Genome Biology

RESEARCH

Open Access

Integrating scRNA-seq and snRNA-seq with spatial transcriptomics to unlock the xylem puzzle

Mingke Wei[1†], Jo-Wei Allison Hsieh[2,3†], Jr-Fong Dang[4†], Botong Tong[5], Hui Li[6], Ronald Sederoff[7], Ying-Chung Jimmy Lin[8,9*] and Quanzi Li[1*]

†Mingke Wei, Jo-Wei Allison Hsieh and Jr-Fong Dang contributed equally to this work.

*Correspondence: YCJimmyLin@ntu.edu.tw; liqz@zafu.edu.cn

[1] National Key Laboratory for Development and Utilization of Forest Food Resources, Zhejiang Key Laboratory of Forest Genetics and Breeding, International Research Center for Plant Cell Wall, College of Forestry and Biotechnology, Zhejiang A&F University, Hangzhou, China
[8] Department of Life Science, College of Life Science, National Taiwan University, Taipei, Taiwan
Full list of author information is available at the end of the article

Abstract

Background: Xylem development is a dynamic, continuous process fundamental to secondary growth in woody plants and to biomass accumulation on earth. While single-cell RNA sequencing (scRNA-seq) enables reconstruction of early xylem differentiation trajectories, its reliance on protoplast isolation excludes late-stage cells with thickened secondary cell walls, leaving key phases such as secondary cell wall deposition and programmed cell death poorly characterized.

Results: We perform single-nucleus RNA sequencing (snRNA-seq) of the stem-developing xylem of *Populus* and integrate previous scRNA-seq datasets to reconstruct a comprehensive developmental landscape of xylem formation. Anatomical validation confirms that scRNA-seq profiles predominantly represent early-stage stem-developing xylem, while snRNA-seq enriches for deeper, secondary cell wall depositing layers. Integrated analysis reveals a spatially and transcriptionally defined secondary cell wall zone, supported by both lignin autofluorescence and its correlation with laser capture microdissection-derived transcriptomes. Differential expression and gene ontology analyses uncover enrichment for lignin biosynthesis and programmed cell death associated genes, suggesting that secondary cell wall formation and programmed cell death initiation are transcriptionally coordinated. Unsupervised clustering and machine learning by support vector machine classification further reveal greater transcriptomic heterogeneity among early-stage xylem cells compared to late-stage cells.

Conclusions: Our findings demonstrate the high compatibility and complementarity of scRNA-seq and snRNA-seq platforms. This integrated approach not only overcomes key limitations of individual technologies but also provides a practical and scalable framework for resolving complex developmental trajectories in plants. It successfully reconstructs the full continuum of xylem development, from cambial progenitors through secondary cell wall formation and to programmed cell death.

Keywords: Single-nucleus RNA-seq, Spatial transcriptome, Single-cell RNA-seq, Xylem development, Secondary cell wall, Programmed cell death, Laser capture microdissection, Machine learning

Background

In multicellular organisms, development is fundamentally a continuous and dynamic process [1]. From the earliest cell fate decisions to terminal differentiation, cells undergo gradual and coordinated transitions in gene expression, morphology, and function. Capturing these transitions with high resolution is essential for uncovering the mechanisms that drive tissue patterning, organogenesis, and environmental adaptation [2–4]. In recent years, single-cell-level technologies have revolutionized our ability to reconstruct these developmental trajectories, providing powerful frameworks to infer lineage relationships and identify intermediate states [5–7]. This principle applies broadly across biological systems—from animal embryogenesis to plant meristem differentiation—where understanding the continuity of developmental states is crucial for constructing accurate models of growth and form.

Xylem is the most abundant tissue on earth in terms of biomass, with the majority of this biomass generated through secondary growth [8–10]. In the context of xylem development in plant secondary growth, once vascular cambium cells divide and commit to a xylem fate, their differentiation proceeds through three major stages: (1) cell expansion, (2) secondary cell wall (SCW) deposition, and (3) programmed cell death (PCD) [11, 12]. Recent advances in single-cell transcriptomics have provided unprecedented insights into the early phases of this process, enabling the reconstruction of dynamic differentiation trajectories from cambial progenitors to specialized xylem cell types such as libriform fibers (fibers) and vessel elements (vessels) [13–19]. These efforts have culminated in a widely accepted working model of xylem development that is shared across the global plant research community, shedding light on key regulatory genes and transitional states involved in early xylem formation [20–24]. In such working model, xylem development is composed of two developmental lineages from axial and radial system. The axial system arises from the fusiform lineage, beginning with fusiform organizer cells that give rise to early fusiform precursors, which then differentiate into intermediate precursors and ultimately mature into either fibers or vessels. The radial system is derived from the ray lineage, originating from ray organizer cells that give rise to ray precursors, which subsequently differentiate into ray parenchyma cells [16, 24]. However, this model is largely constructed from single-cell RNA sequencing (scRNA-seq) data, which inherently exclude cells with extensively thickened SCWs due to the technical limitations of protoplast-based cell isolation. As a result, our understanding remains incomplete—particularly for the later stages of xylem differentiation, including the maturation of SCW architecture and the execution of PCD.

To overcome the limitations of protoplast-based single-cell transcriptomics, single-nucleus RNA sequencing (snRNA-seq) has emerged as a powerful alternative. Unlike scRNA-seq, which requires enzymatic digestion of the cell wall to isolate intact protoplasts, snRNA-seq bypasses this constraint by isolating nuclei directly from frozen, mechanically disrupted tissues [25, 26]. This enables the capture of transcriptomic information from cells that are otherwise inaccessible—particularly those with thickened SCWs, such as mature xylem cells—thereby extending the developmental window that can be profiled [27]. However, snRNA-seq has several inherent limitations. Because it captures only nuclear RNA, it tends to yield sparser transcriptomes, with fewer detected genes and limited representation of cytoplasmic transcripts. This sparsity

can hinder the identification of subtle cell states and regulatory programs, and the absence of cytoplasmic RNA may obscure critical continuous and dynamic process. In contrast, scRNA-seq captures both nuclear and cytoplasmic RNAs, providing more comprehensive transcriptomic coverage [28–30]. It typically detects more genes per cell and reveals finer transcriptional dynamics, particularly in early developmental stages or metabolically active cell types. Given the complementary nature of these two technologies—scRNA-seq offering high-resolution transcriptomic profiles but restricted to thin-walled, wall-degradable cells, and snRNA-seq enabling access to a broader range of cell types, including those with thick SCWs—a combined approach is essential to reconstruct the full developmental trajectory of xylem formation [30].

In this study, we performed snRNA-seq on the stem-developing xylem (SDX) of *Populus alba* × *Populus glandulosa* and integrated the resulting data with our previously generated scRNA-seq data. By jointly analyzing the two transcriptomic modalities, we constructed a more comprehensive and continuous model of xylem development—spanning from early-stage, thin-walled cells to late-stage, lignified elements—and to evaluate the extent to which scRNA-seq and snRNA-seq can reinforce each other in resolving plant cell developmental trajectories.

Results

Complementary sampling properties of protoplast-scRNAseq and SDX-snRNAseq revealed by anatomical analysis

Building upon our previously reported scRNA-seq (protoplast-scRNAseq) datasets, we conducted SDX-snRNAseq in this study to enable direct comparative analyses between the two platforms [31]. To compare the developmental stages captured by single-cell and single-nucleus transcriptomics in xylem tissue, we first examined the sample preparation workflows underlying protoplast-scRNAseq and SDX-snRNAseq. We carefully peeled the bark from stem segments of *P. alba* × *P. glandulosa* to expose the SDX and minimize contamination from phloem and vascular cambium. This debarking procedure ensured that the surface of debarked stem consisted almost exclusively of SDX. For protoplast-scRNAseq, fresh debarked stems were immediately subjected to enzymatic digestion to isolate protoplasts, which were then processed for single-cell transcriptome profiling. For SDX-snRNAseq, the exposed SDX tissue was scraped and flash-frozen, followed by mechanical grinding and nuclear purification to obtain nuclei for library preparation. To precisely evaluate the differences in tissue capture between protoplast-scRNAseq and SDX-snRNAseq at the sample preparation level, we performed comprehensive anatomical analyses.

First, we used agarose-embedded stem samples to compare intact and debarked tissues to check whether the debarking process can effectively remove phloem and cambium layers and expose SDX on the surface of the debarked stem. Compared to intact stems (Fig. 1a), debarked stems exhibited a clear separation (a gap) between the SDX and cambial region, with nearly all the flattened cambium cells located on the phloem side, indicating that the debarking procedure effectively removes potential contamination from both the phloem and cambium tissues (Fig. 1b). Second, the debarked stems were subjected to three different treatments—protoplasting, flash-freezing, and tissue scraping—each representing key steps in the protoplast-scRNAseq, laser capture

