


RESEARCH

Open Access

# Transcriptional landscape of highly lignified poplar stems at single-cell resolution



Yang Chen<sup>1†</sup>, Shaofei Tong<sup>1†</sup>, Yuanzhong Jiang<sup>1†</sup>, Fandi Ai<sup>1</sup>, Yanlin Feng<sup>1</sup>, Junlin Zhang<sup>1</sup>, Jue Gong<sup>1</sup>, Jiajia Qin<sup>1</sup>, Yuanyuan Zhang<sup>1</sup>, Yingying Zhu<sup>2</sup>, Jianquan Liu<sup>1,2</sup> and Tao Ma<sup>1\*</sup> 

\* Correspondence: [matao.yz@gmail.com](mailto:matao.yz@gmail.com)

<sup>†</sup>Yang Chen, Shaofei Tong and Yuanzhong Jiang contributed equally to this work.

<sup>1</sup>Key Laboratory of Bio-resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, China  
Full list of author information is available at the end of the article

## Abstract

**Background:** Plant secondary growth depends on the activity of the vascular cambium, which produces xylem and phloem. Wood derived from xylem is the most abundant form of biomass globally and has played key socio-economic and subsistence roles throughout human history. However, despite intensive study of vascular development, the full diversity of cell types and the gene networks engaged are still poorly understood.

**Results:** Here, we have applied an optimized protoplast isolation protocol and RNA sequencing to characterize the high-resolution single-cell transcriptional landscape of highly lignified poplar stems. We identify 20 putative cell clusters with a series of novel cluster-specific marker genes and find that these cells are highly heterogeneous based on the transcriptome. Analysis of these marker genes' expression dynamics enables reconstruction of the cell differentiation trajectories involved in phloem and xylem development. We find that different cell clusters exhibit distinct patterns of phytohormone responses and emphasize the use of our data to predict potential gene redundancy and identify candidate genes related to vascular development in trees.

**Conclusions:** These findings establish the transcriptional landscape of major cell types of poplar stems at single-cell resolution and provide a valuable resource for investigating basic principles of vascular cell specification and differentiation in trees.

## Background

The evolution of vascular tissue played key roles in plants' colonization and domination of terrestrial environments by enabling long-distance transport of water and nutrients and providing mechanical strength that supports their vertical growth [1, 2]. These tissues mainly consist of xylem and phloem cells, usually respectively produced by inward and outward division of cells in the vascular cambium, namely a process of secondary growth [3]. A typical feature of this process is thickening of the cell wall in the specialized xylem and phloem cell types, a continuous process of carbon accumulation and generation of the most abundant form of biomass in the world: wood [4]. In recent decades, there has been considerable progress in identifying key enzymes involved in cell wall biosynthesis



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

[5], and a hierarchical regulatory network involving several types of transcription factors that regulate cell specification and differentiation during vascular development has been proposed [6–8]. Those studies were mostly based on roots and inflorescence stems of the model plant *Arabidopsis* [9] and stems of the model tree poplar [10]. However, unlike the herbaceous annual *Arabidopsis*, perennial woody plants exhibit extreme secondary growth, characterized by formation of secondary cell walls (SCW) and additional dynamic seasonal changes influenced by various environmental stresses [11, 12]. Thus, partly due to reliance on information and markers from *Arabidopsis*, the molecular-level regulation of vascular tissue development in trees, especially the SCW regulatory network involved, has not been fully elucidated [10, 13].

Quantifying and comparing gene expression in specific cell types is essential for understanding the complex regulatory networks that control vascular development. By combining use of cell-type-specific reporter lines with cell sorting, laser capture microdissection, microarray analysis, and RNA sequencing, extensive studies have characterized the gene expression patterns of certain types of vascular cells in poplars, such as fusiform and ray cambial cells, fiber, and vessel cells [14–17]. Despite these studies, the full diversity of cell types and cell differentiation trajectories in the vascular tissues of woody plants have not yet been comprehensively analyzed, and their degree of heterogeneity is still unclear. However, advances in single-cell RNA sequencing (scRNA-seq) provide unprecedented opportunities for high-resolution characterization of gene expression [18–21]. The technique has been recently used to identify cell types in roots (*Arabidopsis* and rice), shoot apices (*Arabidopsis*, tomato and maize), leaves (*Arabidopsis* and peanut), and both ears and anthers of maize [22–36]. These studies have consistently revealed a wide range of cell heterogeneity and allowed the identification of rare and novel cell types, the characterization of multiple different cell types and states, and the establishment of detailed developmental trajectories during tissue development [20, 21]. However, the technique has not been applied in vascular tissue, especially tissues of woody plants, until its recent application in the differentiating xylem of *Populus alba* × *Populus glandulosa* [37]. This may be due to the presence of SCW with high and varying thickness, which greatly complicates isolation of individual cells without loss of their viability and purity.

In an attempt to circumvent these problems, here we applied an optimized protocol to isolate high-quality protoplasts, combined with 10x Genomic scRNA-seq technology to construct a single-cell transcriptional landscape of the poplar stem. In total, we identified 20 cell clusters covering nearly all major cell types in vascular tissue. As expected, we found a high degree of heterogeneity among these cells. In combination with in situ hybridization experiments, hundreds of cell-type-specific marker genes were identified. The differentiation trajectories of phloem and xylem cells were further reconstructed according to their transcriptome profiles. We also established a web server (<https://scu-populus.shinyapps.io/scRNAPal/>) to facilitate the use of our scRNA-seq data, which will provide a valuable resource for future functional studies of vascular development in trees.

## Results

### Construction of a single-cell transcriptional landscape of the poplar stem

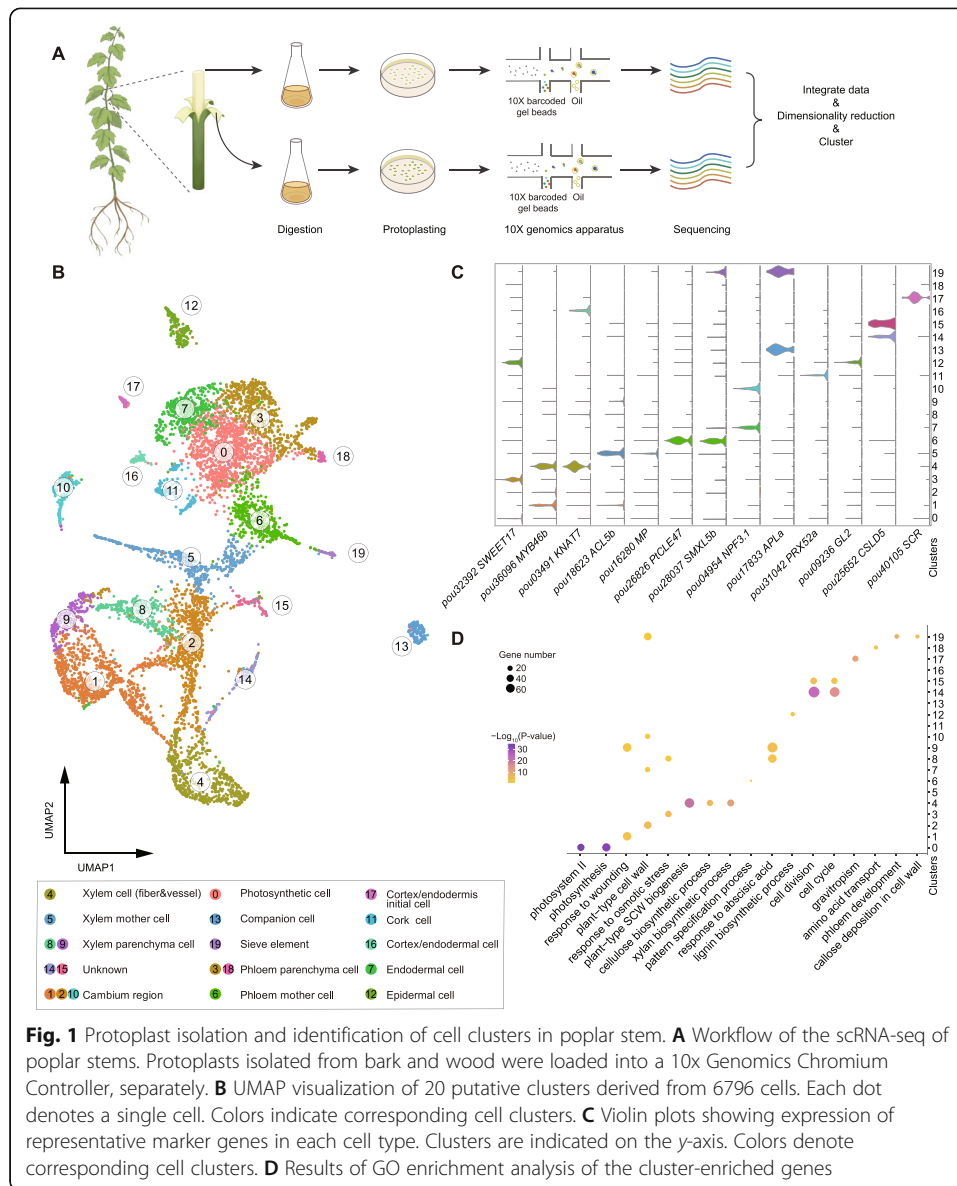
To enable scRNA-seq of poplar stem cells, we harvested stems of three 4-month-old *Populus alba* var. *pyramidalis* plantlets. We then separated bark and wood tissues of

parts below the third node (Fig. 1A) and obtained free protoplasts from them using a previously described digestion method [38]. The cross section of the stem showed that most of the cell types in the vascular tissue have been digested and released (Additional file 1: Fig. S1). As the isolated protoplasts were fragile and the suspension was rich in cell debris, we used an optimized method for washing and cell resuspension to maximize their purity and maintain their viability, before loading into a 10x Genomics Chromium Controller. Then, scRNA-seq libraries were constructed and sequenced using Illumina HiSeq 2000 platform. After prefiltering at both cell and gene levels, we successfully captured 3626 and 3170 cells, with median numbers of 2512 and 3330 expressed genes per cell for bark and wood tissues, respectively (Additional file 2: Table S1). In total, we detected transcripts of 27,297 and 26,653 genes in our scRNA-seq of bark and wood, respectively, covering 92.8% and 83.1% of the genes detected by bulk RNA-seq. In addition, gene expression values obtained from the combined scRNA-seq analysis showed strong correlations with the bulk RNA-seq profiles (Pearson's  $r$ : 0.81 and 0.82 for bark and wood respectively, Additional file 1: Fig. S2), confirming the robustness and high quality of our scRNA-seq data. To construct a comprehensive transcriptional landscape for poplar stem, we combined these two datasets (including 6796 high-quality cells in total) for subsequent analysis.

Using the 2000 genes that showed the highest variation in expression, we scaled and reduced the combined dataset into 30 principal components (PCs) with linear dimensional reduction. The transcriptome profiles of cells were then projected in an unsupervised manner without a priori knowledge of marker genes. In total, these cells were classified into 20 clusters and visualized using  $t$ -distributed stochastic neighborhood embedding (t-SNE) tool and uniform manifold approximation and projection (UMAP) algorithm (Fig. 1B, Additional file 1: Fig. S3-5). Among the detected genes, 6170 specifically expressed in one or two clusters were identified as cluster-enriched (Additional file 3: Table S2, Additional file 1: Fig. S6).

