate downstream ARFs; however, no *AgrARF* genes were identified in the present study. This observation suggests that additional components of the regulatory cascade may remain undetected by the current analysis.

The primary objective of this study was to identify potential key regulators involved in triggering PCW thickening in collenchyma. Therefore, the absence of intermediate regulatory factors from the inferred network does not preclude their participation in the underlying transcriptional program.

Toward a two-tailed regulatory model

Within the proposed “two-tailed” regulatory framework, AgrIAA4- and AgrERF079-centered modules may represent functionally divergent and potentially antagonistic regulatory branches that control PCW development in collenchyma. Specifically, AgrIAA4 may act as a negative regulator of PCW formation by repressing downstream ARF-mediated transcriptional programs, whereas AgrERF079 may function as a positive regulator promoting PCW-associated gene expression. Under this assumption, the balance between Aux/IAA–ARF-mediated repression and ERF-dependent activation could contribute to fine-tuning of cell wall remodeling during collenchyma differentiation.

The hypothesis that *AgrIAA4* acts as a potential negative regulator of PCW thickening is based on its developmental expression pattern. In the tissue-specific RNA-seq dataset, *AgrIAA4* showed a distinctive developmental expression pattern, with significantly higher expression in young collenchyma than in other young petiole tissues and significantly lower expression in mature collenchyma than in other mature tissues (Fig. 2). However, this pattern was not reflected in the organ-level qPCR analysis, in which *AgrIAA4* showed no significant difference in expression between petioles and leaves. This discrepancy cautions against interpreting *AgrIAA4* as a petiole-specific regulator but does not exclude a developmentally restricted role in collenchyma. Within the RNA-seq dataset, the contrast between developmental stages provides a more specific basis for considering a potential role for *AgrIAA4* in collenchyma development. Its expression was approximately fivefold higher in young than in mature collenchyma. This decline may reflect a stage-dependent role in developmental processes that accompany PCW thickening, although it does not establish whether *AgrIAA4* directly represses this process. Such a role could be relevant in collenchyma, where PCW thickening must be coordinated with continued tissue growth.

The increased expression of *AgrIAA4* in mature parenchyma relative to young parenchyma may further support a repressive role. Because parenchyma typically develops thin PCWs and does not undergo SCW formation, elevated AgrIAA4 expression could contribute to the maintenance of limited PCW thickening in the absence of alternative inhibitory mechanisms. A similar pattern is observed in vascular bundles, where *AgrIAA4* expression is higher in young tissues than in mature ones. During early vascular development, when SCW formation has not yet been fully established, repression of excessive PCW thickening may be required. In mature vascular tissues, the presence of SCWs may itself limit further PCW thickening, thereby reducing the need for AgrIAA4-mediated repression. Together with the relatively weak correlation between *AgrIAA4* and genes encoding PCW biosynthetic enzymes, these observations are consistent with the possibility that *AgrIAA4* acts indirectly as a negative regulator within the proposed two-tailed regulatory network. In this model, the *AgrIAA4*-associated branch may

counterbalance the potentially activating role of *AgrERF079* during PCW thickening in celery collenchyma (Syrchina et al. 2025).

Cis-element analysis further supports extensive cross-regulation among TFs. Upstream regions of TFs and candidate target genes contain motifs associated with ERF, ARF, THX, ZF-HD, bHLH, C2H2, and bZIP binding. This architecture is consistent with a densely interconnected regulatory network containing potential feedback and feedforward interactions. The relatively limited distribution of AuxRE motifs compared with the widespread occurrence of ERF-related motifs suggests that auxin-dependent regulation may be more selective, whereas ERF-mediated regulation may affect a broader set of genes. However, because the present analysis was initiated from an *AgrERF*-centered coexpression network, a more detailed investigation of *AgrIAA4*-associated genes may reveal additional auxin-responsive regulatory components.

The functional enrichment analysis of the *AgrERF*-based coexpression network revealed significant enrichment of auxin transport and regulation pathways, providing independent support for the proposed “two-tailed” model involving the AgrIAA4/ARF axis. This is particularly noteworthy given that auxin signaling is a central regulator of plant growth, cell expansion, and cell wall dynamics. The presence of auxin-related genes within the network, together with the identification of AuxRE motifs in the promoters of several TF and target genes, further supports the potential functional relevance of the AgrIAA4-centered branch of our proposed TRN.

The enrichment of lipid metabolic and catabolic processes is also relevant. Lipids are integral to membrane biogenesis and vesicle trafficking, which are essential for the secretion of cell wall polysaccharides and proteins (Gille and Pauly 2012). This finding is consistent with the overrepresentation of genes encoding proteins localized to the endomembrane system and endoplasmic reticulum, compartments where cell wall polymers are synthesized and modified. The enrichment of cutin-related processes, while noteworthy, may reflect broader coordination of secretory pathway outputs, as cutin biosynthesis shares components with general lipid metabolism and the secretion machinery rather than indicating a specific role in PCW formation.

Collectively, these functional enrichments — particularly the signatures of auxin transport and lipid metabolism — provide additional support for a role of the *AgrERF*-centered network in PCW formation. They also suggest avenues for future research into the interplay between auxin signaling, lipid metabolism, and cell wall deposition in collenchyma.

Altogether, coexpression, diffusion pseudotime, and cis-element analyses support a hypothetical two-tailed (ERF- and ARF-associated) TRN model (Fig. 7). Although the proposed network is based almost entirely on in silico analyses and requires experimental validation, it provides a useful framework for future studies and highlights candidate regulatory components beyond the previously characterized ERF-centered pathways.