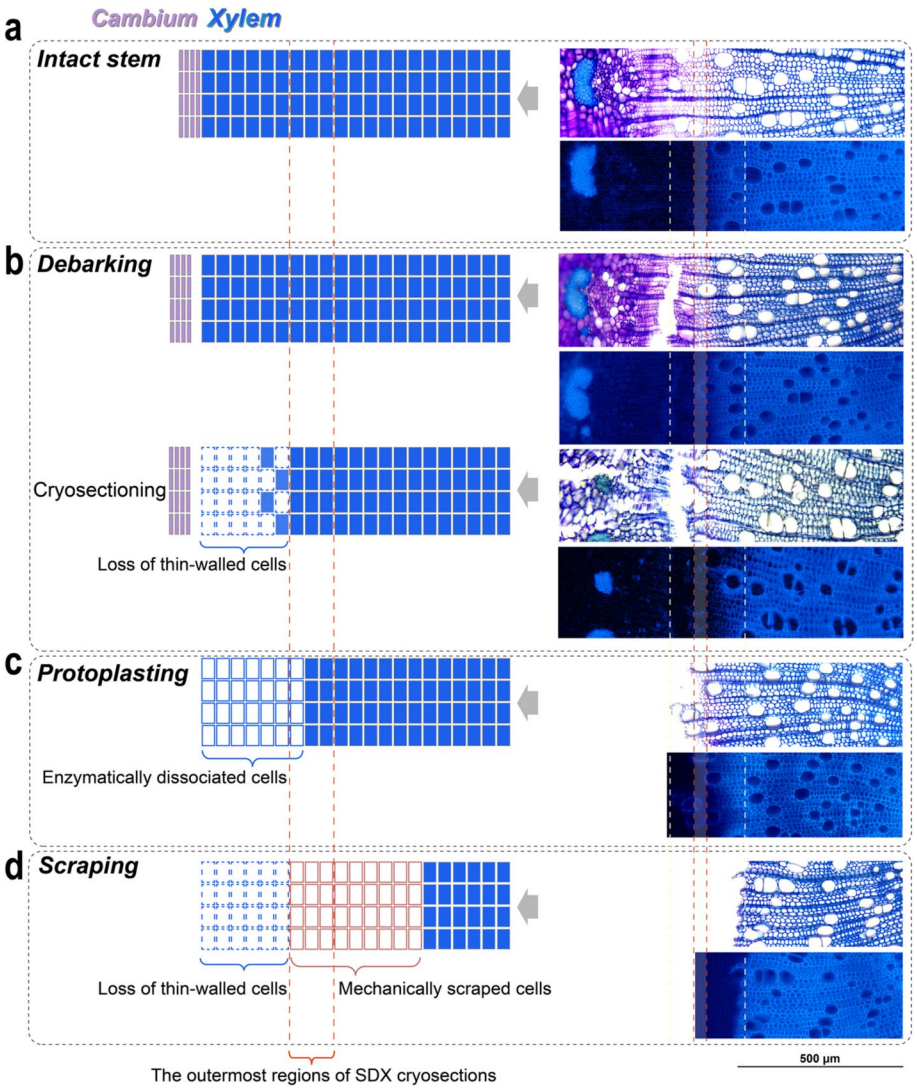

Fig. 1 Comprehensive anatomical analysis of xylem development and tissue processing. Schematic diagrams (left) show the cell capture based on the stem cross-sections (right) on different sample processes. In the right images, the upper panels show TBO-stained stem cross-sections, and lower panels display autofluorescence of the corresponding stem cross-sections under UV excitation. **a** Cross-sectional anatomy of the intact stem. Solid grids in dark purple and blue represent sequential layers of cambium and xylem cells, respectively. **b** Post-peeling stem cross-sections show vibratome-sectioned microstructure and cryosectioned microstructure. Dashed grids indicate regions of parenchyma cell loss. **c** Protoplast isolation from enzymatically digested xylem. Blue hollow grids represent protoplast-enriched cell layers collected after digestion of peeled xylem-facing tissues. **d** Scraping-based nuclear acquisition from xylem. Red hollow grids denote cell layers subjected to nuclear scraping. White dashed lines demarcate post-processing regions (enzymatic digestion/cryotreatment/scraping). Red dashed lines outline LCM target zones. Scale bars: 500 μm

microdissection (LCM) and SDX-snRNAseq workflows, respectively. The samples were subsequently embedded in agarose (protoplasts and scraped samples) and optimal cutting temperature (OCT) compound (flash-frozen samples) for anatomical analysis. In the anatomical sections, regions appearing as empty spaces indicated areas where cells were either enzymatically removed during protoplasting, disrupted by flash-freezing, or physically detached during tissue scraping (Fig. 1b and c). The empty spaces observed

in both the protoplasted and flash-frozen samples appeared similar, suggesting that these two treatments predominantly and respectively collect and destroy thin-walled cells. Following peel-and-freeze treatment, outer parenchyma cells exhibited rupture (Fig. 1b). Similarly, the snap-freezing step during SDX-snRNAseq sample preparation may lead to the loss of xylem cells with thin walls. In contrast, tissue scraping techniques enable targeted isolation of thick-walled cells residing in deeper SDX regions, particularly those undergoing active secondary cell wall biosynthesis, a process biochemically validated through lignin-specific autofluorescence signatures (Fig. 1d). Compared to protoplasting, which typically yields approximately 6 to 8 layers of xylem cells, tissue scraping allowed the recovery of an additional 8 layers, resulting in a total of around 15 xylem cell layers being collected (white double-dashed line in Fig. 1). In other words, protoplast-scRNAseq primarily captures the transcriptomic profiles of early-stage SDX development, characterized by thin-walled, actively expanding cells. On the other hand, SDX-snRNAseq—through the combination of morphological analyses via tissue scraping and flash-freezing—enables access to later stages of xylem differentiation, including thick-walled cells undergoing SCW deposition. Given the distinct yet complementary sampling strategies of the two platforms, we also anticipated a partial overlap in transcriptomic profiles at intermediate, transitioning stages of SDX development.

Integration of protoplast-scRNAseq and SDX-snRNAseq datasets reveals a continuous xylem developmental trajectory

We conducted SDX-snRNAseq on the SDX from two individual poplar trees. A total of 26,925 (13,496 in Bio1 and 13,429 in Bio2) nuclei were recovered, with ~ 26,500 genes detected (26,511 in Bio1 and 26,589 in Bio2) (Additional file 1: Fig. S1a). The numbers of the UMI (Unique molecular identifier) per nucleus obtained in the two replicates were 13,496 and 13,429, respectively (Additional file 1: Fig. S1b, Additional file 2: Supplementary Dataset S1). To assess the reproducibility of our platform, we calculated the transcriptomic overlap between the two samples. The extremely high overlap rate (97%) demonstrated the robustness and reliability of our SDX-snRNAseq workflow (Fig. 2a, Additional file 1: Fig. S2a). Similarly, MetaNeighbor analysis revealed strong concordance in cell type clustering between the two biological replicates (Additional file 1: Fig. S2d).

To obtain a more comprehensive collection of xylem cell types, we integrated the SDX-snRNAseq dataset with our previously generated protoplast-scRNAseq dataset. The two datasets exhibited distinct global distributions on the integrated UMAP (Fig. 2b). Cells from protoplast-scRNAseq and SDX-snRNAseq were clearly positioned on opposite sides of the UMAP space, with partial overlap observed only in the intermediate region. To rule out the possibility that the observed distribution differences between scRNA-seq and snRNA-seq data resulted from platform-specific biases, we isolated protoplasts from the SDX and subsequently extracted nuclei for snRNA-seq (referred to as protoplast-snRNAseq). Integration of the protoplast-scRNAseq, SDX-snRNAseq, and protoplast-snRNAseq datasets showed that the protoplast-snRNAseq data exhibited a transcriptomic profile highly concordant with protoplast-scRNAseq and clustered predominantly with protoplast-scRNAseq in the UMAP embedding (Fig. 3a, b). The cell overlap between protoplast-snRNAseq and protoplast-scRNAseq was strikingly

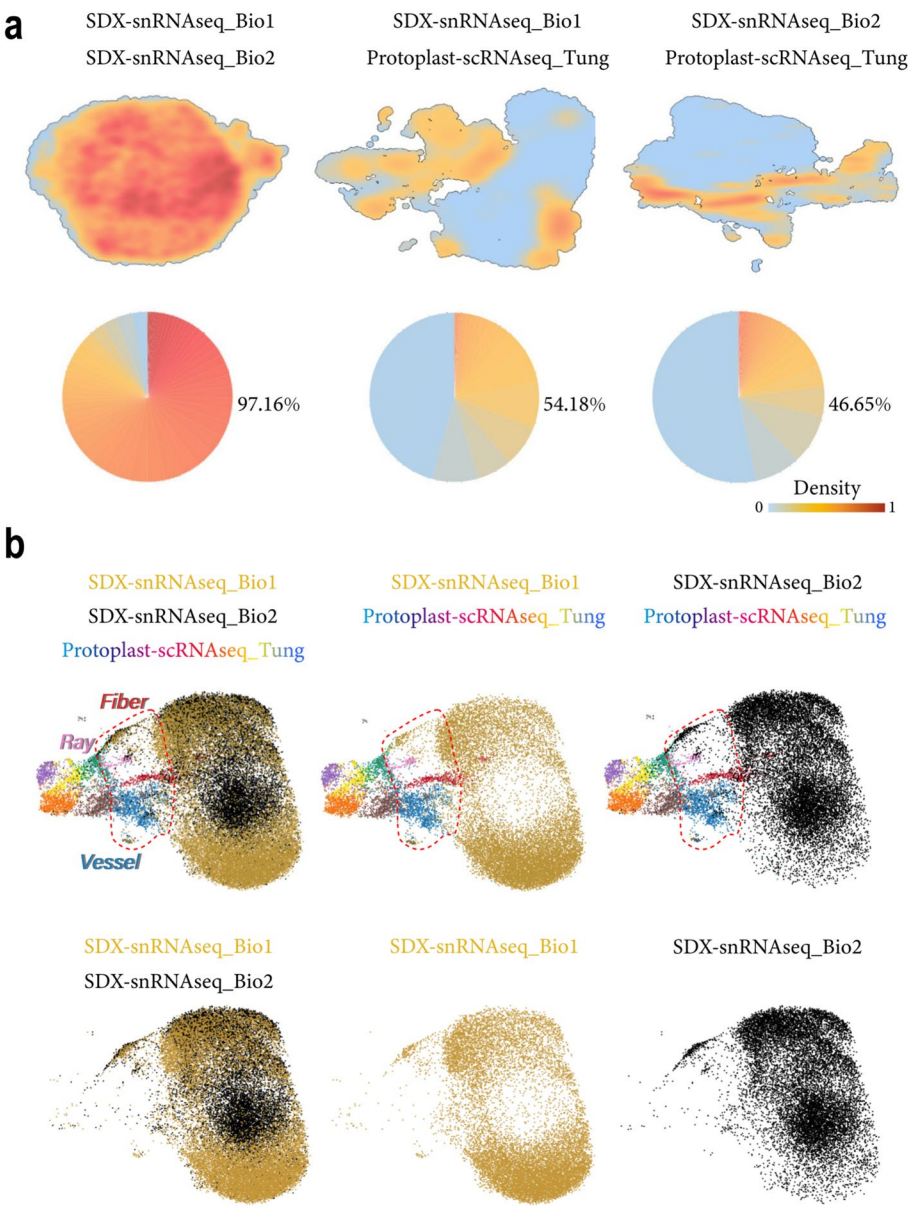

Fig. 2 Integrated analysis of xylem single-cell and single-nucleus transcriptomes. **a** Density overlay of two biological replicates ($n = 2$) in single-nucleus RNA sequencing (SDX-snRNAseq), demonstrating high reproducibility across xylem cell populations. Densities from 0 to 1 are divided into 500 bins with different color shading, with proportions of different densities shown in a pie chart in each panel. **b** Integrated UMAP visualization of protoplast-scRNAseq (rainbow-colored clusters) and SDX-snRNAseq (gold and black clusters, representing two biological replicates). Regions where scRNA-seq-derived fiber cells (red), vessels (blue), and ray parenchyma cells (pink) begin to overlap with SDX-snRNAseq clusters are dash-circled

high, reaching approximately 94% (90.31% and 96.65% for replicate 1 and 2, respectively; Fig. 3c). In contrast, the overlap between protoplast-snRNAseq and SDX-snRNAseq was only 7.5% (4.3% and 10.6% across replicates; Fig. 3c).