#### Identification of cell types with marker genes

Since very few marker genes were known for cell types of tree stems, we first compiled a list of known or predicted wood formation marker genes whose functions and expression patterns have been well studied in poplar and *Arabidopsis* (Additional file 4: Table S3, Additional file 5: Table S4), and then compared them with our identified cluster-enriched genes to annotate the clusters. Analysis of correlations between profiles derived from the scRNA-seq dataset and published high-spatial-resolution expression profiles for secondary growth of *P. tremula* [15] further validated the cell type of each cluster (Additional file 1: Fig. S7). Specifically, we found that several master regulatory genes related to xylem development including *SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN 1 (SND1)*, *NAC SECONDARY WALL THICKENING PROMOTING FACTOR 1 (NST1)*, and *MYB DOMAIN PROTEIN 46* and *83* [6] were highly enriched in clusters 4, 8, and 9 (Fig. 1C, Additional file 1: Fig. S8), and thus were identified as xylem cell populations. Moreover, many genes involved in SCW biosynthesis were predominantly expressed in cluster 4 (Fig. 1C, Additional file 1: Fig. S8), implying that cells in this



cluster may undergo SCW thickening. These genes included the following: the xylan biosynthesis and deposition genes *IRREGULAR XYLEM 9 (IRX9)*, *IRX14-L*, and *IRX15-L* [39, 40]; SCW cellulose and hemicellulose synthase genes *COBRA-LIKE4 (COBL4)*, *CesA4*, *CesA7*, and *CesA8* [41]; genes encoding laccases (*LAC4* and *LAC17*) localized to the thick SCW of xylem vessel elements and fibers [42]; and *KNOTTED-LIKE HOMEBOX OF ARABIDOPSIS THALIANA 7 (KNAT7)*, which negatively regulates SCW biosynthesis [43]. Furthermore, *VASCULAR RELATED NAC-DOMAIN PROTEIN 1 (VND1)*, a transcriptional regulator of SCW biosynthesis in xylem vessels [44], was also overrepresented in some cells of cluster 4, implying that vessel cells were contained in this cell cluster (Additional file 1: Fig. S8). In addition, *CINNAMYL ALCOHOL DEHYDROGENASE 7 (CAD7)*, whose expression was previously detected in xylem parenchyma cells to provide lignin precursors to the adjacent vessels and fibers [37, 45], was highly enriched in clusters

8 and 9, which were further identified as xylem parenchyma cells (Additional file 1: Fig. S8).

For phloem cells, we used *ALTERED PHLOEM DEVELOPMENT (APL)*, a master regulator gene that promotes asymmetric cell division of phloem initials into sieve elements (SEs) and companion cells (CCs) [46], and gene members of the PHLOEM PROTEIN 2 family (*PP2-A1*, *PP2-A4*, and *PP2-A10*) that is specifically expressed in SE-CC complex [47] as markers (Additional file 6: Table S5). We found that these marker genes were highly enriched in cells of clusters 13 and 19, suggesting that they are likely phloem cells (Fig. 1C, Additional file 1: Fig. S9). In addition, we also found that *SUCROSE-PROTON SYMPORTER 2 (SUC2)*, which is essential for sucrose uptake from apoplast to CCs [48], was highly expressed in cluster 13, while several homologous genes of *SIEVE-ELEMENT-OCCLUSION-RELATED 1 (SEOR1)*, which is necessary for the formation of phloem filament in *Arabidopsis* SEs [49], were particularly enriched in cluster 19 (Additional file 1: Fig. S9). Accordingly, the cells in clusters 13 and 19 were identified as CCs and SEs respectively. Recent studies have shown that several members of the UMAMIT amino acid transporter family are specifically expressed in phloem parenchyma (PP) cells of *Arabidopsis* leaves [27]. We found that several members of this gene family (*UMAMIT9*, *UMAMIT12*, *UMAMIT20*, *UMAMIT21*, and *UMAMIT22*) were highly enriched in clusters 3 and 18 (Additional file 1: Fig. S9). Thus, cells in these two clusters were identified as PP cells. Interestingly, we found that a homolog of *SWEET17*, which is specifically expressed in vascular parenchyma cells and acts as fructose transporters in the vasculature of *Arabidopsis* leaves [50], was also highly expressed in cells of cluster 3 (Additional file 1: Fig. S9), suggesting that it may play a role in controlling the transport of fructose in the phloem of poplar.

In order to identify other bark cell populations, the regulatory module composed of *SCARECROW (SCR)* and *SCARECROW-LIKE 23 (SCL23)* (Fig. 1C, Additional file 1: Fig. S10), which is required for controlling asymmetric cell division of the cortex/endodermis initial (CEI) and plays an important role in the endodermis development of *Arabidopsis* shoots and roots [51], was used as markers for the CEI cells (cluster 17). We also selected *NPF3.1 (NRT1/PTR FAMILY 3.1)*, which is specifically expressed in the endodermis [52, 53], and *CYCLIND 6;1 (CYCD6;1)* that is involved in cortex/endodermis asymmetric stem cell division [54], as markers for endodermal cells (cluster 7) and cortex/endodermal cells (cluster 16) respectively (Fig. 1C, Additional file 1: Fig. S10). As an important component of bark, cork is characterized by the presence of both lignin and suberin (a large aliphatic biopolymer composed of long-chain fatty acids and fatty alcohols of various lengths) in the cell wall [55]. We found that the suberin biosynthetic gene, *ASFT (ALIPHATIC SUBERIN FERULOYL TRANSFERASE)* [56, 57], and several genes involved in biosynthesis of lignin and very long-chain fatty acids, including *PEROXIDASE 52 (PRX52)*, *PRX72*, *3-KETOACYL-COA SYNTHASE 11 (KCS11)*, *LONG-CHAIN ACYL-COA SYNTHASE 1 (LACS1)*, and *LACS2* [58, 59], were specifically expressed in cluster 11 (Additional file 1: Fig. S10). Finally, the homeobox gene *GLABRA2 (GL2)*, required for epidermal cell differentiation in *Arabidopsis* roots [60], was highly expressed in cells of cluster 12 (Additional file 1: Fig. S10), which were therefore identified as epidermal cells.

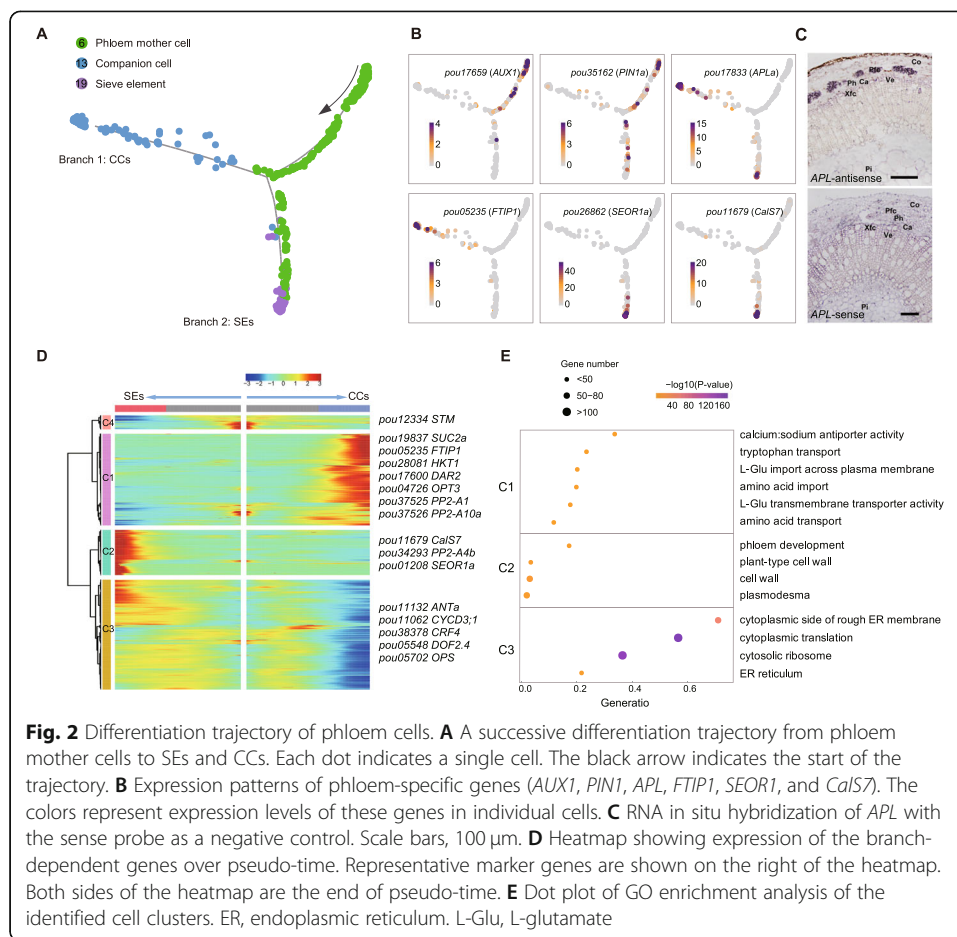
As markers for the cambium and adjacent cell populations, we used multiple genes including the key cambial regulators *WOX4 (WUSCHEL-RELATED HOMEBOX 4)*



and *PXY* (*PHLOEM INTERCALATED WITH XYLEM*), the auxin response transcription factor *MP* (*MONOPTEROS*), and genes encoding the auxin efflux transporter *PIN1* (*PINFORMED 1*), cytokinin-responsive transcription factor gene *ANT* (*AINTEGUMENTA*), and positive regulatory peptide of cambial activity *Populus CLV3/ESR-RELATED 47* (*PttCLE47*) [61, 62]. The results support the cambium and adjacent region identities of cell clusters 1, 2, 5, 6, and 10 (Fig. 1C, Additional file 1: Fig. S11). In more detail, the three *HD-ZIP III* transcription factor genes, homologs of which (*PtrHB4* and *PtrHB7*) play critical roles in regulation of vascular cambium activity and xylem differentiation in poplars [63, 64], were highly expressed in cluster 5 (Additional file 1: Fig. S11). In contrast, *OCTOPUS* (*OPS*), a gene encoding a polarly localized plasma membrane-associated protein that plays a crucial role in protophloem formation [65], and *SMXL5* (*SUPPRESSOR OF MAX2 1-LIKE5*), which is key promoter of phloem differentiation [66], were particularly enriched in cluster 6 (Fig. 1C, Additional file 1: Fig. S11). Accordingly, we designated the cells in clusters 5 and 6 as xylem and phloem mother cells, respectively. Due to the lack of proven marker genes, we could not determine the cell types in clusters 14 and 15. However, transcriptome profiles of the cells in both clusters were highly correlated with published expression profiles of phloem and expanding-xylem tissues [15] (Additional file 1: Fig. S7), indicating that they are probably spatially distributed near the cambium. Further examination of the enriched genes revealed that genes associated with mitosis and cell cycling were strongly expressed in clusters 14 and 15, such as *CSLD5* (*CELLULOSE SYNTHASE-LIKE D5*) [67], *PLE* (*PLEIADE*) [68], *HIK* (*HINKEL*) [69], mitotic cyclin (*CYCA1;1*, *CYCA2;1*, *CYCA2;4*, *CYCB1;4*, *CYCB2;1*, *CYCB2;3*, *CYCB2;4*, *CYCB3;1*) [70], and cyclin-dependent kinase (*CDKB1;2*, *CDKB2;1*) [71] (Fig. 1C, Additional file 1: Fig. S12), indicating that this cluster is rich in dividing cells. Overall, these results confirmed the high degree of cell heterogeneity in poplar stems.

#### Differentiation trajectory of phloem cells

As the major functional domains of stems, phloem and xylem are produced through periclinal cell divisions of vascular cambium toward the outside and inside, respectively. Since the cells in intermediate and terminal developmental states were captured simultaneously by scRNA-seq, we applied Monocle2 (v2.10.0) [72] to explore the continuous differentiation trajectories of the phloem and xylem development through pseudo-time analysis (Fig. 2, Fig. 3). To construct developmental trajectory of phloem cells, phloem mother cells (cluster 6), companion cells (cluster 13), and sieve elements (cluster 19) were selected. A subset of cells from cluster 6 (phloem mother cells) assembled at the beginning of pseudo-time and gradually bifurcated into two ends representing two distinct cell clusters (Fig. 2A). We found that *SMXL5* and *OPS* were prominently expressed at the branching point and neighboring cells (Additional file 1: Fig. S13), in agreement with their key roles in phloem formation and differentiation [65, 66]. Interestingly, the auxin influx and efflux transporters *AUX1* (*AUXIN RESISTANT 1*) and *PIN1*, which are polarly localized at opposite ends of protophloem cells of *Arabidopsis* root tips [73], were also extremely highly expressed around the branching point (Fig. 2B). This result clearly suggests that they may play an important role in establishment of the auxin gradient required for regulation of cambial activity and radial growth.