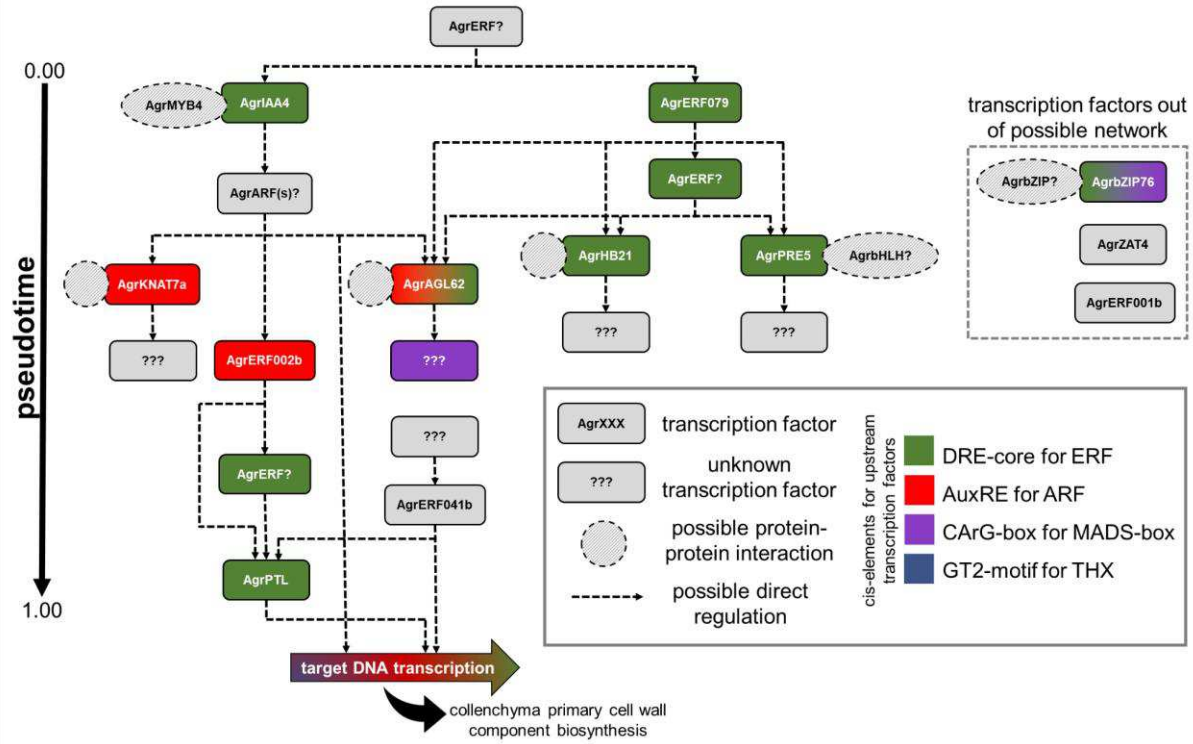


Fig. 7. Hypothetical two-tailed transcriptional regulatory network regulating primary cell wall formation in celery collenchyma. The model integrates the pseudotime-based expression patterns of candidate genes, predicted cis-regulatory elements in their promoter regions, and potential protein-protein interactions.

Conclusion

Our analysis indicates that PCW thickening in celery collenchyma is associated with a regulatory system extending beyond the previously implicated ERF family. Using a collenchyma-associated coexpression network, we identified candidate transcriptional regulators belonging to the Aux/IAA, KNOX, MADS-box, ZF-HD, THX, C2H2, atypical bHLH, and bZIP families. The network was associated with genes involved in cellulose, hemicellulose, and pectin metabolism and showed predominantly negative correlations with SCW-related *CESAs*. Diffusion pseudotime, force-directed network analysis, and cis-element predictions revealed extensive connectivity among the identified regulators and suggested a highly integrated transcriptional system. In particular, the data support a hypothetical two-tailed regulatory framework involving ERF- and Aux/IAA-associated branches that may coordinate the balance between cell wall deposition and maintenance of wall extensibility during collenchyma development. The identification of auxin-related genes within the network, together with AuxRE motifs in TF and target gene promoters, provides independent support for the AgrIAA4-centered branch of this proposed model. Although these conclusions are based on in silico analyses and require experimental validation, the identified candidates provide a valuable resource for future functional studies and establish celery collenchyma as a promising model for investigating the transcriptional regulation of thick primary cell wall formation.

Supplementary information

Table S1 — qPCR primers information
Table S2 — potential target genes information
Table S3 — all *AgrERF*-based network genes information
Table S4 — KNOX family homolog searching

Fig. S1 — summary of Gene Ontology (GO) enrichment analysis of *A. thaliana* orthologs from the *AgrERF*-based coexpression network.

Fig. S2 — transcription factor gene coexpression with potential target genes

References

1. Akoh CC, Lee GC, Liaw YC, Huang TH, Shaw JF (2004) GDSL family of serine esterases/lipases. Progress in Lipid Research 43:534-552. https://doi.org/10.1016/j.plipres.2004.09.002

2. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2002) Molecular Biology of the Cell. 4th ed. Garland Science, New York

3. Allelign Ashagre H, Zaltzman D, Idan-Molakandov A, Romano H, Tzfadia O, Harpaz-Saad S (2021) FASCICLIN-LIKE 18 Is a New Player Regulating Root Elongation in *Arabidopsis thaliana*. Front Plant Sci 12:645286. https://doi.org/10.3389/fpls.2021.645286

4. Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O (2017) DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33:3387-3395. https://doi.org/10.1093/bioinformatics/btx431

5. Amos RA, Atmodjo MA, Huang C, Gao Z, Venkat A, Taujale R, Kannan N, Moremen KW, Mohnen D (2023) Author correction: polymerization of the backbone of the pectic polysaccharide rhamnogalacturonan I. Nat Plants 9:673. https://doi.org/10.1038/s41477-023-01395-z

6. Amos RA, Mohnen D (2019) Critical Review of Plant Cell Wall Matrix Polysaccharide Glycosyltransferase Activities Verified by Heterologous Protein Expression. Front Plant Sci 10:915. https://doi.org/10.3389/fpls.2019.00915

7. Bai MY, Fan M, Oh E, Wang ZY (2012) A Triple Helix-Loop-Helix/Basic Helix-Loop-Helix Cascade Controls Cell Elongation Downstream of Multiple Hormonal and Environmental Signaling Pathways in *Arabidopsis*. Plant Cell 24:4917-4929. https://doi.org/10.1105/tpc.112.105163

8. Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 28:45-48. https://doi.org/10.1093/nar/28.1.45

9. Bakshi M, Oelmüller R (2014) WRKY transcription factors: Jack of many trades in plants. Plant Signal Behav 9:e27700. https://doi.org/10.4161/psb.27700

10. Baumann MJ, Eklof JM, Michel G, Kallas AM, Teeri TT, Czjzek M, Brumer H (2007) Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo-transglycosylases: biological implications for cell wall metabolism. Plant Cell 19:1947-1963. https://doi.org/10.1105/tpc.107.051391