The cell distribution in SDX-snRNAseq and protoplast-scRNAseq is consistent with our anatomical analyses, which showed that protoplast-scRNAseq and SDX-snRNAseq capture largely non-overlapping developmental zones, except for the transitioning

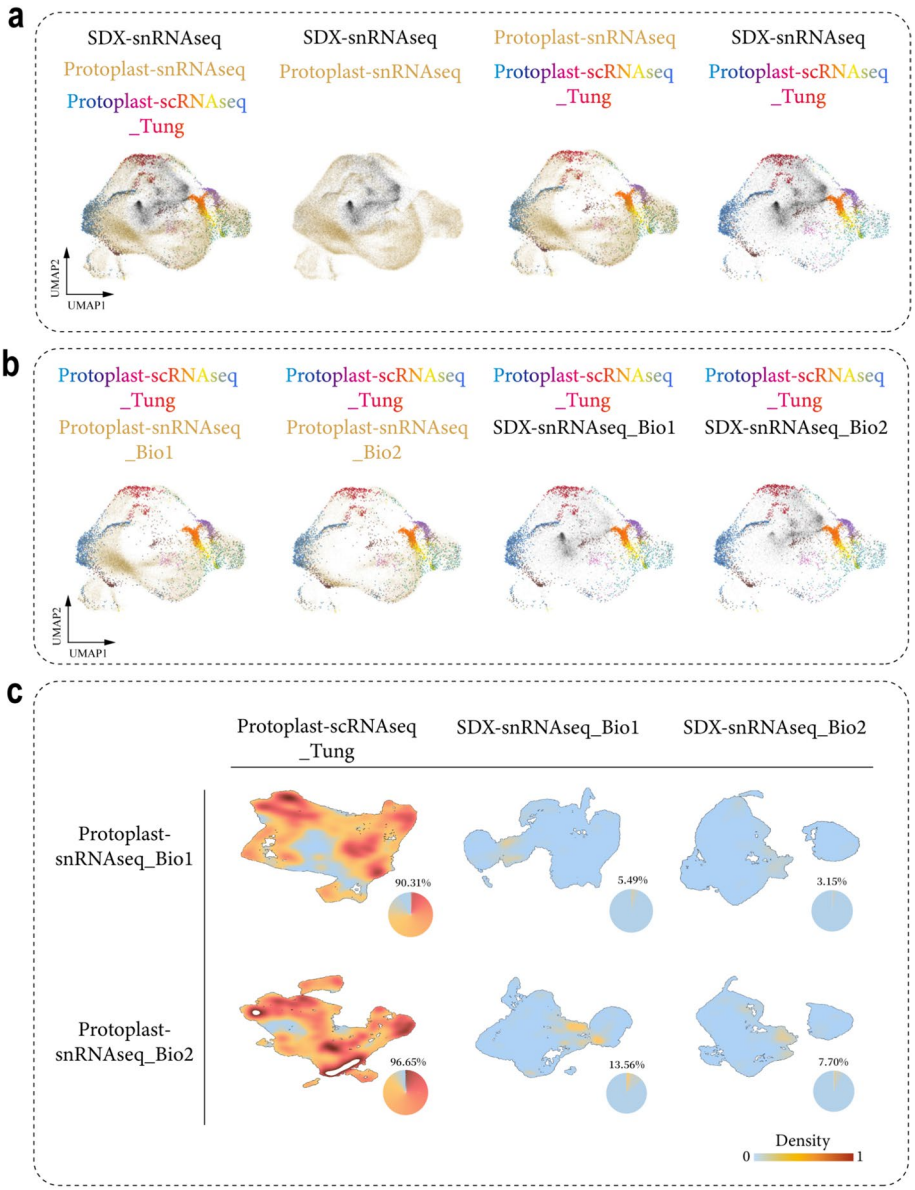

Fig. 3 Integration and clustering of protoplast-scRNAseq, SDX-snRNAseq, and protoplast-snRNAseq data with visualization. **a** UMAP visualization: integration of three datasets and pairwise comparisons. **b** UMAP visualization: integrated analysis of protoplast-scRNAseq, protoplast-snRNAseq replicates ($n = 2$), and SDX-snRNAseq replicates ($n = 2$). **c** Density overlay plots demonstrate high reproducibility between protoplast-scRNAseq and protoplast-snRNAseq, compared with their respective replicates ($n = 2$ for each). Densities ranging from 0 to 1 are divided into 500 bins, each assigned a distinct color shade, while a pie chart in each panel shows the proportion of each density

stages of SDX differentiation. This interpretation is further supported by the previously established cell-type annotations from our protoplast-scRNAseq dataset [24]. In that analysis, the most developmentally advanced cell clusters identified were vessel element cell (the blue cell cluster), libriform fiber cells (red), and ray parenchyma cells (pink). Intriguingly, these three clusters are located near the transitional region in the integrated UMAP, precisely where the protoplast-scRNAseq and SDX-snRNAseq datasets begin to

overlap (Fig. 2b). This spatial arrangement suggests that the overlapping zone represents both the terminal limit of protoplast-scRNAseq detection and the entry point of SDX-snRNAseq coverage, reinforcing the idea that the two platforms capture complementary, yet contiguous, segments of the xylem developmental trajectory.

Identification of the SCW deposition zone captured by SDX-snRNAseq

Given that SDX-snRNAseq preferentially captures thick-walled, late-stage xylem cells, we next sought to investigate how the protoplast-scRNAseq–defined libriform fiber, vessel element, and ray parenchyma cell types further develop at later stages. Specifically, we asked which downstream cell states these early-identified clusters transition into, as revealed by the extended coverage of SDX-snRNAseq. We adapted our previous cell-type-specific transcriptomic data generated from the specific cell types harvested by LCM, the cell-type LCM RNAseq [24]. Correlation analyses were performed between the cell-type-specific transcriptomics and single-nucleus transcriptomics. Consistent with our previous findings, each cell type–specific transcriptome exhibited the highest correlation with the corresponding cell cluster previously defined in the protoplast-scRNAseq dataset (Fig. 4a-d). For example, transcriptomes specific to vessels displayed a markedly high correlation with the vessel cluster (blue) identified in the protoplast-scRNAseq data (the blue dashed circle in Fig. 4c). We noticed a region within the SDX-snRNAseq dataset that exhibited high correlation with both fiber– and vessel–specific transcriptomes, suggesting that the cells in this region are descendants of both lineages (the black dashed circle in Fig. 4b and c). This observation suggests that, during the stage of SCW deposition, the transcriptomes of fiber cells and vessels become increasingly similar, despite their distinct lineage origins. Interestingly, the region showing high correlation with both fiber cell– and vessel–specific transcriptomes exhibited extremely low correlation with the ray parenchyma–specific transcriptome (Fig. 4b-d). This is consistent with expectations, as this region represents an area of intense SCW deposition, whereas ray parenchyma cells are thin-walled and do not engage in SCW biosynthesis. These results further support the notion that this region in the SDX-snRNAseq dataset corresponds to a transcriptionally defined SCW deposition zone.

To further validate the spatial localization of the SCW deposition zone, we examined the anatomical features of the cryosectioned SDX tissue. During the cryosectioning process, thin-walled cells are highly susceptible to rupture due to freezing-induced mechanical stress. As a result, the outermost layer that remains intact in frozen sections likely corresponds to cells with well-developed SCWs—those resistant to freezing-induced damage. This provides anatomical support for the interpretation that the preserved surface layer of the cryosection represents the SCW-enriched region. Supporting this interpretation, we observed intense lignin autofluorescence localized to the outermost layer of the cryosection (Fig. 1b). To further validate this spatial assignment, we leveraged our previously generated dataset in which the outermost regions of SDX cryosections were collected using LCM followed by transcriptomic analysis, serving as a spatial transcriptomic reference [17, 32]. Excitingly, the region showing the highest correlation with the outermost cryosection closely matched the SCW deposition zone identified in the SDX-snRNAseq dataset (Fig. 4e and f, Additional file 1: Fig. S3). Thus, both anatomical observations and spatial transcriptomic data