In accordance with above speculation, cells from clusters 13 and 19 were located along the two branches of pseudo-time respectively, reflecting transcriptional rewiring during phloem development (Fig. 2A). More specifically, the sucrose carrier *SUC2* [48] and an essential regulator of florigen transport from phloem CCs to SEs, *FTIP1* (*FT-INTERACTING PROTEIN 1*) [74], were highly expressed in branch 1 (Fig. 2B, Additional file 1: Fig. S13), supporting the assignment of these cells as phloem CCs. In agreement with CCs' function in providing metabolic support for SEs and promoting the loading and unloading of nutrients, genes related to "amino acid transport" were significantly enriched in this branch (Fig. 2D,E, Additional file 7: Table S6). We also found that several key strigolactone-responsive genes, including *SMXL6*, *SMXL7*, and *DWARF14* (*D14*) [75, 76], were prominently expressed in CCs (Additional file 1: Fig. S13), implying that this phytohormone and/or related signaling pathways may also be involved in differentiation of phloem cells. In contrast, the SEs marker gene *SEOR1* that is required for the formation of phloem filaments [49], and the phloem-specific callose synthase *CalS7* that is responsible for callose deposition in developing SEs [77], were enriched in branch 2 (Fig. 2B, Additional file 1: Fig. S13). The transcriptomic signature of branch 2 was initially enriched in "cytoplasmic translation," "rough endoplasmic reticulum," and "cytoplasmic translational elongation," and eventually enriched in "phloem development" and "plant-type cell wall" (Fig. 2E, Additional file 7: Table S6). These observations are consistent with SEs' differentiation process, which reportedly

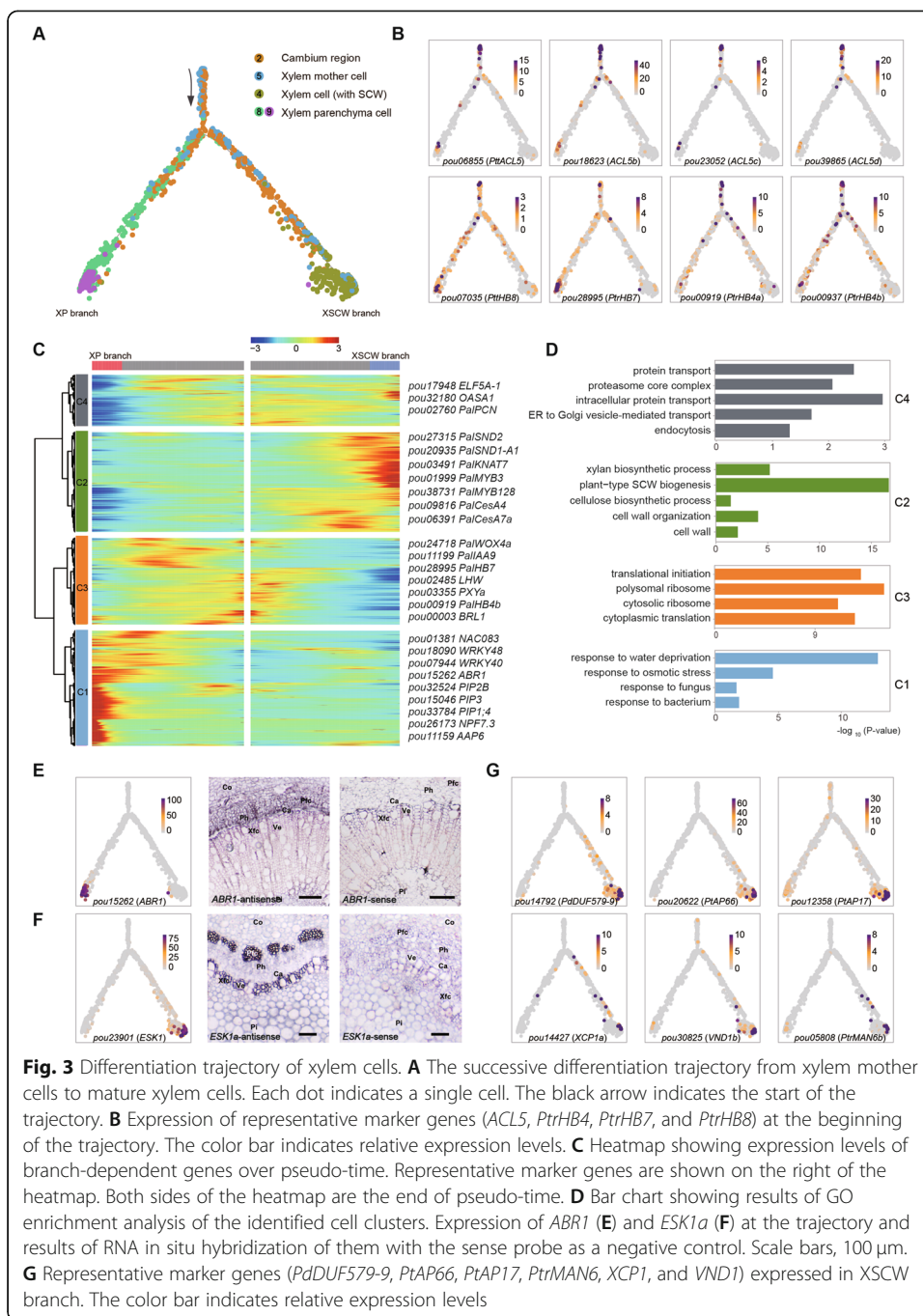
involves loss of ribosomes from cytoplasm, modification of endoplasmic reticulum, and cell wall thickening [78]. Interestingly, we noted that several genes homologous to *ENODL9* (*EARLY NODULIN-LIKE PROTEIN 9*), which encodes an extracellular ATP-binding protein attaching to the SE plasma membrane by a glycosylphosphatidylinositol anchor in *Arabidopsis* [79, 80], were also highly expressed in branch 2 (Additional file 1: Fig. S13), suggesting that they may play a role in SEs.

In addition, we found that many genes involved in phloem development, including the marker gene *APL* required for specific division of phloem cells [46], and *CHER1* (*CHOLINE TRANSPORTER-LIKE 1*) required for sieve plate and sieve pore development [81], were highly expressed in both branches (Fig. 2B, Additional file 1: Fig. S13). The *APL* expression was further validated by in situ hybridization assays (Fig. 2C), supporting the ordered developmental process of phloem cells. Interestingly, we found that different gene members of *PP2* gene family showed clear branch specificity in their expression patterns: *PP2-A1* and *PP2-A10* were preferentially expressed in branch 1, while *PP2-A4* was more preferentially expressed in branch 2 (Additional file 1: Fig. S13). Further molecular experiments are needed to verify their roles in the differentiation of SEs and CCs, but collectively, these results provide insights into the differentiation trajectory of phloem cells and transcriptome profiles during cell state transitions.

#### Differentiation trajectory of xylem cells

Construction of the developmental trajectory of xylem, using clusters 2, 4, 5, 8, and 9, resulted in a bifurcate pseudo-time backbone representing two distinct final states (Fig. 3A), with the clusters arranged at different branch sites. As expected, most cells from clusters 2 and 5 (cambium and xylem mother cells respectively) assembled at the beginning of pseudo-time, while xylem parenchyma cells (clusters 8 and 9) and xylem cells with SCW (cluster 4) grouped into different branches, namely XP and XSCW respectively. We found that multiple *HD-ZIP III* transcription factor genes (such as *PtrHB4*, *PtrHB7*, and *PtrHB8*, all of which play crucial roles in regulation of vascular cambium activity and xylem development in poplars [63, 64, 82]) were prominently expressed at the beginning of pseudo-time (Fig. 3B, Additional file 1: Fig. S14). Similarly, four homologs of the thermospermine synthase *ACAULISS* (*ACL5*), which forms a negative feedback loop with *HB8* that is required for proper vascular development and xylem cell specification [82], were prominently expressed around the branching point (Fig. 3B). We further investigated the differences in gene expression patterns between XP and XSCW branches (Fig. 3C). The results showed that the transcripts preferentially expressed in XP branch were initially enriched in functional categories related to “cytoplasmic translation” and eventually enriched in “response to various biotic and abiotic stresses” (Fig. 3D, Additional file 8: Table S7). Specifically, several genes encoding plasma membrane intrinsic protein (PIP) aquaporins, including *PIP1;4*, *PIP2B*, and *PIP3* were highly expressed in XP cells (Fig. 3C). In addition, we also found that multiple transcription factors associated with abiotic stress response, such as *WRKY40*, *WRKY48*, and *ABR1* (AP2-like ABA repressor 1), were highly enriched in XP cells. The expression of *ABR1* in the parenchyma cells was further validated by in situ hybridization assay (Fig. 3E). These characteristics are consistent with functions of the XP cells in carbohydrate and fat storage, water conduction, defense against pathogens, and healing and regeneration under stress conditions [83, 84].





In contrast, the XSCW branch is mainly composed of cells from cluster 4. In these cells, genes associated with “plant SCW biogenesis,” “xylan and cellulose biosynthetic process,” and “cell wall organization” were preferentially expressed (Fig. 3D), indicating that they were undergoing SCW thickening and lignification. In detail, multiple transcription factors in the core regulatory network of SCW formation, such as *SND1*, *SND2*, *KNAT7*, *MYB83*, *46*, and *103* [6], and genes encoding enzymes involved in SCW biosynthesis, such as SCW-associated cellulose synthase genes (*CesA4*, *CesA7*, and *CesA8*) [41], xylan biosynthesis genes (*IRX8*, *IRX9*, *IRX10*, *IRX14*, *IRX15*, and *IRX15-L*) [6], and lignin biosynthesis-related laccase (*LAC4* and *LAC17*) [42] were all