11. Bhargava A, Ahad A, Wang S, Mansfield SD, Haughn GW, Douglas CJ, Ellis BE (2013) The interacting MYB75 and KNAT7 transcription factors modulate secondary cell wall deposition both in stems and seed coat in *Arabidopsis*. Planta 237:1199-1211. https://doi.org/10.1007/s00425-012-1821-9

12. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech 2008:P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

13. Brewer PB, Howles PA, Dorian K, Griffith ME, Ishida T, Kaplan-Levy RN, Kilinc A, Smyth DR (2004) PETAL LOSS, a trihelix transcription factor gene, regulates perianth architecture in the *Arabidopsis* flower. Development 131:4035-4045. https://doi.org/10.1242/dev.01279

14. Cao Y, Chen X, Wei B, Wang B, Zhang Y, Xu L, Zhao Y, Song C (2025) Genome-wide identification and analysis of the FLA gene family in maize (*Zea mays* L.) and its expression in response to abiotic stresses. BMC Plant Biol 25:982. https://doi.org/10.1186/s12870-025-06981-1

15. Chen C, Wu Y, Li J, Wang X, Zeng Z, Xu J, Liu Y, Feng J, Chen H, He Y, Xia R (2023) TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. Mol Plant 16:1733-1742. https://doi.org/10.1016/j.molp.2023.09.010

16. Chen D, Harris PJ, Sims IM, Zujovic Z, Melton LD (2017) Polysaccharide compositions of collenchyma cell walls from celery (*Apium graveolens* L.) petioles. BMC Plant Biol 17:104. https://doi.org/10.1186/s12870-017-1046-y

17. Chen D, Melton LD, Zujovic Z, Harris PJ (2019) Developmental changes in collenchyma cell-wall polysaccharides in celery (*Apium graveolens* L.) petioles. BMC Plant Biol 19:81. https://doi.org/10.1186/s12870-019-1648-7

18. Chow CN, Yang CW, Wu NY, Wang HT, Tseng KC, Chiu YH, Lee TY, Chang WC (2024) PlantPAN 4.0: updated database for identifying conserved non-coding sequences and exploring dynamic transcriptional regulation in plant promoters. Nucleic Acids Res 52:D1569-D1578. https://doi.org/10.1093/nar/gkad945

19. Coculo D, Lionetti V (2022) The plant invertase/pectin methylesterase inhibitor superfamily. Front Plant Sci 13:863892. https://doi.org/10.3389/fpls.2022.863892

20. Coifman RR, Lafon S (2006) Diffusion maps. Appl Comput Harmon Anal 21:5-30. https://doi.org/10.1016/j.acha.2006.04.006

21. Collins PP, O'Donoghue EM, Rebstock R, Tiffin HR, Sutherland PW, Schröder R, McAtee PA, Prakash R, Ireland HS, Johnston JW, Atkinson RG, Schaffer RJ, Hallett IC, Brummell DA (2019) Cell type-specific gene expression underpins remodelling of cell wall pectin in exocarp and cortex during apple fruit development. J Exp Bot 70:6085-6099. https://doi.org/10.1093/jxb/erz370

22. Coomey JH, MacKinnon KJM, McCahill IW, Hazen SP (2024) Mechanically induced localisation of SECONDARY WALL INTERACTING bZIP is associated with thigmomorphogenic and secondary cell wall gene expression. Quant Plant Biol 5:e5. https://doi.org/10.1017/qpb.2024.5

23. Cosgrove DJ, Bedinger P, Durachko DM (1997) Group I allergens of grass pollen as cell wall-loosening agents. Proc Natl Acad Sci USA 94:6559-6564. https://doi.org/10.1073/pnas.94.12.6559

24. Crow E, Murphy RJ (2000) Microfibril orientation in differentiating and maturing fibre and parenchyma cell walls in culms of bamboo (Phyllostachys viridi-glaucescens (Carr.) Riv. & Riv.). Bot J Linn Soc 134:339-359. https://doi.org/10.1111/j.1095-8339.2000.tb02357.x

25. Csardi G, Nepusz T (2006) The igraph software package for complex network research. InterJournal Complex Systems 1695

26. Davis J, Brandizzi F, Liepman AH, Keegstra K (2010) *Arabidopsis* mannan synthase CSLA9 and glucan synthase CSLC4 have opposite orientations in the Golgi membrane. Plant J 64:1028-1037. https://doi.org/10.1111/j.1365-313X.2010.04392.x

27. Del Corpo D, Fullone MR, Miele R, Lafond M, Pontiggia D, Grisel S, Kieffer-Jaquinod S, Giardina T, Bellincampi D, Lionetti V (2020) AtPME17 is a functional *Arabidopsis thaliana* pectin methylesterase regulated by its PRO region that triggers PME activity in the resistance to Botrytis cinerea. Mol Plant Pathol 21:1620-1633. https://doi.org/10.1111/mpp.13002

28. Derbyshire P, Ménard D, Green P, Saalbach G, Buschmann H, Lloyd CW, Pesquet E (2015) Proteomic Analysis of Microtubule Interacting Proteins over the Course of Xylem Tracheary Element Formation in *Arabidopsis*. Plant Cell 27:2709-2726. https://doi.org/10.1105/tpc.15.00314

29. Dröge-Laser W, Snoek BL, Snel B, Weiste C (2018) The *Arabidopsis* bZIP transcription factor family — an update. Curr Opin Plant Biol 45:36-49. https://doi.org/10.1016/j.pbi.2018.05.001

30. Du Q, Avci U, Li S, Gallego-Giraldo L, Pattathil S, Qi L, Hahn MG, Wang H (2015) Activation of miR165b represses AtHB15 expression and induces pith secondary wall development in *Arabidopsis*. Plant J 83:388-400. https://doi.org/10.1111/tpj.12897

31. Eddy SR (2011) Accelerated Profile HMM Searches. PLoS Comput Biol 7:e1002195. https://doi.org/10.1371/journal.pcbi.1002195

32. Fan M, Herburger K, Jensen JK, Zemelis-Durfee S, Brandizzi F, Fry SC, Wilkerson CG (2018) A trihelix family transcription factor is associated with key genes in mixed-linkage glucan accumulation. Plant Physiol 178:1207-1221. https://doi.org/10.1104/pp.18.00978