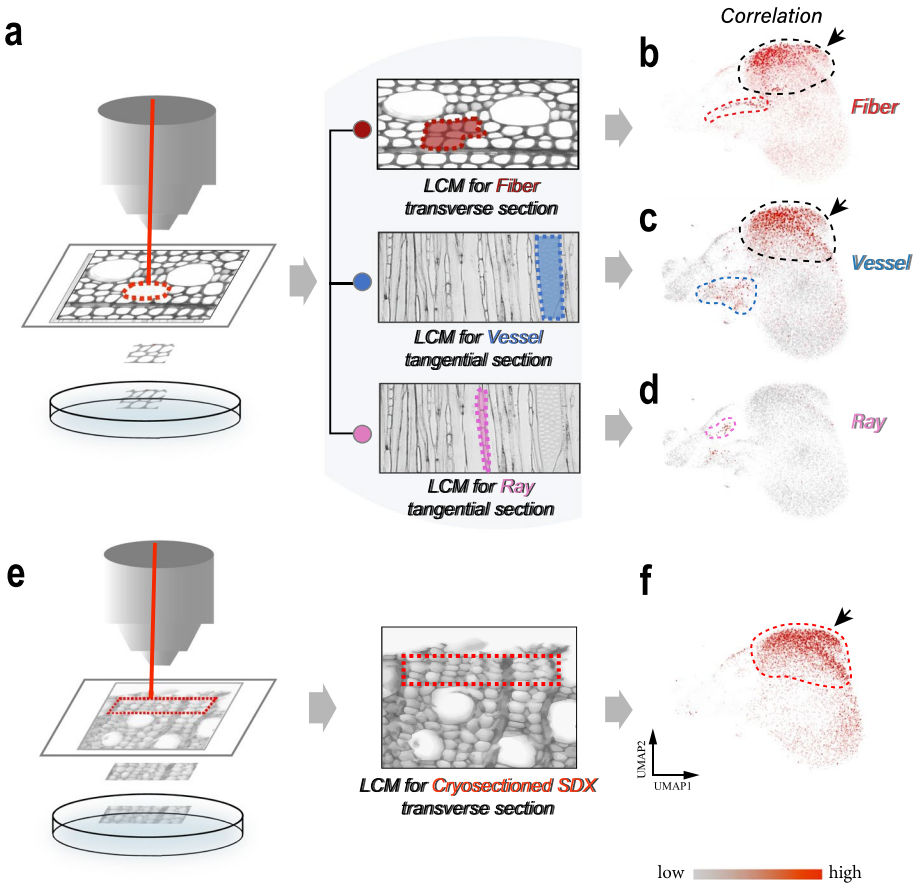

Fig. 4 Correlation analysis between LCM RNAseq cell-type subpopulations and SDX-snRNAseq. **a-d** Correlation profiles of LCM-isolated fiber cells, vessels, and ray parenchyma cells with SDX-snRNAseq subpopulations. Regions outlined in red, blue, and pink denote libriform fiber cell- (b), vessel element- (c), and ray parenchyma-enriched (d) correlations, respectively. The black dashed circles denote regions exhibiting higher correlations with libriform fiber and vessel element cells. **e–f** Correlation map between LCM-captured outermost cell layers (cryosectioned) and SDX-snRNAseq. The red dashed circles denote regions exhibiting higher correlations with LCM-derived cell populations. Color intensity (deep red = high correlation) reflects transcriptional concordance

consistently identified a population of xylem cells captured by SDX-snRNAseq that are undergoing extensive SCW deposition.

Transcriptomic signature of SCW deposition and onset of programmed cell death

To gain further biological insights into the cells within the SCW deposition region, we identified the genes that were differentially expressed in this population and subsequently performed Gene Ontology (GO) enrichment analysis. By integrating LCM RNAseq and SDX-snRNAseq data through correlation analysis, we annotated the subpopulations in the SDX-snRNAseq dataset and identified two distinct clusters (Clusters 11 and 12). These two subpopulations, together with the ten clusters identified through protoplast-scRNAseq [24], form a comprehensive set of 12 cell clusters representing different stages of xylem cell development (Fig. 5a). Among the upregulated genes in the SCW deposition region, enriched GO terms included aromatic amino acid lyase, phenylalanine ammonia-lyase, and phenylpropanoid metabolism—all of which are key

components of the lignin biosynthetic pathway and serve as hallmark indicators of SCW deposition (Fig. 5b). Additionally, GO terms related to flavonoid biosynthesis were also enriched. As flavonoids and monolignols share the same upstream phenylpropanoid pathway, this further supports the notion that these cells are transcriptionally committed to lignin production and SCW deposition.

The final stage of xylem development is PCD. While the detailed molecular mechanisms that trigger PCD during xylem development remain largely unresolved [33], several genes and pathways—such as aspartase, ethylene response factors (ERFs), and extensin—have previously been implicated in plant PCD [34–39]. Notably, we found that these genes were also differentially regulated in the SCW-depositing cell population and were enriched among the identified GO terms (Additional file 3: Supplementary Dataset S2), suggesting that the onset of PCD may occur concurrently with SCW deposition.

We further selected the genes that were upregulated in different clusters and performed RNA in situ hybridization for comparison of their expression patterns with those obtained by the protoplast-scRNAseq/SDX-snRNAseq analysis (Fig. 5c, d, Additional file 1: Fig. S4). For the genes upregulated in Clusters 11 and 12, their in situ hybridization signals were predominantly observed in inner xylem cells, which are undergoing cell wall thickening and PCD (Fig. 5c). For instance, *ACL5* (*ACAULIS5*), encoded a S-adenosyl-L-methionine-dependent methyltransferase, was strongly expressed in the layers 5–15, *FLA12* (*FASCICLIN-like arabinogalactan protein 12*) showed pronounced expression in layers 12–18, and both *RLK* (*Receptor-like protein kinase-related family protein*) and *XCP2* (*xylem cysteine peptidase 2*)—a gene involved in vacuolar collapse during PCD—were expressed in the mature xylem region. While for the genes selected from clusters 1–10, their expression patterns were primarily localized to the outmost cell layers adjacent to the cambium (Fig. 5d). However, it was difficult to distinguish hybridization signals between fusiform initial precursors and ray initial precursors. For example, Potri.014G161200 (enriched in fusiform intermediate precursor, Cluster 2) and Potri.T011801 (ray organizer, Cluster 3) were uniformly expressed across both fiber and ray cells. Similarly, Potri.014G022200 (fusiform intermediate precursor, Cluster 2) was detected in vessels, and fibers and ray cells. The corresponding UMAP expression profiles of these genes showed broad distribution across multiple clusters, which was consistent with the RNA in situ hybridization. The expression of these genes could not distinguish the cell populations at early stage of xylem development, which is related

(See figure on next page.)

Fig. 5 Expression of cell-type-specific genes in clusters associated with identified early and late developmental stages. **a** Twelve cell clusters, 1 to 12, were obtained through unsupervised *K*-means clustering and visualized by UMAP. 1-V, vessel element. 2-FulP, fusiform intermediate precursor. 3-RO, ray organizer. 4-FuEP, fusiform early precursor. 5-RP, ray precursor. 6-FuO, fusiform organizer. 7-F, libriform fiber. 8-R, ray parenchyma cell. 9–10, unknown cell. 11-L1, late stage cell1. 12-L2, late stage cell2. **b** GO enrichment terms in SCW deposition region of SDX-snRNAseq data. **c** RNA in situ hybridization of six selected genes at the late stage, including *ACL5*, *FLA12*, *GH1*, *Psp*, *RLK* and *XCP2* in Cluster 11 and 12. **d** RNA in situ hybridization of eight selected genes at the early stage, including *RIN4* in Cluster 6, *DUF761* in Cluster 4, Potri.014G161200 and Potri.014G022200 in Cluster 2, Potri.T011801 in Cluster 3, *UGT73B4* in Cluster5, *ATCTL2* in Cluster 7, *WAT1* in Cluster 1. **c** and **d** show the hybridization images using antisense probes, and the corresponding hybridization images using sense probes are provided in Additional file 1: Fig. S4. Fc: libriform fiber cell, Rc: ray parenchyma cell, Ve: vessel element cell. Arrow, site of gene expression. The red dashed circles denote gene expression regions. Scale bars: 100 μm

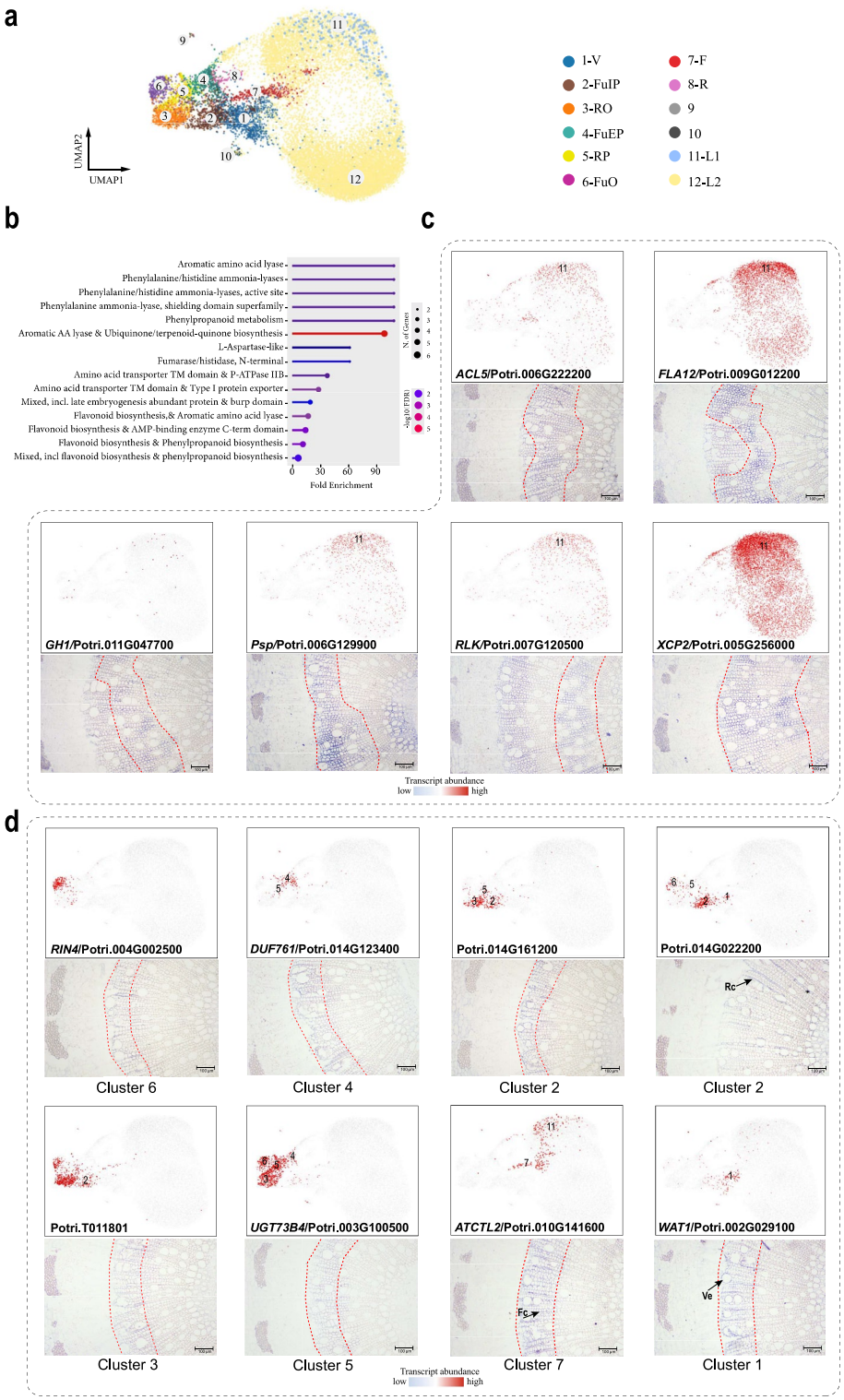

Fig. 5 (See legend on previous page.)

to the fact that these genes were expressed across the populations. Further functional studies on these gene candidates by gene-editing or silencing would provide insights on these cell populations.