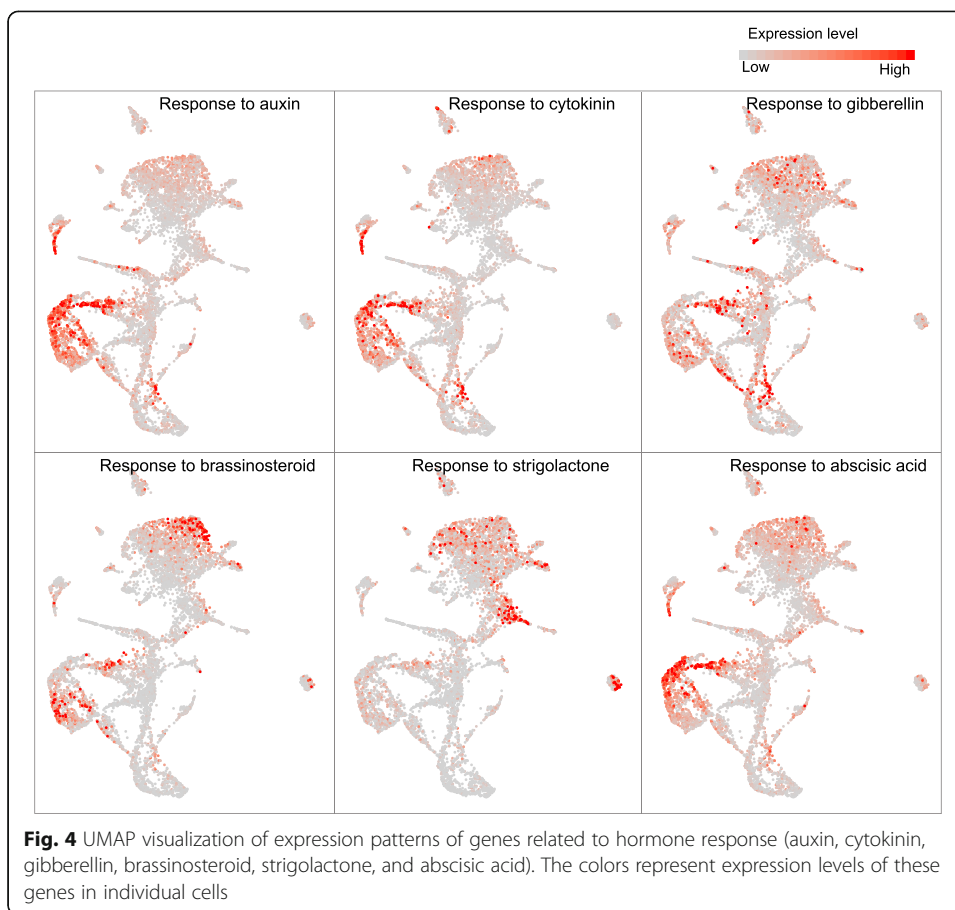
prominently expressed in these cells (Fig. 3C, Additional file 1: Fig. S15). These genes' expression patterns were strongly positively correlated, and binding motifs of MYB and NAC transcription factors are highly enriched in the 2-kb sequences upstream of them (Additional file 1: Fig. S16). These observations corroborate the hypothesis that members of the NAC and MYB gene families are master transcriptional regulators of SCW biosynthesis. Moreover, our results revealed high correlations in expression of homologs of *ESK1* (*ESKIMO1*), which plays an essential role in xylan acetylation during SCW biosynthesis in *Arabidopsis* [85], with these SCW-associated genes (Additional file 1: Fig. S17). The expression of *ESK1* in cells with SCW was also validated by in situ hybridization assays (Fig. 3F). Further analysis showed that multiple genes previously reported to be specifically expressed in poplar fiber cells, including *IRX15-L* (*PdDUF579-9*), *ASPARTIC PROTEASE 66* (*PtAP66*), and *PtAP17* [86–88], were particularly expressed in cells from XSCW branch (Fig. 3G). Interestingly, we found that a gene orthologous to *PtrMAN6* (*ENDO-BETA-MANNASE 6*), which is specifically expressed in poplar vessel cells [89], as well as *XCP1* (*XYLEM CYSTEINE PROTEASES 1*) and *XCP2*, both of which are reportedly vessel-specifically expressed in *Arabidopsis* [90], were also highly enriched in some cells associated with XSCW branch (Fig. 3G, Additional file 1: Fig. S15). These results imply that the XSCW branch is composed of at least two cell types, vessels and fibers. Taken together, our results clearly reveal a high degree of heterogeneity and complexity in the differentiation and fate determination of poplar xylem cells. They also confirm the robustness of our dataset and its potential utility for mining to detect new regulatory genes.

#### Hormonal biosynthesis and response patterns in poplar stem cells

Plant hormones play key, highly interactive roles in vascular tissue development and cambium maintenance [2, 61]. By plotting the spatiotemporal expression patterns of the genes related to hormone biosynthesis and response pathways, we next analyzed the cell-specific patterns of phytohormone regulation in poplar stem. We found that biosynthesis genes of most hormones, including gibberellin acid (GA), abscisic acid (ABA), and strigolactone (SL), were mainly overrepresented in cortex and endodermal cells, while auxin, cytokinin, and brassinosteroid (BR) biosynthesis genes were also enriched in xylem cells, especially parenchyma cells (Additional file 1: Fig. S18). Specifically, all of them were not expressed in cells with SCW. Like SL biosynthesis genes, SL response genes were overrepresented in the cortex and endodermal clusters, indicating that these cells have an autocrine function (Fig. 4). However, the other hormone response genes, especially genes responsive to auxin, cytokinin and GA, were overrepresented in cambium and adjacent cells, such as cells from clusters 1 and 10. In addition, auxin-, cytokinin-, and ABA-responsive genes were also clearly overrepresented in xylem parenchyma cells (clusters 8 and 9) (Fig. 4). Collectively, these results revealed complex patterns and interactions of various hormones during poplar vascular development and enabled investigation of potential hormonal signaling at single-cell resolution.

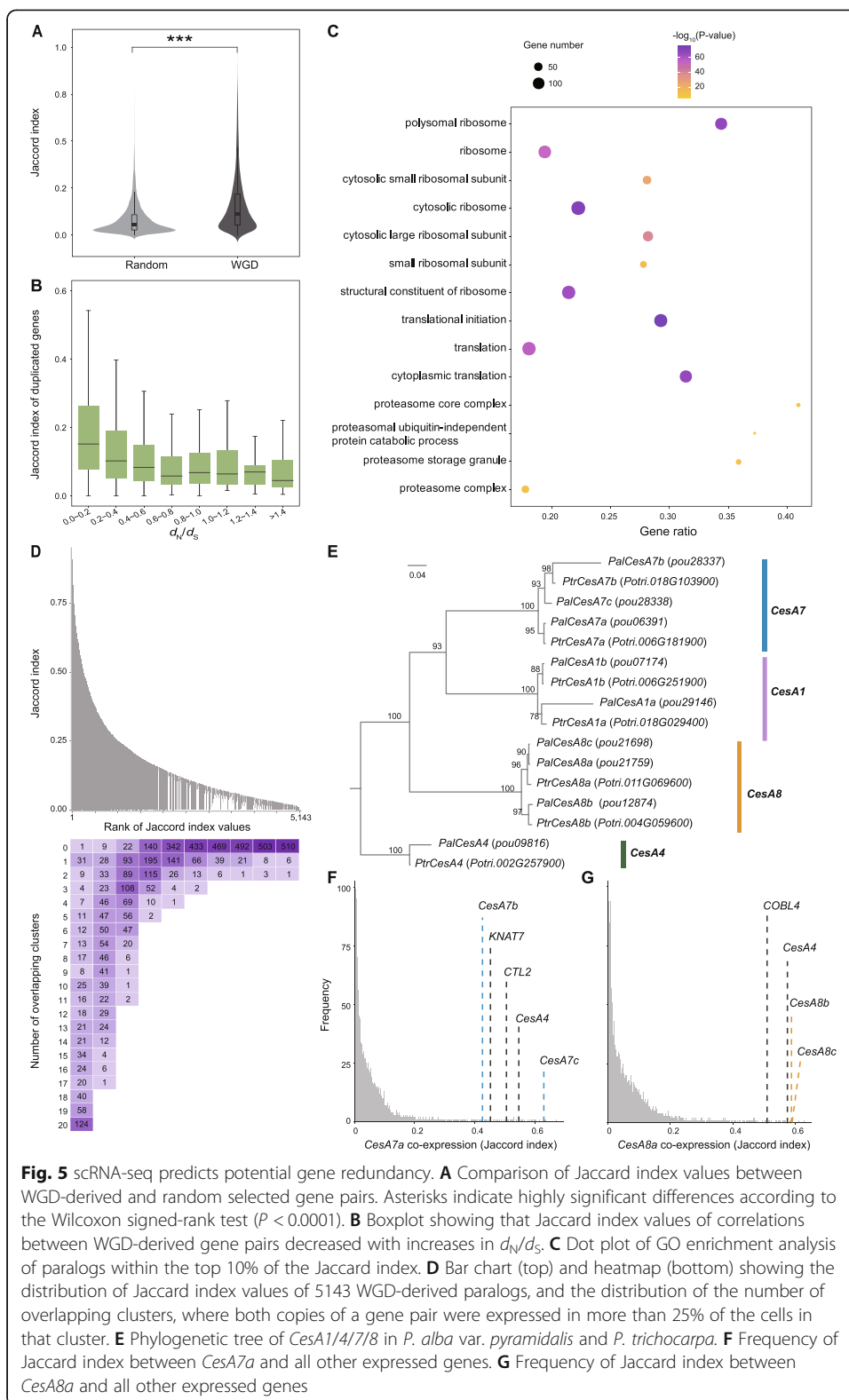
#### scRNA-seq predicts potential gene redundancy in poplar vascular development

A recent whole-genome duplication (WGD) resulted in two copies of each gene in the poplar genome [91, 92]. These paralogous genes usually exhibit functional redundancy,



and knocking out one of them will not result in a clear phenotypic change, which limits functional study of these genes. To predict functional redundancy in poplar vascular development, we first identified a total of 9147 paralogous gene pairs derived from WGD in the *P. alba* var. *pyramidalis* genome. Among these, both copies of 5143 gene pairs were detected to be expressed in at least 1% of the cells captured by our scRNA-seq. In-depth analysis showed that expression patterns of these WGD-derived paralogs were significantly more strongly correlated (higher Jaccard co-expression index values) than the genomic background ( $P$  value  $< 0.0001$ , Wilcoxon signed-rank test), and the correlation declined with increasing protein sequence differentiation (Fig. 5A,B), indicating that the functional redundancy between these paralogs is modulated by protein sequence evolution and expression pattern differentiation. Functional enrichment analysis revealed that the paralogs falling within the top 10% of the Jaccard index distribution were associated with basic cellular component and maintenance, such as “small/large ribosomal subunit,” “cytosolic ribosome,” “proteasome complex,” and “translation initiation” (Fig. 5C, Additional file 9: Table S8, Additional file 10: Table S9), which is consistent with the observation that the duplicated genes related to basic cellular machinery in *Arabidopsis* have a slower divergence rate at both sequence and expression levels [93].

In order to further investigate functional redundancy of the paralogs across cell clusters, we defined that if both copies of a gene pair were expressed in more than 25% of the cells in a certain cluster, they were considered to be overlapping in that cluster. Approximately 56.8% (2921) of the paralogs did not show any overlapping clusters, which is consistent



**Fig. 5** scRNA-seq predicts potential gene redundancy. **A** Comparison of Jaccard index values between WGD-derived and random selected gene pairs. Asterisks indicate highly significant differences according to the Wilcoxon signed-rank test ( $P < 0.0001$ ). **B** Boxplot showing that Jaccard index values of correlations between WGD-derived gene pairs decreased with increases in  $d_N/d_S$ . **C** Dot plot of GO enrichment analysis of paralogs within the top 10% of the Jaccard index. **D** Bar chart (top) and heatmap (bottom) showing the distribution of Jaccard index values of 5143 WGD-derived paralogs, and the distribution of the number of overlapping clusters, where both copies of a gene pair were expressed in more than 25% of the cells in that cluster. **E** Phylogenetic tree of *Cesa*A1/4/7/8 in *P. alba* var. *pyramidalis* and *P. trichocarpa*. **F** Frequency of Jaccard index between *Cesa7a* and all other expressed genes. **G** Frequency of Jaccard index between *Cesa8a* and all other expressed genes

with their low expression correlation (Fig. 5D). However, we found that many paralogs with overlapping clusters were also detected as cluster-enriched genes, especially in clusters 4, 9, 14, and 15 (Additional file 1: Fig. S19), indicating that functional redundancy

may be more frequent in these cell types. Consistent with the basic cellular functions of the paralogs in the top 10% of the Jaccard index, most of them exhibited overlapping expression patterns in almost all cell clusters (Fig. 5D). Despite this, we identified 40 gene pairs with the highest expression similarity but showing only one or two overlapping clusters, strongly suggesting that their functions are redundant in these cell types (Additional file 9: Table S8). Interestingly, we noted that most of these paralogs were specifically expressed in xylem cells with SCW (cluster 4), including *IQD10*, *MYB46*, *SND2*, *CesA7*, and *CesA8*. Previous studies have shown that both *CesA7* (*PtrCesA7A* and *7B*) and *CesA8* (*PtrCes8A* and *8B*) gene pairs in the *P. trichocarpa* genome are functionally redundant, so the *ptrcesa7a/b* or *ptrcesa8a/b* double mutants, but not single mutants, have lower cellulose contents than wild-type counterparts and abnormal morphology [94, 95]. In the *P. alba* var. *pyramidalis* genome, we detected three copies of both *CesA7* and *CesA8* (Fig. 5E). In addition to the paralogs derived from the WGD, there is another copy of both of these genes due to independent gene duplication events in this species. Our data revealed strong co-expression of the three copies of *CesA7* and three copies of *CesA8* (Fig. 5F,G), supporting that functional redundancy in poplar vascular development could be predicted by co-expression networks in scRNA-seq. Overall, these results highlight the utility of poplar stem scRNA-seq data for predicting gene redundancy.