33. Fruchterman TMJ, Reingold EM (1991) Graph drawing by force-directed placement. Softw Pract Exper 21:1129-1164. https://doi.org/10.1002/spe.4380211102

34. Gantulga D, Turan Y, Bevan DR, Esen A (2008) The *Arabidopsis* At1g45130 and At3g52840 genes encode beta-galactosidases with activity toward cell wall polysaccharides. Phytochemistry 69:1661-1670. https://doi.org/10.1016/j.phytochem.2008.01.023

35. Gao Y, Cui Y, Li S, Feng M, Liu L, Xu W, Zhang J (2023) GhERF41, a group IIIe ERF protein, participates in regulating cell wall synthesis in cotton fiber. Ind Crops Prod 192:116001. https://doi.org/10.1016/j.indcrop.2022.116001

36. Ge SX, Jung D, Yao R (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics 36:2628-2629. https://doi.org/10.1093/bioinformatics/btz931

37. Gille S, Pauly M (2012) O-Acetylation of Plant Cell Wall Polysaccharides. Front Plant Sci 3:12. https://doi.org/10.3389/fpls.2012.00012

38. Girard AL, Mounet F, Lemaire-Chamley M, Gaillard C, Elmorjani K, Vivancos J, Runavot JL, Quemener B, Petit J, Germain V, Rothan C, Marion D, Bakan B (2012) Tomato GDSL1 is required for cutin deposition in the fruit cuticle. Plant Cell 24:3119-3134. https://doi.org/10.1105/tpc.112.101055

39. González-Grandío E, Pajoro A, Franco-Zorrilla JM, Tarancón C, Immink RG, Cubas P (2017) Abscisic acid signaling is controlled by a BRANCHED1/HD-ZIP I cascade in *Arabidopsis* axillary buds. Proc Natl Acad Sci USA 114:E245-E254. https://doi.org/10.1073/pnas.1613199114

40. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40:D1178-D1186. https://doi.org/10.1093/nar/gkr944

41. Gorshkova T, Chernova T, Mokshina N, Ageeva M, Mikshina P (2018) Plant 'muscles': fibers with a tertiary cell wall. New Phytol 218:66-72. https://doi.org/10.1111/nph.14997

42. Griffith ME, da Silva Conceição A, Smyth DR (1999) PETAL LOSS gene regulates initiation and orientation of second whorl organs in the *Arabidopsis* flower. Development 126:5635-5644. https://doi.org/10.1242/dev.126.24.5635

43. Guo L, Luo X, Li M, Zhang Y, Zhao J, Deng Z, Cao Y, Zhang X, Zhao H, Li Z (2022) Mechanism of fertilization-induced auxin synthesis in the endosperm for seed and fruit development. Nat Commun 13:3985. https://doi.org/10.1038/s41467-022-31656-y

44. Gurumoorthy V, Hicks A, Tiruvadi-Krishnan S, Leite WC, Chodankar S, Kolape JL, Smith JC, Lamichhane R, O'Neill H (2025) The N-terminal domain of COMPANION OF CELLULOSE SYNTHASE1 promotes microtubule array formation in *Arabidopsis*. Plant Physiol 199:kiaf392. https://doi.org/10.1093/plphys/kiaf392

45. Harris PJ (2006) Primary and secondary plant cell walls: A comparative overview. N Z J For Sci 36:36-53

46. He JB, Zhao XH, Du PZ, Zeng W, Beahan CT, Wang YQ, Li HL, Bacic A, Wu AM (2018) KNAT7 positively regulates xylan biosynthesis by directly activating IRX9 expression in *Arabidopsis*. J Integr Plant Biol 60:514-528. https://doi.org/10.1111/jipb.12638

47. Huala E, Dickerman AW, Garcia-Hernandez M, Weems D, Reiser L, LaFond F, Hanley D, Kiphart D, Zhuang M, Huang W, Mueller LA, Bhattacharyya D, Bhaya D, Sobral BW, Beavis W, Meinke DW, Town CD, Somerville C, Rhee SY (2001) The *Arabidopsis* Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res 29:102-105. https://doi.org/10.1093/nar/29.1.102

48. Ilievska M, Chong SL, Lim KJ, Immanen J, Nieminen K, Maaheimo H, Helariutta Y, Wurman-Rodrich J, Dupree P, Ord J, Tenkanen M, Salojärvi J (2025) Gene regulatory network analysis of silver birch reveals the ancestral state of secondary cell wall biosynthesis in core eudicots. New Phytol 246:2059-2074. https://doi.org/10.1111/nph.70126

49. Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. EMBO J 19:6150-6161. https://doi.org/10.1093/emboj/19.22.6150

50. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587-589. https://doi.org/10.1038/nmeth.4285

51. Kang C, Zhai H, He S, Zhao N, Liu Q (2019) A novel sweetpotato bZIP transcription factor gene, IbbZIP1, is involved in salt and drought tolerance in transgenic *Arabidopsis*. Plant Cell Rep 38:1373-1382. https://doi.org/10.1007/s00299-019-02441-x

52. Kang IH, Steffen JG, Portereiko MF, Lloyd A, Drews GN (2008) The AGL62 MADS domain protein regulates cellularization during endosperm development in *Arabidopsis*. Plant Cell 20:635-647. https://doi.org/10.1105/tpc.107.055137

53. Kaplan-Levy RN, Quon T, O'Brien M, Sappl PG, Smyth DR (2014) Functional domains of the PETAL LOSS protein, a trihelix transcription factor that represses regional growth in *Arabidopsis thaliana*. Plant J 79:477-491. https://doi.org/10.1111/tpj.12574

54. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-780. https://doi.org/10.1093/molbev/mst010

55. Kesten C, Wallmann A, Schneider R, McFadden A, Kirik A, Spink K, Taylor I, Zhang Y, Zhang Y, Persson S (2019) The companion of cellulose synthase 1 confers salt tolerance through a Tau-like mechanism in plants. Nat Commun 10:857. https://doi.org/10.1038/s41467-019-08780-3

56. Ko JH, Kim WC, Han KH (2009) Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in *Arabidopsis*. Plant J 60:649-665. https://doi.org/10.1111/j.1365-313X.2009.03989.x