Late-stage xylem cells exhibit reduced transcriptomic heterogeneity compared to early stages

To illustrate the developmental lineage from early stages to the SCW deposition and PCD stages, we performed unsupervised clustering to identify transcriptionally distinct cell clusters within the SDX-snRNAseq dataset. To ensure an unbiased assessment, we applied *K*-means clustering and systematically explored a range of cluster numbers, presenting results from $K=2$ to 20. Cells within the SCW deposition and PCD regions consistently formed a coherent grouping across different *K* values. While this cluster occasionally split into two subgroups at higher *K* values, it remained topologically stable—never merging with or being absorbed into neighboring clusters (Additional file 1: Fig. S5). This pattern indicates a well-defined transcriptomic identity with internally consistent structure and distinct boundaries from other cell populations. We then identified the cell clusters corresponding to the SCW deposition and PCD regions and integrated them with our previous scRNA-seq (protoplast-scRNAseq) results to reconstruct a comprehensive developmental lineage of xylem differentiation (Fig. 6e). This lineage refines prior scRNA-seq (protoplast-scRNAseq) trajectories through integration of late-stage molecular profiles, overcoming sampling biases that mask terminal differentiation states [24]. By combining SDX-snRNAseq and protoplast-scRNAseq data, our multimodal strategy enhances temporal resolution and biological reproducibility. The framework resolves developmental discontinuities while providing a critical resource for single-cell analysis of plant vascular ontogeny.

While performing *K*-means clustering, we observed a notable pattern: transcriptomic diversity appeared to be higher among early-stage xylem cells than in late-stage cells (Fig. 6a, Additional file 1: Fig. S5). This was evidenced by the tendency of early-stage regions to be subdivided into multiple clusters as *k* increased, whereas late-stage cells within the SCW deposition and PCD zones remained relatively consolidated. Since early-stage cells were primarily captured by protoplast-scRNAseq and late-stage cells by SDX-snRNAseq, one potential concern is that the observed differences in transcriptomic heterogeneity may simply reflect platform-specific biases. To investigate this possibility, we focused exclusively on the SDX-snRNAseq dataset. Because SDX-snRNAseq inherently enriches for late-stage xylem cells and contains relatively fewer early-stage cells, we performed stratified sampling to equalize the number of early- and late-stage cells, allowing for a balanced comparison of transcriptomic variability within a single platform (Fig. 6c). To assess transcriptomic diversity between early and late stages of xylem development, we applied a linear function to classify cells into early-stage and late-stage groups. We then compared the number of clusters formed among early-stage cells versus late-stage cells across a range of *K* values. Early-stage cells consistently exhibited a significantly higher number of clusters than late-stage cells, indicating greater transcriptional heterogeneity during the initial phases of xylem differentiation (Fig. 6b, Additional file 1: Fig. S6). To objectively validate the

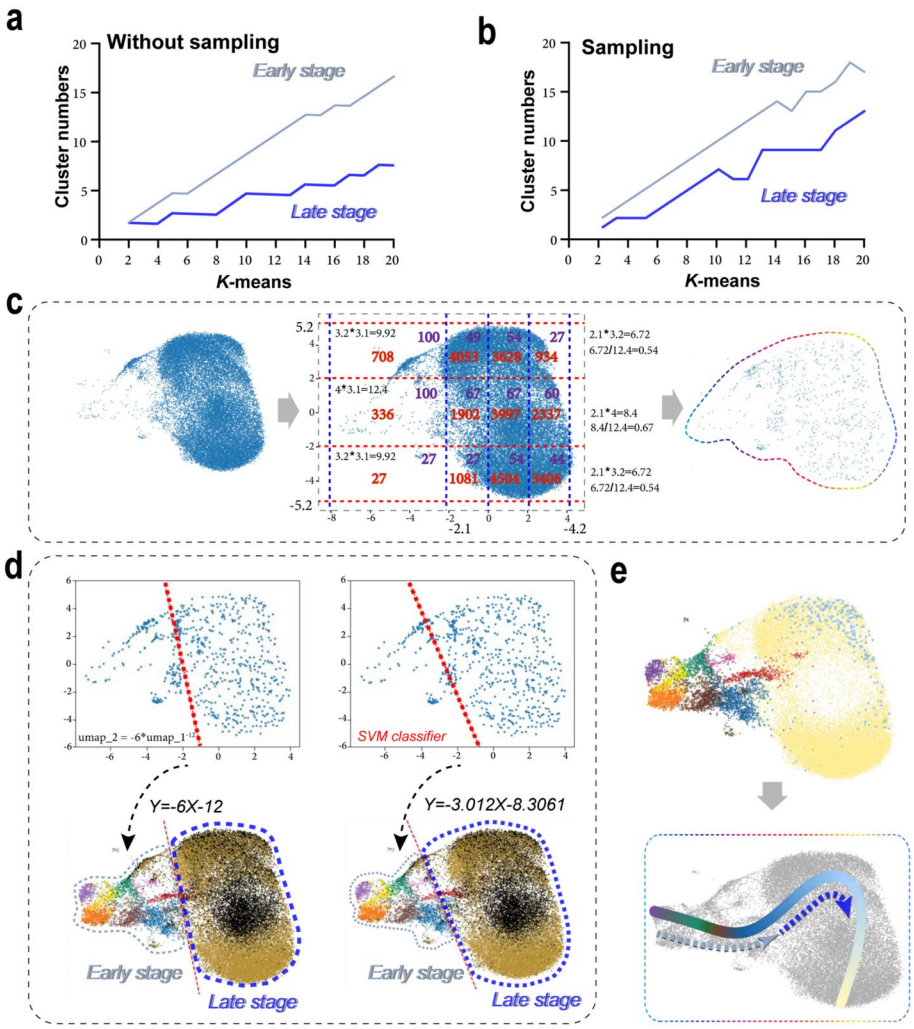

Fig. 6 Functional enrichment and transitional zone identification during xylem development. **a** Cell clusters trends under varying *K*-means clusters, comparing protoplast-scRNAseq (early-stage enriched) and SDX-snRNAseq (late-stage inferred). **b** *K*-means cluster trends post-sampling normalization for early- and late-stage cells. **c** Sampled cell distribution across UMAP transcriptional space. **d** Transition zone demarcation between early- and late-stage xylem development. Left, random linear segmentation; right, machine learning-based stable partitioning. e Slingshot-inferred differentiation trajectory of full xylem ontogeny

separation between early- and late-stage cells, we further employed a machine learning approach using a SVM classifier (Fig. 6c and d). This allowed us to derive an unbiased linear decision boundary that distinguishes early- and late-stage transcriptomic profiles based on gene expression features. Excitingly, the results based on the SVM-derived linear function were highly consistent with our previous clustering-based observations, further supporting the distinction in transcriptomic diversity between early- and late-stage xylem cells. Importantly, this also demonstrates the high degree of compatibility between protoplast-scRNAseq and SDX-snRNAseq platforms when applied to xylem developmental analysis, reinforcing the validity of integrating both datasets to reconstruct a continuous developmental trajectory.

Discussion

In this study, we conducted SDX-snRNAseq using more than 13,000 nuclei from SDX on each biological replicate. The detected median gene number in our study is comparable to some published data. For example, a snRNAseq study on Arabidopsis roots [40], their median genes/cell is 500–1,000. Considering that cellular heterogeneity is substantially higher in roots than in SDX, a median of 500–1,000 genes per cell is already regarded as sufficient for root studies, which indicates that the gene detection depth in our SDX dataset is more than adequate for robust downstream analyses. Compared to recent studies of snRNA-seq in poplar stems [41, 42], the detected gene number per nucleus in our study is relatively low (approximately 600–800), and the resulting clustering patterns differ from those in recent nuclei-based studies. Several factors may contribute to these discrepancies, including species-specific variation, differences in stem developmental stages of the sampled stems, and variability in nuclei isolation efficiency and nuclear RNA integrity. In addition, our nuclei isolation protocol preserved structurally mature, thick-walled cells, which may preferentially retain transcripts characteristic of later stages of xylem differentiation while limiting the capture of low-abundance or transient transcripts associated with early cambial or rapidly transitioning cell states. Such restricted transcript capture could reduce the resolution for fine-grained sub-clustering and might bias cell-type proportions toward transcriptionally robust, late-stage populations. Nevertheless, the core conclusions of this study are supported by the consistent identification of major xylem lineages, clear differentiation trajectories toward secondary wall-forming cells, and concordance with anatomical, spatial transcriptomics and marker-based validation. Integration of SDX-snRNAseq and protoplast-scRNAseq dataset provides practical insights into the compatibility between scRNAseq and snRNAseq in plants and underscores the importance of careful sample preparation and integration strategies. As cell type annotation remains a major challenge in plant single-cell transcriptomics—due to limited marker availability and the lack of reference atlases—physical separation of tissues during sample preparation can significantly reduce ambiguity [22, 24]. For instance, performing snRNAseq on whole stems eliminates the need for immediate processing or tissue dissection, but substantially increases the difficulty of downstream cell type annotation due to the presence of multiple, closely associated tissue types. Furthermore, during the nuclei preparation, a significant amount of impurities is generated from phloem, which affects the efficiency of nuclear purification. In contrast, physically enriching for specific regions—such as debarking to isolate the SDX—can greatly simplify annotation by reducing cellular heterogeneity [28]. However, this approach requires flash-freezing of the tissue to prevent artificial perturbations introduced by physical separation, which in turn can result in the loss of fragile, thin-walled cells during sample collection. Therefore, we advocate for a combined approach, as demonstrated in this study, where protoplast-scRNAseq and SDX-snRNAseq are jointly applied to the same anatomically defined region. This strategy enables comprehensive coverage of the xylem developmental trajectory while simultaneously reducing the ambiguity of cell identity assignments, offering a practical and scalable framework for future single-cell studies in plant systems.