## Discussion

scRNA-seq is a revolutionary technology with far greater power to identify new cell types and reveal cellular heterogeneity at high resolution than bulk RNA-seq [18, 19]. Recently, it has been widely used to elucidate cell differentiation and development processes of herbs or crops, including *Arabidopsis*, rice, maize, and tomato, particularly in young and/or easily dissociated tissues, such as root tips, shoot apices, and leaves [23, 26, 27]. However, similar studies on tree species or highly lignified plant tissues are still challenging, due to the presence of thick SCW that is difficult to digest to release single cells. However, the optimized protocol applied in this study enabled protoplast isolation and subsequent generation of a comprehensive single-cell transcriptional landscape of the poplar stem. We identified numerous cell clusters and state changes associated with vascular development, including xylem, phloem, cambium, epidermis, and endodermis cells (Fig. 1B). These results confirmed that our method is sensitive enough to detect most poplar genes and suggest that our optimized or further improved protocol can be potentially applied to more tree species or lignified plant tissues.

Nevertheless, it should be noted that processes involved in cell fate determination and pattern formation in poplar stems are still poorly understood, and some cell types were defined by marker genes that are either homologs of *Arabidopsis* genes or genes associated with related biological processes. Moreover, the cells in poplar vascular tissue are surrounded by cell walls with large variations in thickness and components, making it difficult to both obtain sufficient single cells and avoid changes in gene expression during protoplast isolation. For example, we have not detected a cell type or state in which genes related to programmed cell death are strongly expressed. This may be due to the thickened SCW of these cells hindering efficient isolation of protoplasts from them. Similarly, we found that only a few cork cells (cluster 11) were captured by our scRNA-seq (Additional file 1: Fig. S1), which may be due to the presence of suberin in their cell wall that can strongly affect the release of protoplasts [96]. In addition,



although we have excluded the effect of the cell-cycle genes before cell clustering, we still found that the genes involved in cell cycling and division were specifically expressed in clusters 14 and 15, without any other confirmed marker genes to determine their cell types (Fig. 1D, Additional file 1: Fig. S12). Therefore, further improvement of the protocol for preparing protoplasts from these cell types is needed, and future results of scRNA-seq analysis must be carefully verified using other techniques, such as promoter reporter analysis and cell lineage tracing. Alternatively, application of single nucleus RNA sequencing and/or spatial transcriptomics technologies to vascular tissue may address these issues [97, 98].

Our results clearly confirm the power of scRNA-seq to reveal cellular heterogeneity, discover new cell types, and characterize cell states along the developmental trajectory of poplar stems that would be missed by lower resolution sequencing. For example, our scRNA-seq data revealed two subtypes of xylem parenchyma cells and their state transitions along the differentiation trajectory of xylem cells (Fig. 3). Our results also revealed that the cells within and/or around the vascular cambium are highly heterogeneous, although we cannot faithfully construct a continuous differentiation trajectory of these cells, possibly due to the relatively low representation of procambium and precursor cells, or failure of our current protocol to capture some intermediate cell types. Nevertheless, the identified multiple cell-type-specific marker genes (Additional file 6: Table S5), together with the gene expression dynamics along the developmental trajectories of phloem and xylem cells, provide insights with unprecedented resolution into vascular cell differentiation in poplar stems. Finally, co-expression networks derived from scRNA-seq data such as the set obtained in this study have clear utility for identifying new regulatory genes and predicting the functional redundancy of homologous genes. In the future, molecular functional verification of these candidate markers will help to understand the cell differentiation and fate determination during the vascular development of trees at single-cell resolution.

## Conclusions

In this study, we applied scRNA-seq technology to highly lignified poplar stems and found that they are composed of highly heterogeneous cells. We identified a total of 20 putative cell clusters with a series of novel cell-type-specific marker genes, whose expression dynamics were further used to reconstruct the cell differentiation trajectories involved in phloem and xylem development. We also investigated the complex hormonal response patterns of these cell types during poplar vascular development, and emphasized the use of our data to predict potential gene redundancy and identify candidate genes related to vascular development in trees. In conclusion, our research established a transcriptional landscape of the major cell types of poplar stems at single-cell resolution and provided a valuable resource for studying the basic principles of vascular cell specification and differentiation in trees.

## Methods

### Poplar growth conditions and protoplast isolation

*Populus alba* var. *pyramidalis* plantlets were grown in a greenhouse with a photoperiod of 16 h light/8 h darkness and 60% humidity. The stems below the third internode of

three 4-month-old poplars were harvested, separated into bark and wood, and cut into 1-cm segments. The segments from bark and wood tissues were mixed separately, and then respectively submerged in cell wall digestion solution consisting of 1.5% (wt/vol) cellulase R-10, 0.4% (wt/vol) macerozyme R-10, 0.5 M mannitol, 20 mM KCl, 10 mM CaCl<sub>2</sub>, and 0.1% (wt/vol) bovine serum albumin [38]. And the mixtures were placed in a shaker at 50–55 rpm for 2.5 h in the dark at room temperature to release protoplasts respectively. The protoplasts were filtered with a 40- $\mu$ m cell strainer to eliminate any clumped cells, and collected by centrifugation at 100 $\times$ g for 4 min. Since the protoplasts were very fragile in the washing buffers recommended by 10x Genomics, we used W5 solution (5 mM glucose, 2 mM MES (pH 5.7), 154 mM NaCl, 125 mM CaCl<sub>2</sub> and 5 mM KCl) to gently wash and resuspend the protoplasts three times. After removing the supernatant, the protoplasts were resuspended by adding a small amount of digestion solution without enzyme to avoid the influence of excessive calcium on the reverse transcription reaction. The concentration and viability of the protoplasts were estimated by trypan blue staining, and high-quality protoplasts with viability  $\geq$  75% were immediately processed with the 10x Genomics Single Cell Protocol (CG00052, RevC).

#### scRNA-seq library construction and sequencing

The scRNA-seq libraries were constructed using the 10x Chromium Single Cell 3' Platform according to the manufacturer's instructions (10x Genomics). Briefly, the protoplasts isolated from bark or wood tissues were loaded into a Chromium microfluidic chip with 3' v2 chemistry and barcoded with a 10x Chromium Controller to generate single cell GEMs (gel beads in emulsion). scRNA-seq libraries were subsequently prepared using a Chromium Single Cell 3' reagent kit v2 (10x Genomics, Pleasanton, CA). The DNA libraries were qualitatively analyzed using an Agilent 2100 Bioanalyzer, and paired-end reads were produced by an Illumina HiSeq 2000 platform.

#### scRNA-seq data pre-processing

The raw scRNA-seq reads from wood and bark tissues were separately aligned to the *P. alba* var. *pyramidalis* reference genome [99], using the Cell Ranger pipelines (<https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger>) with default parameters (v2.0, 10x Genomics), and expression matrices for each gene and each cell were generated. The gene-cell matrices were then loaded into the Seurat package (v3.1.0) for further analysis [100]. To remove the low-quality genes and cells, only the genes that were expressed in at least three cells were considered, and we filtered the cells with expressed genes over 9000 or fewer than 200. Cells with less than 500 or more than 70,000 unique molecular identifiers (UMIs) were also filtered out, as well as cells with more than 5% mitochondrial genes. In addition, doublets (two or multiple cells in one oil droplet) in each scRNA-seq dataset were detected with DoubletFinder (v2.0.3) [101] using the number of artificial doublets (pN) of 0.25. To identify the optimal neighborhood size (pK), the function "paramSweep\_v3" was executed using parameter "PCs = 1:20" and the maximum of pK value was selected as an optimal pK parameter. For the estimate of the number of expected real doublets (nExp), the doublet formation rate was assumed as 7.5% and nExp value was then adjusted according to homotypic doublet

proportion. Finally, the parameters of nExp, pN, and pK were set to “223, 0.25, 0.005” and “265, 0.25, 0.3” for wood and bark dataset respectively, and only cells marked with “Singlets” were retained for further analysis. After filtering, 3626 and 3170 cells from bark and wood tissues, respectively, remained in the matrix. We integrated these two scRNA-seq datasets into an experiment-wide feature-barcode matrix for subsequent analysis, using combined canonical correlation analysis (CCA) [102] and mutual nearest neighbors (MNN) [103] methods in the Seurat package (v3.1.0) [100].

#### Cell clustering and identification of cluster-enriched genes

The scaled gene-cell matrix obtained as described above was normalized by “LogNormalize” in Seurat (v3.1.0) [100]. To mitigate the effects of protoplasting and cell-cycle heterogeneity on cell clustering [26], the cell cycle score and the proportion of protoplasting genes in each cell were calculated by the “CellCycleScoring” and “PercentageFeatureSet” function provided in Seurat. The cell-cycle and protoplasting genes were identified as the closest homologous genes to *Arabidopsis*, which are provided by two published papers [104, 105]. The variations resulted from cell-cycle and protoplasting genes were regressed out by the “ScaleData” function using “vars.to.regress.” Then, we identified the highly variable genes using “FindVariableFeatures” with “selection.method = “vst”” parameter, and the top 2000 genes were used for linear dimensionality reduction by principal component analysis (PCA). The first 30 PCs were used as inputs for a graph-based approach to identify cell clusters [106], which were further visualized and explored by *t*-SNE (with parameter “dims = 20”) [107] and UMAP (with parameter “umap.method = “umap-learn””) [108]. We also identified cluster-enriched genes using “FindAllMarkers” with parameters “min.pct = 0.25” and “logfc.threshold = 0.58” implemented in Seurat (v3.1.0) [100]. Cluster-specific marker genes were finally selected from these cluster-enriched genes that were expressed in at least 25% of the cells in a cluster, but less than 25% of the cells in all other clusters.

#### Sequencing and analysis of bulk RNA-seq

The stems below the third internode of a 4-month-old poplar were harvested, separated into bark and wood, and frozen immediately for bulk RNA sequencing. Briefly, total RNA extracted from each tissue was used for library construction and deep sequencing by BGISEQ-500 (BGI-Shenzhen, China). The obtained clean reads were aligned to the reference genome [99] using HISAT2 (v2.1.0) [109], and then the expression levels were calculated and normalized by StringTie (v1.3.3b) [110]. Finally, the Pearson correlations between bulk RNA-seq and scRNA-seq datasets were calculated in R. In addition, the transcriptome data of cryosections from the secondary phloem, vascular cambium, and wood-forming tissues of *P. tremula* were downloaded [15]. MScanX [111] software was used to identify 17,403 orthologs between *P. tremula* and *P. alba* var. *pyramidalis*. Then Pearson correlations between gene expression levels of our scRNA-seq and the RNA-seq of cryosections were calculated.