57. Korolev AV, Buschmann H, Doonan JH, Lloyd CW (2007) AtMAP70-5, a divergent member of the MAP70 family of microtubule-associated proteins, is required for anisotropic cell growth in *Arabidopsis*. J Cell Sci 120:2241-2247. https://doi.org/10.1242/jcs.007393

58. Korolev AV, Chan J, Naldrett MJ, Doonan JH, Lloyd CW (2005) Identification of a novel family of 70 kDa microtubule-associated proteins in *Arabidopsis* cells. Plant J 42:547-555. https://doi.org/10.1111/j.1365-313X.2005.02393.x

59. Lampugnani ER, Kilinc A, Smyth DR (2012) PETAL LOSS is a boundary gene that inhibits growth between developing sepals in *Arabidopsis thaliana*. Plant J 71:724-735. https://doi.org/10.1111/j.1365-313X.2012.05023.x

60. Lampugnani ER, Kilinc A, Smyth DR (2013) Auxin controls petal initiation in *Arabidopsis*. Development 140:185-194. https://doi.org/10.1242/dev.084582

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

62. Leroux O (2012) Collenchyma: a versatile mechanical tissue with dynamic cell walls. Ann Bot 110:1083-1098. https://doi.org/10.1093/aob/mcs186

63. Letunic I, Bork P (2024) Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res 52:W78-W82. https://doi.org/10.1093/nar/gkae268

64. Li E, Bhargava A, Qiang W, Friedmann MC, Forneris N, Savidge RA, Johnson LA, Mansfield SD, Ellis BE, Douglas CJ (2012) The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in Populus. New Phytol 194:102-115. https://doi.org/10.1111/j.1469-8137.2011.04016.x

65. Li M, Nie C, He S, Xue Z, Li J, Li Z, He C, Zheng X, Tan B, Cheng J, Wang W, Li J, Ye X, Feng J (2025) VvARF19 represses VvLBD13-mediated cell wall degradation to delay softening of grape berries. Hortic Res 12:uhae322. https://doi.org/10.1093/hr/uhae322

66. Li X, Duan X, Jiang H, Sun Y, Tang Y, Yuan Z, Guo J, Liang W, Chen L, Yin J, Ma H, Wang J, Zhang D (2006) Genome-Wide Analysis of Basic/Helix-Loop-Helix Transcription Factor Family in Rice and *Arabidopsis*. Plant Physiol 141:1167-1184. https://doi.org/10.1104/pp.106.080580

67. Li X, Qin G, Chen Z, Gu H, Qu LJ (2008) A gain-of-function mutation of transcriptional factor PTL results in curly leaves, dwarfism and male sterility by affecting auxin homeostasis. Plant Mol Biol 66:315-327. https://doi.org/10.1007/s11103-007-9272-6

68. Liepman AH, Cavalier DM (2012) The CELLULOSE SYNTHASE-LIKE A and CELLULOSE SYNTHASE-LIKE C families: recent advances and future perspectives. Front Plant Sci 3:109. https://doi.org/10.3389/fpls.2012.00109

69. Liu E, MacMillan CP, Shafee T, Ma Y, Ratcliffe J, van de Meene A, Bacic A, Humphries J, Johnson KL (2020) Fasciclin-Like Arabinogalactan-Protein 16 (FLA16) Is Required for Stem Development in *Arabidopsis*. Front Plant Sci 11:615392. https://doi.org/10.3389/fpls.2020.615392

70. Liu Y, Douglas CJ (2015) A role for OVATE FAMILY PROTEIN1 (OFP1) and OFP4 in a BLH6-KNAT7 multi-protein complex regulating secondary cell wall formation in *Arabidopsis thaliana*. Plant Signal Behav 10:e1033126. https://doi.org/10.1080/15592324.2015.1033126

71. Liu Y, You S, Taylor-Teeples M, Li WL, Schuetz M, Brady SM, Douglas CJ (2014) BEL1-LIKE HOMEODOMAIN6 and KNOTTED ARABIDOPSIS THALIANA7 interact and regulate secondary cell wall formation via repression of REVOLUTA. Plant Cell 26:4843-4861. https://doi.org/10.1105/tpc.114.128322

72. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8

73. Ma G, Hou B, Wang Y, Zhang L, Liu C, Wang J, Wang X (2024) Camellia sinensis CsMYB4a participates in regulation of stamen growth by interaction with auxin signaling transduction repressor CsAUX/IAA4. Crop J 12:188-201. https://doi.org/10.1016/j.cj.2023.11.008

74. MacMillan CP, Birke H, Chuah A, Smith MOR, Besson X, Taylor J, Southerton S (2017) Tissue and cell-specific transcriptomes in cotton reveal the subtleties of gene regulation underlying the diversity of plant secondary cell walls. BMC Genomics 18:539. https://doi.org/10.1186/s12864-017-3902-4

75. McCarthy RL, Zhong R, Ye ZH (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. https://doi.org/10.1093/pcp/pcp139

76. Minh BQ, Nguyen MA, von Haeseler A (2013) Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol 30:1188-1195. https://doi.org/10.1093/molbev/mst024

77. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A (2021) Pfam: The protein families database in 2021. Nucleic Acids Res 49:D412-D419. https://doi.org/10.1093/nar/gkaa913

78. Mitsuda N, Iwase A, Yamamoto H, Yoshida M, Seki M, Shinozaki K, Ohme-Takagi M (2007) NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. Plant Cell 19:270-280. https://doi.org/10.1105/tpc.106.047043

79. Mokshina N (2025a) Cellulose synthase machinery in thickened primary cell walls of celery collenchyma via prizm of transcriptome analysis. J Plant Res 138:887-903. https://doi.org/10.1007/s10265-025-01662-4

80. Mokshina N (2025b) Functional specialization of cellulose synthase-like proteins (CSLs) in celery: tissue-specific roles in hemicellulose biosynthesis and beyond. Plant Physiol Rep 30:548-557. https://doi.org/10.1007/s40502-025-00890-z

81. Mokshina N, Panina A, Galinousky D, Mikshina P, Gorshkova T (2023) Transcriptome profiling of celery petiole tissues reveals peculiarities of the collenchyma cell wall formation. Planta 257:18. https://doi.org/10.1007/s00425-022-04042-7