Our finding that SCW deposition and PCD are transcriptionally co-regulated is further supported by known molecular pathways [8, 43] (Fig. 7a and b). In particular,

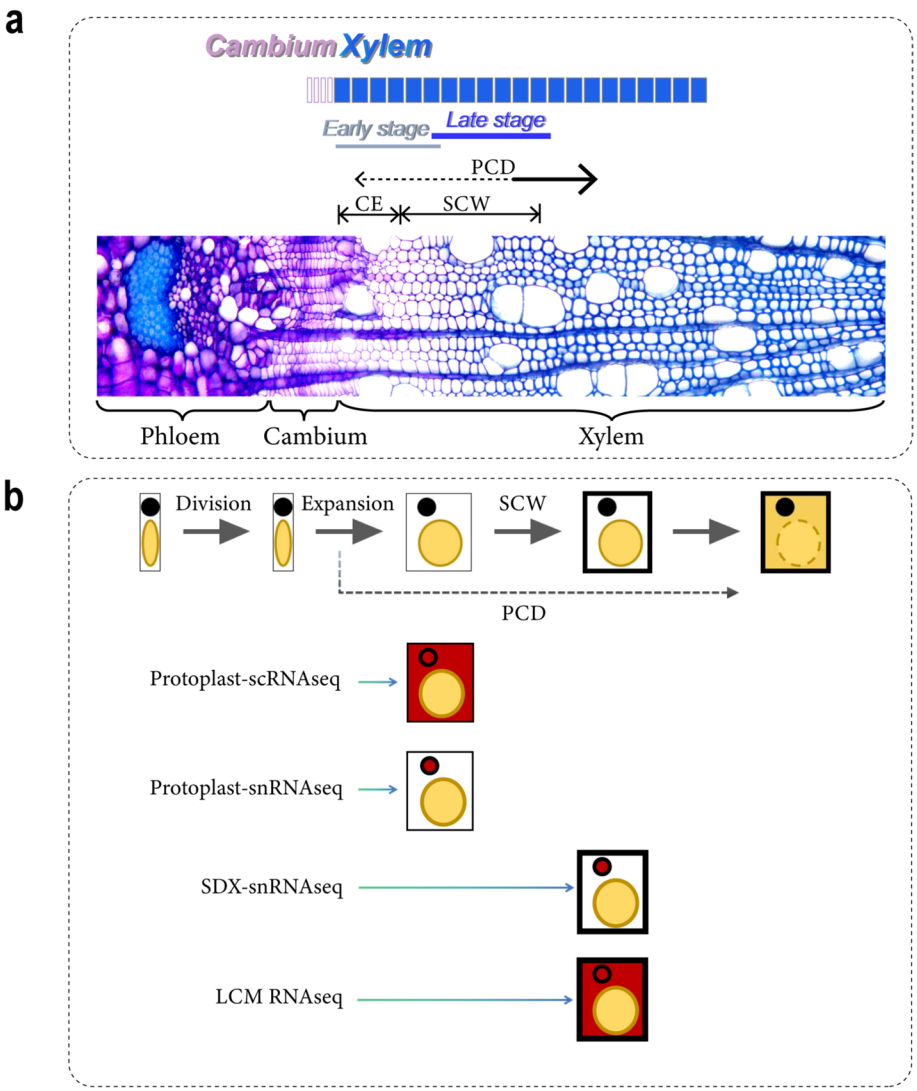

Fig. 7 Developmental zoning of xylem differentiation. **a** Spatiotemporal map of early and late xylem developmental stages. Key zones: CE (cell expansion), SCW (secondary cell wall deposition), and PCD (putative programmed cell death initiation). **b** Spatial alignment of transcriptional profiling techniques-protoplast-scRNAseq (early stages), protoplast-snRNAseq (early stages), SDX-snRNAseq (mid-late stages), and LCM RNAseq (terminal stages)-across the differentiation continuum. In the schematic diagram of xylem cells, the square represents the overall outline of a xylem cell. During late-stage xylem development, the vacuole ruptures and programmed cell death is completed. Within the square, the small black circles illustrate the nucleus, the large yellow circle represents the vacuole. Red color represents the content harvested for protoplast-scRNAseq, SDX-snRNAseq or LCM RNAseq. For example, protoplast-scRNAseq captures both nuclear and cytosolic RNA for sequencing

VASCULAR-RELATED NAC-DOMAIN6 (VND6) transcription factor, which is the master regulator of xylem differentiation and SCW biosynthesis, has been shown to directly activate *XCP* genes—key executors of vacuolar collapse during PCD [44–47]. A hallmark of xylem PCD is vacuolar collapse, which triggers the release of hydrolytic enzymes—including proteases, DNases, and RNases—into the cytoplasm, initiating autolytic degradation [10, 46, 48, 49]. For instance, mutations in vacuolar processing enzymes (VPEs) delay the clearance of cellular contents in fiber cells, concomitant with

abnormal SCW thickening, while loss of cysteine protease1 (CEP1) disrupts PCD progression in both fiber cells and vessels [50, 51]. Furthermore, mitochondria play a pivotal role in initiating plant PCD. Prior to vacuolar rupture, mitochondria release cytochrome c into the cytosol, a conserved signal amplifying PCD execution [43, 52–54]. In *Populus*, mitochondrial ascorbate peroxidase (PtomtAPX), which localizes to mitochondria under normal conditions, translocates to the cell wall following mitochondrial membrane permeabilization [55]. This spatial redistribution positions PtomtAPX to potentially mediate oxidative crosslinking during SCW lignification, thereby coupling PCD execution with wall reinforcement.

The dual regulatory roles of these genes strongly support the notion that the initiation of PCD is tightly coupled to the onset of SCW deposition, rather than occurring as strictly sequential events [9, 56, 57]. The co-expression of SCW-related biosynthetic genes and PCD-associated genes within the same transcriptomic space provides functional evidence that xylem cells undergo wall thickening and cell death programming in parallel, as part of a coordinated differentiation process [58, 59].

The timing of cell fate determination in fusiform cells—whether it is specified early in the initial cells or arises later during the differentiation of xylary derivatives—remains a subject of debate. One proposed model, based on the ubiquitous early expression of VNS family transcription factors in fusiform cells, suggests that all fusiform cells are initially primed toward vessel differentiation, but may later be redirected toward libriform fiber fate in response to external or positional cues [60–64]. For example, under mechanical stress, tension wood exhibits a dramatic increase in libriform fiber formation, supporting the idea of fate plasticity in response to environmental stimuli [65–67]. In our previous work, we hypothesized that early (the green cell cluster) and intermediate (brown) fusiform cells may possess the potential to transition into libriform fibers, based on their spatial proximity and transcriptomic similarities. However, due to a lack of direct supporting evidence at the time, this hypothesis was removed following reviewers' recommendations (see review history). In the present study, the expanded SDX-snRNAseq dataset provides compelling support for a broader version of this idea. As shown in the integrated UMAP, cells classified as early (green) and intermediate (brown) fusiform precursors are spatially adjacent to—and in some regions intermixed with—the SCW deposition zone. This spatial continuity suggests a developmental continuum in which multiple fusiform-derived populations may directly contribute to SCW-forming cells. These findings update our original interpretation: rather than specifically transitioning into libriform fibers, fusiform precursors may retain plasticity that allows direct differentiation into SCW-depositing cells. This observation aligns with a more flexible developmental model in which SCW deposition functions as a terminal program recruited from diverse upstream fates—including both vessel- and fiber-associated lineages—depending on positional cues or stress responses. The presence of shared SCW and PCD markers further supports the idea of a common transcriptional endpoint.

Conclusions

In this study, we integrated single-cell and single-nucleus RNA sequencing to construct a comprehensive developmental lineage of xylem development in *Populus* stems. By combining the high resolution of protoplast-scRNAseq for early-stage, thin-walled cells with the broader developmental coverage of SDX-snRNAseq for late-stage, thick-walled cells, we were able to capture the full trajectory of xylem differentiation—from cambial derivatives through SCW deposition and to PCD. This dual-platform strategy not only extended the developmental window that can be profiled, but also provided a practical solution to overcome the challenges of cell type annotation in plant single-cell transcriptomics. Further integration of protoplast-scRNAseq, SDX-snRNAseq and LCM RNAseq, including phloem tissue or using whole stem, is worth exploring to construct a complete developmental trajectory from cambium toward both xylem and phloem.

Methods

SnRNA-seq library construction and sequencing

Stem fragments below 15th internode from two 5-month-old *P. alba* × *P. glandulosa* cv. "84 K" trees were used as two biological replicates. After debarking, the tissues on the xylem side were scraped using a single-edged razor and frozen immediately in liquid nitrogen. Nuclei were prepared as described in a previous study [26], and further purified on the CytoFLEX SRT (Beckman, USA). For protoplast-snRNAseq, protoplasts were prepared from the SDX of two individual trees as we previously described [24]. Libraries were constructed using the Chromium Single Cell 3'Reagent Kits (V3.1) and sequenced on the Illumina NovaSeq 6000 platform at Gene Denovo (Guangzhou, China).

SnRNA-seq analysis

Raw reads of SDX-snRNAseq and protoplast-snRNAseq datasets were aligned to the *P. trichocarpa* reference genome using Cell Ranger (v7.2.0, 10 × Genomics) with default parameters. The resulting raw count matrix was then imported into R (v4.3.2) for further analysis using the Seurat (v5.0.1) package. Both datasets were processed through a standardized analytical pipeline encompassing dimensionality reduction, unsupervised clustering, and differential gene expression analysis, as previously described methodologies [24]. Integration of SDX-snRNAseq and protoplast-scRNAseq datasets was achieved through the identification of anchors using *K*-nearest neighbors with *K*.anchor = 5. The integrated UMAP was then used to compare the cell distribution patterns between SDX-snRNAseq and protoplast-scRNAseq datasets. MetaNeighbor analysis was performed to assess the reproducibility [68]. To validate cellular annotations, Pearson's correlation coefficients were computed between SDX-snRNA-seq-derived transcriptomes and LCM RNA-seq-generated cell-type reference profiles.