#### RNA in situ hybridization

We conducted mRNA in situ hybridization as previously described [112]. Briefly, fifth internode of 4-month-old poplar were collected and immediately fixed in FAA (10%

formaldehyde, 5% acetic acid, 47.5% ethanol) for 48 h. The fixed tissue was dehydrated through a graded alcohol series (70, 85, 90, 95, and 100%) and a HistoClear series, then embedded in paraplast. Then, 10- $\mu$ m sections were cut and mounted on Pro-beOn Plus Slides (Fisher Scientific, 22-230-900). *APL*, *ABRI*, and *ESK1a* probes were labeled using a DIG RNA Labeling Kit (Roche). Primer sequences for all genes were listed in Additional file 11: Table S10. And then hybridization and detection of hybridization signals were performed as previously described [112] using a DIG Nucleic Acid Detection Kit (Roche).

### Pseudo-time analysis

To explore the developmental trajectories of specific cell clusters, the R package Monocle2 (v2.10.0) [72] was used. Briefly, the subset of raw data from target cell clusters were first extracted, and the variance in each gene's expression across cells was calculated using "dispersionTable." Variable genes were then determined based on average expression level (with parameters "mean\_expression > = 0.1" and "dispersion\_empirical > = 1 \* dispersion\_fit") to define a developmental progress. The dimensionality of the data was subsequently reduced to two components using the "DDRTree" method, and the trajectory was inferred with the "orderCells" function. The pseudo-time trajectory was then plotted using the "plot\_cell\_trajectory" function in Monocle2 (v2.10.0) [72]. Finally, we used "BEAM" to identify the branch-dependent or pseudo-time-dependent genes, which were visualized by the "plot\_genes\_branched\_heatmap" function. The dynamical expression of genes along the pseudo-time were visualized by the "plot\_pseudotime\_heatmap" function in Monocle2 (v2.10.0) [72].

### Functional and motif enrichment analysis

We performed Gene Ontology (GO) enrichment analysis to detect and evaluate cluster-specific or branch-dependent genes with an in-house perl script ([https://github.com/wk8910/bio\\_tools/blob/master/15.GO\\_analysis/02.go\\_enrichment/](https://github.com/wk8910/bio_tools/blob/master/15.GO_analysis/02.go_enrichment/)). For GO enrichment, all genes, bark-expressed genes, and wood-expressed genes were selected as background sets in cluster-specific (Fig. 1D), phloem-branch-dependent (Fig. 2E), and xylem-branch-dependent (Fig. 3D) analysis, respectively. *P* value was obtained by Fisher's exact test and adjusted by Benjamini-Hochberg false discovery rate. Motif enrichment analysis of interesting gene sets was performed using a built-in script of Homer (findmotif.pl) [113], with promoters of wood-expressed genes as the background and "*Arabidopsis*" as "Promoter Sets."

### scRNA-seq co-expression analysis

To perform co-expression analysis using our scRNA-seq data, Pearson's correlation coefficient between the expression level of each gene pair was calculated to identify genes with similar expression patterns. And the gene correlation network was displayed using Cytoscape (v3.7.1) [114]. In addition, we calculated Jaccard index value (also known as Jaccard similarity coefficient) for each pair of genes expressed in at least 1% of the captured cells [28]. For each gene, the cells with at least 1 UMI were assigned a value of 1,

and all other cells were assigned a value of 0. To explore functional redundancy between paralogs, we first identified WGD-derived paralogs as previously described [99], and filtered the gene pairs if one or both copies lacked UMI values. In total, 5143 WGD-derived paralogs were finally retained to compare their Jaccard index values. The non-synonymous ( $d_N$ ) and synonymous ( $d_S$ ) substitution rate and their ratio were estimated using the “yn00” function in PAML (v4.9e) [115].

**Review history** The review history is available as Additional file 12.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-021-02537-2>.

**Additional file 1: Fig. S1-S19.** Supplementary figure legends and supplementary figures (Fig. S1-S19).  
**Additional file 2: Table S1.** Summary statistics of scRNA-seq data for bark and wood tissues.  
**Additional file 3: Table S2.** List of the cluster-enriched genes.  
**Additional file 4: Table S3.** List of genes that have been functionally studied in poplars.  
**Additional file 5: Table S4.** Gene list used for cell-type identification in this study.  
**Additional file 6: Table S5.** Cluster-specific marker genes.  
**Additional file 7: Table S6.** Results of GO enrichment analysis for genes in Fig. 2D.  
**Additional file 8: Table S7.** Results of GO enrichment analysis for gene clusters in Fig. 3C.  
**Additional file 9: Table S8.** Information of the WGD-derived gene pairs falling within the top 10% of the Jaccard index distribution.  
**Additional file 10: Table S9.** GO enrichment analysis for paralogs falling within the top10% of the Jaccard index distribution.  
**Additional file 11: Table S10.** Primers used for RNA in situ hybridization in this study.  
**Additional file 12.** Review history.

### Peer review information

Wenjing She was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

### Authors' contributions

T.M. conceived the project. Y.J., Y. Zhu, and J.L. made suggestions for the project. J.Q. and Y. Zhang prepared the materials for protoplasts. S.T., Y.C., and J.Z. prepared protoplasts. Y.C. analyzed the data. F.A. and Y.F. participated in the figures' design and visualization. S.T. and J.G. performed the experiments of RNA in situ hybridization. T.M., Y.J., and Y.C. wrote the paper. All authors discussed the results and commented on the manuscript. The authors read and approved the final manuscript.

### Funding

This work was supported by the National Key Research and Development Program of China (2021YFD2201100), National Natural Science Foundation of China (31922061, 41871044, 31561123001, 31500502), and Fundamental Research Funds for the Central Universities (SCU2021D006 and 2020SCUNL103).

### Availability of data and materials

All sequencing data generated in this study have been submitted to the National Genomics Data Center (NGDC; <https://bigd.big.ac.cn/bioproject>) under BioProject accession number PRJCA005543 [116]. A supplementary online web server (<https://scu-populus.shinyapps.io/scrNAPal/>) was also developed to facilitate the use of our datasets.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.



**Author details**

<sup>1</sup>Key Laboratory of Bio-resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, China. <sup>2</sup>State Key Laboratory of Grassland Agro-Ecosystem, Institute of Innovation Ecology, Lanzhou University, Lanzhou, China.

Received: 21 July 2021 Accepted: 10 November 2021

Published online: 22 November 2021

**References**

- Lucas WJ, Groover A, Lichtenberger R, Furuta K, Yadav SR, Helariutta Y, et al. The plant vascular system: evolution, development and functions. *J Integr Plant Biol.* 2013;55(4):294–388. <https://doi.org/10.1111/jipb.12041>.
- Miyashima S, Sebastian J, Lee JY, Helariutta Y. Stem cell function during plant vascular development. *EMBO J.* 2013;32(2):178–93. <https://doi.org/10.1038/emboj.2012.301>.
- De Rybel B, Mähönen AP, Helariutta Y, Weijers D. Plant vascular development: from early specification to differentiation. *Nat Rev Mol Cell Biol.* 2016;17(1):30–40. <https://doi.org/10.1038/nrm.2015.6>.
- Bar-On YM, Phillips R, Milo R. The biomass distribution on Earth. *PNAS.* 2018;115(25):6506–11. <https://doi.org/10.1073/pnas.1711842115>.
- Meents MJ, Watanabe Y, Samuels AL. The cell biology of secondary cell wall biosynthesis. *Ann Bot.* 2018;121(6):1107–25. <https://doi.org/10.1093/aob/mcy005>.
- Schuetz M, Smith R, Ellis B. Xylem tissue specification, patterning, and differentiation mechanisms. *J Exp Bot.* 2013;64(1):11–31. <https://doi.org/10.1093/jxb/ers287>.
- Wang H-Z, Dixon RA. On-off switches for secondary cell wall biosynthesis. *Mol Plant.* 2012;5(2):297–303. <https://doi.org/10.1093/mp/srr098>.
- Zhong R, McCarthy RL, Haghight M, Ye Z-H. The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PLoS One.* 2013;8(7):e69219. <https://doi.org/10.1371/journal.pone.0069219>.
- Zhang J, Elo A, Helariutta Y. *Arabidopsis* as a model for wood formation. *Curr Opin Biotechnol.* 2011;22(2):293–9. <https://doi.org/10.1016/j.copbio.2010.11.008>.
- Zhang J, Xie M, Tuskan GA, Muchero W, Chen J-G. Recent advances in the transcriptional regulation of secondary cell wall biosynthesis in the woody plants. *Front Plant Sci.* 2018;9:1535. <https://doi.org/10.3389/fpls.2018.01535>.
- Zhao C, He Y, Yu Y, Zhou M, Zhao L, Xia X, et al. Transcriptomic analysis of seasonal gene expression and regulation during xylem development in “Shanxin” hybrid Poplar (*Populus davidiana* × *Populus bolleana*). *Forests.* 2021;12(4):451. <https://doi.org/10.3390/f12040451>.
- Arend M, Fromm J. Seasonal change in the drought response of wood cell development in poplar. *Tree Physiol.* 2007;27(7):985–92. <https://doi.org/10.1093/treephys/27.7.985>.
- Camargo EL, Ployet R, Cassan-Wang H, Mounet F, Grima-Pettenati J. Digging in wood: new insights in the regulation of wood formation in tree species. *Adv Bot Res.* 2019;89:201–33. <https://doi.org/10.1016/bs.abr.2018.11.007>.
- Shi R, Wang JP, Lin Y-C, Li Q, Sun Y-H, Chen H, et al. Tissue and cell-type co-expression networks of transcription factors and wood component genes in *Populus trichocarpa*. *Planta.* 2017;245(5):927–38. <https://doi.org/10.1007/s00425-016-2640-1>.
- Sundell D, Street NR, Kumar M, Mellerowicz EJ, Kucukoglu M, Johnsson C, et al. AspWood: high-spatial-resolution transcriptome profiles reveal uncharacterized modularity of wood formation in *Populus tremula*. *Plant Cell.* 2017;29(7):1585–604. <https://doi.org/10.1105/tpc.17.00153>.
- Chen H-C, Song J, Wang JP, Lin Y-C, Ducoste J, Shuford CM, et al. Systems biology of lignin biosynthesis in *Populus trichocarpa*: heteromeric 4-coumaric acid: coenzyme A ligase protein complex formation, regulation, and numerical modeling. *Plant Cell.* 2014;26(3):876–93. <https://doi.org/10.1105/tpc.113.119685>.
- Goué N, Lesage-Descauses MC, Mellerowicz EJ, Magel E, Label P, Sundberg B. Microgenomic analysis reveals cell type-specific gene expression patterns between ray and fusiform initials within the cambial meristem of *Populus*. *New Phytol.* 2008;180(1):45–56. <https://doi.org/10.1111/j.1469-8137.2008.02556.x>.
- Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods.* 2014;11(2):163–6. <https://doi.org/10.1038/nmeth.2772>.
- Macosko EZ, Basu A, Satija R, Nemes J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell.* 2015;161(5):1202–14. <https://doi.org/10.1016/j.cell.2015.05.002>.
- Seyfferth C, Renema J, Wendrich JR, Eekhout T, Seurinck R, Vandamme N, et al. Advances and opportunities in single-cell transcriptomics for plant research. *Annu Rev Plant Biol.* 2021;72(1):847–66. <https://doi.org/10.1146/annurev-arplant-081720-010120>.
- Shaw R, Tian X, Xu J. Single-cell transcriptome analysis in plants: advances and challenges. *Mol Plant.* 2020;14(1):115–26. <https://doi.org/10.1016/j.molp.2020.10.012>.
- Liu Q, Liang Z, Feng D, Jiang S, Wang Y, Du Z, et al. Transcriptional landscape of rice roots at the single-cell resolution. *Mol Plant.* 2021;14(3):384–94. <https://doi.org/10.1016/j.molp.2020.12.014>.
- Zhang T-Q, Xu Z-G, Shang G-D, Wang J-W. A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root. *Mol Plant.* 2019;12(5):648–60. <https://doi.org/10.1016/j.molp.2019.04.004>.
- Satterlee JW, Strable J, Scanlon MJ. Plant stem-cell organization and differentiation at single-cell resolution. *PNAS.* 2020;117(52):33689–99. <https://doi.org/10.1073/pnas.2018788117>.
- Tian C, Du Q, Xu M, Du F, Jiao Y. Single-nucleus RNA-seq resolves spatiotemporal developmental trajectories in the tomato shoot apex. *BioRxiv.* 2020.
- Zhang T-Q, Chen Y, Wang J-W. A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Dev Cell.* 2021;56(7):1056–74. <https://doi.org/10.1016/j.devcel.2021.02.021>.
- Kim J-Y, Symeonidi E, Pang TY, Denyer T, Weidauer D, Bezruczyk M, et al. Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *Plant Cell.* 2021;33(3):511–30. <https://doi.org/10.1093/plcell/koaa060>.

28. Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, et al. Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Dev Cell*. 2021;56(4):557–68. <https://doi.org/10.1016/j.devcel.2020.12.015>.
29. Nelms B, Walbot V. Defining the developmental program leading to meiosis in maize. *Science*. 2019;364(6435):52–6. <https://doi.org/10.1126/science.aav6428>.
30. Shulze CN, Cole BJ, Ciobanu D, Lin J, Yoshinaga Y, Gouran M, et al. High-throughput single-cell transcriptome profiling of plant cell types. *Cell reports*. 2019;27(7):2241–7. <https://doi.org/10.1016/j.celrep.2019.04.054>.
31. Denyer T, Ma X, Klesen S, Scacchi E, Nieselt K, Timmermans MC. Spatiotemporal developmental trajectories in the *Arabidopsis* root revealed using high-throughput single-cell RNA sequencing. *Dev Cell*. 2019;48(6):840–52. <https://doi.org/10.1016/j.devcel.2019.02.022>.
32. Ryu KH, Huang L, Kang HM, Schiefelbein J. Single-cell RNA sequencing resolves molecular relationships among individual plant cells. *Plant Physiol*. 2019;179(4):1444–56. <https://doi.org/10.1104/pp.18.01482>.
33. Liu Z, Zhou Y, Guo J, Li J, Tian Z, Zhu Z, et al. Global dynamic molecular profiling of stomatal lineage cell development by single-cell RNA sequencing. *Mol Plant*. 2020;13(8):1178–93. <https://doi.org/10.1016/j.molp.2020.06.010>.
34. Lopez-Anido CB, Vatén A, Smoot NK, Sharma N, Guo V, Gong Y, et al. Single-cell resolution of lineage trajectories in the *Arabidopsis* stomatal lineage and developing leaf. *Dev Cell*. 2021;56(7):1043–55. <https://doi.org/10.1016/j.devcel.2021.03.014>.
35. Liu H, Hu D, Du P, Wang L, Liang X, Li H, et al. Single-cell RNA-seq describes the transcriptome landscape and identifies critical transcription factors in the leaf blade of the allotetraploid peanut (*Arachis hypogaea* L.). *Plant Biotechnol J*. 2021;19(11):2261–76.
36. Gala HP, Lanctot A, Jean-Baptiste K, Guizou S, Chu JC, Zemke JE, et al. A single-cell view of the transcriptome during lateral root initiation in *Arabidopsis thaliana*. *Plant Cell*. 2021;33(7):2197–220. <https://doi.org/10.1093/plcell/koab101>.
37. Li H, Dai X, Huang X, Xu M, Wang Q, Yan X, et al. Single-cell RNA sequencing reveals a high-resolution cell atlas of xylem in *Populus*. *J Integr Plant Biol*. 2021;00(11):1–16. <https://doi.org/10.1111/jipb.13159>.
38. Lin Y-C, Li W, Chen H, Li Q, Sun Y-H, Shi R, et al. A simple improved-throughput xylem protoplast system for studying wood formation. *Nat Protoc*. 2014;9(9):2194–205. <https://doi.org/10.1038/nprot.2014.147>.
39. Wu A-M, Hörnblad E, Voxeur A, Gerber L, Rihouey C, Lerouge P, et al. Analysis of the *Arabidopsis* *IRX9/IRX9-L* and *IRX14/IRX14-L* pairs of glycosyltransferase genes reveals critical contributions to biosynthesis of the hemicellulose glucuronoxylan. *Plant Physiol*. 2010;153(2):542–54. <https://doi.org/10.1104/pp.110.154971>.
40. Lee C, O'Neill MA, Tsumuraya Y, Darvill AG, Ye Z-H. The *irregular xylem9* mutant is deficient in xylan xylosyltransferase activity. *Plant Cell Physiol*. 2007;48(11):1624–34. <https://doi.org/10.1093/pcp/pcm135>.
41. Lasserre E, Jobet E, Llauro C, Delseny M. *AtERF38* (*At2g35700*), an AP2/ERF family transcription factor gene from *Arabidopsis thaliana*, is expressed in specific cell types of roots, stems and seeds that undergo suberization. *Plant Physiol Biochem*. 2008;46(12):1051–61. <https://doi.org/10.1016/j.plaphy.2008.07.003>.
42. Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Cézard L, Le Bris P, et al. Disruption of *LACCASE4* and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell*. 2011;23(3):1124–37. <https://doi.org/10.1105/tpc.110.082792>.
43. Li E, Bhargava A, Qiang W, Friedmann MC, Forneris N, Savidge RA, et al. The Class II *KNOX* gene *KNAT7* negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in *Populus*. *New Phytol*. 2012;194(1):102–15. <https://doi.org/10.1111/j.1469-8137.2011.04016.x>.
44. Tan TT, Endo H, Sano R, Kurata T, Yamaguchi M, Ohtani M, et al. Transcription factors VND1-VND3 contribute to cotyledon xylem vessel formation. *Plant Physiol*. 2018;176(1):773–89. <https://doi.org/10.1104/pp.17.00461>.
45. Feuillet C, Lauergeat V, Deswarte C, Pilate G, Boudet A, Grima-Pettenati J. Tissue- and cell-specific expression of a cinnamyl alcohol dehydrogenase promoter in transgenic poplar plants. *Plant Mol Biol*. 1995;27(4):651–67. <https://doi.org/10.1007/BF00020220>.
46. Bonke M, Thitamadee S, Mähönen AP, Hauser M-T, Helariutta Y. APL regulates vascular tissue identity in *Arabidopsis*. *Nature*. 2003;426(6963):181–6. <https://doi.org/10.1038/nature02100>.
47. Dinant S, Clark AM, Zhu Y, Vilaine F, Palauqui J-C, Kusiak C, et al. Diversity of the superfamily of phloem lectins (phloem protein 2) in angiosperms. *Plant Physiol*. 2003;131(1):114–28. <https://doi.org/10.1104/pp.013086>.
48. Sauer N, Stolz J. SUC1 and SUC2: two sucrose transporters from *Arabidopsis thaliana*; expression and characterization in baker's yeast and identification of the histidine-tagged protein. *Plant J*. 1994;6(1):67–77. <https://doi.org/10.1046/j.1365-3113.1994.6010067.x>.
49. Pelissier HC, Peters WS, Collier R, Bel AJv, Knoblauch M: GFP tagging of sieve element occlusion (SEO) proteins results in green fluorescent forisomes. *Plant Cell Physiol*. 2008;49(11):1699–710. <https://doi.org/10.1093/pcp/pcn141>.
50. Chardon F, Bedu M, Calenge F, Klemens PA, Spinner L, Clement G, et al. Leaf fructose content is controlled by the vacuolar transporter SWEET17 in *Arabidopsis*. *Curr Biol*. 2013;23(8):697–702. <https://doi.org/10.1016/j.cub.2013.03.021>.
51. Long Y, Goedhart J, Schneijderberg M, Terpstra I, Shimotohno A, Bouchet BP, et al. SCARECROW-LIKE 23 and SCARECROW jointly specify endodermal cell fate but distinctly control SHORT-ROOT movement. *Plant J*. 2015;84(4):773–84. <https://doi.org/10.1111/tbj.13038>.
52. Hassan H, Scheres B, Blilou I. JACKDAW controls epidermal patterning in the *Arabidopsis* root meristem through a non-cell-autonomous mechanism. *Development*. 2010;137(9):1523–9. <https://doi.org/10.1242/dev.048777>.
53. David LC, Berquin P, Kanno Y, Seo M, Daniel-Vedele F, Ferrario-Méry S. N availability modulates the role of NPF3. 1, a gibberellin transporter, in GA-mediated phenotypes in *Arabidopsis*. *Planta*. 2016;244(6):1315–28. <https://doi.org/10.1007/s00425-016-2588-1>.
54. Sozzani R, Cui H, Moreno-Risueno M, Busch W, Van Norman J, Vernoux T, et al. Spatiotemporal regulation of cell-cycle genes by SHORTROOT links patterning and growth. *Nature*. 2010;466(7302):128–32. <https://doi.org/10.1038/nature09143>.
55. Geldner N. The endodermis. *Annu Rev Plant Biol*. 2013;64(1):531–58. <https://doi.org/10.1146/annurev-arplant-050312-120050>.
56. Molina I, Li-Beisson Y, Beisson F, Ohlrogge JB, Pollard M. Identification of an *Arabidopsis* feruloyl-coenzyme A transferase required for suberin synthesis. *Plant Physiol*. 2009;151(3):1317–28. <https://doi.org/10.1104/pp.109.144907>.
57. Naseer S, Lee Y, Lapierre C, Franke R, Nawrath C, Geldner N. Casparian strip diffusion barrier in *Arabidopsis* is made of a lignin polymer without suberin. *PNAS*. 2012;109(25):10101–6. <https://doi.org/10.1073/pnas.1205726109>.