82. Mokshina N, Sautkina O, Gorshkov O, Mikshina P (2025) A Fresh Look at Celery Collenchyma and Parenchyma Cell Walls Through a Combination of Biochemical, Histochemical, and Transcriptomic Analyses. Int J Mol Sci 26:738. https://doi.org/10.3390/ijms26020738

83. Mokshina N, Syrchina N (2025) "Shadow government": how transcription factors regulate plant cell wall formation. Plant Physiol Biochem 229:110589. https://doi.org/10.1016/j.plaphy.2025.110589

84. Nakano T, Suzuki K, Fujimura T, Shinshi H (2006) Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. Plant Physiol 140:411-432. https://doi.org/10.1104/pp.105.073783

85. Nakano Y, Nishikubo N, Goué N, Ohtani M, Yamaguchi M, Katayama Y, Demura T (2010) MYB transcription factors orchestrating the developmental program of xylem vessels in *Arabidopsis* roots. Plant Biotechnol 27:267-272. https://doi.org/10.5511/plantbiotechnology.27.267

86. Nakano Y, Yamaguchi M, Endo H, Rejab NA, Ohtani M (2015) NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. Front Plant Sci 6:288. https://doi.org/10.3389/fpls.2015.00288

87. Nakata MT, Sakamoto S, Nuoendagula, Kajita S, Mitsuda N (2021) Fiber Cell-Specific Expression of the VP16-Fused Ethylene Response Factor 41 Protein Increases Biomass Yield and Alters Lignin Composition. Front Plant Sci 12:654655. https://doi.org/10.3389/fpls.2021.654655

88. Nguema-Ona E, Vicré-Gibouin M, Gotté M, Plancot B, Lerouge P, Bardor M, Driouich A (2014) Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. Front Plant Sci 5:499. https://doi.org/10.3389/fpls.2014.00499

89. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268-274. https://doi.org/10.1093/molbev/msu300

90. Nixon BT, Mansouri K, Singh A, Du J, Davis LM, Lee JG, Slabaugh E, Vandavasi VG, O'Neill HM, Roberts AW, Roberts EM, Yingling YG, Haigler CH (2016) Comparative structural and computational analysis supports eighteen cellulose synthases in the plant cellulose synthesis complex. Sci Rep 6:28696. https://doi.org/10.1038/srep28696

91. Nookaraju A, Pandey SK, Ahlawat YK, Joshi CP (2022) Understanding the modus operandi of class II KNOX transcription factors in secondary cell wall biosynthesis. Plants 11:493. https://doi.org/10.3390/plants11040493

92. Oh E, Zhu JY, Wang ZY (2012) Interaction between BZR1 and PIF4 integrates brassinosteroid and environmental responses. Nat Cell Biol 14:802-809. https://doi.org/10.1038/ncb2545

93. Pařenicová L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, Angenent GC, Colombo L (2003) Molecular and Phylogenetic Analyses of the Complete MADS-Box Transcription Factor Family in *Arabidopsis* : New Openings to the MADS World. Plant Cell 15:1538-1551. https://doi.org/10.1105/tpc.011544

94. Park YB, Cosgrove DJ (2012) A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. Plant Physiol 158:1933-1943. https://doi.org/10.1104/pp.111.192880

95. Peaucelle A, Wightman R, Hofte H (2015) The control of growth symmetry breaking in the *Arabidopsis* hypocotyl. Curr Biol 25:1746-1752. https://doi.org/10.1016/j.cub.2015.05.022

96. Pei Y, Xue Q, Shu P, Xu W, Du X, Wu M, Liu K, Pirrello J, Bouzayen M, Hong Y, Liu M (2024) Bifunctional transcription factors SlERF.H5 and H7 activate cell wall and repress gibberellin biosynthesis genes in tomato via a conserved motif. Dev Cell 59:1345-1359.e6. https://doi.org/10.1016/j.devcel.2024.03.006

97. Persson E, Sonnhammer ELL (2022) InParanoid-DIAMOND: faster orthology analysis with the InParanoid algorithm. Bioinformatics 38:2918-2919. https://doi.org/10.1093/bioinformatics/btac194

98. Persson S, Paredez A, Carroll A, Palsdottir H, Doblin M, Poindexter P, Khitrov N, Auer M, Somerville CR (2007) Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis*. Proc Natl Acad Sci USA 104:15566-15571. https://doi.org/10.1073/pnas.0706592104

99. Pesquet E, Korolev AV, Calder G, Lloyd CW (2010) The microtubule-associated protein AtMAP70-5 regulates secondary wall patterning in *Arabidopsis* wood cells. Curr Biol 20:744-749. https://doi.org/10.1016/j.cub.2010.02.057

100. Polko JK, van Zanten M, van Rooij JA, Marée AFM, Voesenek LACJ, Peeters AJM, Pierik R (2012) Ethylene-induced differential petiole growth in *Arabidopsis thaliana* involves local microtubule reorientation and cell expansion. New Phytol 193:339-348. https://doi.org/10.1111/j.1469-8137.2011.03920.x

101. Puentes-Romero AC, González SA, González-Villanueva E, Figueroa CR, Ruiz-Lara S (2022) AtZAT4, a C2H2-Type Zinc Finger Transcription Factor from *Arabidopsis thaliana*, Is Involved in Pollen and Seed Development. Plants 11:1974. https://doi.org/10.3390/plants11151974

102. Qiao L, Tan H, Sun M, Zhang X, Xiao S, Ye Z, Chen Z, Zhang X, Tu L (2025) Bifunctional transcription factor GhMYB4 orchestrates transition from elongation to secondary cell wall synthesis trade-off in cotton fiber. Dev Cell 60:2503-2517.e5. https://doi.org/10.1016/j.devcel.2025.05.004

103. Qin W, Yin Q, Chen J, Zhao X, Yue F, He J, Yang L, Liu L, Zeng Q, Lu F, Mitsuda N, Ohme-Takagi M, Wu AM (2020) The class II KNOX transcription factors KNAT3 and KNAT7 synergistically regulate monolignol biosynthesis in *Arabidopsis*. J Exp Bot 71:5469-5483. https://doi.org/10.1093/jxb/eraa266

104. Radonić A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A (2004) Guideline to reference gene selection for quantitative real-time PCR. Biochem Biophys Res Commun 313:856-862. https://doi.org/10.1016/j.bbrc.2003.11.177