Cell density analysis of distribution overlapping cells between the three datasets was conducted using the minimal spanning tree (MST) method. Pseudotime lineage trajectory was implemented via Slingshot (v2.10.0), utilizing UMAP embeddings generated through Seurat-based data integration, following the computational framework established by Tung et al. [24]. Through correlation analysis with the outermost region isolated by LCM RNAseq, we clustered the cell populations in the SDX-snRNAseq data and identified distinct subpopulations. Differential expression

analysis (DEGs) was performed using the edgeR package with thresholds of $\log_2$-fold change ($\log_2$FC) ≥ 1 and a p-value < 0.05. GO annotation analysis was performed by extracting GO terms from the Arabidopsis TAIR database and comparing them with genes from the *P. trichocarpa* phytozome database [69, 70].

Histological analyses

Transverse stem Sects. (50 μm thick) were prepared from the 15th internode of 3-month-old *P. alba* × *P. glandulosa* hybrid plants using a vibratome (Leica VT1200S, Nussloch, Germany). Four distinct tissue treatments were applied: (1) intact stems, (2) debarked stems, (3) enzymatically digested stems (protoplast-isolated), and (4) blade-scraped stems. Similarly, stem samples subjected to the three experimental treatments were embedded in OCT (Sakura Finetek, USA) compound, flash-frozen in liquid nitrogen, and cryosectioned at −20 °C using a Leica CM1950 cryostat (Leica Biosystems, Germany). The sections were stained with 0.05% (w/v) toluidine blue (TBO) and observed under an Olympus BX53 light microscope (Olympus, Japan).

RNA in situ hybridization

RNA in situ hybridization was performed as previously described [71]. Stem fragments of the 13th internode were fixed in FAA (50% [v/v] ethanol, 5% [v/v] acetic acid, and 3.7% [v/v]), embedded in paraffin, and sectioned into 10 μm slices using a rotary microtome (Leica, RM2255). Sense and antisense RNA probes were synthesized by in vitro transcription using SP6 and T7 RNA polymerases from a transcription kit (Roche, 11,175,025,910), respectively. Hybridization was carried out with digoxigenin-labeled sense/antisense RNA probes targeting candidate genes, followed by incubation with anti-digoxigenin-AP conjugate (Roche, Germany) and chromogenic development using NBT/BCIP (Thermo, USA). Bright-field images were taken under a microscope (Olympus BX53, Japan). Primer sequences used for probe amplification are listed in Additional file 4: Supplementary Dataset S3.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-026-04007-z.

Additional file 1: Supplementary Figures S1-S6. Fig. S1. SDX-snRNAseq sequencing metrics for biological replicates. Fig. S2. Reproducibility between two replicates of SDX-snRNAseq. Fig. S3. Correlation map between LCM-captured libriform fiber, vessel element, ray parenchyma, outermost cell layer and two SDX-snRNAseq replicates. Fig. S4. RNA in situ hybridization images using sense probes for 14 genes. Fig. S5. Cluster number determination for integrated xylem transcriptomes. Fig. S6. *K*-means cluster robustness validation in subsampled protoplast-scRNAseq and SDX-snRNA-seq data.

Additional file 2: Supplementary Dataset S1. Number of UMIs per nucleus, summarized by barcode.

Additional file 3: Supplementary Dataset S2. Identification of differentially upregulated genes in SCW deposition region via SDX-snRNAseq.

Additional file 4: Supplementary Dataset S3. Primers used for RNA in situ hybridization in this study.

Acknowledgements
We thank Xiaojing Yan for helpful discussion.

Peer review information
Jixian Zhai and Wenjing She were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors' contributions
Q.L. and Y-C.L. conceived and designed the project. M.W, J-W.A.H. and J-F.D. analyzed the data. B.T. collected the samples. Y-C.L., M.W. and Q.L. wrote the manuscript. H.L. and R.R.S gave suggestions on the manuscript. All authors read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (32572084) and Zhejiang A & F University Scientific Research Development Fund (2024LFR088).

Data availability
The data generated in this study have been deposited in the Genome Sequence Archive under accession numbers CRA038282 (SDX-snRNAseq) [72] and CRA038285 (protoplast-snRNAseq) [73], at the National Genomics Data Center, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences [74, 75]. The previously published protopolast-scRNA-seq and LCM RNA-seq data from this article can be found in the NCBI GEO under accession number GSE180121 [31]. All custom code generated during this study is available in the Zenodo repository [76].

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]National Key Laboratory for Development and Utilization of Forest Food Resources, Zhejiang Key Laboratory of Forest Genetics and Breeding, International Research Center for Plant Cell Wall, College of Forestry and Biotechnology, Zhejiang A&F University, Hangzhou, China. [2]The Genome Center, University of California, Davis, CA, USA. [3]Department of Life Sciences, National Cheng Kung University, Tainan, Taiwan. [4]Graduate Institute of Intelligent Manufacturing Technology, National Taiwan University of Science and Technology, Taipei, Taiwan. [5]Institute of Industrial Crops, Heilongjiang Academy of Agricultural Sciences, Harbin, China. [6]College of Forestry, Inner Mongolia Agricultural University, Hohhot, China. [7]Forest Biotechnology Group, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, USA. [8]Department of Life Science, College of Life Science, National Taiwan University, Taipei, Taiwan. [9]Institute of Plant Biology, College of Life Science, National Taiwan University, Taipei, Taiwan.

Received: 15 May 2025 Accepted: 10 February 2026
Published online: 23 February 2026

References

1. Li P, He Y, Xiao L, Quan M, Gu M, Jin Z, et al. Temporal dynamics of genetic architecture governing leaf development in *Populus*. New Phytol. 2024;242:1113–30.
2. Wagner DE, Klein AM. Lineage tracing meets single-cell omics: opportunities and challenges. Nat Rev Genet. 2020;21:410–27.
3. Wang L, Wan M-C, Liao R-Y, Xu J, Xu Z-G, Xue H-C, et al. The maturation and aging trajectory of *Marchantia polymorpha* at single-cell resolution. Dev Cell. 2023;58:1429–44.e1426.
4. Wang H, Xu Y, Zhang Z, Zhang G, Tan C, Ye L. Development and application of transcriptomics technologies in plant science. Crop Design. 2024;3:100057.
5. Stuart T, Butler A, Hoffman P, Hafemeister C, Satija R. Comprehensive integration of single-cell data. Cell. 2019;177:1888-1902.e1821.
6. Saelens W, Cannoodt R, Todorov H, Saeys Y. A comparison of single-cell trajectory inference methods: towards more accurate and robust tools. Nat Biotechnol. 2018;37:547–54.
7. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Satija R. Integrated analysis of multimodal single-cell data. Cell. 2020;184:3573–87.
8. Turner S, Gallois P, Brown D. Tracheary element differentiation. Annu Rev Plant Biol. 2007;58:407–33.
9. Sauter JJ. Xylem: structure and function. In: Behnke HD, Esser K, Kubitzki K, Runge M, Ziegler H, editors. Progress in Botany; 1986//; Berlin, Heidelberg. Berlin Heidelberg: Springer; 1986. p. 388–405.
10. Fukuda H. Xylogenesis: initiation, progression, and cell death. Annu Rev Plant Physiol Plant Mol Biol. 1996;47:299–325.
11. Luo L, Li L. Molecular understanding of wood formation in trees. For Res. 2022;2:35–45.
12 Li W, Lin Y-C, Chen Y-L, Zhou C, Li S, De Ridder N, et al. Woody plant cell walls: fundamentals and utilization. Mol Plant. 2024;17:112–40.
13. Borthakur D, Busov V, Cao XH, Du Q, Gailing O, Isik F, et al. Current status and trends in forest genomics. For Res. 2022;2:99–120.
14. Shaw R, Tian X, Xu J. Single-cell transcriptome analysis in plants: advances and challenges. Mol Plant. 2021;14:115–26.