58. Joubès J, Raffaele S, Bourdenx B, Garcia C, Laroche-Traineau J, Moreau P, et al. The VLCFA elongase gene family in *Arabidopsis thaliana*: phylogenetic analysis, 3D modelling and expression profiling. *Plant Mol Biol*. 2008;67(5):547–66. <https://doi.org/10.1007/s11103-008-9339-z>.
59. Lü S, Song T, Kosma DK, Parsons EP, Rowland O, Jenks MA. *Arabidopsis CER8* encodes LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1) that has overlapping functions with LACS2 in plant wax and cutin synthesis. *Plant J*. 2009;59(4):553–64. <https://doi.org/10.1111/j.1365-3113.2009.03892.x>.
60. Masucci JD, Rerie WG, Foreman DR, Zhang M, Galway ME, Marks MD, et al. The homeobox gene *GLABRA2* is required for position-dependent cell differentiation in the root epidermis of *Arabidopsis thaliana*. *Development*. 1996;122(4):1253–60. <https://doi.org/10.1242/dev.1224.1253>.
61. Fischer U, Kucukoglu M, Helariutta Y, Bhalerao RP. The dynamics of cambial stem cell activity. *Annu Rev Plant Biol*. 2019;70(1):293–319. <https://doi.org/10.1146/annurev-arplant-050718-100402>.
62. Kucukoglu M, Chaabouni S, Zheng B, Mähönen AP, Helariutta Y, Nilsson O. Peptide encoding *Populus CLV3/ESR-RELATED 47* (*PtCLE47*) promotes cambial development and secondary xylem formation in hybrid aspen. *New Phytol*. 2020;226(1):75–85. <https://doi.org/10.1111/nph.16331>.
63. Zhu Y, Song D, Xu P, Sun J, Li L. A *HD-ZIP III* gene, *PtRHB4*, is required for interfascicular cambium development in *Populus*. *Plant Biotechnol J*. 2018;16(3):808–17. <https://doi.org/10.1111/pbi.12830>.
64. Zhu Y, Song D, Sun J, Wang X, Li L. *PtRHB7*, a class III HD-Zip gene, plays a critical role in regulation of vascular cambium differentiation in *Populus*. *Mol Plant*. 2013;6(4):1331–43. <https://doi.org/10.1093/mp/sss164>.
65. Breda AS, Hazak O, Hardtke CS. Phosphosite charge rather than shootward localization determines OCTOPUS activity in root protophloem. *PNAS*. 2017;114(28):E5721–30. <https://doi.org/10.1073/pnas.1703258114>.
66. Wallner ES, Tonn N, Shi D, Jouannet V, Greb T. *SUPPRESSOR OF MAX2 1-LIKE 5* promotes secondary phloem formation during radial stem growth. *Plant J*. 2020;102(5):903–15. <https://doi.org/10.1111/tpj.14670>.
67. Gu F, Bringmann M, Combs JR, Yang J, Bergmann DC, Nielsen E. *Arabidopsis* CSLD5 functions in cell plate formation in a cell cycle-dependent manner. *Plant Cell*. 2016;28(7):1722–37. <https://doi.org/10.1105/tpc.16.00203>.
68. Müller S, Fuchs E, Ovecka M, Wysocka-Diller J, Benfey PN, Hauser M-T. Two new loci, *PLEIADE* and *HYADE*, implicate organ-specific regulation of cytokinesis in *Arabidopsis*. *Plant Physiol*. 2002;130(1):312–24. <https://doi.org/10.1104/pp.004416>.
69. Strompen G, El Kasmi F, Richter S, Lukowitz W, Assaad FF, Jürgens G, et al. The *Arabidopsis* *HINKEL* gene encodes a kinesin-related protein involved in cytokinesis and is expressed in a cell cycle-dependent manner. *Curr Biol*. 2002;12(2):153–8. [https://doi.org/10.1016/S0960-9822\(01\)00655-8](https://doi.org/10.1016/S0960-9822(01)00655-8).
70. Boruc J, Mylle E, Duda M, De Clercq R, Rombauts S, Geelen D, et al. Systematic localization of the *Arabidopsis* core cell cycle proteins reveals novel cell division complexes. *Plant Physiol*. 2010;152(2):553–65. <https://doi.org/10.1104/pp.109.148643>.
71. Boudolf V, Rombauts S, Naudts M, Inzé D, De Veylder L. Identification of novel cyclin-dependent kinases interacting with the CKS1 protein of *Arabidopsis*. *J Exp Bot*. 2001;52(359):1381–2. <https://doi.org/10.1093/jexbot/52.359.1381>.
72. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14(10):979–82. <https://doi.org/10.1038/nmeth.4402>.
73. Kleine-Vehn J, Dhonukshe P, Swarup R, Bennett M, Friml J. Subcellular trafficking of the *Arabidopsis* auxin influx carrier AUX1 uses a novel pathway distinct from PIN1. *Plant Cell*. 2006;18(11):3171–81. <https://doi.org/10.1105/tpc.106.042770>.
74. Liu L, Liu C, Hou X, Xi W, Shen L, Tao Z, et al. FTIP1 is an essential regulator required for florigen transport. *PLoS Biol*. 2012;10(4):e1001313. <https://doi.org/10.1371/journal.pbio.1001313>.
75. Yao R, Ming Z, Yan L, Li S, Wang F, Ma S, et al. DWARF14 is a non-canonical hormone receptor for strigolactone. *Nature*. 2016;536(7617):469–73. <https://doi.org/10.1038/nature19073>.
76. Wang L, Wang B, Yu H, Guo H, Lin T, Kou L, et al. Transcriptional regulation of strigolactone signalling in *Arabidopsis*. *Nature*. 2020;583(7815):277–81. <https://doi.org/10.1038/s41586-020-2382-x>.
77. Barratt DP, Kölling K, Graf A, Pike M, Calder G, Findlay K, et al. Callose synthase GSL7 is necessary for normal phloem transport and inflorescence growth in *Arabidopsis*. *Plant Physiol*. 2011;155(1):328–41. <https://doi.org/10.1104/pp.110.1.66330>.
78. Sjolund RD. The phloem sieve element: a river runs through it. *Plant Cell*. 1997;9(7):1137–46. <https://doi.org/10.1105/tpc.9.7.1137>.
79. Khan JA, Wang Q, Sjolund RD, Schulz A, Thompson GA. An early nodulin-like protein accumulates in the sieve element plasma membrane of *Arabidopsis*. *Plant Physiol*. 2007;143(4):1576–89. <https://doi.org/10.1104/pp.106.092296>.
80. Song, N. Characterization of the *Arabidopsis* sieve element-specific early nodulin protein. Master's thesis. Oklahoma State University, Department of Biochemistry and Molecular Biology; 2010.
81. Dettmer J, Ursache R, Campilho A, Miyashima S, Belevich I, O'Regan S, et al. CHOLINE TRANSPORTER-LIKE1 is required for sieve plate development to mediate long-distance cell-to-cell communication. *Nat Commun*. 2014;5(1):1–11. <https://doi.org/10.1038/ncomms5276>.
82. Milhinhos A, Prestele J, Bollhöner B, Matos A, Vera-Sirera F, Rambla JL, et al. Thermospermine levels are controlled by an auxin-dependent feedback loop mechanism in *Populus* xylem. *Plant J*. 2013;75(4):685–98. <https://doi.org/10.1111/tpj.12231>.
83. Morris H, Jansen S. Secondary xylem parenchyma—from classical terminology to functional traits. *Lawa Journal*. 2016;37(1):1–15. <https://doi.org/10.1163/22941932-20160117>.
84. Secchi F, Pagliarani C, Zwieniecki MA. The functional role of xylem parenchyma cells and aquaporins during recovery from severe water stress. *Plant Cell Environ*. 2017;40(6):858–71. <https://doi.org/10.1111/pce.12831>.
85. Yuan Y, Teng Q, Zhong R, Ye Z-H. The *Arabidopsis* DUF231 domain-containing protein ESK1 mediates 2-O- and 3-O-acetylation of xylosyl residues in xylan. *Plant Cell Physiol*. 2013;54(7):1186–99. <https://doi.org/10.1093/pcp/pct070>.
86. Cao S, Guo M, Wang C, Xu W, Shi T, Tong G, et al. 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):1–17. <https://doi.org/10.1186/s12870-019-1865-0>.
87. Song D, Sun J, Li L. Diverse roles of *PtDUF579* proteins in *Populus* and *PtDUF579-1* function in vascular cambium proliferation during secondary growth. *Plant Mol Biol*. 2014;85(6):601–12. <https://doi.org/10.1007/s11103-014-0206-9>.

88. Gui J, Lam PY, Tobimatsu Y, Sun J, Huang C, Cao S, et al. Fibre-specific regulation of lignin biosynthesis improves biomass quality in *Populus*. *New Phytol.* 2020;226(4):1074–87. <https://doi.org/10.1111/nph.16411>.
89. Zhao Y, Song D, Sun J, Li L. *Populus* endo-beta-mannanase *PtrMAN6* plays a role in coordinating cell wall remodeling with suppression of secondary wall thickening through generation of oligosaccharide signals. *Plant J.* 2013;74(3):473–85. <https://doi.org/10.1111/tpj.12137>.
90. 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(1):84–94. <https://doi.org/10.1104/pp.010514>.
91. Ma T, Wang J, Zhou G, Yue Z, Hu Q, Chen Y, et al. Genomic insights into salt adaptation in a desert poplar. *Nat Commun.* 2013;4(1):1–9. <https://doi.org/10.1038/ncomms3797>.
92. Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science.* 2006;313(5793):1596–604. <https://doi.org/10.1126/science.1128691>.
93. Blanc G, Wolfe KH. Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell.* 2004;16(7):1679–91. <https://doi.org/10.1105/tpc.021410>.
94. Xu W, Cheng H, Zhu S, Cheng J, Ji H, Zhang B, et al. Functional understanding of secondary cell wall cellulose synthases in *Populus trichocarpa* via the Cas9/gRNA-induced gene knockouts. *New Phytol.* 2021;231(4):1478–95. <https://doi.org/10.1111/nph.17338>.
95. Abbas M, Peszlen I, Shi R, Kim H, Katahira R, Kafle K, et al. Involvement of *CesA4*, *CesA7-A/B* and *CesA8-A/B* in secondary wall formation in *Populus trichocarpa* wood. *Tree Physiol.* 2020;40(1):73–89. <https://doi.org/10.1093/treephys/tpz020>.
96. 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(3):372–83. <https://doi.org/10.1016/j.molp.2021.01.001>.
97. Ding J, Adiconis X, Simmons SK, Kowalczyk MS, Hession CC, Marjanovic ND, et al. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nat Biotechnol.* 2020;38(6):737–46. <https://doi.org/10.1038/s41587-020-00465-8>.
98. Grindberg RV, Yee-Greenbaum JL, McConnell MJ, Novotny M, O'Shaughnessy AL, Lambert GM, et al. RNA-sequencing from single nuclei. *PNAS.* 2013;110(49):19802–7. <https://doi.org/10.1073/pnas.1319700110>.
99. Zhang L, Zhao J, Bi H, Yang X, Zhang Z, Su Y, et al. Bioinformatic analysis of chromatin organization and biased expression of duplicated genes between two poplars with a common whole-genome duplication. *Hortic Res.* 2021;8(1):1–12. <https://doi.org/10.1038/s41438-021-00494-2>.
100. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* 2015;33(5):495–502. <https://doi.org/10.1038/nbt.3192>.
101. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell systems.* 2019;8(4):329–37 e324.
102. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* 2018;36(5):411–20. <https://doi.org/10.1038/nbt.4096>.
103. Haghverdi L, Lun AT, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol.* 2018;36(5):421–7. <https://doi.org/10.1038/nbt.4091>.
104. Gutierrez C. The *Arabidopsis* cell division cycle. *Arabidopsis* Book, 2009:e0120, 7. <https://doi.org/10.1199/tab.0120>.
105. Yadav RK, Girke T, Pasala S, Xie M, Reddy GV. Gene expression map of the *Arabidopsis* shoot apical meristem stem cell niche. *PNAS.* 2009;106(12):4941–6. <https://doi.org/10.1073/pnas.0900843106>.
106. Villani A-C, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science.* 2017;356(6335):eaah4573.
107. Van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* 2008;9(11):2579–605.
108. Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IW, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol.* 2019;37(1):38–44. <https://doi.org/10.1038/nbt.4314>.
109. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357–60. <https://doi.org/10.1038/nmeth.3317>.
110. Perteu M, Perteu GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol.* 2015;33(3):290–5. <https://doi.org/10.1038/nbt.3122>.
111. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-H, Jin H, Marler B, Guo H. MCSScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012;40(7):e49–e49. <https://doi.org/10.1093/nar/gkr1293>.
112. Li G-S, Meng Z, Kong H-Z, Chen Z-D, Theissen G, Lu A-M. Characterization of candidate class A, B and E floral homeotic genes from the perianthless basal angiosperm *Chloranthus spicatus* (Chloranthaceae). *Dev Genes Evol.* 2005;215(9):437–49. <https://doi.org/10.1007/s00427-005-0002-2>.
113. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell.* 2010;38(4):576–89. <https://doi.org/10.1016/j.molcel.2010.05.004>.
114. Kohl M, Wiese S, Warscheid B. Cytoscape: software for visualization and analysis of biological networks. *Methods Mol Biol.* 2011;696:291–303. [https://doi.org/10.1007/978-1-60761-987-1\\_18](https://doi.org/10.1007/978-1-60761-987-1_18).
115. Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Bioinformatics.* 1997;13(5):555–6. <https://doi.org/10.1093/bioinformatics/13.5.555>.
116. Chen Y, Tong S, Jiang Y, Ai F, Feng Y, Zhang J, Gong J, Qin J, Zhang Y, Zhu Y, Liu J, Ma T. Transcriptional landscape of highly lignified poplar stems at single-cell resolution. *NGDC. BioProject.* <https://ngdc.cnbc.ac.cn/bioproject/browse/PRJCA005543> (2021).

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.