105. Robinson DF, Foulds LR (1981) Comparison of phylogenetic trees. Math Biosci 53:131-147. https://doi.org/10.1016/0025-5564(81)90043-2

106. Rose JK, Braam J, Fry SC, Nishitani K (2002) The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. Plant Cell Physiol 43:1421-1435. https://doi.org/10.1093/pcp/pcf171

107. Saelim L, Akiyoshi N, Tan TT, Ihara A, Yamaguchi M, Hirano K, Matsuoka M, Demura T, Ohtani M (2019) *Arabidopsis* Group IIId ERF proteins positively regulate primary cell wall-type CESA genes. J Plant Res 132:117-129. https://doi.org/10.1007/s10265-018-1074-1

108. Sakamoto S, Somssich M, Nakata MT, Unda F, Atsuzawa K, Kaneko Y, Wang T, Bågman AM, Gaudinier A, Yoshida K, Brady SM, Mansfield SD, Persson S, Mitsuda N (2018) Complete substitution of a secondary cell wall with a primary cell wall in *Arabidopsis*. Nat Plants 4:777-783. https://doi.org/10.1038/s41477-018-0260-4

109. Sánchez-Gerschon V, Martínez-Fernández I, González-Bermúdez MR, de la Hoz-Rodríguez S, González FV, Lozano-Juste J, Ferrándiz C, Balanzà V (2024) Transcription factors HB21/40/53 trigger inflorescence arrest through abscisic acid accumulation at the end of flowering. Plant Physiol 195:2743-2756. https://doi.org/10.1093/plphys/kiae234

110. Schwechheimer C, Schröder PM, Blaby-Haas CE (2022) Plant GATA Factors: Their Biology, Phylogeny, and Phylogenomics. Annu Rev Plant Biol 73:123-148. https://doi.org/10.1146/annurev-arplant-072221-092913

111. Shen H, Cao K, Wang X (2007) A conserved proline residue in the leucine zipper region of AtbZIP34 and AtbZIP61 in *Arabidopsis thaliana* interferes with the formation of homodimer. Biochem Biophys Res Commun 362:425-430. https://doi.org/10.1016/j.bbrc.2007.08.026

112. Shore P, Sharrocks AD (1995) The MADS-box family of transcription factors. Eur J Biochem 229:1-13. https://doi.org/10.1111/j.1432-1033.1995.tb20430.x

113. Siriwardana CL (2024) Plant Nuclear Factor Y (NF-Y) Transcription Factors: Evolving Insights into Biological Functions and Gene Expansion. Int J Mol Sci 26:38. https://doi.org/10.3390/ijms26010038

114. Slowikowski K (2024) ggrepel: Automatically Position Non-overlapping Text Labels with 'ggplot2'. R Package Version 0.9.8. https://CRAN.R-project.org/package=ggrepel

115. Smalle J, Kurepa J, Haegman M, Gielen J, Van Montagu M, Van Der Straeten D (1998) The trihelix DNA-binding motif in higher plants is not restricted to the transcription factors GT-1 and GT-2. Proc Natl Acad Sci USA 95:3318-3322. https://doi.org/10.1073/pnas.95.6.3318

116. Song SK, Jang HU, Kim YH, Lee BH, Lee MM (2020) Overexpression of three related root-cap outermost-cell-specific C2H2-type zinc-finger protein genes suppresses the growth of *Arabidopsis* in an EAR-motif-dependent manner. BMB Rep 53:160-165. https://doi.org/10.5483/BMBRep.2020.53.3.286

117. Syrchina N, Mikshina P, Mokshina N (2025) Ethylene response factors as potential regulators of cell wall formation in celery (*Apium graveolens* L.) collenchyma: phylogeny and gene expression analysis. Protoplasma 262:1-20. https://doi.org/10.1007/s00709-025-02042-4

118. Tan QKG, Irish VF (2006) The *Arabidopsis* Zinc Finger-Homeodomain Genes Encode Proteins with Unique Biochemical Properties That Are Coordinately Expressed during Floral Development. Plant Physiol 140:1095-1108. https://doi.org/10.1104/pp.105.070565

119. Tang D, Chen M, Huang X, Zhang G, Zeng L, Zhang G, Wang Y (2023) SRplot: a free online platform for data visualization and graphing. PLoS ONE 18:e0294236. https://doi.org/10.1371/journal.pone.0294236

120. Tian GW, Chen MH, Zaltsman A, Citovsky V (2006) Pollen-specific pectin methylesterase involved in pollen tube growth. Dev Biol 294:83-91. https://doi.org/10.1016/j.ydbio.2006.02.026

121. Tukey JW (1949) Comparing individual means in the analysis of variance. Biometrics 5:99-114. https://doi.org/10.2307/3001913

122. UniProt Consortium (2023) UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res 51:D523-D531. https://doi.org/10.1093/nar/gkac1052

123. Van Leene J, Blomme J, Kulkarni SR, Cannoot B, De Winne N, Eeckhout D, Persiau G, Van De Slijke E, Vercruysse L, Vanden Bossche R, Heyndrickx KS, Vanneste S, Goossens A, Gevaert K, Vandepoele K, Gonzalez N, Inzé D, De Jaeger G (2016) Functional characterization of the *Arabidopsis* transcription factor bZIP29 reveals its role in leaf and root development. J Exp Bot 67:5825-5840. https://doi.org/10.1093/jxb/erw347

124. Vega-Sánchez ME, Verhertbruggen Y, Christensen U, Chen X, Sharma V, Varanasi P, Jobling SA, Talbot M, White RG, Joo M, Singh S, Auer M, Scheller HV, Ronald PC (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. https://doi.org/10.1104/pp.112.195495

125. Vogel J (2008) Unique aspects of the grass cell wall. Curr Opin Plant Biol 11:301-307. https://doi.org/10.1016/j.pbi.2008.03.002

126. Wachananawat B, Kuroha T, Takenaka Y, Kajiura H, Naramoto S, Yokoyama R, Ishizaki K, Nishitani K, Ishimizu T (2020) Diversity of pectin rhamnogalacturonan I rhamnosyltransferases in glycosyltransferase family 106. Front Plant Sci 11:997. https://doi.org/10.3389/fpls.2020.00997