15. Liang S, Li Y, Chen Y, Huang H, Zhou R, Ma T. Application and prospects of single-cell and spatial omics technologies in woody plants. For Res. 2023;3:27.
16. Chen Y-L, Hsieh J-W, Kuo S-C, Kao C-T, Tung C-C, Yu J-H, et al. Merit of integrating in situ transcriptomics and anatomical information for cell annotation and lineage construction in single-cell analyses of *Populus*. Genome Biol. 2024;25:85.
17. Hsieh J-WA, Liou P-C, Lin C-C, Dai X, Hu C-W, Wang IF, et al. Single-cell and spatial multiomics identifies heterogeneous xylem development driven by mechanical stress in *Populus*. Dev Cell. 2025:s1543–5807(1525)00436–00438.
18. Shuai P, Hsieh J-WA, Kao C-T, Hu C-W, Wang R, Kuo S-C, et al. Single-cell and spatial omics reveal progressive loss of xylem developmental complexity across seed plants. Plant Cell. 2025;37:koaf253.
19 Yu JA-O, Hsieh JA, Wang Z, Wei J, Li QA-O, Chen YA-O, et al. Decoding xylem development in flowering plants: Insights from single-cell transcriptomics. Plant Cell Environ. 2025;48:8756–70.
20. Chen Y, Tong S, Jiang Y, Ai F, Feng Y, Zhang J, et al. Transcriptional landscape of highly lignified poplar stems at single-cell resolution. Genome Biol. 2021;22:319.
21. Li H, Dai X, Huang X, Xu M, Li Q. Single cell RNA sequencing reveals a high-resolution cell atlas of xylem in *Populus*. J Integr Plant Biol. 2021;63:1906–21.
22. Xie J, Li M, Zeng J, Li X, Zhang D. Single-cell RNA sequencing profiles of stem-differentiating xylem in poplar. Plant Biotechnol J. 2022;20:417–9.
23. Li R, Wang Z, Wang J-W, Li L. Combining single-cell RNA sequencing with spatial transcriptome analysis reveals dynamic molecular maps of cambium differentiation in the primary and secondary growth of trees. Plant Commun. 2023;4:100665.
24. Tung CC, Kuo SC, Yang CL, Yu JH, Huang CE, Liou PC, et al. Single-cell transcriptomics unveils xylem cell development and evolution. Genome Biol. 2023;24:3.
25. Xin X, Du F, Jiao Y. Plant nuclei isolation for single-nucleus RNA sequencing. Methods Mol Biol. 2023;2686:307–11.
26. Long Y, Liu Z, Jia J, Mo W, Fang L, Lu D, et al. FlsnRNA-seq: protoplasting-free full-length single-nucleus RNA profiling in plants. Genome Biol. 2021;22:66.
27. Wang K, Zhao C, Xiang S, Duan K, Chen X, Guo X, et al. An optimized FACS-free single-nucleus RNA sequencing (snRNA-seq) method for plant science research. Plant Sci. 2023;326:111535.
28. Lin YC, Li W, Chen H, Li Q, Sun YH, Shi R, et al. A simple improved-throughput xylem protoplast system for studying wood formation. Nat Protoc. 2014;9:2194–205.
29. Wang T, Wang F, Deng S, Wang K, Feng D, Xu F, et al. Single-cell transcriptomes reveal spatiotemporal heat stress response in maize roots. Nat Commun. 2025;16:177.
30. Liu Q, Ma W, Chen R, Li ST, Wang Q, Wei C, et al. Multiome in the same cell reveals the impact of osmotic stress on Arabidopsis root tip development at single-cell level. Adv Sci. 2024;11:e2308384.
31. Tung CC, Kuo SC, Yang CL, Yu JH, Huang CE, Liou PC, et al. Single-cell transcriptomics unveils xylem cell development and evolution: NCBI. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE180121. 2022.
32. Hsieh J-WA, Liou P-C, Lin C-C, Dai X, Hu C-W, Wang IF, et al. Single-cell and spatial multiomics identifies heterogeneous xylem development driven by mechanical stress in *Populus*: NCBI. Gene Expression Omnibus. 2025. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE295131.
33. Liu AC, Cheng CP. Pathogen-induced ERF68 regulates hypersensitive cell death in tomato. Mol Plant Pathol. 2017;18:1062–74.
34. Cao S, Guo M, Cheng J, Cheng H, Liu X, Ji H, et al. Aspartic proteases modulate programmed cell death and secondary cell wall synthesis during wood formation in poplar. J Exp Bot. 2022;73:6876–90.
35. Carlos Faro SG. Aspartic proteinase content of the Arabidopsis genome. Curr Protein Pept Sci. 2005;6:493–500.
36. Heyman J, Cools T, Canher B, Shavialenka S, Traas J, Vercauteren I, et al. The heterodimeric transcription factor complex ERF115-PAT1 grants regeneration competence. Nat Plants. 2016;2:16165.
37. Bucher M, Brunner S, Zimmermann P, Zardi GI, Amrhein N, Willmitzer L, et al. The expression of an extensin-like protein correlates with cellular tip growth in tomato. Plant Physiol. 2002;128:911–23.
38. Kirschner GK. A matter of life and death: networks regulating programmed cell death in Arabidopsis. Plant J. 2023;115:2.
39. Elsheery NI. Genome-wide characterization of aspartic protease (AP) gene family in *Populus trichocarpa* and identification of the potential PtAPs involved in wood formation. BMC Plant Biol. 2019;19:1–17.
40. Farmer A, Thibivilliers S, Ryu KH, Schiefelbein J, Libault M. Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in Arabidopsis roots at the single-cell level. Mol Plant. 2021;14:372–83.
41. Schmidt HW, Conde D, Pereira WJ, Triozzi PM, Balmant KM, Dervinis C, et al. Deep tissue profiling of *Populus* stem at single nucleus level reveals uncharacterized cell types and cell-specific gene regulatory networks. Genome Biol. 2025;26:258.
42. Gómez-Soto D, Pereira WJ, Piedrabuena-Díaz A, Dervinis C, Kirst M, Allona I, et al. Single-nucleus transcriptomics revealed auxin-driven mechanisms of wood plasticity to enhance severe drought tolerance in poplar. Genome Biol. 2025;26:312.
43. Turner SR, Hall M. The gapped xylem mutant identifies a common regulatory step in secondary cell wall deposition. Plant J. 2010;24:477–88.
44. Ohashi-Ito K, Fukuda OH. *Arabidopsis* VASCULAR-RELATED NAC-DOMAIN6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. Plant Cell. 2010;22:3461–73.
45. Zhong R, Lee C, Ye ZH. Global analysis of direct targets of secondary wall NAC master switches in *Arabidopsis*. Mol Plant. 2010;3:1087–103.
46. Funk V, Kositsup B, Zhao C, Beers EP. The *Arabidopsis* xylem peptidase XCP1 is a tracheary element vacuolar protein that may be a papain ortholog. Plant Physiol. 2002;128:84–94.
47. Avci U, Petzold HE, Ismail IO, Beers EP, Haigler CH. Cysteine proteases XCP1 and XCP2 aid micro-autolysis within the intact central vacuole during xylogenesis in *Arabidopsis* roots. Plant J. 2010;56:303–15.

48. Fukuda IH. ZEN1 is a key enzyme in the degradation of nuclear DNA during programmed cell death of tracheary elements. Plant Cell. 2002;14:3201–11.
49. Lehmann K, Hause B, Altmann D, Köck M. Tomato ribonuclease LX with the functional endoplasmic reticulum retention motif HDEF is expressed during programmed cell death processes, including xylem differentiation, germination, and senescence. Plant Physiol. 2001;127:436.
50. Cheng Z, Zhang J, Yin B, Liu Y, Wang B, Li H, et al. γVPE plays an important role in programmed cell death for xylem fiber cells by activating protease CEP1 maturation in *Arabidopsis thaliana*. Int J Biol Macromol. 2019;137:703–11.
51. Zhang D, Liu D, Lv X, Wang Y, Lu H. The cysteine protease CEP1, a key executor involved in tapetal programmed cell death, regulates pollen development in *Arabidopsis*. Plant Cell. 2014;26:2939.
52. Balk J, Chew SK, Leaver CJ, Mccabe PF. The intermembrane space of plant mitochondria contains a DNase activity that may be involved in programmed cell death. Plant J. 2003;34:573–83.
53. Yu XH, Perdue TD, Heimer YM, Jones AM. Mitochondrial involvement in tracheary element programmed cell death. Cell Death Differ. 2002;9:189–98.
54. Mitsuda N. The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. Plant Cell. 2005;17:2993–3006.
55. Zhang J, Liu Y, Li C, Yin B, Liu X, Guo X, et al. PtomtAPX is an autonomous lignification peroxidase during the earliest stage of secondary wall formation in *Populus tomentosa* Carr. Nat Plants. 2022;8:828–39.
56. Moreau C, Aksenov N, Lorenzo MG, Segerman B, Funk C, Nilsson P, et al. A genomic approach to investigate developmental cell death in woody tissues of *Populus* trees. Genome Biol. 2005;6:R34.
57. Locato V, Gara LD. Programmed cell death in plants: an overview. Methods Mol Biol. 2018;1743:1–8.
58. Zhao C, Avci U, Grant E, Haigler C, Beers E. XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulose synthesis and programmed cell death in xylem. Plant J. 2008;53:425–36.
59. Courtois-Moreau C, Pesquet E, Sjödin A, Muñiz L, Bollhöner B, Kaneda M, et al. A unique program for cell death in xylem fibers of *Populus* stem. Plant J. 2009;58:260–74.
60. Yamaguchi M, Goué N, Igarashi H, Ohtani M, Nakano Y, Mortimer J, et al. VASCULAR-RELATED nac-domain6 and VASCULAR-RELATED nac-domain7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. Plant Physiol. 2010;153:906–14.
61. Taylor-Teeples M, Lin L, de Lucas M, Turco G, Toal TW, Gaudinier A, et al. An Arabidopsis gene regulatory network for secondary cell wall synthesis. Nature. 2015;517:571–5.
62. Ramachandran P, Augstein F, Mazumdar S, Van Nguyen T, Minina EA, Melnyk CW, et al. Abscisic acid signaling activates distinct VND transcription factors to promote xylem differentiation in Arabidopsis. Curr Biol. 2024;31:3153-3161.e3155.
63. Xu B, Ohtani M, Yamaguchi M, Toyooka K, Wakazaki M, Sato M, et al. Contribution of NAC transcription factors to plant adaptation to land. Science. 2014;343:1505–8.
64. Zhong R, Ye Z-H. Regulation of cell wall biosynthesis. Curr Opin Plant Biol. 2007;10:564–72.
65. Liu B, Liu J, Yu J, Wang Z, Wang JP. Transcriptional reprogramming of xylem cell wall biosynthesis in tension wood. Plant Physiol. 2021;186:250–69.
66. Kerwin RE. Under pressure: transcriptional regulation of tension wood in *Populus trichocarpa* (California poplar). Plant Physiol. 2021;186:212–4.
67. Felten J, Sundberg B. Biology, chemistry and structure of tension wood. Springer, Berlin Heidelberg. 2013;20:203–24.
68. Fischer S, Crow M, Harris BD, Gillis J. Scaling up reproducible research for single-cell transcriptomics using MetaNeighbor. Nat Protoc. 2021;16:4031–67.
69. The Arabidopsis information resource. https://www.arabidopsis.org. Accessed 30 Jan 2021.
70. Phytozome v14. https://phytozome-next.jgi.doe.gov/. Accessed 18 Feb 2026.
71. Wang Z, Mao Y, Guo Y, Gao J, Liu X, Li S, et al. MYB transcription factor161 mediates feedback regulation of secondary wall-associated NAC-domain1 family genes for wood formation. Plant Physiol. 2021;184:1389–406.
72. Wei MK, Hsieh J-WA, Dang J-F, Tong BT, Li H, Sederoff R, et al. *Populus* developing xylem snRNA-seq dataset: China National Center for Bioinformation. 2026. https://ngdc.cncb.ac.cn/gsa/browse/CRA038282.
73. Wei MK, Hsieh J-WA, Dang J-F, Tong BT, Li H, Sederoff R, et al. Single-nucleus RNA sequencing (snRNA-seq) of protoplast-isolated nuclei from *Populus* developing xylem (SDX): China National Center for Bioinformation. 2026. https://ngdc.cncb.ac.cn/gsa/browse/CRA038285.
74. The GSA Family in 2025: A broadened sharing platform for multi-omics and multimodal data. Geno Proteomics Bioinform. 2025;23(4):qzaf072.
75. Database resources of the National Genomics Data Center. China National Center for Bioinformation in 2025. Nucleic Acids Res. 2025;53(D1):D30–44.
76. Wei MK, Hsieh J-WA, Dang J-F, Tong BT, Li H, Sederoff R, et al. Integrating scRNA-seq and snRNA-seq with spatial transcriptomics to unlock the xylem puzzle: Zenodo. 2025. https://doi.org/10.5281/zenodo.17095636.

Publisher's Note