127. Wan Y, Gritsch C, Tryfona T, Ray MJ, Andongabo A, Hassani-Pak K, Jones HD, Dupree P, Karp A, Shewry PR, Mitchell RAC (2014) Secondary cell wall composition and candidate gene expression in developing willow (Salix purpurea) stems. Planta 239:1041-1053. https://doi.org/10.1007/s00425-014-2034-1

128. Wang H, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA (2010) Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. Proc Natl Acad Sci USA 107:22338-22343. https://doi.org/10.1073/pnas.1016436107

129. Wang S, Yamaguchi M, Grienenberger E, Martone PT, Samuels AL, Mansfield SD (2020) The Class II KNOX genes KNAT3 and KNAT7 work cooperatively to influence deposition of secondary cell walls that provide mechanical support to *Arabidopsis* stems. Plant J 101:293-309. https://doi.org/10.1111/tpj.14541

130. Wickham H (2016) ggplot2: Elegant graphics for data analysis. 2nd ed. Springer, New York. https://doi.org/10.1007/978-3-319-24277-4

131. Wolters-Arts AMC, Van Amstel T, Derksen J (1993) Tracing cellulose microfibril orientation in inner primary cell walls. Protoplasma 175:102-111. https://doi.org/10.1007/BF01385007

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

133. Yao S, Chen P, Yu Y, Zhang M, Wang D, Liu J, Hao Q, Ji K (2021) PmMYB4, a Transcriptional Activator from *Pinus massoniana*, Regulates Secondary Cell Wall Formation and Lignin Biosynthesis. Forests 12:1618. https://doi.org/10.3390/f12121618

134. Yong Y, Zhang Y, Lyu Y (2021) Functional characterization of *Lilium lancifolium* cold-responsive Zinc Finger Homeodomain (ZFHD) gene in abscisic acid and osmotic stress tolerance. PeerJ 9:e11508. https://doi.org/10.7717/peerj.11508

135. Yu H, Soler M, San Clemente H, Mila I, Paiva JAP, Myburg AA, Bouzayen M, Grima-Pettenati J, Cassan-Wang H (2015) Comprehensive Genome-Wide Analysis of the Aux/IAA Gene Family in *Eucalyptus*: Evidence for the Role of EgrIAA4 in Wood Formation. Plant Cell Physiol 56:700-714. https://doi.org/10.1093/pcp/pcu215

136. Yu L, Liu D, Yin F, Zhang L, Cheng X, Wang L, Jiang J, Liu S, Zhou Y (2023) Interaction between phenylpropane metabolism and oil accumulation in the developing seed of Brassica napus revealed by high temporal-resolution transcriptomes. BMC Biol 21:202. https://doi.org/10.1186/s12915-023-01705-z

137. Yu Y (2019) OsKNAT7 Bridges Secondary Cell Wall Formation and Cell Growth Regulation. Plant Physiol 181:385-386. https://doi.org/10.1104/pp.19.01018

138. Yu Y, Liu H, Zhang N, Gao C, Qi L, Wang C (2021) The BpMYB4 Transcription Factor From *Betula platyphylla* Contributes Toward Abiotic Stress Resistance and Secondary Cell Wall Biosynthesis. Front Plant Sci 11:606062. https://doi.org/10.3389/fpls.2020.606062

139. Zar JH (2010) Biostatistical analysis. 5th edn. Prentice-Hall/Pearson, Upper Saddle River

140. Zhang B, Zhang L, Li F, Zhang D, Liu X, Wang H, Xu Z, Chu C, Zhou Y (2017) Control of Secondary Cell Wall Patterning Involves Xylan Deacetylation by a GDSL Esterase. Nat Plants 3:17017. https://doi.org/10.1038/nplants.2017.17

141. Zhang C, Xiong AT, Ren MY, Xu L, Zhang Y, Xiang H, Xia Y, Yu C, Wang J, Gu H (2024) An epigenetically mediated double negative cascade from EFD to HB21 regulates anther development. Nat Commun 15:7796. https://doi.org/10.1038/s41467-024-52114-x

142. Zhang M, Henquet MGL, Chen Z, Zhang H, Zhang Y, Ren X, van der Krol AR, Gonneau M, Bosch HJ, Gong Z (2009) LEW3, encoding a putative a-1,2-mannosyltransferase (ALG11) in N-linked glycoprotein, plays vital roles in cell-wall biosynthesis and the abiotic stress response in *Arabidopsis thaliana*. Plant J 60:983-999. https://doi.org/10.1111/j.1365-313X.2009.04013.x

143. Zhang N, Julian JD, Zabotina OA (2025) Multiprotein Complexes of Plant Glycosyltransferases Involved in Their Function and Trafficking. Plants 14:350. https://doi.org/10.3390/plants14030350

144. Zhao T, Piñeyro-Nelson A, Yu Q, Zhang J, Wang Y, Gao Y, Fan Y (2021) Transcriptome-Wide Analyses Provide Insights into Development of the *Hedychium coronarium* Flower, Revealing Potential Roles of PTL. J Plant Biol 64:431-445. https://doi.org/10.1007/s12374-021-09312-7

145. Zhong R, Demura T, Ye ZH (2006) SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*. Plant Cell 18:3158-3170. https://doi.org/10.1105/tpc.106.047399

146. Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. Plant Cell 20:2763-2782. https://doi.org/10.1105/tpc.108.061325

147. Zhong R, Richardson EA, Ye ZH (2007) Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis*. Planta 225:1603-1611. https://doi.org/10.1007/s00425-007-0498-y

148. Zhong R, Zhou D, Chen L, Rose JP, Wang BC, Ye ZH (2024) Plant Cell Wall Polysaccharide O-Acetyltransferases. Plants 13:2304. https://doi.org/10.3390/plants13162304

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

150. Zhuang Z, Bian J, Ren Z, Ta W, Peng Y (2025) Plant Aux/IAA Gene Family: Significance in Growth, Development and Stress Responses. Agronomy 15:1228. https://doi.org/10.3390/agronomy15051228

151. Zujovic Z, Chen D, Melton LD (2016) Comparison of celery (*Apium graveolens* L.) collenchyma and parenchyma cell wall polysaccharides enabled by solid-state 13C NMR. Carbohydr Res 420:51-57. https://doi.org/10.1016/j.carres.2015.11.006

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryinformation.xlsx